Probing the binding of interstrand crosslinked DNA containing a fluorescent base analog with human and *E. coli* homologs of O^6-alkyl-guanine-DNA-alkyltransferase

Jordan Vergara

A Thesis

In

The Department

Of

Chemistry and Biochemistry

Presented in partial fulfillment of the requirements

For the degree of Master of Science at

Concordia University

Montreal, Quebec, Canada

October 2012

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: **Jordan Vergara**

Entitled: **Probing the binding of interstrand crosslinked DNA containing a fluorescent analog with human and *E. coli* homologs of O^6-alkylguanine-DNA-alkyltransferase**

and submitted in partial fulfillment of the requirements for the degree of

Master of Science

complies with the regulations of the University and m
eets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair
Dr. Joanne Turnbull Examiner
Dr. Judith Kornblatt Examiner
Dr. Christopher Wilds Supervisor

Approved by Dr. Heidi Muchall
Chair of Department or Graduate Program Director
Dr. Brian Lewis
Dean of Faculty

Date October 12, 2012

Abstract

O^6-alkylguanine-DNA-alkyltransferase (AGT) has been implicated in reducing the therapeutic efficacy of alkylating agents that cause mutations and apoptosis. Certain chemotherapeutic treatments involve the use of alkylating agents; *bis*-chloroethylnitrosourea (BCNU) is an example of an agent that has been investigated which has been shown to introduce DNA interstrand crosslinks (ICL). AGT directly repairs O^6-alkylated guanines by flipping the damaged base into the active site and irreversibly transferring the alkyl group to an active site cysteine. Previous studies revealed that human AGT could repair DNA containing an O^6-2'-deoxyguanosine-heptylene-O^6-2'-deoxyguanosine (O^6-dG-heptylene-O^6-dG) ICL in a 5'-GNC-3' motif mimicking a hepsulfam lesion. It is not yet understood, however, how the enzyme could access the bulky ICL lesion.

We synthesized and studied DNA duplexes containing O^6-dG-heptylene-O^6-dG ICL lesion in a 5'-GNC-3' motif (where N is any base) and a fluorescent base incorporated at various positions to observe the extent to which DNA is denatured upon AGT binding. 6-Methyl-3-(β-D-2-deoxyribofuranosyl)pyrrolo[2,3-d] pyrimidin-2-one (pyrrolo-dC), is a base analog that fluoresces when not base-paired and exhibits little effect on DNA stability, structure and AGT repair. All DNA substrates were synthesized and AGT homologs purified and characterized using various biophysical techniques including CD, UV spectroscopy, gel electrophoresis and MS-Q-TOF. Radioactivity binding and repair assays were used to characterize ICL-AGT association. The DNA substrates exhibited B-form structure regardless of pyrrolo-dC or ICL incorporation. Furthermore, C145S and R128A variants of AGT and wild type Ada-C (from *E. coli*)

display no repair of the ICL DNA. The arginine "finger" (R128) appears to play an essential role in repair but not in damage detection. Incorporation of pyrrolo-dC did not specifically demonstrate the extent of DNA denaturation due to the observation of increased fluorescence when positioned near the end of the DNA double helix relative to when placed in the center of the duplex. Fluorescence studies provide a base with which to develop new AGT variants with more complex ICL DNA substrates to study their interaction and possibly provide a new method for determining binding dissociation constants without hazardous radioactive material. Better understanding of the interaction of AGT and its homologs with ICL DNA provides greater knowledge about AGT function, an enzyme that plays a role in both restoring genomic integrity and therapeutic resistance.

Acknowledgments

First and foremost I would like to thank Dr. Christopher Wilds for giving me the opportunity to work in his lab, his patience, guidance and support. As well as the hard work he put into making the crosslinked amidites. I would also like to thank my committee members Dr. Joanne Turnbull and Dr. Judith Kornblatt for their support, valuable suggestions and their help.

I would like to thank Dr. Anne Noronha for her support and hard work in synthesizing the oligonucleotides, as well as training on HPLC techniques. She provided valuable insight into DNA preparation and purification. I would like to thank Francis McManus for his training in several biochemistry techniques, always willing to share his experience and useful insight. I would like to thank Gang Sun for his help with the chemistry of de-silylation and de-alloc reactions. I would like to thank all the lab members as a whole for their support and making a pleasant working environment.

Finally, I would like to thank my mother for her support and love over the years that have made me who I am today and helped me to overcome any obstacle life had to send my way.

Table of Contents

Abbreviations

ACN: Acetonitrile

AGT: O^6-alkylguanine-DNA-alkyltransferase

Alloc: allyloxycarbonyl

AMP: Ampicillin

APS: Ammonium persulphate

BG: Benzyl guanine

CD: Circular dichroism

CIAP: Calf intestinal alkaline phosphatase

CPG: Controlled pore glass

C145S: Cysteine 145 to serine

dC: 2'-deoxycytidine

dG: 2'-deoxyguanosine

Da: Daltons

DCM: Dichloromethane

DMT: Dimethoxytrityl

DNA: Deoxyribonucleic acid

DTT: Dithiothreitol

dsDNA: Double stranded deoxyribonucleic acid

EDTA: Ethylenediamine tetraacetate

eq.: equivalence

HT29 cells: Human colon tumour cells

ICL: Interstrand crosslink

IPTG: Isopropyl β-D-1-thiogalactopyranoside

LB: Luria broth

Lcaa: Long chain alkylamine

OD: Optical density

PCHRV: Proline-Cysteine-Histidine-Arginine-Valine

PCR: Polymerase chain reaction

QTOF: Quadrupole mass analyzer - time-of-flight

R128A: Arginine 128 to alanine

SDS-PAGE: Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SPS: Solid-phase synthesis

ssDNA: Single stranded deoxyribonucleic acid

SVP: Snake venom phosphodiesterase I

TBS: *tert*-butyldimethylsilyl

TCA: Trichloroacetic acid

TEA: Triethylamine

TEMED: N,N,N,N-tetramethylethylenediamine

THF: Tetrahydrofuran

wt: Wild type

List of figures

List of tables

Chapter 1: Introduction

1.1 General background

Organisms are constantly under threat of damaging events that undermine their ability to function, one of the major targets being deoxyribonucleic acid (DNA). Genomic DNA contains all the genetic material required for cell development and is the template for transcription of mRNA, which is required for the synthesis of various proteins that perform essential biological functions. DNA is vulnerable to a variety of damaging events: UV radiation, chemicals and various endogenous and exogenous species[1, 2]. Since Muller's discovery of the mutagenic effects of X-rays in 1927 and subsequent interest in the effects of mutagenic chemicals, it was not until the 1940s that repair mechanisms were beginning to be understood[3]. Ever since, researchers have been interested in the study of the effects of damaging events on cell survival and their ability to reverse it. This has led to the discovery and in-depth study of repair systems over the last 70 years[4, 5].

In order to maintain cell integrity, various DNA repair pathways exist. These involve proteins whose specific function is to repair DNA acting on damaged or modified bases; one such class of proteins is the alkyltransferases, such as O^6-alkylguanine-DNA-alkyltransferase (AGT). AGT repairs the alkylated O^6 position of guanine reversing the effects of certain alkylating agents, reducing the incidence of mutation and apoptosis[6-55].

Greater knowledge of cellular repair machinery allows for increased comprehension of how the cell functions and improves our abilities to stave off and combat cancers and diseases. More specifically, a better understanding of the interaction of AGT and its homologs with various forms of alkylated DNA will provide greater

knowledge about AGT function, an enzyme that plays a role in both restoring genomic integrity and therapeutic resistance.

1.2 DNA and DNA damaging events

Although considered more chemically stable than proteins and RNA, DNA is subject to spontaneous depurination and chemical alteration by oxidative species, UV radiation and alkylating agents[1, 2]. These DNA damaging events may cause double strand DNA breaks, depurination, oxidized bases (*eg*: 8-oxo-2'-deoxyguanosine), thymine-thymine cyclobutane pyrimidine photodimers, alkylated bases (*eg*: O^6-methyl-deoxyguanosine) and interstrand crosslinks to name a few[1, 2, 11, 56, 57]. Due to the range of damage that can be caused, researchers have realized the potential of employing synthetic alkylating agents as a treatment to induce cell death of fast-replicating cancer cells.

1.2.1 Alkylating agents

Alkylating agents are chemically electrophilic and target the nucleophilic sites of DNA. DNA has multiple nucleophilic sites but not all sites are equally susceptible to alkylation (see **Table 1**). Using mono-functional alkylating chemicals, methyl methanesulphonate (MMS) and N-methyl-N-nitrosourea (MNU), researchers were able to observe the sites of DNA that are most prone to alkylation. They observed that the most nucleophilic sites appear to be N3 of adenine, O^6 and N7 of guanine and the non-bridging oxygen atoms of the phosphodiester group. Mono-functional alkylating agents can only affect one nucleophilic site, whereas a bi-functional agent can affect two sites and may form crosslinks, such as an N7-heptylene-dG-N7-heptylene-dG interstrand linkage of hepsulfam.

Site of methylation	MMS + dsDNA	MMU + dsDNA
Adenine		
N1-	3.8	1.3
N3-	10.4	9
N7-	1.8	1.7
Guanine		
N3-	0.6	0.8
O^6-	0.3	6.3
N7-	83	67
Uracil/thymine		
O^2-	ND	0.11
N3-	ND	0.3
O^4-	ND	0.4
Cytosine		
O^2-	—	0.1
N3-	<1	0.6
Phosphodiester backbone	0.8	17

Table 1 Methylation patterns of double stranded nucleic acids upon reaction with MMS and MNU. Adapted from Drabløs F *et al* 2004[11]. The values herein are expressed as a percent of the total observed methylation. ND means not determined.

There exist both endogenous and exogenous alkylating species, some naturally occurring while some are synthesized as a therapeutic agent. A well characterized endogenous species, S-adenosylmethionine, has been implicated in minor non-enzymatic methylation of DNA. Predominant sites of methylation are at the sites N7 of guanine and N3 of adenine[11, 58]. In addition, it has been observed that starving *E. coli* cells, deficient of alkyltransferase genes, were more predisposed to spontaneous mutations (G:C to A:T transitions) possibly due to the presence of O^6-methyl-dG. O^6-methyl-dG damage can induce altered amino acid sequences upon transcription leading to mutant proteins that may be rendered non-functional or over-active, which could lead to apoptosis and cancer[70]. Furthermore, studies showed that addition of sodium nitrite enabled the nitrosation of various chemical species (*eg*: amides), by bacterial enzymes only, producing endogenous methylation that caused an increase in mutations and was cytotoxic[59].

There are naturally occurring exogenous alkylating agents such as halocarbons[11] and by-products of combustion of plant materials, such as acrolein and crotonaldehyde[56]. Various synthetic agents are employed or have been studied as a therapeutic treatment for cancers. The first generation of therapeutic alkylating agents were mono-functional, adding a methyl group or larger alkyl groups to nucleic acids such as methyl methanesulphonate (MMS), N-methyl-N-nitrosourea (MNU) and temozolomide. Realising the potential of creating interstrand crosslinks and intrastrand crosslinks, bi-functional alkylating agents such as *bis*-chloroethylnitrosourea (BCNU), busulfan and hepsulfam were developed (see **Figure 1** for all chemical structures). ICLs are highly potent as they covalently link each DNA strand, blocking replication and transcription, possibly leading to apoptosis[60]. Although not specifically an alkylating agent, Cisplatin (*cis*-diamminedichloroplatinum (II)), a platinum-based chemotherapeutic treatment, predominantly forms intrastrand crosslinks that distort the DNA double helix, inhibiting transcription and replication much like bi-functional alkylating agents[61, 62].

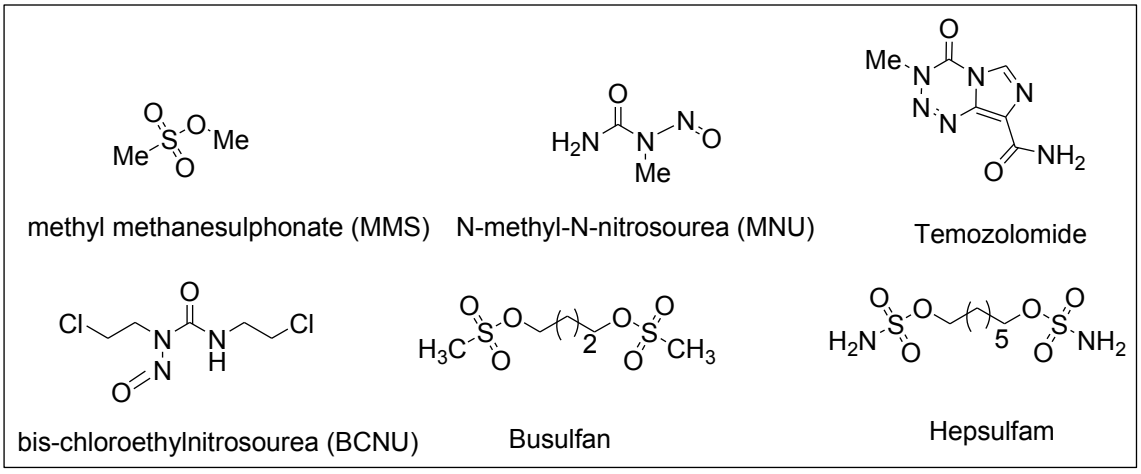


Figure 1 Chemical structures of various alkylating agents. Mono-functional alkylating agents (top) and bi-functional alkylating agents (bottom) are shown.

Despite the high potency of alkylating agents their mutagenic and cytotoxicity effects are significantly decreased in *in vivo* studies. The cell has a variety of defense repair mechanisms that reduce the therapeutic efficiency of alkylating agents. As such, researchers have synthesized a variety of modified DNA containing alkylated bases or ICLs in order to better understand their repair and develop new more effective agents[60, 62-69].

1.3 Cellular DNA repair systems

The cell has developed various DNA repair pathways to maintain integrity and to reverse mutagenesis. Some of the major pathways include: direct repair (DR), base excision repair (BER), mismatch repair (MMR), nucleotide excision repair (NER) and homologous recombination (HR). All of these pathways require complexes of protein, except direct repair, and a complicated signalling system. Interestingly, cells have evolved in such a manner that there is an overlap of repair systems for a given mutagenic DNA damage (see **Figure 2**). For instance, alkylated guanine can be repaired by certain direct repair proteins, NER and BER[2, 11, 34, 41, 60, 70].

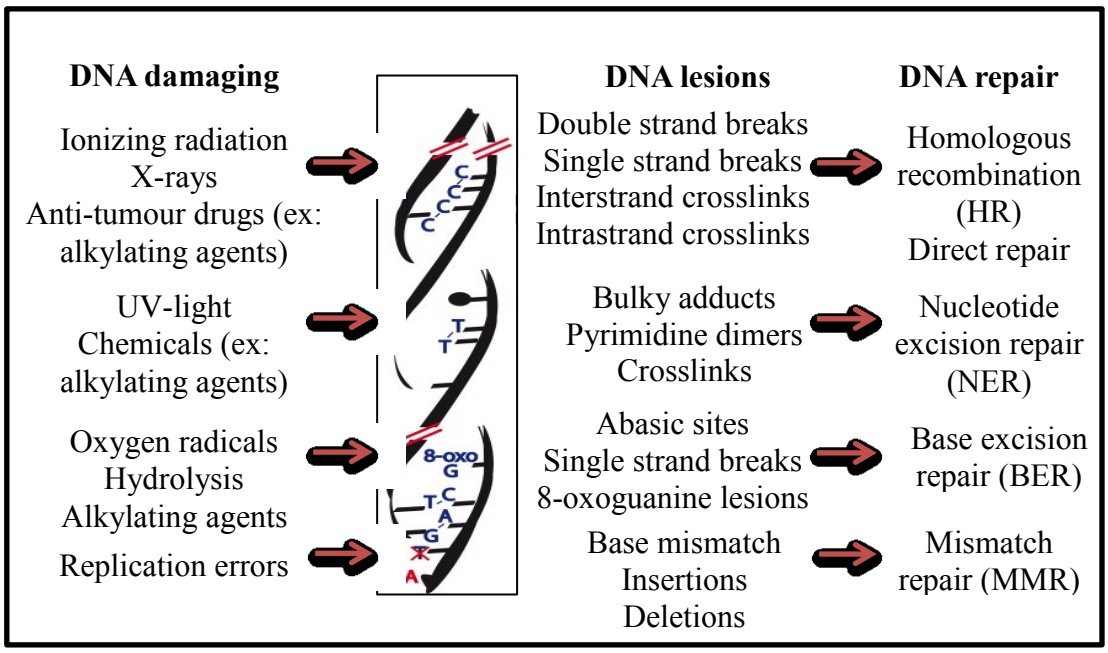


Figure 2 General scheme of DNA damaging events with corresponding damage and cellular repair pathway (modified from Blanpain C *et al* 2011)[70].

1.3.1 Direct Repair

Direct repair is one of the simplest repair pathways, often requiring only one protein. DR involves stoichiometric reversal of damage without significant disruption of the DNA double helix in a one-step mechanism such as alkyltransferases and photolyase. This system enables for fast and simple repair without the recruitment of large complexes.

When DNA is exposed to UV light, thymine-thymine cyclobutane pyrimidine photodimers and pyrimidine-pyrimidone (6-4) photoproduct (Pyr [6-4] Pyr) can form, which are mutagenic and cytotoxic. In order to reverse this damage, some organisms use photolyase. Over 50 homologs have been sequenced, showing sequence identities from 17-70%. Photolyases are monomeric and require two cofactors, a photocatalyst and a photoantenna; FAD and methenyltetrahydrofolate (MTHF) or 8-hydroxy-7,8-didemethyl-5-deazariboflavin (8-HDF) respectively. The FAD cofactor is required for interacting and

breaking the cyclobutane ring. The photoantenna absorbs light energy to transfer to the FAD cofactor. *E. coli* photolyase has been crystallized and characterized, yet there does not appear to be a human homolog. Human cells must use the more complex repair pathway NER to remove photodimer damage[57].

Another major direct repair class of proteins are the alkyltransferases. There are two main types that will be discussed: O^6-alkylguanine-DNA-alkyltransferase (AGT) and AlkB proteins. AGT will be discussed in further detail in the following section. AlkB was first discovered in 1983, when mutation of its gene resulted in an increase in mutagenesis[34]. However, it was only in 2002 that AlkB's mechanism of action was discovered as the removal of alkyl damage through oxidation. More specifically, AlkB utilizes an iron (II), α-ketoglutarate and oxygen complex to catalyze the repair of N1-methyl-adenine and N3-methyl-cytosine. This complex enables AlkB to form a hydroxylated methyl group intermediate that degrades to formaldehyde and an unmodified base (see **Figure 3**). Although AlkB repair results in an unmodified base similar to that of AGT, the mechanism of action is completely dissimilar[34].

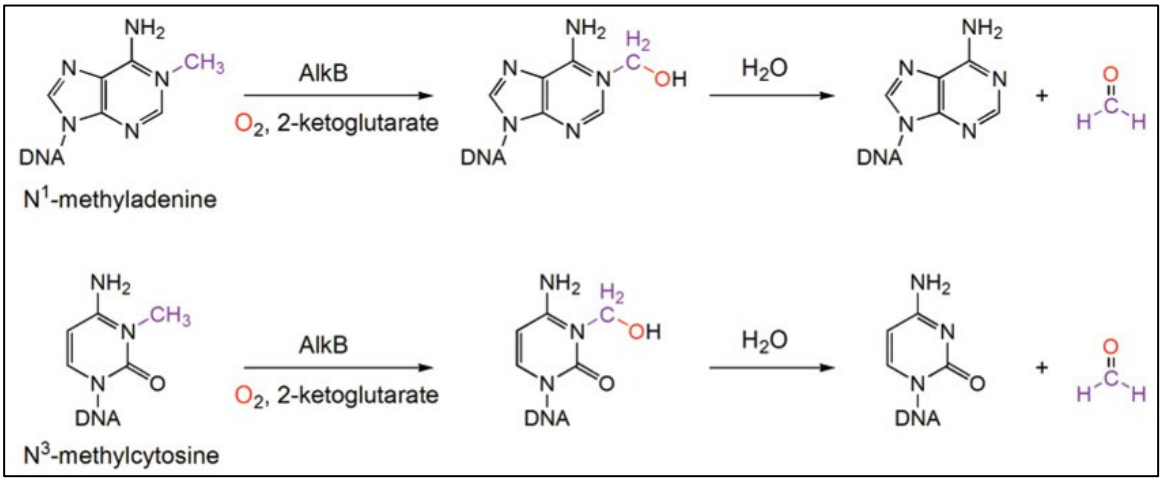


Figure 3 AlkB oxidative dealkylation reaction scheme of N1-methyl-adenine and N3-methyl-cytosine[34].

1.4 AGT background

The first alkyltransferase protein discovered was the inducible *E. coli* Ada protein. It was determined that it could be induced in the presence of mutagenic agents[34]. Ada is composed of two distinct regions, the N- and C-terminal domains. The N-terminal was shown to repair alkylated phophodiester backbone, while the C-terminal reversed alkyl damage at the O^6 position of guanine[19, 39, 49]. The C-terminal domain was isolated and purified (as Ada-C), its repair substrates and mechanism elucidated and crystallized[19, 35, 45, 50]. Ada-C is a monomeric suicidal protein that irreversibly and stoichiometrically (1:1) transfers the alkyl damage to a cysteine in the active site. Although initial reports suggested that the protein had to make a significant conformational change in order to repair the modified DNA, it was later determined that the DNA double helix substrate was distorted upon Ada-C binding[50]. Ada-C appeared to be specific to methyl groups at the O^6 position of guanine, with repair of O^6-ethyl-dG being 1000 times less efficient and almost negligible repair of larger substrates or substituents at the O^4 position of thymine[19].

In addition to Ada-C, a constitutively expressed *E. coli* homolog Ogt was discovered. Ogt has never been crystallized due to poor yields of purification; however, repair studies have been conducted. Importantly, Ogt was not only able to repair O^6-methyl-dG and O^6-ethyl-dG faster than Ada-C, but could also reverse methylation damage at O^4 of thymine[52]. In addition, studies showed Ogt's ability to repair larger groups such as a heptyl group with a either a terminal hydroxyl or a 2'-deoxyguanosine (repaired to a lesser extent)[42].

It was later discovered that humans possessed a protein homologous to both Ogt and Ada-C, namely O^6-alkylguanine-DNA-alkyltransferase (AGT). The AGT human protein is a 22kDa protein found in many organisms with the exception of plants[39]. AGT is of interest as it has been implicated in reducing the therapeutic efficacy of alkylating agents that cause mutations and apoptosis.

1.4.1 AGT structure

The human protein is characterized by two domains; the N- and C-terminal. The C-terminal contains a conserved -Proline-Cysteine-Histidine-Arginine- (-PCHR-) active site, an arginine "finger" (Arg128, see **Figure 4**), and a helix-turn-helix DNA binding motif. The N-terminal contains a zinc binding site. According to crystal structures of AGT with modified DNA (see **Figures 5** and **6**) the Arg128 residue is required to aid in flipping out the modified base into the active site and potentially stabilizing the double helix. The residue Tyr 114 also appears to assist in rotation of the alkylated base by associating with the phosphodiester backbone[8].

```
hAGT_[Homo_sapiens]              ------------PVFQQESFTRQVLWKLLKVVKFGEVISYQQLAALAGNP
AGT_[Canis_lupus_familiaris]     ------------PIFQRGSFTARVLRKLLELVKFGDTVSYQQLAALAGNP
AGT_[Mus_Musculus]_MAETCKMKYSV   ------------PVFQQDSFTRQVLWKLLKVVKFGETVSYQQLAALAGNP
Ada_[E._coli_str._K-12_substr.   ASLNQRDTPLTLPLDIRGTAFQQQVWQALRTIPCGETVSYQQLANAIGKP
Ogt_[E._coli_str._K-12_substr.   ----NLSIIDTLPTATGGTPFQREVWKTLRTIPCGQVMHYGQLAEQLGRP
                                             *     :   : : : *. :  *:.: * ***   *.*

hAGT_[Homo_sapiens]              KAARAVGGAMRGNPVPILIPCHRVVCSSGAVGNYSGG-LAVKEWLLAHEG
AGT_[Canis_lupus_familiaris]     KAARAVGGAMRSNPVPILIPCHRVVCSSGAMGNYTGG-LATKEWLLAHEG
AGT_[Mus_Musculus]_MAETCKMKYSV   KAARAVGGAMRSNPVPILIPCHRVVRSDGAIGHYSGGGQAVKEWLLAHEG
Ada_[E._coli_str._K-12_substr.   KAVRAVASACAANKLAIIIPCHRVVRGDGTLSGYRWG-VSRKAQLLRREA
Ogt_[E._coli_str._K-12_substr.   GAARAVGAANGSNPISIVVPCHRVIGRNGTMTGYAGG-VQRKEWLLRHEG
                                  *.***..*  .* :.*::*****:  .*::  *  *    *  ** :*.

hAGT_[Homo_sapiens]              HRL-----------------------------------
AGT_[Canis_lupus_familiaris]     RLAGKPTYPGGSYLAGTRRGALGGTPSPHPAGRDGCVRE
AGT_[Mus_Musculus]_MAETCKMKYSV   IPTGQPASKGLGLTGTWLKSSFESTSSEPSGRN------
Ada_[E._coli_str._K-12_substr.   ENEER-------------------------------------
Ogt_[E._coli_str._K-12_substr.   YLLL-------------------------------------
```

Figure 4 Sequence alignment of AGTs from various organisms with the active site and the arginine "finger" highlighted in yellow.

Furthermore, AGT binds DNA in the minor groove, where most DNA binding proteins bind the major groove. The binding of zinc appears to favour protein stability and increase its rate of repair[8, 9, 14]. It is also believed that with increasing stability of the protein, specifically the N-terminal, that a hydrophobic patch is formed which may play a role in protein-protein interactions[13]. Although the precise function of the zinc-binding site in the N-terminal of AGT is not understood, a study showed that the C-terminal could repair alkylation damage, but could do so more efficiently in the presence of its N-terminal[14]. In other homologs the N-terminal demonstrates endonuclease V activity (in *Ferroplasma acidarmanus*) or can act as a signalling molecule (in *E. coli* Ada protein)[27, 34]. There also exists a *Caenorhabditis elegans* homolog, cAGT-2, that lacks the characteristic N-terminal, and yet is able to repair alkylated guanine[26].

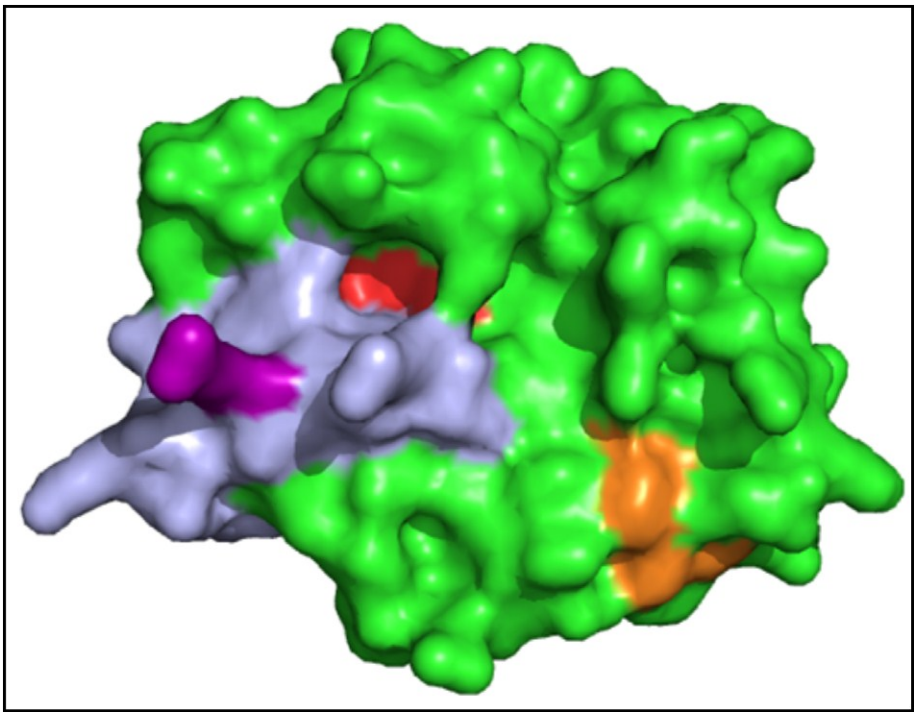


Figure 5 Crystal structure (IT39) of AGT. The helix-turn-helix motif (light blue), active site (red), zinc binding site (orange) and the arginine finger (purple) are highlighted[8].

1.4.2 AGT and mono-adducts

Mono-adducts of guanine, such as O^6-methyl-dG, can be repaired stoichiometrically and irreversibly by AGT. By means of an arginine "finger" the damaged base is flipped into the active site where the alkyl group is irreversibly transferred to a highly reactive cysteine residue (Cys145 in human AGT, see **Figure 6 and 7**)[20]. The Cys145 is an active nucleophilic species due to its placement within a glutamate-histidine-water-cysteine hydrogen bond network.

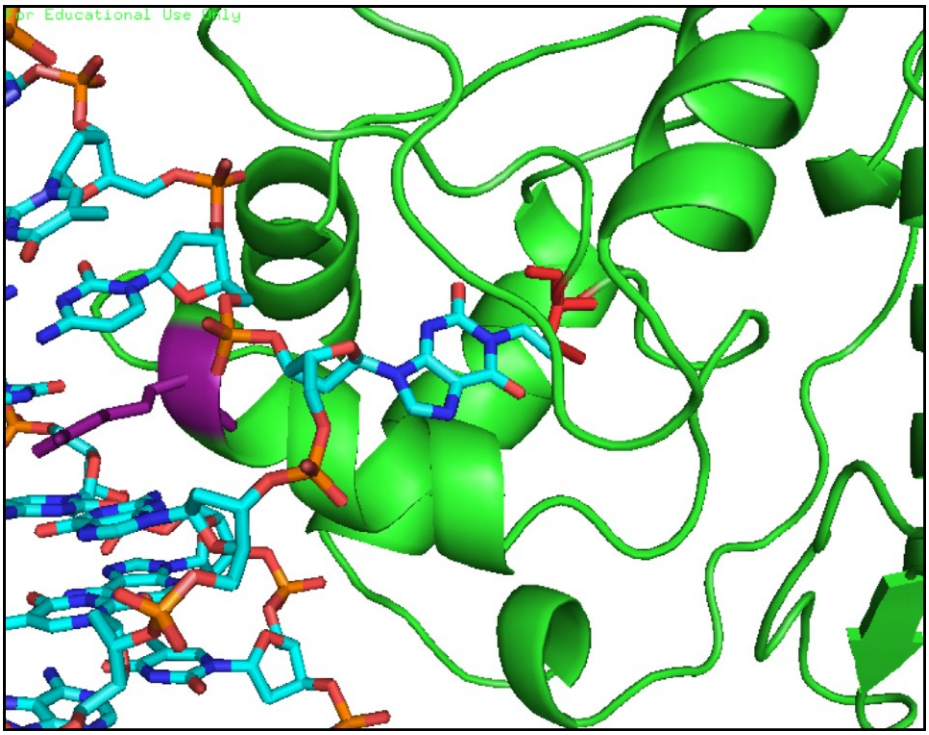


Figure 6 An enlarged image, using pymol, of the IT39 AGT crystal structure, where AGT was co-crystallized with a modified base. Red denotes the active site cysteine 145 and the purple residue is the arginine "finger" that has placed itself within the DNA duplex[8].

The crystal structure clearly illustrates AGT's mechanism of repair by flipping out the alkylated base with no major distortion or melting of the DNA double helix. In addition, it appears that the arginine "finger" plays an essential role in damage reversal,

as was also verified by *in vitro* repair studies. Interestingly, AGT can repair an O^6-alkyl-guanine in dsDNA regardless of the base directly opposite the alkylated guanine adduct[43, 47]. In addition, AGT has demonstrated a lack of sequence dependence of the flanking residues when repairing alkylated guanine and reverses damage in a cooperative manner[3].

Studies have shown that once alkylated, AGT becomes ubiquitinated and consequently degraded, in addition to becoming more sensitive to protease digestion *in vitro*[25, 44, 54]. Furthermore, a novel function of AGT as a signalling molecule was discovered, where it acted as a negative regulator of estrogen-receptor mediated transcription[73]. Although not considered inducible as in the case of the Ada protein, studies showed that partial hepatectomy and an increase in glucocorticoids results in elevated levels of AGT expression[21, 36].

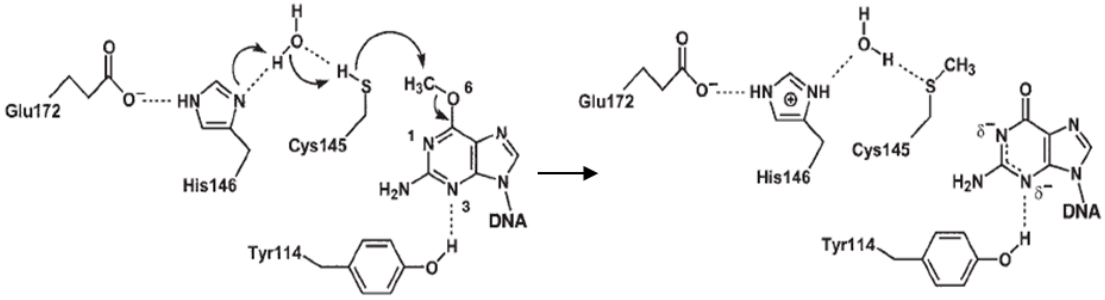


Figure 7 AGT repair mechanism of O^6-methyl-dG (modified from Daniels DS 2004) [9].

AGT, unlike its *E. coli* counterparts, is able to repair a wide variety of alkyl damage both at the O^6-atom of guanine (see **Figure** 7) and the O^4-atom of thymine, but to a lesser extent. AGT can repair substrates as small as methyl groups to much larger benzyl groups and even ICLs (see **Figure 8**). The human homolog is able to accommodate a greater range of substrates because of its larger active site and hydrophobic pocket[39].

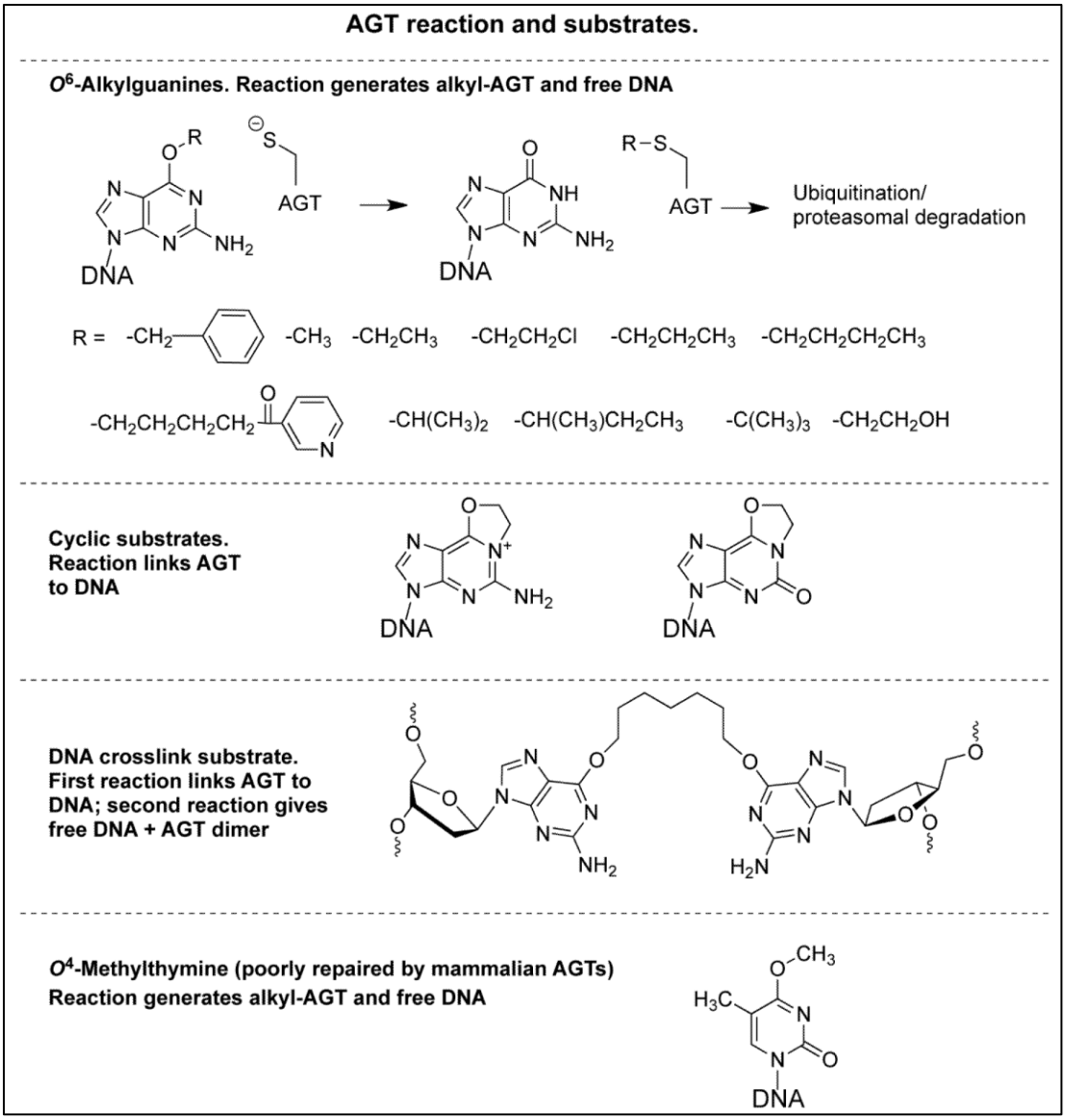


Figure 8 Simplified AGT reaction and examples of repairable substrates[39].

Due to AGT's ability to repair a wide variety of alkyl damage, it became apparent that inhibition of this protein could play a therapeutic role in cancer treatment[16, 17]. AGT's role in reducing therapeutic effectiveness of various alkylating agents was supported with *in vivo* assays[7, 10, 46]. As a result of AGT's importance several detection assays have been developed, from real-time fluorescence assays, to restriction endonuclease inhibition and magnetic bead separation and action as a labelling molecule for small molecules[23, 53, 80].

One of the major questions though, was how does AGT detect damage to repair it? In addition, was there a molecule that could be used to inhibit AGT to reduce its effectiveness?

1.4.3 Damage detection

A long standing question has been: How does AGT detect modified bases? Initial hypotheses believed that AGT actively flipped out each base as a means to detect repair (see **Figure 9A**), yet without the requirement of a cofactor or an energetic molecule (ATP) this seemed unlikely. Another hypothesis believes that AGT transiently detects extrahelical lesions (see **Figure 9C**) or is able to detect DNA intrahelical lesions based on DNA instability or distortion (see **Figure 9B**).

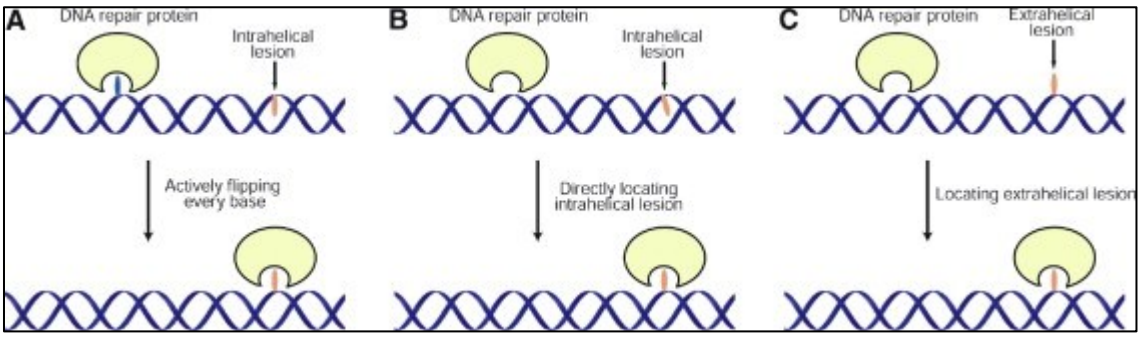


Figure 9 Illustration of the various hypotheses of DNA damage detection by AGT[12].

Duguid *et al* approached the problem with a chemical crosslinking experiment using a modified base that was linked to a disulphide bond. On repairing the modified base, AGT would be chemically linked to the DNA strand (see **Figure 10**)[12]. Using a disulphide linking approach, they determined that AGT must sense modified bases by either detecting adducts that are hidden intrahelically or capturing transient extrahelical lesions, but not by actively searching for damaged bases. Conversely, the E. coli homolog Ada-C was only able to capture transient extrahelical lesions[12].

Hu *et al*, using a computational approach, showed that AGT is able to scan DNA for damage. Based on their calculations they determined that O^6-methyl-dG flipped 10^5 times faster into the active site than dG, and that AGT could move along the DNA strands 10^3 times slower than O^6-methyl-dG flipping, suggesting that AGT is able to scan the DNA and flip predominantly only O^6-methyl-dG into the active site, whereas dG would be too slow to flip[22].

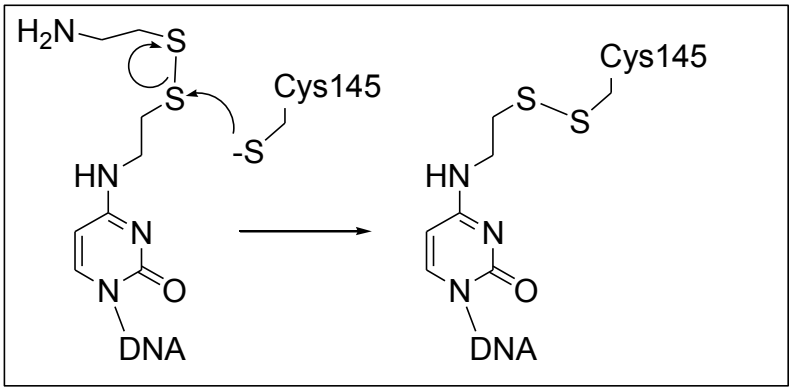


Figure 10 Illustration of a modified cytosine to form a disulfide linkage between AGT and DNA (modified from Duguid *et al* 2003)[12].

Computational and experimental data have yet to fully agree on AGT's ability to detect and scan for alkylated bases. Much work is still required before the exact mechanism of damage detection is elucidated.

1.4.4 AGT and interstrand crosslinked DNA

Certain chemotherapeutic treatments involve the use of alkylating agents that introduce DNA interstrand crosslinks (ICL); hepsulfam is an example. Although mono-adduct repair has been researched in-depth, it is still unclear how AGT can access bulky ICL lesions. McManus *et al* demonstrated that AGT is able to repair an O^6-heptylene-dG-O^6-heptylene-dG lesion directly opposed and in a 5'-GNC-3' motif (N is any base), and hypothesized that two AGT molecules are required (see **Figure 11**)[33]. MS studies

demonstrated that the 5'-GNC-3' motif was the most favourable sequence motif for alkylation crosslink formation upon hepsulfam treatment[68].

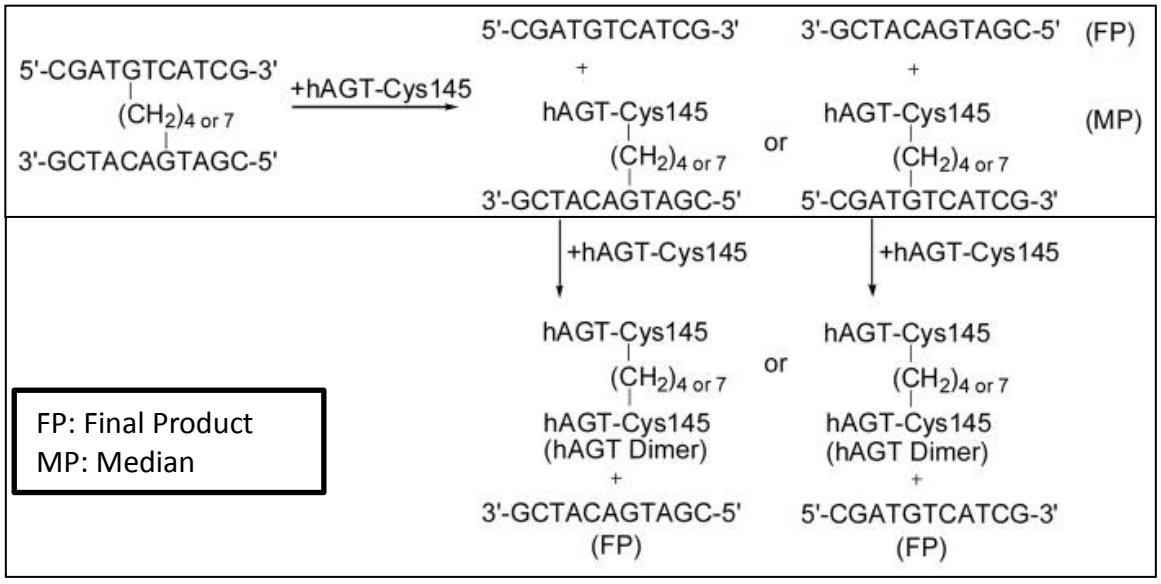


Figure 11 Proposed repair scheme of ICL in a 5'-GNC-3'motif by AGT[33].

A model of repair has suggested that the DNA duplex is denatured and the repaired strand released, while the alkyl adduct and attached DNA remain covalently bound to AGT[15]. The model illustrates a completely denatured DNA double helix, where this varies drastically from mono-adduct repair. Development of binding assays to determine AGT's mode of binding and repairing of ICLs is required to confirm the model and improve our understanding.

1.4.5 AGT inhibitors

AGT's ability to diminish the therapeutic efficiency of various alkylating agents has inspired the development of inhibitors. Initial work began in 1990 when O^6-benzyl-guanine was used to deplete cellular AGT as a means to verify its function as a protectant against mutagenesis and alkylating agents[10]. Since that time, new inhibitors that mimic substrates have been developed (see **Figure 12**)[39].

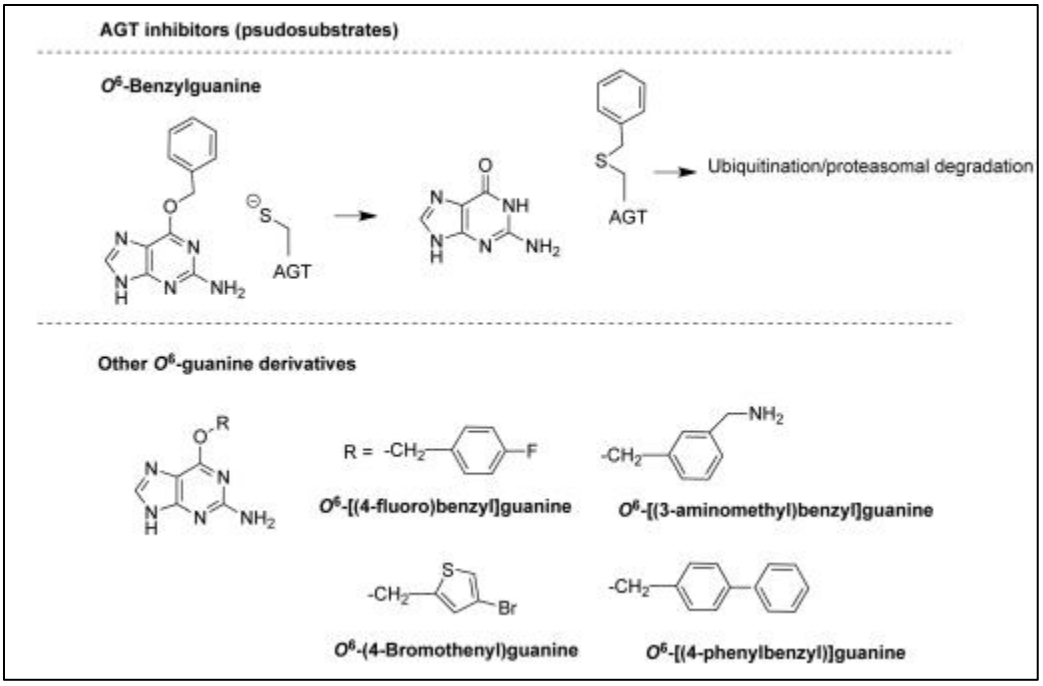


Figure 12 Inhibitors of AGT[39].

Inhibition of AGT is often monitored using mass spectrometry (MS) techniques or gels. Thus far, one of the major problems with inhibitor drug development has been poor solubility. Further development of AGT inhibition is still required before a treatment regimen can be commercialized.

1.5 Fluorescent base analog to detect DNA denaturation

There are several well-known fluorescent base analogs from 2-amino purine to 6-methyl-3-(β-D-2-deoxyribofuranosyl)pyrrolo[2,3-d] pyrimidin-2-one (pyrrolo-dC). We will be focusing our discussion on the development of pyrrolo-dC and its properties. In the 1980's the first pyrrolo derivative was discovered, but it was not until 2004 that the pyrrolo-dC analog was developed and begun to be incorporated in DNA to study its properties (see **Figure 13**)[84].

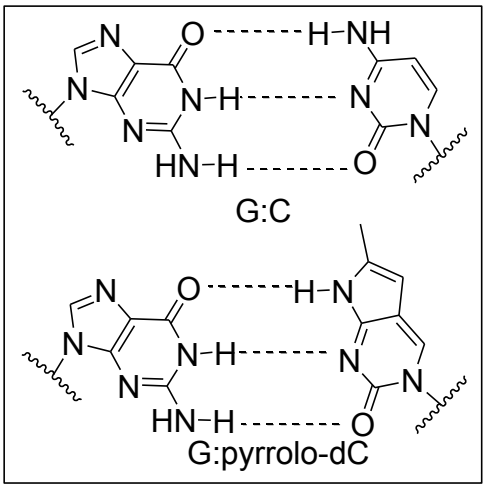


Figure 13 Representation of base pairing between guanine and cytosine (top) and guanine and pyrrolo-dC structure (bottom).

Pyrrolo-dC is commercially available and characteristically stable under DNA solid-phase synthesis, with little apparent effect on DNA structure and stability[55, 74, 83, 86]. Pyrrolo-dC's ability to maintain DNA structure and stability may be due in large part to its ability to base pair with guanine in much the same way as cytosine. In addition, pyrrolo-dC appears to discriminate well between the various bases, where it associates best with guanine, which is useful in DNA repair and base mismatch detection[76, 77, 82, 85]. An important property is that pyrrolo-dC fluoresces when not base paired, while it is quenched when base paired, more precisely in B-form DNA. Based on computational results, it is postulated that the observed quenching of pyrrolo-dC fluorescence is not mainly because it is base paired but due to the pi-pi stacking within B-form DNA[75, 83]. One study demonstrated that the position of pyrrolo-dC within a DNA double helix will have a differing fluorescence profile, where the position near the middle will show lower overall fluorescence emission as compared to a position near the end[76]. Pyrrolo-dC has been used for various assays, from being used in repair assays of AGT with benzylated or methylated guanine residues to a probe for the detection of silver ions[55, 81]. This

fluorescent base analog may allow for the ability to monitor DNA melting due to protein (AGT) binding.

1.6 Research Objective

Although the ability of AGT to repair an ICL has been investigated, the influence of AGT binding and effect on the structure of the ICL DNA duplex to access the damage is unknown. Therefore, to better understand the effect of AGT binding, a base that fluoresces when not base-paired (pyrrolo-dC), will be incorporated into oligonucleotides at various positions relative to the lesion, to observe whether DNA melting is occurring and the distance from the site of the lesion. The lesion will be an O^6-2'-deoxyguanosine-heptylene-O^6-2'-deoxyguanosine (O^6-dG-heptylene-O^6-dG) ICL in a 5'-GNC-3' motif. This ICL DNA mimics the lesion formed by hepsulfam (a heptylene linkage between two guanines) and the sequences of the oligonucleotide substrates are shown in **Figure 14**.

In order to achieve our goal, we will synthesize the ICL DNA and its controls, purify them to high purity and biophysically characterize the DNA. In addition, proteins (AGT wild type, C145S variant, R128A variant and Ada-C) will be expressed and fully characterized to ensure purity and proper folding. The C145S variant is inactive and will be used for binding assays while R128A variant has the arginine "finger" replaced with an alanine in order to understand its role in ICL DNA repair and binding. Once complete, we can then begin to study their association using radioactivity repair and binding studies. Those proteins not able to repair and that bind strongly to the DNA substrates will then be selected for fluorescence binding assays. The fluorescence binding assays

will be monitoring the increase in fluorescence as emitted by pyrrolo-dC, and will give an idea as to the extent to which ICL DNA is melted or distorted upon AGT binding.

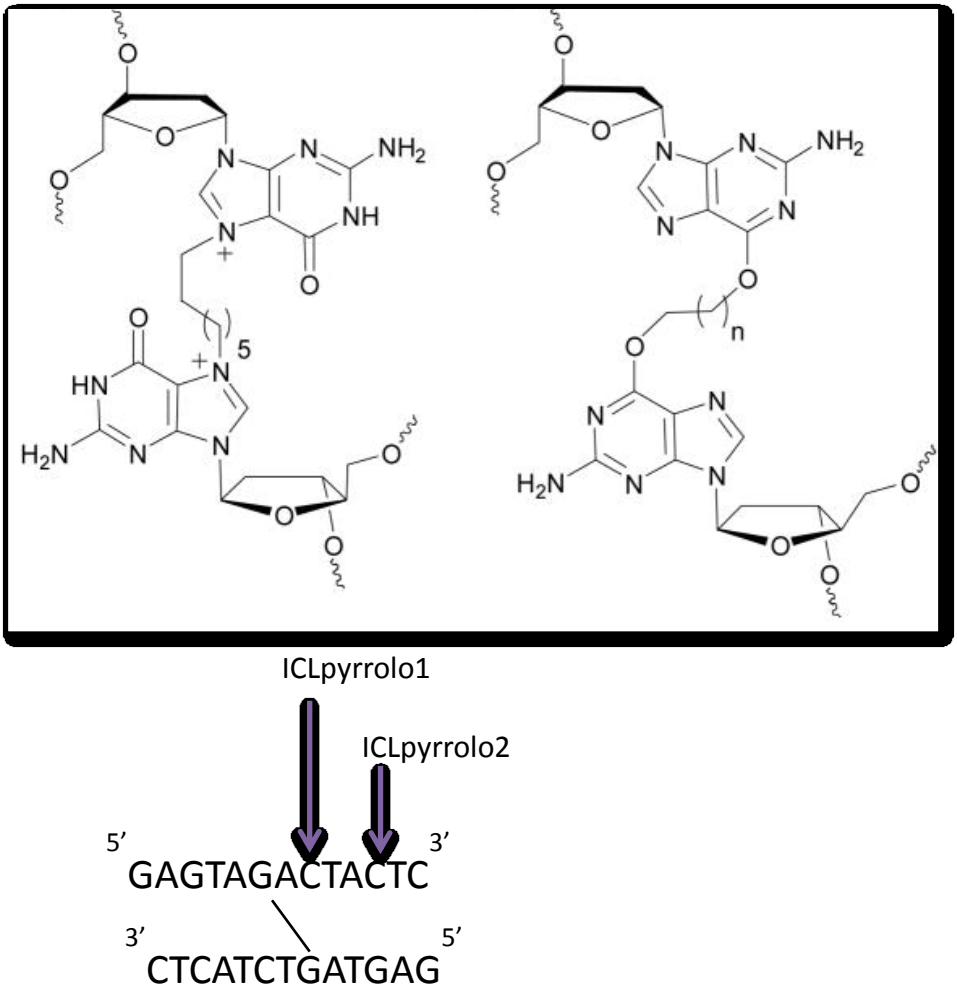


Figure 14 Structures of O^6-dG-heptylene-O^6-dG (top right) and N7-dG-heptylene-N7-dG (top left) linkages, and the oligonucleotide sequence (bottom) to be used with the ICL represented by the line between two guanines and the locations of pyrrolo-dC.

Greater understanding of AGT's association with ICL DNA as compared to unmodified DNA may shed some light in AGT's ability to access these bulky ICL lesions. This will allow for the development of new ICL forming drugs that are resistant to AGT repair, increasing the therapeutic efficiency of certain treatments.

1.7 Summary

DNA plays an essential role in cell survival and is protected by several repair pathways against both endogenous and exogenous DNA damaging events. Alkylating agents have been exploited for their mutagenic and cytotoxic effects by using them as chemotherapeutic agents. Even though alkylation damage can potentially be very potent and lethal, repair pathways have been shown to reduce the damaging effect and increase cell survival rates. One important repair pathway explored is direct repair, the simplest of all repair systems. Alkyltransferases, specifically AGT, have shown their ability to reduce the efficiency of various alkylating agents. AGT stoichiometrically and irreversibly repair O^6-alkyl-dG in an efficient manner, after which it is ubiquitinated and degraded. In addition, AGT has shown to accommodate a wide variety of substrates including ICLs. As a consequence AGT inhibitors have been developed in the interest of improving therapeutic regimens.

Of interest is AGT's ability to repair ICLs, as it is still unknown how this occurs. A fluorescence based binding assay will be explored in the hope of gaining a greater understanding as to how AGT associates with ICL and the extent to which DNA is distorted upon binding. Greater understanding of ICL-AGT association may allow for the development of new therapeutic crosslinking agents that are invulnerable to AGT repair.

Chapter 2: Materials and methods

2.1 Materials

All biochemical reagents were purchased from Bioshop Canada Inc (Burlington, ON). Ni-NTA Superflow Resin was purchased from Qiagen (Mississauga, ON). Complete, Mini, EDTA-free Protease Inhibitor Cocktail Tablets were obtained from Roche (Laval, QC). Nanosep 3K Omega and Sep-Pak® C-18 cartridges were obtained from Pall Corporation (Ann Arbor, Michigan) and Waters (Wexford, Ireland), respectively. Unstained Protein Molecular Weight Marker was purchased from Fermentas Life Sciences (Burlington, ON).

2.2 Site-directed mutagenesis

The AGT wild type plasmid was previously prepared by Francis McManus (Ph.D. candidate). The AGT gene was inserted into a pQE30 vector, IPTG inducible and AMP resistant, between *EcoRI* and *KpnI* cleavage sites. In addition, all primers (sequences shown in **Table 2**) were designed based upon the modification of a single residue, and were purchased from BioCorp (Montreal, Quebec). Phusion DNA polymerase and *DpnI* were obtained from Fermentas Life Sciences.

The R128 residue of AGT wild type was replaced with an alanine (see **Table 2** for primer sequences) using the QuikChange® site-directed mutagenesis kit from Stratagene (Agilent, Mississauga, Ontario). As per the Stratagene recommendation, the reaction mixture (50 μL) contained 1 μL (5-10ng) of pQE30-AGTwt template, 2 μL dNTPs, 10 μL of 5X HF buffer, 0.5 μL of Phusion DNA polymerase and 1μL (10 pmol) of both forward and reverse primers. The mixture was then placed in a GeneAmp PCR

System9700 for 2-2.5 hours under the following temperature cycling conditions: 98°C, 2 minutes; followed by 18 cycles of 98°C, 1 minute; 65°C, 1 minute; 72°C, 4 minutes; ending with 72°C for 10 minutes and cooling to 4°C. PCR sample was subsequently digested with 1 µL *DpnI* at 37°C for 2-2.5 hours. *DpnI* digests methylated DNA, resulting in digestion of the template DNA with only the PCR products remaining that may then be used to transform bacteria.

Protein	Oligonucleotide sequence	Amino acid sequence
hAGT wild type	CCC AAA GCC GCG CGA GCA GTG GGA GGA GG	PKAARAVGG
hAGT R128A	CCC AAA GCC GCG GCC GCA GTG GGA GGA GG	PKAAAAVGG
Forward primer (29 bases)	CCT CCT CCC ACT GCT GCC GCG GCT TTG GG	
Reverse primer (29 bases)	CCC AAA GCC GCG GCA GCA GTG GGA GGA GG	

Table 2 Summary of modifications made (underlined), based on amino acid sequence (N- to C- terminal) and oligonucleotide sequence (5' to 3'), and the primers- using *E. coli* codons for optimum expression- employed for site-directed mutagenesis.

2.3 Cell transformation

The *DpnI* digests of PCR based site-directed mutagenesis, 2 µL, is incubated with 50 µL of competent XL-10 gold *E. coli* cells on ice for 2-5 minutes. The sample is subsequently heat shocked at 42°C for 40 seconds and cooled on ice for 2 minutes. The entire content were plated on LB agar/AMP plates and incubated at 37°C overnight.

2.4 Extraction of plasmid DNA

Colonies formed from the bacterial transformation were selected and used to inoculate 10 mL 2XYT (16 g tryptone, 10 g yeast extract and 5 g NaCl in 1 L of water)

and 10 µL AMP (0.1 g/mL) shaking at 225rpm overnight at 37°C. Plasmid DNA was extracted *via* a GeneJET™ Plasmid Miniprep Kit (Fermentas Life Sciences) stored in distilled water at -20°C as per the manufacturer's instructions. The plasmid DNA was quantified using UV spectroscopy at 260 nm with the assumption that double stranded DNA has an extinction coefficient of 50 ng μL^{-1} cm^{-1}. Pure plasmid DNA was then sent for sequencing (Genome Quebec) and to be used to transform bacteria to express protein.

2.5 Glycerol stock preparation

Colonies from transformed XL-10 Gold *E. coli* cells were used to inoculate 2XYT solution containing AMP (0.1 g/mL) that was incubated overnight at 37°C shaking at 225 rpm. A 700 µL aliquot of *E. coli* cells and 300 µL of a 50% glycerol solution were then mixed and subsequently stored at -80°C.

2.6 Protein purification

Glycerol stocks of AGT wild type and variant were prepared by Francis McManus (Ph.D. candidate). Overnight cultures were grown in 10 mL of 2XYT and 10 µL of AMP (0.1 g/mL) in a 50 mL falcon tube at 37^{o}C in a shaker at 225 rpm. These were then used to inoculate 1 L of 2XYT and 0.1 g AMP at 37^{o}C for 2.5 hours. When the culture reached an absorbance reading between 0.4-0.6 at 600nm, protein expression was induced by adding 4 mM IPTG and the culture was allowed to grow for another 4 hours. Cells were then centrifuged at 8500g using a Beckman J2-HS centrifuge for 20 minutes at 4^{o}C. Pellets were re-suspended in 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 20% glycerol, 20 mM β-mercaptoethanol and 1 Complete Mini EDTA-free Protease Inhibitor Cocktail tablet. The sample was homogenized using a glass homogenizer and the cells were

disrupted by passing the sample twice through a French Press. Genomic DNA was sheared by applying the sample through a small syringe and the lysate was subsequently centrifuged at 17000g for 45 minutes at 4^{o}C. All debris was discarded, while supernatant containing soluble histidine tagged protein was retained.

The sample was applied to a column of Ni-NTA affinity resin (3 mL) and the column was washed with 300 mL of buffer containing 20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 20 mM imidazole and 20 mM β-mercaptoethanol to remove non-specific binding proteins. The desired AGT protein was eluted in 1 mL fractions with 20 mM Tris-HCl (pH8.0), 250 mM NaCl, 200 mM imidazole and 20 mM β-mercaptoethanol. Detection of protein present in each fraction was determined *via* a Bradfords reagent. All fractions containing protein were pooled and dialysed overnight in 2 L of buffer containing 50 mM Tris-HCl (pH 7.6), 250 mM NaCl, 10% glycerol, 5 mM DTT and 0.1 mM EDTA. Protein concentration was determined using UV/vis spectroscopy and the calculated molar extinction coefficient (26470 M^{-1} cm^{-1} assuming all cysteine residues are reduced) was determined by the ProtParam tool on the ExPASy site. The pure protein was stored at -80^{o}C.

2.7 SDS-PAGE

SDS-PAGEs were performed on Biorad Mini-PROTEAN Tetra Cell system. Resolving mixture (5 mL) composed of 0.375 M Tris-HCl (pH 8.8), 0.1% SDS, 11.4% acrylamide, 0.6% bis-acrylamide, 25 µL 20% ammonium persulphate (APS) and 5 µL tetramethylethylenediamine (TEMED) was mixed and poured into the gel apparatus where it was allowed to polymerize. Ethanol was added on top to smooth out the polymerizing gel. Once polymerized, 2 mL of stacking solution containing 0.125 M Tris-

HCl (pH 6.8), 0.1% sodium dodecyl sulphate (SDS), 3.8% acrylamide, 0.2% bis-acrylamide, 25 µL 20% APS and 5 µL TEMED was added. Samples, 20 µL, were prepared by mixing approximately 5 µg of protein, distilled water and 2X SDS-PAGE loading dye followed by boiling for 10 minutes. The 2X SDS-PAGE loading dye was composed of 0.06 M Tris-HCl (pH 6.8), 2% SDS, 10% glycerol, 0.025% bromophenol blue and 1.2 M β-mercaptoethanol. The gel was run at 120 V in glycine running buffer- 0.025 M Tris-HCl (pH8.3), 0.192 M Glycine and 1% SDS- until the tracking dye almost migrates to the bottom. The gel was removed from the gel apparatus and incubated for about 30 minutes in staining solution, which consisted of 50% methanol, 10% acetic acid and 0.25% coomassie blue. The gel was subsequently destained in 50% methanol and 10% acetic acid for 30 minutes or until optimally visualized. The gel was then cast in cellophane and left to dry.

2.8 Protein mass spectrometry

All protein samples were analysed using ESI-MS, where 25 µg of protein was precipitated using one volume of trichloroacetic acid (TCA) to four volumes of protein solution. The mixture was then incubated for 10 minutes on ice and subsequently centrifuged at 14000 rpm at 4^{o}C for 5 minutes using a Sorvall Legend Micro 21R centrifuge. The supernatant was removed and the pellet washed with acetone several times to remove impurities. Residual acetone was removed by heating the sample in a sand bath at 60^{o}C for 10 minutes. Once completely dry, the precipitate was dissolved in 100 µL of distilled water and 100 µL of acetonitrile containing 0.1% formic acid. Samples were then analysed with a Waters Micromass Q-TOF-2 mass spectrometer operating in positive-ion mode.

2.9 Protein far UV circular dichroism spectroscopy

Protein samples, approximately 5 μM, were buffer exchanged with 50 mM KH_2PO_4, 75 mM NaCl, 1 mM DTT, pH 7.5 buffer using Nanosep 3K Omega spin columns at 10,000 rpm at 4^{o}C using a Sorvall Legend Micro 21R centrifuge. Circular dichroism spectra were obtained on a JASCO-810 spectropolarimeter using a 0.2 cm path length cell. The sensitivity was on standard mode, and the data pitch was 0.2 nm with a response time of 1 sec. The samples were measured from 260-200 nm with a scan speed of 20 nm/minute at a constant temperature of 20^{o}C, maintained with a Pelletier temperature controller. The band width used was 1.0 nm and the scanning mode continuous. The data (the average of five accumulations) obtained was then smoothed and converted to molar ellipticity using software provided by JASCO [87].

2.10 Protein thermal denaturation.

Protein samples were buffer exchanged exactly as those prepared for regular far UV circular dichroism spectra. Thermal denaturation was measured at 222 nm on a JASCO-810 spectropolarimeter using a 0.2 cm path length cell. The decrease of α-helical content as a function of increasing temperature was monitored at a rate of 20^{o}C per hour between 20 to 90^{o}C. The data was then smoothed and plotted as percent folded as a function of temperature. The melting temperature (T_m) value was determined from the first derivative maxima of the thermal denaturation curves.

2.11 Protein quantitation

After protein dialysis, the absorbance of the pure protein sample was determined at 280 nm using a Varian Cary3000 Bio UV-visible Spectrophotometer, with the dialysis buffer as the blank. The absorbance observed was used to calculate the protein

concentration using Beer-Lambert's Law and the molar extinction coefficient of the AGT variant (26470 M^{-1} cm^{-1} assuming all cysteine residues are reduced) calculated by the ProtParam program on the ExPASy site.

A secondary protein quantitation method (Bradfords assay) was employed to confirm the concentration of samples with volumes less than 1 mL. A 10 µL protein sample was incubated with 1000 µL of Bradfords reagent to a total of 1100 µL solution at room temperature for 10 minutes. Absorbance was measured at 595 nm and compared to a standard curve previously prepared using BSA protein at various concentrations.

2.12 DNA solid-phase synthesis

Solid-phase synthesis (SPS) was performed on an Applied Biosystems 3400 DNA synthesizer using cyanoethyl-N,N-diisopropyl-phosphoramidites (phosphoramidites) as per standard protocols[88-90]. Unlike biological systems, oligonucleotide solid-phase synthesis is typically performed in a 3' to 5' direction. SPS is an efficient and simple method to quantitatively synthesize short DNA sequences.

All control DNA from either the bi-partite or tri-partite projects- unmodified and those containing pyrrolo-dC (procured from Glen Research)- were synthesized using long chain alkylamine (lcaa) controlled pore glass (CPG) solid support on a 1 µmole scale, whereas ICL and O^6-methyl-dG modified DNA were prepared using polystyrene solid support on a 2 µmole scale (see **Table 3** for complete list of sequences). All 3'-phosphoramidites, containing standard protecting groups, were purchased from Glen Research. Furthermore, 3'-amidites were diluted to 0.1 M in 100% acetonitrile (ACN), while 5'-amidites from ChemGenes Inc. were diluted to 0.2 M in 100%ACN.

Synthesis was conducted under the following cycle parameters, with an ACN wash in between each step: (1) detritylation with 3.0% TCA in dichloromethane (DCM) for 120 seconds; (2) coupling with ethyl thiotetrazole as activator for 120 seconds for 3'-phosphoramidites or 180 seconds for 5'-phosphoramidites or 600 seconds for pyrrolo-dC (0.1 M) and O^6-dG-heptylene-O^6-dG amidites (0.15 M, prepared by Dr. Christopher Wilds); (3) oxidation with 0.02 M iodine in THF-water-pyridine (2.5:2:1, *v/v/v*); (4) capping of uncoupled oligonucleotides with 1:1 (*v/v*) acetic anhydride-2,6-lutidine-tetrahydrofuran (THF) (1:1:8, *v/v/v*, solution A) for control sequences or 1:1 (*v/v*) phenoxyacetyl anhydride-2,6-lutidine-THF (1:1:8, *v/v/v*, solution A) for ICL or O^6-methyl-dG DNA with 1-methyl-1-H-imidazole-THF (16:84, *w/v*, solution B). The cycle was repeated until the desired oligonucleotide sequence was completed and finally detritylated. The oligonucleotide extension in the ICL synthesis varied, in that additional on column deprotection to remove allyloxycarbonyl (alloc) and/or silyl protecting groups was required prior to completion of the final desired oligonucleotide product (see section **2.13** and **2.14** following). Once synthesis was complete, the DNA was ready for cleavage from the solid support and complete deprotection.

Bi-partite project		
Code name	Oligonucleotide sequence (5'-3')	Modification
CJW322	GAG TAG ACT ACT C	None
CJW323	GAG TAG TCT ACT C	None
CJW324	GAG TAG A**C**T ACT C	pyrrolo-dC (**C**)
CJW325	GAG TAG ACT A**C**T C	pyrrolo-dC (**C**)
CJW326 ICL DNA (ICLpyrrolo1)	GAG TA**X** A**C**T ACT C GAG TA**X** TCT ACT C	O^6-dG-heptylene-O^6-dG (**XX**) pyrrolo-dC (**C**)
CJW327 ICL DNA (ICLpyrrolo2)	GAG TA**X** ACT A**C**T C GAG TA**X** TCT ACT C	O^6-dG-heptylene-O^6-dG (**XX**) pyrrolo-dC (**C**)
CJW328 ICL DNA (ICL)	GAG TA**X** ACT ACT C GAG TA**X** TCT ACT C	O^6-dG-heptylene-O^6-dG (**XX**) pyrrolo-dC (**C**)
Tri-partite project (5'-3')		
Code name	Oligonucleotide sequence (5'-3')	Modification
CJW239	GCA TGA CTA GTC GTG CTG CT	None
CJW243	AGA CGC ACG ACT AGT CAT GC	None
CJW269	GAT CAC TGA CTA CGC TAC	None
CJW270	GTA GCG TAG TCA GTG ATC	None
CJW271	GTA GCG TA**G*** TCA GTG ATC	O^6-methyl-dG
CJW272	GAT CAC TGA **C**TA CGC TAC	pyrrolo-dC
CJW273	GAT CAC TGA CTA **C**GC TAC	pyrrolo-dC
CJW274	GAT CAC TGA CTA CG**C** TAC	pyrrolo-dC
CJW293 ICL DNA	GTA GCG TA**X** TCA GTG ATC GAT CAC T**X**A **C**TA CGC TAC	O^6-dG-heptylene-O^6-dG (**XX**) pyrrolo-dC
CJW294 ICL DNA	GTA GCG TA**X** TCA GTG ATC GAT CAC T**X**A CTA **C**GC TAC	O^6-dG-heptylene-O^6-dG (**XX**) pyrrolo-dC
CJW295 ICL DNA	GTA GCG TA**X** TCA GTG ATC GAT CAC T**X**A CTA CG**C** TAC	O^6-dG-heptylene-O^6-dG (**XX**) pyrrolo-dC

Table 3 List of sequences prepared using SPS for both bi-partite and tri-partite projects. X designates an O^6-dG-heptylene-O^6-dG modification; C designates a substitution of deoxycytidine (dC) with pyrrolo-dC; and G* refers to a deoxyguanosine (dG) replaced with an O^6-methyl-dG.

2.13 De-alloc reaction

Removal of the allyloxycarbonyl (alloc) group was performed on the solid support. Treatment began with overnight incubation with triethylamine (TEA), followed by washing with 20-30 mL THF and 20 mL 100% ACN, and subsequent drying by high

vacuum. Once dry, the sample was transferred to a glass vial and incubated with palladium tetrakis-triphenylphosphine (0.0029 g, 10 eq.), triphenylphosphine (0.0013 g, 20 eq.) and 0.25 mL of buffer solution comprised of butyl amine (100 eq.) and formic acid (100 eq.) at 35°C for 3 hours; mixing every 30 minutes. The mixture was then transferred to a clean column for solid phase synthesis and was washed with 20-30 mL THF and 20 mL ACN and dried by high vacuum. Upon complete dryness, the column was treated with a palladium scavenger, N,N-diethyldithio-carbamate saturated in 3 mL ACN, for 30 minutes and re-treated until no yellow colour was present. The scavenger step limited the possibility of prolonged exposure to palladium that may lead to possible strand breakage. The column was washed again with 20-30 mL THF and 20 mL ACN and dried by high vacuum, after which it was ready for solid-phase synthesis extension.

2.14 De-silylation reaction

The silyl protecting group, *tert*-butyldimethylsilyl (TBS), was removed while the oligonucleotide was on the solid support at room temperature. First, the sample was treated with 1-2 mL TEA overnight to remove cyanoethyl groups, making the phosphate backbone less prone to nucleophilic attack and strand breakage. The column was later washed with 20-30 mL THF and 20 mL ACN and dried by high vacuum for 20 minutes. Once dry, 1 mL of TEA/HF (1:3) was applied and was left incubating for 1 hour with periodic mixing. Sample was then washed with 20-30 mL THF and 20 mL ACN and dried by high vacuum; ready to extend using solid-phase synthesis.

2.15 DNA deprotection

Deprotection of unmodified oligonucleotides containing the most common protecting groups (benzoyl, isobutyryl, 2-cyanoethyl and 4-isopropylphenoxyacetyl (i-Pr-

PAC)) typically involves incubation at 55°C for 4 hours in 500µL ammonia/ethanol (3:1). However, due to the presence of pyrrolo-dC, milder conditions were used for all DNA samples. Oligonucleotides on the solid support were transferred to a 2 mL screw-cap tube and incubated at room temperature over 36-48 hours in the dark in 500 µL ammonia/ethanol (3:1). Once deprotection was complete, supernatant was extracted and the solid support washed twice with 50% ACN, all of which was pooled and dried down for further analysis.

2.16 DNA quantitation

DNA quantitation was performed using absorbance detection at 260nm with a Varian Cary300 Bio UV-visible Spectrophotometer. 18 Mohm water was present in the reference cell and the instrument blanked with 18 Mohm water. DNA samples were diluted in 1 mL 18Mohm water and the absorbance/optical density (OD) was measured, providing a concentration of OD/µL. Quantitation was performed for crude samples, desalted samples and pure samples. The OD/µL concentration can be converted to a µM value by using Beer-lambert's Law and a calculated extinction coefficient (tri-partite ssDNA and dsDNA: 130700 M^{-1} cm^{-1} containing pyrrolo-dC and 127200 M^{-1} cm^{-1} for complement strand with 21021.7 M^{-1} cm^{-1} for dsDNA; bi-partite ssDNA and dsDNA: 170600 M^{-1} cm^{-1} containing pyrrolo-dC and 180600 M^{-1} cm^{-1} for complement strand with and without O^6-methyl-dG with 290442 M^{-1} cm^{-1} for dsDNA), which was determined using OligoAnalyzer for single stranded DNA and IDT Biophysics for double stranded DNA. It must be noted that the molar extinction coefficients for ICL DNA, O^6-methyl-dG and pyrrolo-dC are not exact values, because the extinction coefficient of the

modified base is not always known, and hence only approximations of the μM concentrations can be made.

2.17 DNA purification

A 40mL 20% polyacrylamide gel solution of 19:1 (*w/w*) acrylamide: bis-acrylamide (UltraPure and BioUltraPure grade respectively, Bioshop), 1X TBE- 5.45wt% (*w/v*) Tris-HCl, 2.75wt% (*w/v*) boric acid, 0.186wt% (*w/v*) EDTA (ACS grade, Fisher Scientific) in distilled water- and 7M urea (UltraPure molecular biology grade, Bioshop) was mixed with 100 μL 20wt% (*w/v*) APS and 100μL TEMED and allowed to polymerize for 30 minutes. Once the gel was polymerized and wells washed with 1X TBE buffer, approximately 20-40 ODs of DNA was dissolved in 100 μL of formamide, loaded into the well and the gel was performed in 1X TBE buffer with a SE600X chroma deluxe dual cooled vertical Hoefer unit with a power maximum at 10 W for 3-3.5 hours or 4-5 hours for better resolution of DNA products. Migration was monitored with 6X denaturing loading dye- 0.001wt% (*w/v*) bromophenol blue (ACS Fisher Scientific) and 0.001wt% (*w/v*) xylene cyanol (Bioshop) dissolved in formamide. Upon completion, the gel was removed from the apparatus, visualized under UV light and the desired band was excised. DNA was extracted from the gel by incubating the gel bands in 10 mL 0.1 M sodium acetate (Sigma Aldrich) in the dark, while shaking at room temperature for approximately 16 hours.

2.18 DNA desalting

DNA extracted from gels contains a high concentration of salts, which needed to be removed prior to biophysical analysis of the DNA. A Sep-Pak® C-18 cartridge (Waters) was equilibrated by applying 5 mL 100% ACN, followed by 8 mL 50% ACN

and finally flowing through 10 mL 0.1M sodium acetate. The cartridge was never allowed to dry. Quantitation of the DNA extract was performed and then passed twice through the cartridge at a drop-wise flow rate. Once the DNA was adsorbed to the C-18 cartridge, it is washed with 20-30 mL 18 Mohm water to remove residual salt. Once the flow-through was void of salt, the DNA was eluted with 3-4 mL of 1:2:1 (*v/v/v*) 100% ACN: 100% methanol: 18 Mohm water into a 5 mL screw cap tube and later dried down by speed vacuum, re-suspended in water and quantitated to determine overall yield and concentration of stock samples.

2.19 DNA HPLC analysis

Two forms of HPLC stationary phases were used: Reverse phase (RP) and ion exchange (IEX). Before use of the HPLC instrument, the lines were primed and cleared of the storage solution (50% methanol for RP or 10% methanol for IEX). Analytical IEX-HPLC was performed on a Dionex BioLC® DNAPac® PA-100 column using a gradient from 0-0.5 M NaCl in a buffer of 0.1 M Tris-HCl (pH 7.8) and 10% ACN at a flow rate of 1 mL/minute for 45 minutes with a loading volume of 500 µL containing 0.1-0.5 ODs of DNA in 18 Mohm water for analytical runs with a detection wavelengths of 260 nm and 280 nm. RP-HPLC was used to analyse DNA composition and monitoring removal of TBS protecting groups. A 500 µL sample of approximately 0.1 OD of DNA was loaded on a Water Symmetry® C-18 5 µm column and analyzed with a gradient from 0-45% ACN in a 50 mM sodium phosphate buffer (pH 5.8) at a flow rate of 1 mL/minute for 45 minutes.

2.20 DNA nuclease digestion and compositional analysis

Approximately 0.1OD of a pure DNA sample was dried down and incubated with 0.5 µL snake venom phosphodiesterase I (SVP) (P-L Biochemicals Inc.) and 0.5 µL calf intestinal alkaline phosphatase (CIAP) (Fermentas Life Sciences) in 19 µL of SVP buffer (10 mM Tris-HCl (pH 8.1) and 2 mM $MgCl_2$) at 37°C overnight or 48 hours for ICL DNA. Half of the digested sample was then diluted to 500 µL with 18 Mohm water and analyzed by RP-HPLC. The ratios of the areas for the nucleoside peaks divided by the respective extinction coefficients were used to determine the nucleotide composition of the oligomers.

2.21 DNA thermal denaturation

DNA thermal denaturation was conducted on a Varian Cary300 Bio UV-visible Spectrophotometer using a Cary temperature controller. DNA samples, dried down from stock solutions, were re-suspended in 1 mL T_m buffer (90 mM NaCl, 10 mM sodium phosphate (pH 7.0), and 1 mM EDTA) to a concentration of approximately 2.4 µM for ICL or unmodified duplexes. The sample was heated to 90°C for 10 minutes, cooled at 4°C to anneal the DNA overnight and transferred to a 1 cm quartz cuvette for analysis. The increase in absorbance with increasing temperature was measured from 15-95°C with a 0.5°C increment over 1 hour. The observed thermal denaturation curves were normalized to increasing hyperchromicity and the T_m value was obtained *via* the first derivative maxima.

2.22 DNA far-UV circular dichroism spectroscopy

DNA CD spectra were obtained on a JASCO-810 spectropolarimeter using a 1cm path length cell, with approximately 2.4 μM DNA sample in 1 mL T_m buffer (90 mM NaCl, 10 mM sodium phosphate (pH 7.0), and 1mM EDTA). The sensitivity was on standard mode, and the data pitch was 0.2 nm with a response time of 1 sec. The samples were analyzed from 340-200 nm with a scan speed of 20 nm/minute at a constant temperature of 20^{o}C, maintained with a Pelletier temperature controller. The band width used was 1.0 nm and the scanning mode continuous. The data (the average of five accumulations) obtained was then smoothed and converted to molar ellipticity using the software provided by JASCO[91-93].

2.23 Radioactivity labelling assay

DNA samples were radioactively labelled by incubating 10-100 pmol of DNA (depending if it was for repair or binding assays) with 10X polynucleotide kinase (PNK) buffer, 1.0 μL γ-^{32}P ATP (6000 Ci/mmol, Perkin Elmer) and 0.5-1.0 μL T4 PNK (Fermentas Life Sciences) at 37°C for 1 hour. Samples were then boiled for 5-10 minutes to denature the PNK. Once the ssDNA was labelled, it was incubated with 1.05 times of its complementary strand to ensure near complete incorporation of the labelled strand into dsDNA. Labelled DNA prepared for binding assays were diluted to a 1 mL solution using a salt binding buffer (10 mM Tris-HCl (pH 7.6), 100 mM NaCl, 1 mM DTT, 200 μg/mL BSA and 50% glycerol in water).

2.24 Radioactivity repair assay

Overnight repair analysis was conducted with 60 pmol protein, 2 pmol of DNA and 6.67 mM EDTA in an activity buffer of 50 mM Tris-HCl (pH 7.6), 100 mM NaCl and 5 mM DTT solution at 37°C. Time course assays were conducted using 600 pmol protein, 12 pmol of DNA and 6.67 mM EDTA in activity buffer over 120 minutes at 37°C, with samples collected at various times (minutes): 0, 5, 15, 30, 45, 60, 75, 90, 105 and 120. At each time point, the reaction was terminated with the addition of a solution containing 1%SDS, 72% formamide, 0.9X TBE, 0.001wt% (*w/v*) bromophenol blue and 0.001wt% (*w/v*) xylene cyanol- and the sample boiled to denature the repair protein. Samples were then analyzed on a 20% polyacrylamide gel. The solution was prepared as a 19:1 (*w/w*) acrylamide: bis-acrylamide and 7 M urea mixture dissolved in 1X TBE (5.45wt% (*w/v*) Tris-HCl, 2.75wt% (*w/v*) boric acid, 0.186wt% (*w/v*) EDTA in distilled water), and polymerized using 100 µL 20wt% (*w/v*) APS and 100 µL TEMED. Gels were imaged and a radioactivity count was obtained using a PhosphorImager TyphoonTrio instrument.

2.25 Radioactivity binding assay

Protein concentrations were varied below and above the predicted dissociation constant ($K_{monomeric}$) and were incubated with 0.5 nM DNA in activity buffer (50 mM Tris-HCl (pH 7.6), 100 mM NaCl and 5 mM DTT) at room temperature for 20 minutes. Samples were then applied to a polymerized 15% native polyacrylamide gel composed of 75:1 (*w/w*) acrylamide: bis-acrylamide in 0.5X TBE. The polyacrylamide gels were performed at 250 V in 0.5X TBE for 25 minutes. Gels were imaged and a radioactivity

count was determined using a PhosphorImager TyphoonTrio instrument. K_d's and $K_{monomeric}$'s were obtained using a modified Hill Plot of the data and repeated in duplicate.

The following is the calculations involved to determine the dissociation constants of a given data set plotted in a modified Hill plot[94, 95]:

Dissociation of protein (E) and substrate (S) from the protein-substrate complex (ES) can be defined as:

$$E_nS \leftrightarrow nE + S$$

Where n represents cooperativity and K_d is representative of the overall binding constant. The overall dissociation constant (K_d), where we cannot measure intermediates, can be defined as:

$$K_d = [E]^n [S]/[E_nS]$$

Certain approximations and assumptions are made based on experimental protocol:

(1) $[S] <<<< [E_{total}]$

(2) $[ES] <<<< [E_{total}]$

(3) $[E]= [E_{total}] – [ES] \approx [E_{total}]$; since $[ES] <<<< [E_{total}]$

(4) Therefore, $K_d = [E]([S]/[ES]) \approx [E] \approx [E_{total}]$

(5) $[S] \approx [ES]$ when $[E_{total}] = K_d$

The K_d equation can be manipulated such that the n is no longer an exponent:

$$\log(K_d) = \log([E]^n\ [S]/[E_nS])$$

$$\log(K_d) = \log([E]^n) + \log([S]/[E_nS])$$

$$\log(K_d) - \log([E]^n) = \log([S]/[E_nS])$$

$$\log(K_d) - n\log([E]) = \log([S]/[E_nS])$$

$$\boxed{-\log(K_d) + n\log([E]) = \log([E_nS]/\ [S])}$$

From the boxed equation we can designate log[E] as the x variable and log([E_nS]/ [S]) as the y-variable and n is the slope. As such we can then plot the data obtained as log([E_nS]/ [S]) against log([E_{total}]) where the y-intercept is the -log(K_d). The boxed equation can then be simplified to the following equation:

$$y = nx - \log(K_d)$$

$$\text{y-intercept} = -\log(K_d)$$

$$K_d = 10^{\wedge -\log(Kd)}$$

As an example the following equation will be used as seen in **Figure 34**:

$$y = 1.8644x + 0.2409\ (R^2 = 0.991)$$

$$\text{y-intercept} = 0.2409 = -\log(K_d)$$

$$K_d = 10^{\wedge -0.2409}$$

$$K_d = 0.5742$$

2.26 Fluorescence binding assay

Fluorescence binding assays were conducted using 391 nM DNA and 0-7 μM protein in 6.67 mM EDTA and activity buffer (50 mM Tris-HCl (pH 7.6), 100 mM NaCl and 5 mM DTT) in a volume of 600 μL (a modified protocol from Zhang *et al* 2005[55]). Samples were equilibrated for 30 minutes and fluorescence was subsequently monitored in a 10 mm quartz cuvette using a Varian Cary Eclipse Fluorescence Spectrophotometer at 25^{o}C. Spectra were observed under the following parameters: Excitation at 346 nm with a slit width of 10nm, emission from 400-600 nm with a slit width of 5 nm, medium voltage and a CAT mode of 10.

Chapter 3: Results and discussion

3.1 Protein characterization

3.1.1 General Outline

In order to study binding and repair of ICL DNA several AGT variants were investigated. AGT variants cysteine 145 to serine (C145S) and arginine 128 to alanine (R128A) were developed from AGT wild type plasmids using site-directed mutagenesis (see section **2.2**). The C145S variant, an inactive variant that is unable to repair damaged DNA, was prepared as a standard for binding assays. R128A variant was prepared to observe the importance of the arginine "finger" in repairing and binding of ICL DNA. A control AGT homolog, Ada-C, was employed as it is well known that it cannot repair ICL DNA, but its binding mode may aid in understanding AGT-ICL association. All proteins were purified as is indicated in section **2.6**. Once purified, proteins underwent several biophysical techniques that are required for identification, monitoring stability and changes in secondary structure.

3.1.2 Mass spectrometry

Identification of AGT wild type, C145S, R128A and Ada-C was verified using Waters Micromass Q-TOF-2 mass spectrometer (see section **2.8**). Experimentally determined masses were compared to calculated masses based on the ProtParam tool on the ExPASy site (see **Table 4**). All experimental masses correlated closely to the expected masses with AGT wild type, C145S, R128A and Ada-C having mass differences of 1.2Da, 2.5Da, 0.4Da and 1.4Da respectively. Identity was further

confirmed by analysis of purified proteins using SDS-PAGE gel electrophoresis (see **Figure 15**).

Proteins	Expected mass (Daltons)	Observed mass (Daltons)
hAGT wild type	21876.2	21875.0
C145S	21860.1	21862.6
R128A	21791.0	21789.6
Ada-C (*E.coli* homolog)	20709.5	20708.1

Table 4 MS analysis results of AGT variants, comparing expected against observed masses.

The observed and calculated masses correlate well confirming the identity of each AGT variant. There were peaks of greater mass observed, with mass increases of approximately 23Da, which is representative of sodium ions that are bound to the protein.

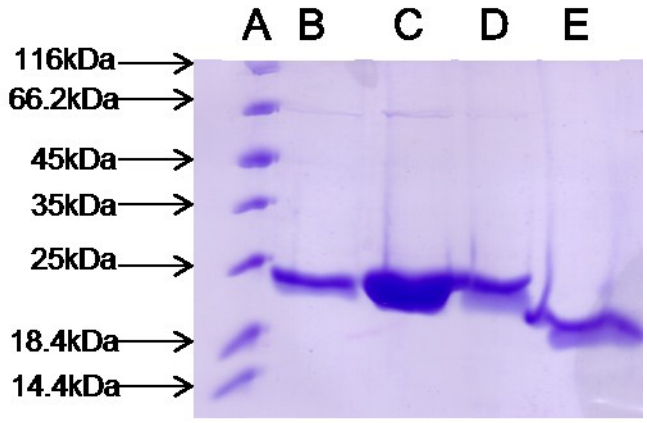


Figure 15 SDS-PAGE analysis of purified proteins wild type AGT (lane B), C145S (lane C), R128A (lane D) and Ada-C (lane E) with a low molecular weight marker (lane A).

3.1.3 Structural analysis comparison and stability

Far-UV circular dichroism was used to compare differences in secondary structure and stability between the AGT variants. Using mean residue ellipticity, we can compare the CD signal observed for proteins of differing number of amino acids of

varying concentration. We observed an increase in CD signal for R128A mutant and Ada-C homologs. In addition, there appears to be a minor loss of CD signal for C145S (see **Figure 16**). It must be noted that although there are some differences between AGT (207 amino acids) and its human variants, the shape of the curve remains very similar, suggesting similar conformation. Ada-C, *E. coli* homolog (188 amino acids), shows significant difference in CD signal as compared to AGT. It must be noted that circular dichroism is used as a comparative tool in this instance, and as such only AGT and its variants (C145S and R128A) can be compared. However, we can say that all proteins are clearly showing secondary structure and are therefore folded and stable at room temperature in the buffer system.

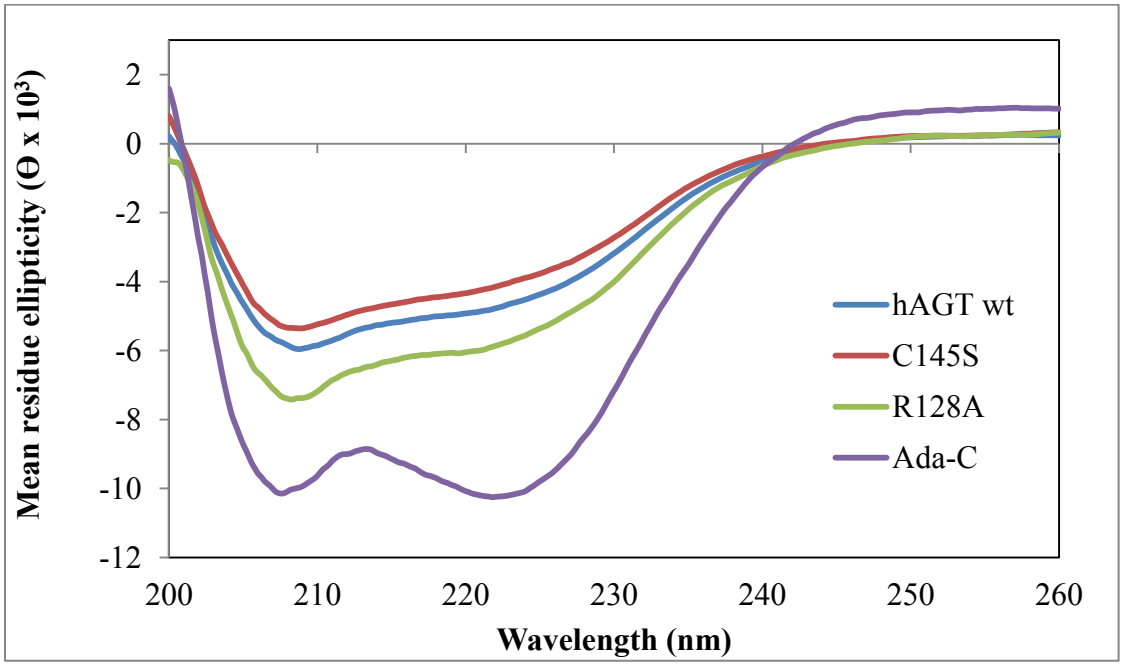


Figure 16 Far-UV CD signal of AGT variants.

In addition to comparing secondary structure signal, we can determine the relative stability of a protein based on thermal denaturation studies. CD signal was measured at 222nm with increasing temperature to measure the decrease in secondary structure signal

over time. This allows us to monitor the effect a mutation may have on the stability of the protein variant. We observed a significant difference in stabilities due to single point mutations. Among the AGT human variants we determined that C145S showed the lowest stability with a T_m of 55.2°C, followed by AGT wild type with a T_m of 59.0°C and the most thermally stable protein was R128A with a T_m of 62.2°C (see **Figure 17**). Ada-C showed the lowest overall thermal stability with a T_m of 49.2°C; however this only represents a fraction of unfolded protein (approximately 60%).

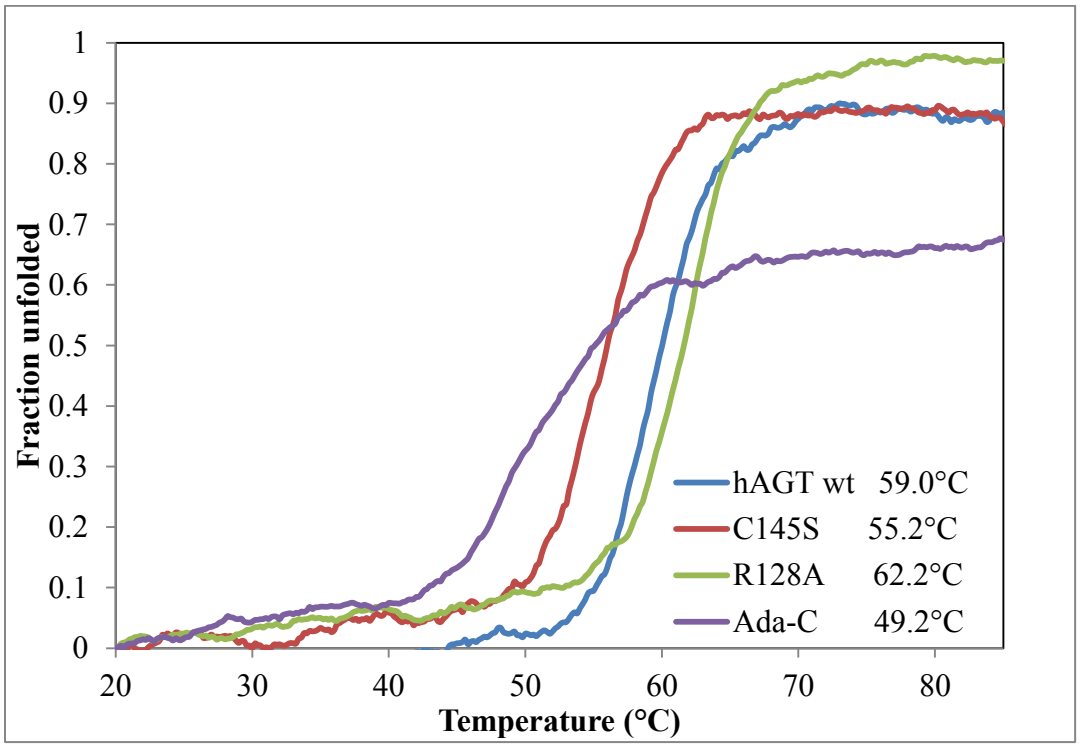


Figure 17 Thermal denaturation studies of AGT and its variants, normalized to fraction unfolded.

3.1.4 Summary

AGT proteins were purified and clearly identified using MS Q-TOF. In addition, all proteins show a folded state in the buffer systems used, with thermal stability studies showing all proteins to be stable for repair and binding assay conditions (37°C). Upon the

completion of characterization of the AGT wild type and variants, we could begin to characterize the DNA substrates, bringing us closer to binding and repair assays.

3.2 Preparation of ICL DNA probes

3.2.1 Tri-partite DNA

3.2.1.1 General

ICL DNA sequences were designed such that the O^6-2'-deoxyguanosine-heptylene-O^6-2'-deoxyguanosine (O^6-dG-heptylene-O^6-dG) crosslink was centrally positioned and in a 5'-GNC-3' motif. In addition, an oligonucleotide containing an asymmetrical diverse base composition was chosen to simulate the environment of genomic DNA. Although, cut-sites are often included to observe repair of mono-adducts, no such sites were incorporated into the ICL DNA as repair could be easily monitored using PAGE gels because there is a major shift in band migration when comparing repaired ssDNA and ICL DNA. Furthermore cut-sites increased symmetry resulting in possible undesired hairpin formation. The sequence also took into consideration specific placement of dC residues that could be replaced with a pyrrolo-dC fluorescent analog, with sites positioned centrally and near the 5'-end. In order to incorporate 3 sites for pyrrolo-dC insertion, a relatively long oligonucleotide was desired (18 bases ssDNA and 36 bases dsDNA). However if the DNA sequence was too long, yields for ICL DNA would be decreased as compared to synthesis of a shorter sequence, and the ability to form competing hairpin structures also increases.

For intermediates **A**, **B**, **C** and **D** (see **Figure 19**), it was necessary to introduce a protective group that could be retained under conditions to remove the 5'-OTBS and 3'-

OAlloc groups in order to avoid undesired chain extension during solid phase synthesis, avoiding various oligonucleotide side products (see **Figure 18 bottom graph**). The protective group chosen was levulenyl (Lev), which was resistant to F^- and palladium treatment. This group was introduced using 3'-deoxyphosphoramidites containing a 5'-Olev protecting group. Initial attempts used 5'-OLev-dG and 5'-OLev-dT, which failed due to the mixture of levunilated and non-levunilated amidites (see **Figure 18 top graph**), where 100% coupling of the crosslinked amidite and no mixture of Lev amidites would result in one peak. Through optimization we were able to determine that 5'-OLev-dC was able to cap the 5'-ends and avoid extension. All ICL DNA were synthesized, purified and biophysically analyzed to reveal purity, structure and stability.

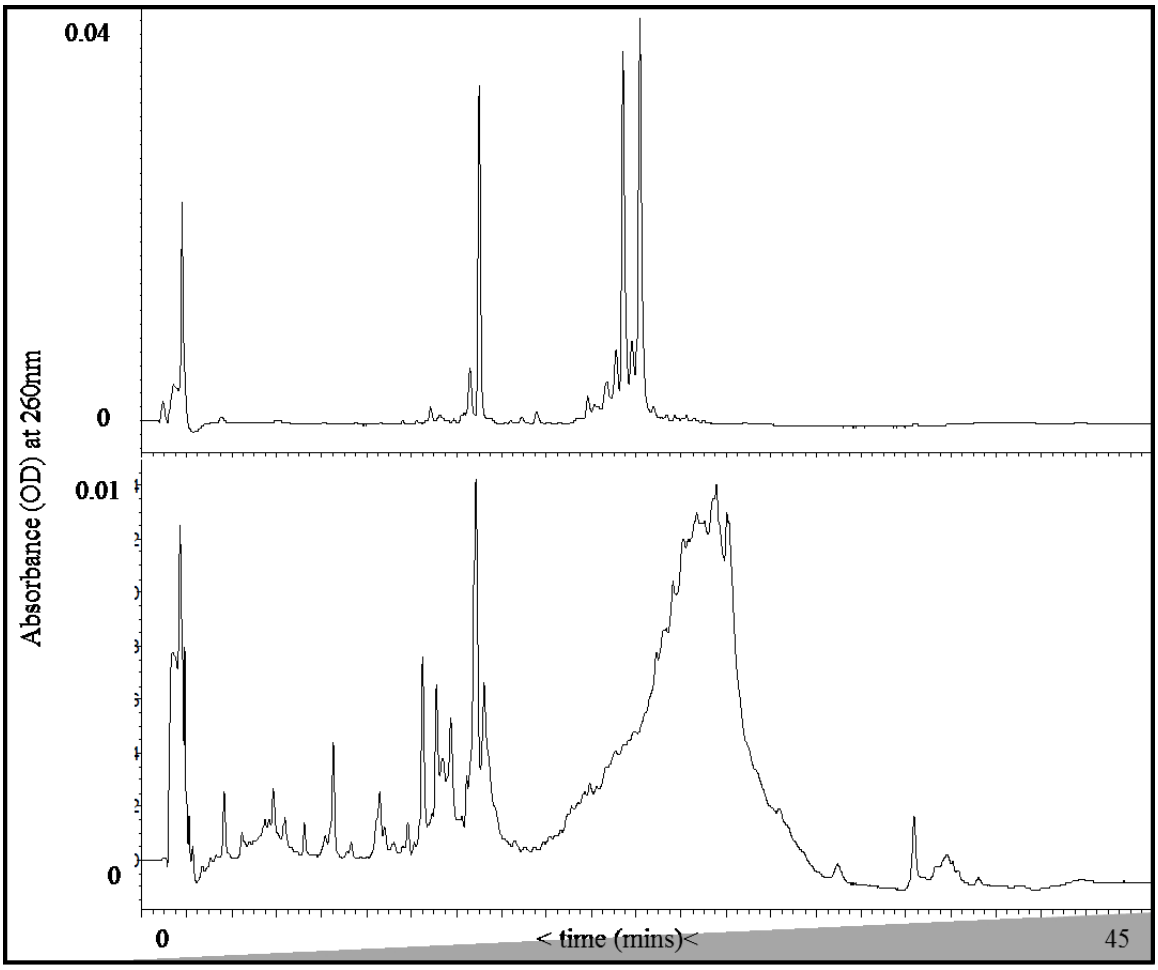


Figure 18 IEX-HPLC analysis of T-form ICL DNA after treatment with TEA/HF for TBS deprotection (top), with the two major peaks representative of the coupling of levulinyl containing dG amidite, and analysis of H-form ICL DNA with the desired product situated within the large broad peak (bottom).

3.2.1.2 ICL DNA Synthesis

Controls were synthesized with high yields and well behaved, and hence will not be discussed in this section. The major issue with ICL DNA synthesis was step (ii) (coupling of dG-dG amidite, see **Figure 19**). Initially coupling was as poor as 10%, resulting in extremely low yield of final product. Extra purification of the amidite was conducted to ensure an extra pure sample. Due to the scarcity of the crosslink amidite we were unable to attempt synthesis with a higher concentration. However, extra dry reagents, increased coupling time and a highly pure amidite were used to improve

coupling. Unfortunately coupling was only able to improve to a maximum of 50% resulting in relatively low overall yields of ICL DNA. It is possible that due to the amount of protecting groups and the bulky crosslinked dG group that steric clashing could have reduced coupling yield.

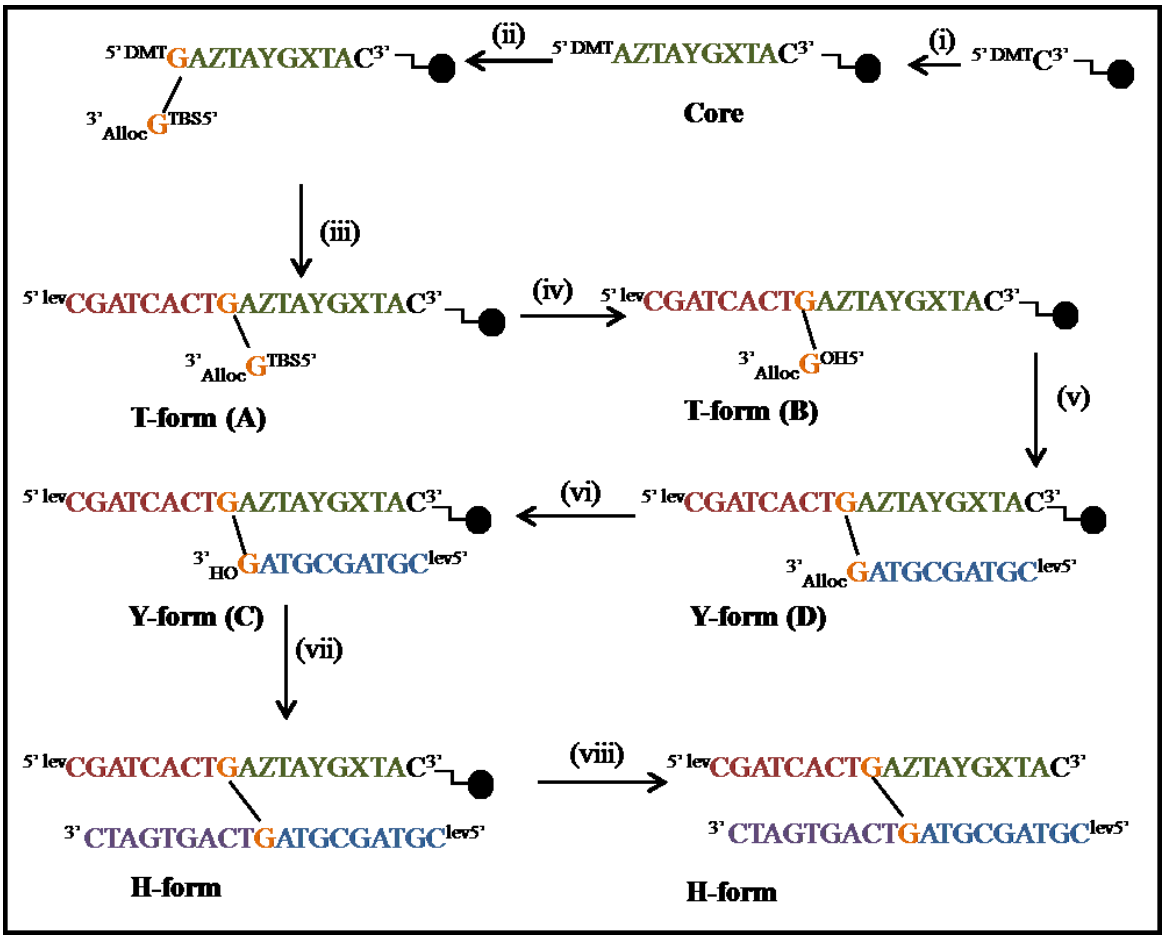


Figure 19 ICL DNA synthesis scheme, with X, Y and Z designating possible sites of incorporation of pyrrolo-dC, where X, Y, Z= dC for wild type, X, Y= dC and Z= pyrrolo-dC for sequence CJW293, X, Z= dC and Y= pyrrolo-dC for sequence CJW294, and Y, Z= dC and X= pyrrolo-dC for sequence CJW295. Steps: (i) extension to the core, (ii) ICL amidite coupling, (iii) extension to the T-form, (iv) deprotection of the silyl group TBS, (v) extension to the Y-form, (vi) deprotection of the alloc group, (vii) extension using 5'-amidites to the final H-form DNA and (viii) deprotection from solid support and all protecting groups with ammonia/ethanol.

3.2.1.3 Purification

ICL DNA purification was performed using polyacrylamide gels. Surprisingly two major peaks were observed which shall be referred to as top and bottom bands, while the lowest band is representative of the "core" that did not couple with the dG-dG amidite (see **Figure 20**). Interestingly, the top band is not detectable under normal IEX-HPLC conditions and does not co-elute with the bottom band which is detectable, hence purification by gel. The top and bottom bands were excised for further analysis, because unlike simple DNA synthesis the upper most bands are not necessarily the desired product due to the size of the DNA and the complexity of its synthesis. The reason for the two major products is unclear, as the shift is more than one uncoupled base and yet they also migrate too closely for one to be the Y-form over the H-form of DNA. It was also postulated that perhaps due to its size it was able to form two overall distinct secondary structures that resulted in two bands. However, an analytical polyacrylamide gel was attempted with denatured and renatured ICL DNA strands with little success, as the same shift in band migration was observed. Isolated yields ranged from 6-10% and further biophysical analysis was completed to discover the true identity of the top and bottom bands.

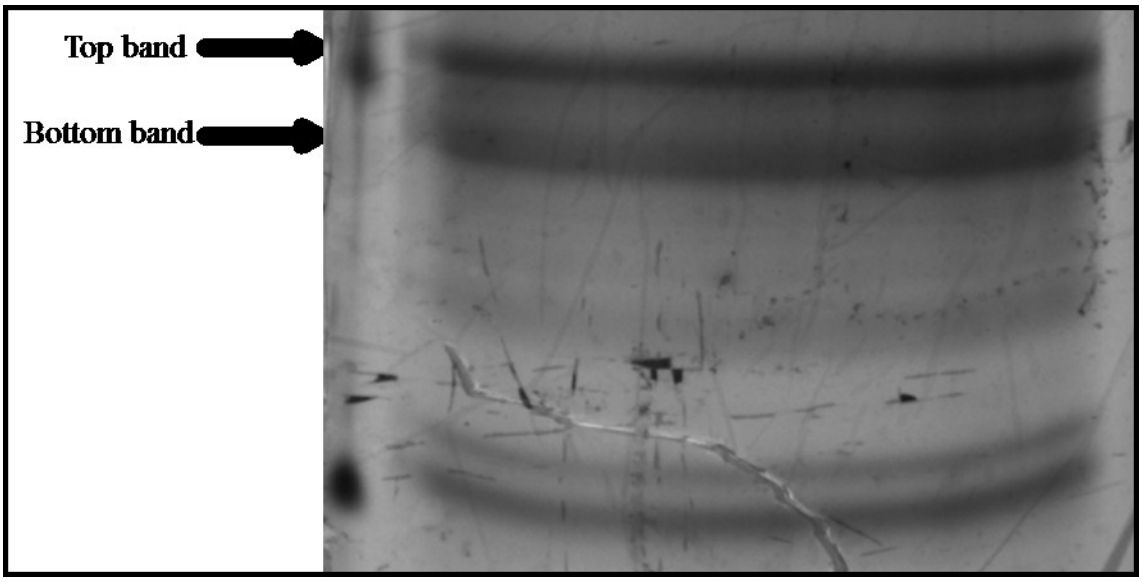
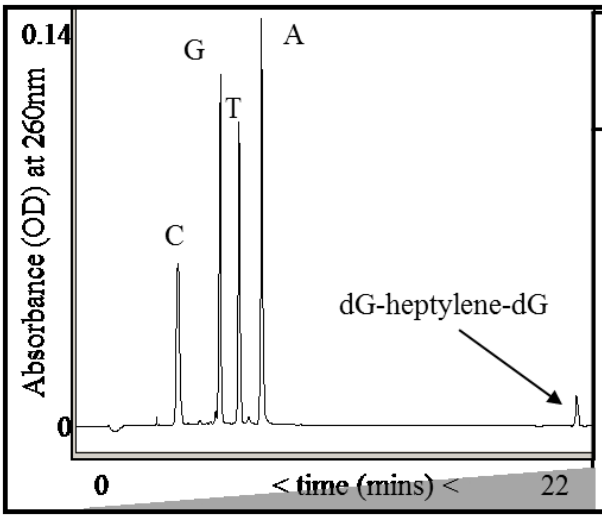


Figure 20 Polyacrylamide purification gel of ICL DNA with the desired product as either the top or bottom bands as designated, where the lowest band is designated as core DNA that did not couple with the crosslinked amidite.

3.2.1.4 DNA nucleoside composition

Nuclease digestion and subsequent RP-HPLC analysis was used to observe if the nucleoside ratios agreed to expected values and more importantly the presence of the desired pyrrolo-dC and crosslinked nucelosides (at approximately 22 minutes). DNA composition analysis showed similar nucleoside ratios for all ICL DNA of both top and bottom bands. In addition, pyrrolo-dC did not appear in DNA composition analyses when it should have been present in ICL DNA (see **Figure 21**). This presented two problems, why was pyrrolo-dC not appearing and was there some other characterization method that could identify the desired product. Firstly, pyrrolo-dC not appearing could be due to possible breakdown during ICL DNA synthesis, most likely at either the deprotection of TBS or alloc groups. However, based on the bi-partite project pyrrolo-dC is stable under the de-silylation conditions (see **Section 3.2.2**). It is uncertain if the deprotection of alloc caused a breakdown of the pyrrolo-dC base; more analysis is

required. Secondly, MS technique LTQ Orbitrap Velos was considered as a method of characterization. However, due to the size of the ICL DNA we were unable to determine a mass, where we were able to determine the mass of smaller dsDNA or ssDNA molecules. Nucleoside DNA composition analysis did not aid in discriminating the identity of either the top or bottom bands. It was then believed that thermal denaturation and possibly CD could help to elucidate the identities of ICL DNA top and bottom bands.



Nucleoside	Nucleoside Ratios	
	Expected	Observed
dC	10.00	9.37
dG	6.00	7.35
dT	9.00	9.00
dA	9.00	7.76
Pyrrolo-dC	1.00	0
dG-heptylene-dG	1.00	1.28

Figure 21 ICL DNA composition analysis. All ICL DNA (top and bottom bands) with and without pyrrolo-dC showed similar ratios.

3.2.2 Bi-partite DNA

3.2.2.1 General

Being unable to have a fully characterized tri-partite ICL DNA, we decided to design a bi-partite ICL DNA sequence. This new shorter oligonucleotide retained the core of the tri-partite sequence including the O^6-2'-deoxyguanosine-heptylene-O^6-2'-deoxyguanosine (O^6-dG-heptylene-O^6-dG) crosslink that was centrally positioned, with only two pyrrolo-dC incorporation sites. Furthermore, the base composition was modified to become symmetrical, except at the 5'-GNC-3' motif. This simpler bi-partite synthesis

did not require levulenyl capping at the 5'-end. All ICL DNA were synthesized, purified and biophysically analyzed to reveal purity, structure and stability.

3.2.2.2 ICL DNA Synthesis

ICL DNA and controls were synthesized as described in section **2.12**. Control DNA synthesis was straightforward with all oligonucleotides purified by gel and characterized, and hence will not be covered in this section. ICL DNA synthesis was completed in several stages as is outlined in **Figure 22** with steps (i) to (v). The steps that proved most difficult due to poor coupling yield were steps (ii) and (v). Step (ii) is the coupling of the modified O^6-dG-heptylene-O^6-dG amidite with what we refer to as the core of the ICL DNA. Coupling tended to be initially low, however, with extra purification of the amidite (completed by Dr. C. Wilds) coupled with use of extra dry reagents and longer coupling time we were able to increase coupling yields as impurities and moisture could react with the prepared amidite Coupling was improved to above 90%.

Step (v) involves the use of 5'-amidites, which usually require longer coupling time than regular 3'-amidites. However, even with increased time there appeared to be poor coupling during extension resulting in increased amount of side products. We attempted to increase coupling with a higher concentration of 5'-amidite, dry reagents and a 30 second increase in coupling time, yet little improvement was observed. It is unclear why the 5'-amidites did not couple at 90% and higher, yet it appears that it could be sequence dependent as poor coupling tended to occur only with the sequences highlighted in this project, but not in other oligonucleotides with differing sequences (not

shown). 5'-amidite coupling could also be hampered due to the steric hindrance of the opposing arm of DNA that could potentially block the coupling site.

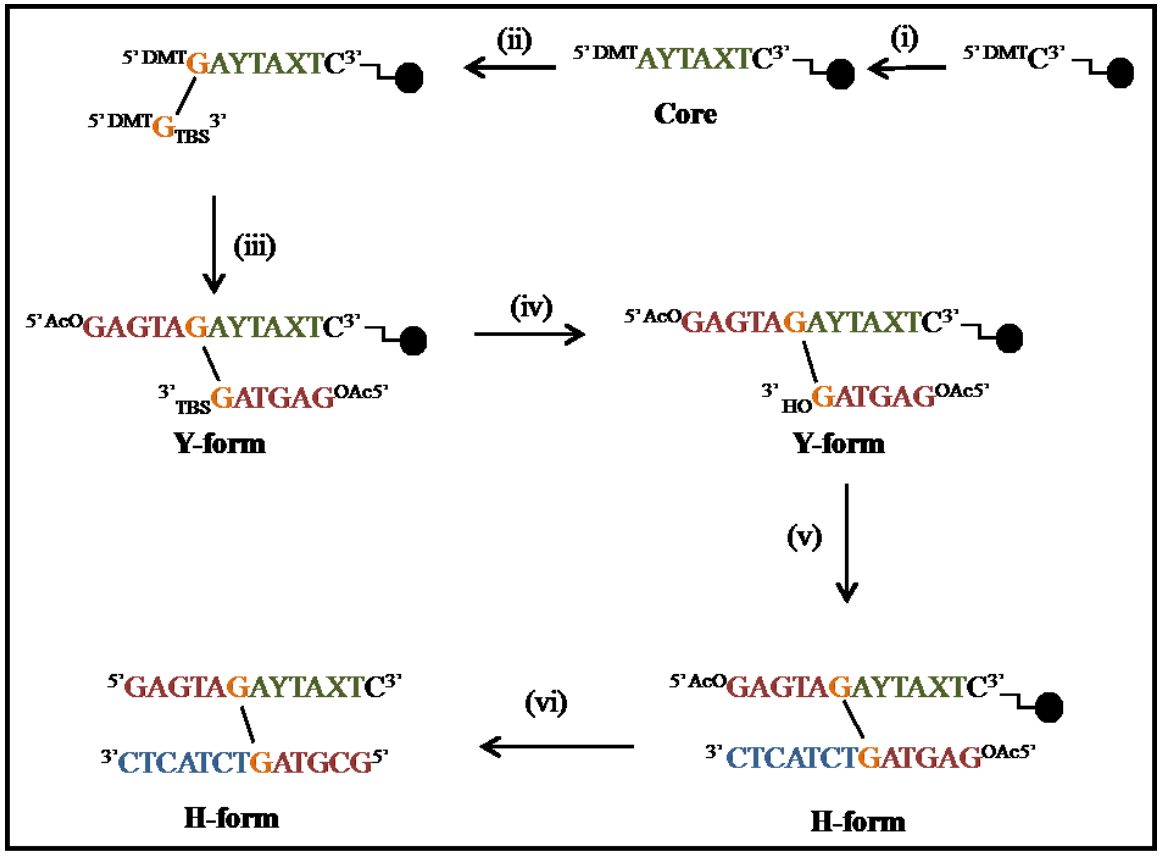


Figure 22 Synthesis scheme of ICL DNA, where X designates positions of pyrrolo-dC incorporation, where X, Y= dC for wild type CJW328, X= dC and Y= pyrrolo-dC for sequence CJW326 (ICL pyrrolo1), and Y= dC and X= pyrrolo-dC for sequence CJW327 (ICLpyrrolo2). Steps: (i) extension to the core, (ii) ICL amidite coupling, (iii) extension to the Y-form, (iv) deprotection of the silyl group TBS, (v) extension using 5'-amidites to the final H-form DNA and (vi) deprotection from solid support and all protecting groups with ammonia/ethanol.

3.2.2.3 De-silylation reaction

An intermediate step of ICL DNA synthesis was the removal of a silyl group on the 5'-end of the dG-heptylene-dG group. Removal of the *tert*-butyldimethyl silyl group (TBS) was monitored using HPLC. First batch of ICL DNA used RP-HPLC to observe

TBS removal, while IEX-HPLC was used in the analysis of the second batch of ICL DNA synthesis.

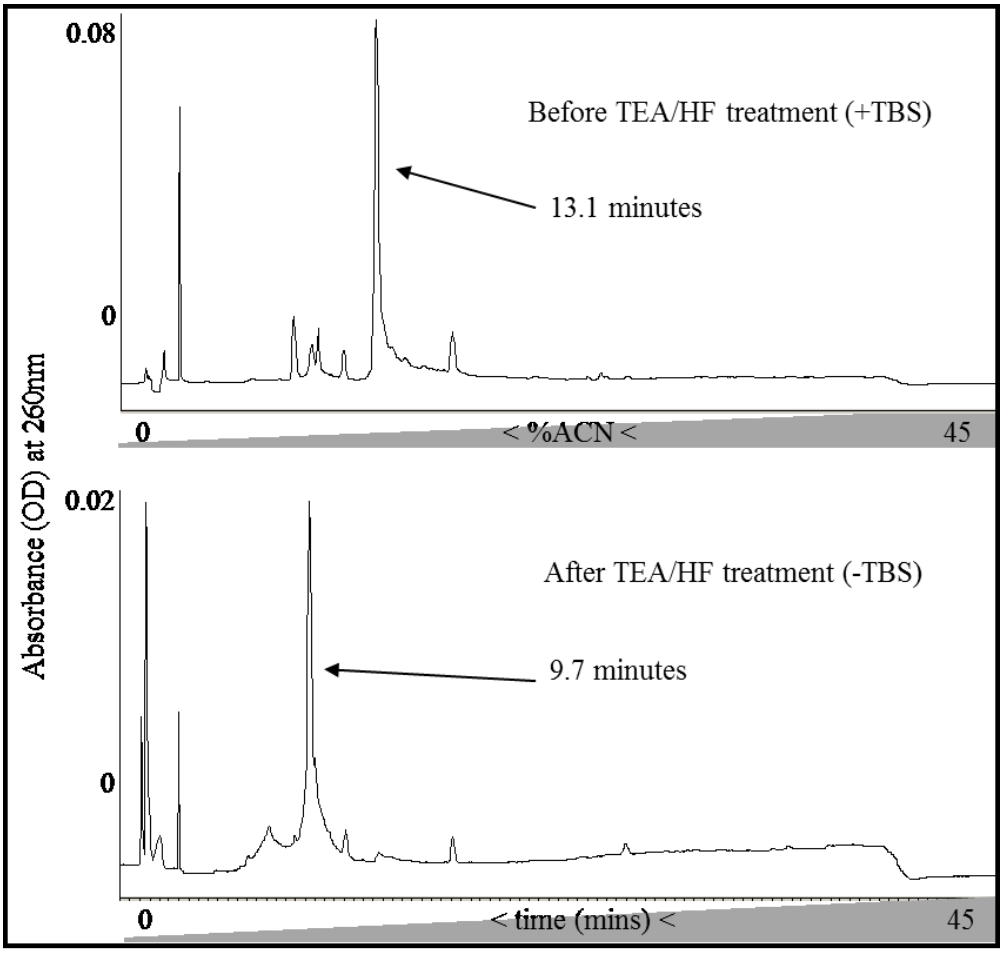


Figure 23 RP-HPLC chromatograms of oligonucleotides of ICLpyrrolo1 DNA in the presence of TBS protecting group and after the treatment with TEA/HF to remove TBS.

RP-HPLC chromatograms showed a clear shift in peak from 13.11 minutes to 9.71 minutes (see **Figure 23**) after treatment with TEA/HF to remove the TBS group. This was consistent for all 3 ICL DNAs: ICL, ICLpyrrolo1 and ICLpyrrolo2. Due to the consistent shift in peak we believed this confirmed TBS removal. In addition, the decrease in retention time upon TBS removal is expected as TBS increases the hydrophobicity of the oligonucleotide. IEX-HPLC chromatograms, on the other hand, displayed a minor shift in peak elution time in samples with and without TBS (see **Figure**

24). Although the shift was minor (0.6 minute difference), which is not unexpected as retention times will not greatly increase or decrease with the removal of one TBS group per ICL molecule as the association between the analyte and the stationary phase is primarily based upon ionic interactions. The change was consistent for all ICL samples.

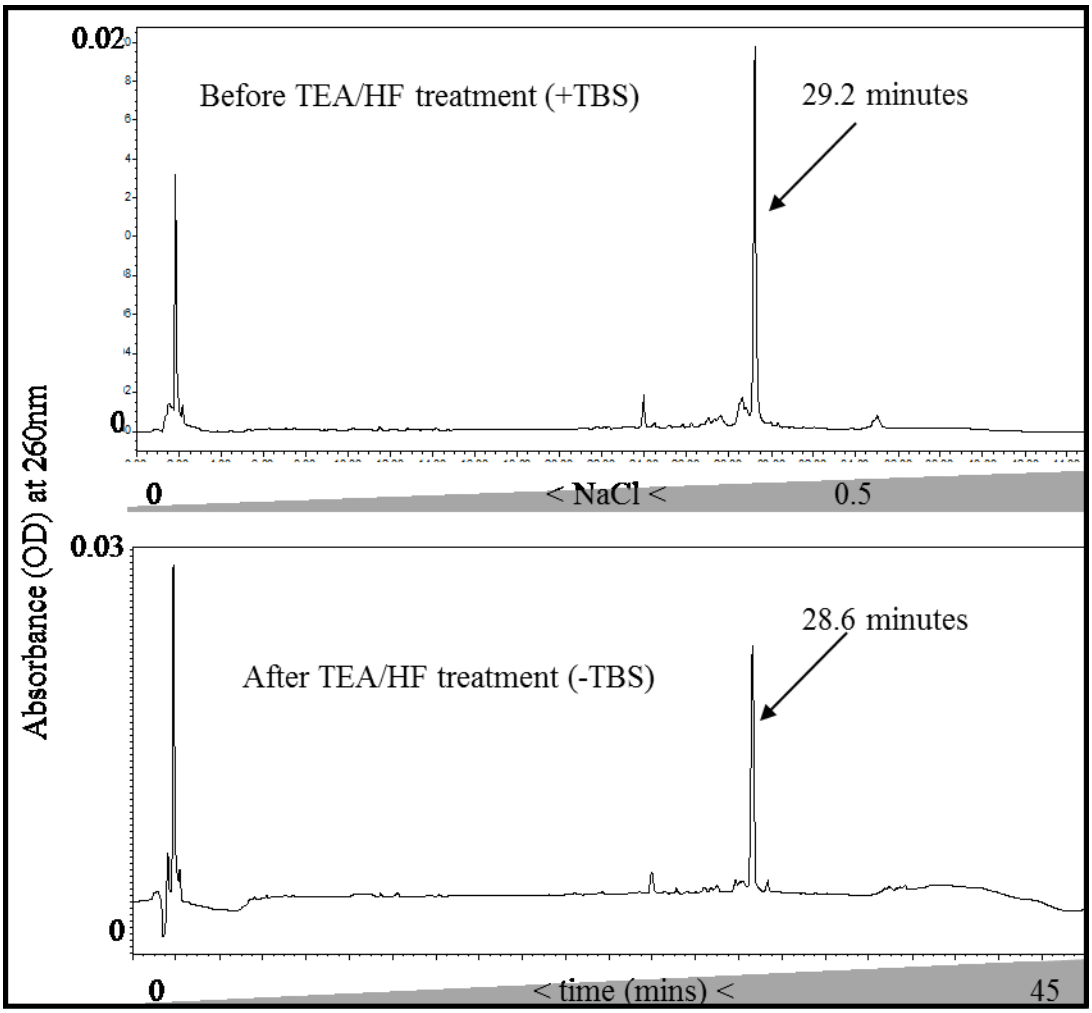


Figure 24 IEX-HPLC chromatograms of oligonucleotides of ICLpyrrolo1 DNA in the presence of TBS protecting group and after the treatment with TEA/HF to remove TBS.

3.2.2.4 Purification

Upon completion of ICL DNA synthesis and subsequent deprotection the samples were analyzed by IEX-HPLC to evaluate the overall efficiency of the synthesis. HPLC

analysis on an analytical scale yields information regarding the overall synthesis and whether modifications to the HPLC conditions (buffer, gradient etc.) must be made. The IEX-HPLC chromatograms revealed the presence of by-products (see **Figure 25**). Thus this method could not be used for initial purification because of the close proximity of highly abundant side-products and the challenge in separating them as in preparation runs the amount of sample loaded is increased 30 fold, resulting in peak broadening and overlap. Fortunately, analysis by PAGE of the crude demonstrated improved resolution of the products. Thus the oligonucleotides were purified by gel electrophoresis.

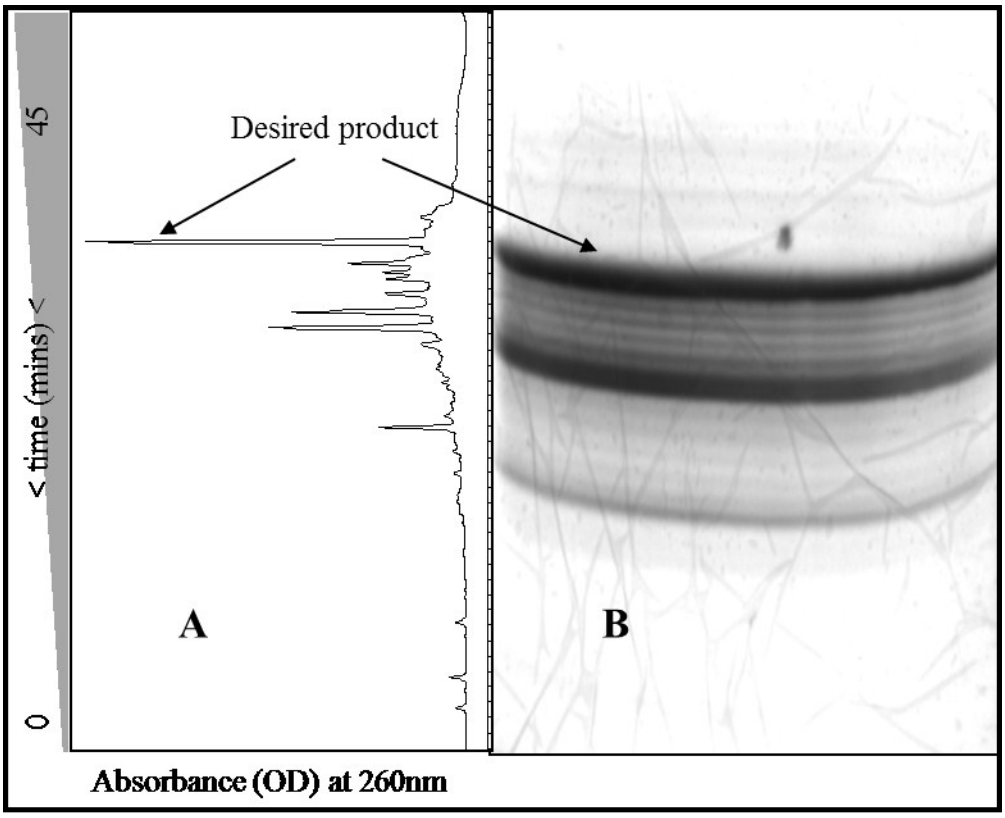


Figure 25 Comparison of IEX-HPLC analytical analysis of crude ICL DNA sample (A) and 20% polyacrylamide purification gel (B). The final product or H-form of the ICL DNA is designated by the arrows.

After purification by PAGE and desalting the ICL DNA samples were analyzed by HPLC. The chromatograms (see **Figure 26**) indicated the presence of impurities that

eluted slightly before and after the desired product. Additional purification was attempted with a pH12.5 buffer, however, it has not yet been confirmed if the crosslink was unaffected. Furthermore, multiple attempts at improving overall synthesis did not significantly reduce by-product impurities, resulting in the constant presence of minor impurities. Moreover, yields purified were low, ranging from 11.7%-23.3%, making a second round of purification to all samples problematic due to the need for relatively high amounts of DNA sample (391 nM DNA/600 μL sample) needed for fluorescent binding studies.

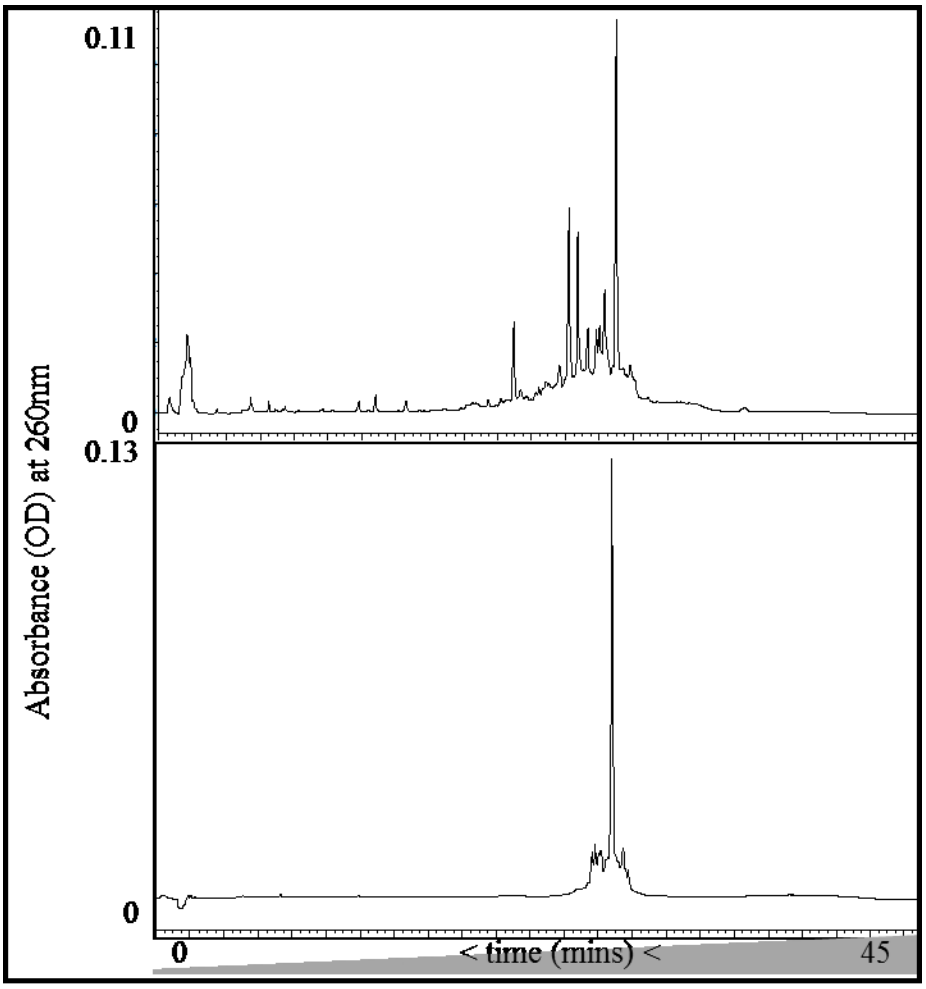
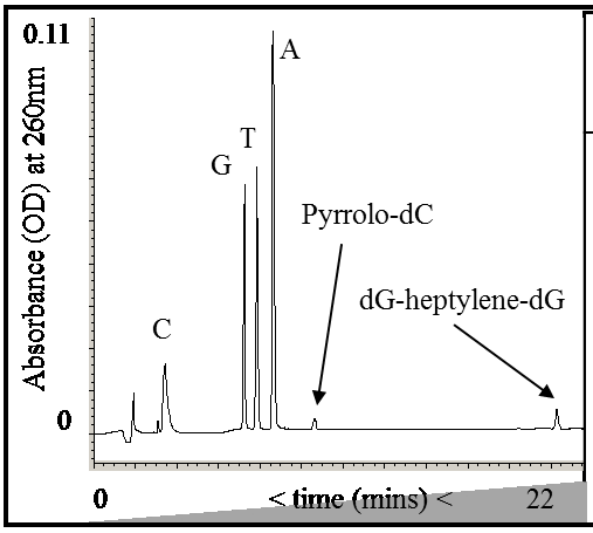


Figure 26 Comparison of IEX-HPLC analytical analysis of crude ICL DNA sample (top) and analytical analysis of pure ICL DNA sample (bottom).

3.2.2.5 DNA nucleoside composition

Verification of the presence of the crosslink and pyrrolo-dC nucleosides, in addition to the overall nucleoside content, was verified by nuclease digest and by RP-HPLC analysis. In addition to the natural four nucleosides, the digests of the sequences displayed the presence of both pyrrolo-dC and dG-heptylene-dG crosslinked nucleosides with peaks at approximately 10.5 and 22 minutes respectively (see **Figure 27**). The absence of additional species suggests the fluorescent base analog pyrrolo-dC showed chemical stability under DNA synthesis and deprotection conditions used for ICL DNA synthesis. The digest data for all ICL sequences revealed they contained the correct nucleoside composition with good agreement between the expected and observed ratios. All linear sequences that were synthesized were also found to have the correct nucleoside composition and the presence of an additional nucleoside corresponding to pyrrolo-dC (not shown).



Nucleoside	Nucleoside Ratios	
	Expected	Observed
dC	5.00	5.03
dG	4.00	4.34
dT	7.00	7.00
dA	7.00	6.82
Pyrrolo-dC	1.00	1.17
dG-heptylene-dG	1.00	0.80

Figure 27 DNA composition of ICL DNA containing pyrrolo-dC. RP-HPLC chromatogram (left) and nucleoside ratios (right)

3.2.3 Summary

Synthesis of the tri-partite ICL DNA was complex with the dG-dG coupling step being the most difficult followed by the capping residue containing a levulinyl group. Although the crosslinked amidite coupling was increased to 50% and levulinyl capping problem solved, purification revealed the presence of two major products (top and bottom bands). Characterisation was challenging as all ICL DNA containing pyrrolo-dC showed similar nucleoside composition.

Bi-partite ICL DNA was synthesized and purified to relatively high purity with a major set-back in yields due to poor 5'-amidite coupling. Composition analysis showed the presence of crosslinked amidite and pyrrolo-dC with the expected ratios.

3.3 Biophysical analysis of DNA duplexes

3.3.1 Tri-partite DNA

3.3.1.1 Thermal denaturation studies

Thermal denaturation studies were used to observe thermal stability and possible changes that that may occur with the incorporation of modified bases. ICL DNA samples revealed significant variations of melting temperatures for bottom band ICL DNAs with values ranging from 54.2^{o}C to 66.0^{o}C, while top band ICL DNA displayed relatively constant melting temperatures (see **Table 5**). The bottom band ICL DNA T_m variations are surprising. Could the differences be due to the change in position of the pyrrolo-dC? However, when looking at simpler ICL DNA from the bi-partite synthesis we do not observe such variances. On the other hand the top bands do show high agreement that may be more characteristic of the desired ICL DNA. However, without a complementary

method of characterization we could not continue to postulate further the reasons for such variance.

ICL DNA code	Melting temperature, T_m (^{o}C)	
	Top band	Bottom band
CJW293	61.1	54.2
CJW294	61.1	62.1
CJW295	62.0	66.0

Table 5 Thermal denaturation studies of tri-partite ICL DNA.

A special note must be given to the control sequences containing O^6-methyl-dG in contrast to unmodified DNA and sequences containing pyrrolo-dC (see **Figure 28**). Firstly, incorporation of O^6-methyl-dG decreased thermal stability of the DNA double helix from 62.1^{o}C to 54^{o}C, where both unmodified DNA and DNA containing pyrrolo-dC have the same T_m. A decrease in stability was expected as the modified base cannot base pair as efficiently as its naturally occurring dG. As expected pyrrolo-dC incorporation in the opposing strand did not result in a change of thermal stability. However, surprisingly when pyrrolo-dC is directly opposed to O^6-methyl-dG (CJW272/271) we observed an increase of 3^{o}C in thermal stability. It appears that pyrrolo-dC is able to partially reduce the instability caused by the modified dG. This possibly could be due to the extra aromatic ring that pyrrolo-dC has and dC does not, which may lead to increased stacking interactions, however more analysis is required.

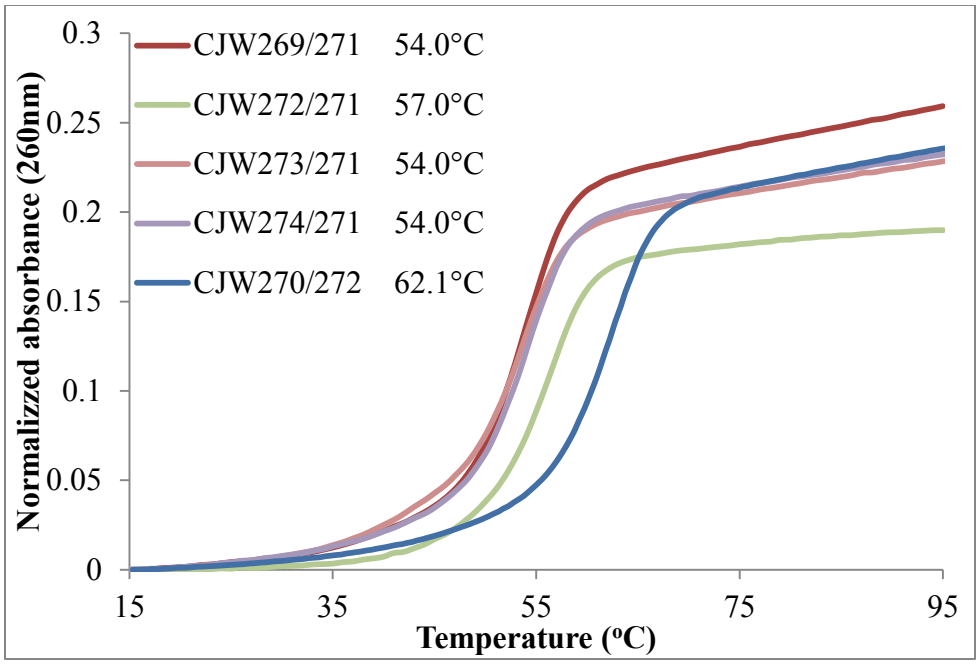


Figure 28 Thermal denaturation studies of DNA controls containing O^6-methyl-dG (CJW271), unmodified DNA (CJW269, CJW270) and pyrrolo-dC directly opposed to O^6-methyl-dG (CJW272) or positions flanking O^6-methyl-dG (CJW273, CJW274).

3.3.1.2 Structural analysis

Circular dichroism was used to determine DNA structure and to observe any changes that may have occurred upon incorporation of modified nucleosides. All controls with and without pyrrolo-dC or the O^6-methyl-dG showed comparable B-form conformations. Although the height of the peaks and valleys vary in size (at 250nm and 280 nm respectively), the shape remains similar and therefore all DNA are considered to be B-form (see **Figure 29**). The differences between the three samples are not considerable enough to suggest a major change in conformation, for example: a transition from B-form to A-form DNA. Structural analysis confirmed that all ICL DNA (both top and bottom bands) retained B-form conformation and that incorporation did not result in major structural double helix distortion. Unfortunately CD analysis of the ICL DNAs did not help to elucidate the desired product between top and bottom bands.

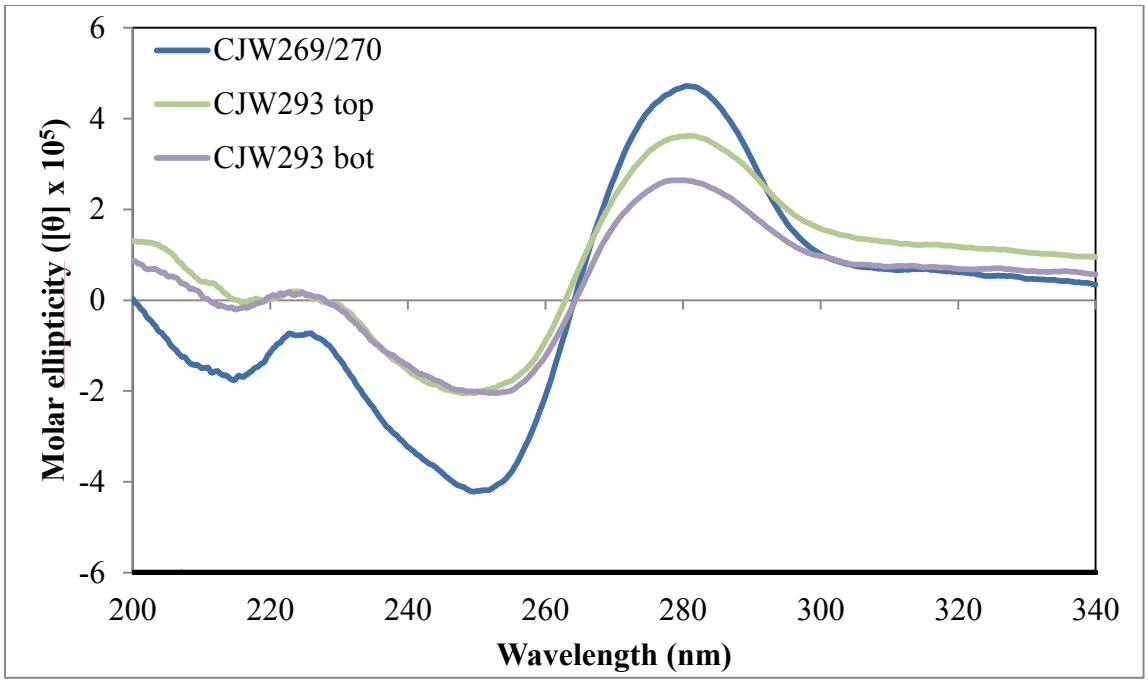


Figure 29 CD comparison of ICL DNA top and bottom bands with unmodified DNA for CJW293.

3.3.2 Bi-partite DNA

3.3.2.1 Thermal denaturation studies

In order to observe the effect of the pyrrolo-dC modification in O^6-dG-heptylene-O^6-dG ICL DNA, thermal denaturation studies were conducted. We observed an increase in thermal stability with the presence of the crosslink, with an increase of 4-6^{o}C relative to the corresponding unmodified (not crosslinked) DNA duplex. This increase may be due to the covalent linkage that pre-organizes the duplex (see **Figure 30**). The ICL duplexes containing the pyrrolo-dC exhibit T_m values, using the first derivative maxima, similar to the ICL duplex (≈ ±1^{o}C) on thermal stability of dsDNA with and without a crosslink (not shown). The T_m curves suggest that the ICL duplexes are not denatured under binding and repair assay conditions of 37^{o}C.

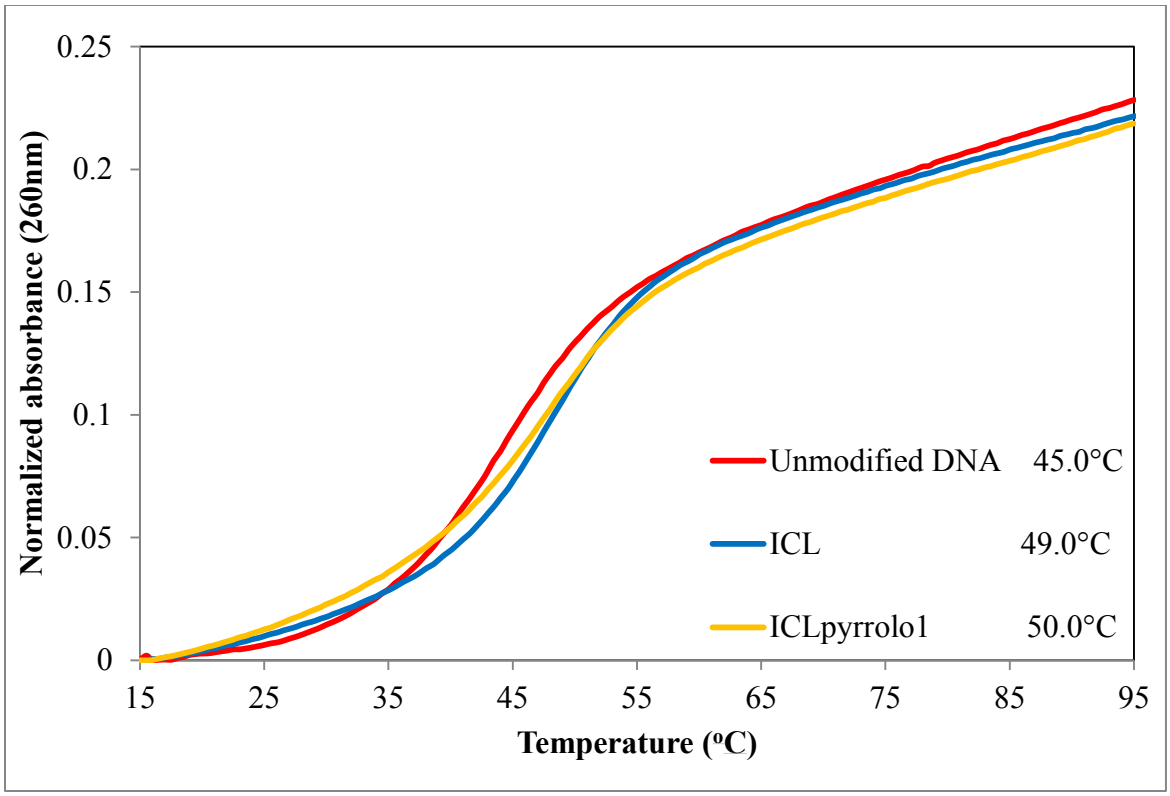


Figure 30 Thermal denaturation studies of unmodified DNA, ICL and ICLpyrrolo1.

3.3.2.2 Structural comparison analysis

To confirm that the various duplexes retained B-form conformation and that incorporation did not result in major structural double helix distortion, CD was used. The control (unmodified) DNA exhibited a large positive signal at 280nm diagnostic of a B-form duplex. Presence of the ICL and pyrrolo-dC does not appear to have any significant effect on structure. All DNA showed comparable B-form conformation CD signals (see **Figure 31**).

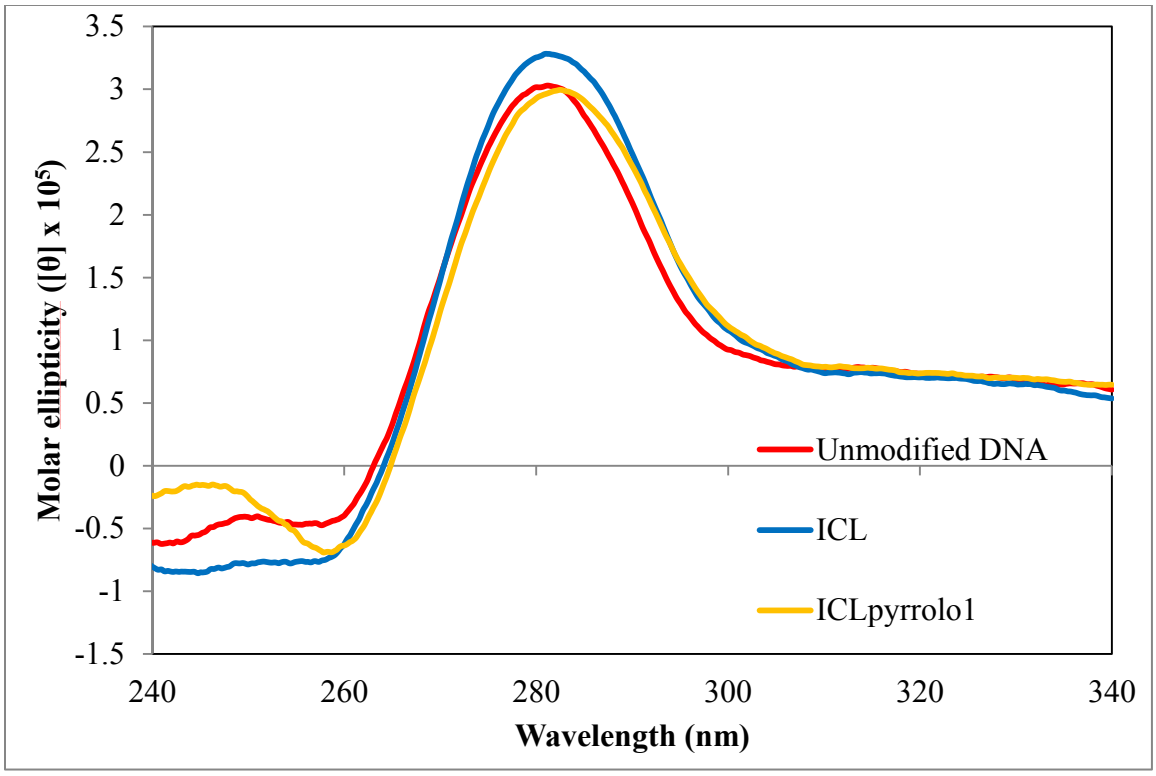


Figure 31 Far-UV CD signal comparison between unmodified DNA, ICL and ICLpyrrolo1.

3.3.3 Summary

Characterisation was challenging as all tri-partite ICL DNA containing pyrrolo-dC showed similar CD data (all B-form). Furthermore, thermal denaturation studies did not conclusively reveal which major product contains the desired ICL, with major variance in melting temperature for bottom band samples. Therefore, another mode of identification of ICL DNA is needed. In addition, a special note was taken for pyrrolo-dC's ability to reduce instability caused by O^6-methyl-dG when directly opposed by showing an increase in thermal stability of 3°C. Further optimization of the tri-partite synthesis is required and could lead to the development of more complex DNA substrates. In addition, all bi-partite ICL oligonucleotides demonstrated favourable and increased thermal stability without major disruption of the DNA double helix as determined by CD.

3.4 Repair, Binding and fluorescence studies of bi-partite ICL DNA with AGTs

3.4.1 Repair studies

Radioactivity repair assays were conducted to observe whether the presence of the pyrrolo-dC affected ICL repair. Initial overnight repair studies indicated that Ada-C does not repair ICL DNA or form any median products. In addition, there was an unexpected issue with the C145S protein variant showing an apparent degradation of the ICL DNA. It was hypothesized that the C145S mutant was able to scavenge a divalent cation (possibly zinc or magnesium) that increased the protein's activity. To eliminate this, assays were performed with the addition of EDTA. Analysis of the labelled products by PAGE revealed that EDTA did in fact remove C145S's apparent activity without significantly affecting wild type AGT's ability to repair ICL DNA (O^6-dG-heptylene-O^6-dG, see **Figure 32**).

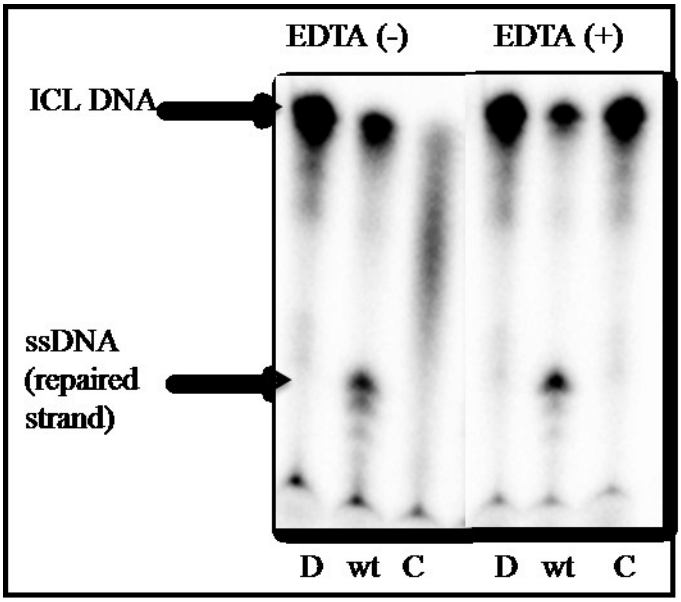


Figure 32 20% polyacrylamide gel of overnight incubation of DNA (D), AGT wild type (wt) and C145S (C) in the presence and absence of EDTA.

Time course assays of AGT wild type and variants with the substrates ICL, ICLpyrrolo1 and ICLpyrrolo2 were conducted. Percent repaired values were obtained by selecting and comparing the regions corresponding to unrepaired ICL DNA, repaired products and median products. C145S and R128A clearly were unable to repair any ICL DNA (middle and bottom graph of **Figure 33**, ICL repair data not shown). Also, the arginine "finger" (R128A) appears not only to be essential for mono-adduct repair but also ICL damage. The apparent 5% repaired ICLpyrrolo1 of the middle graph for C145S and R128A is due to slight degradation of the ICL DNA and not due to the protein. As for AGT wild type protein, pyrrolo-dC does not appear to have a significant effect on the amount of overall repair, as observed error bars tend to overlap around 120 minutes (top graph **Figure 33**).

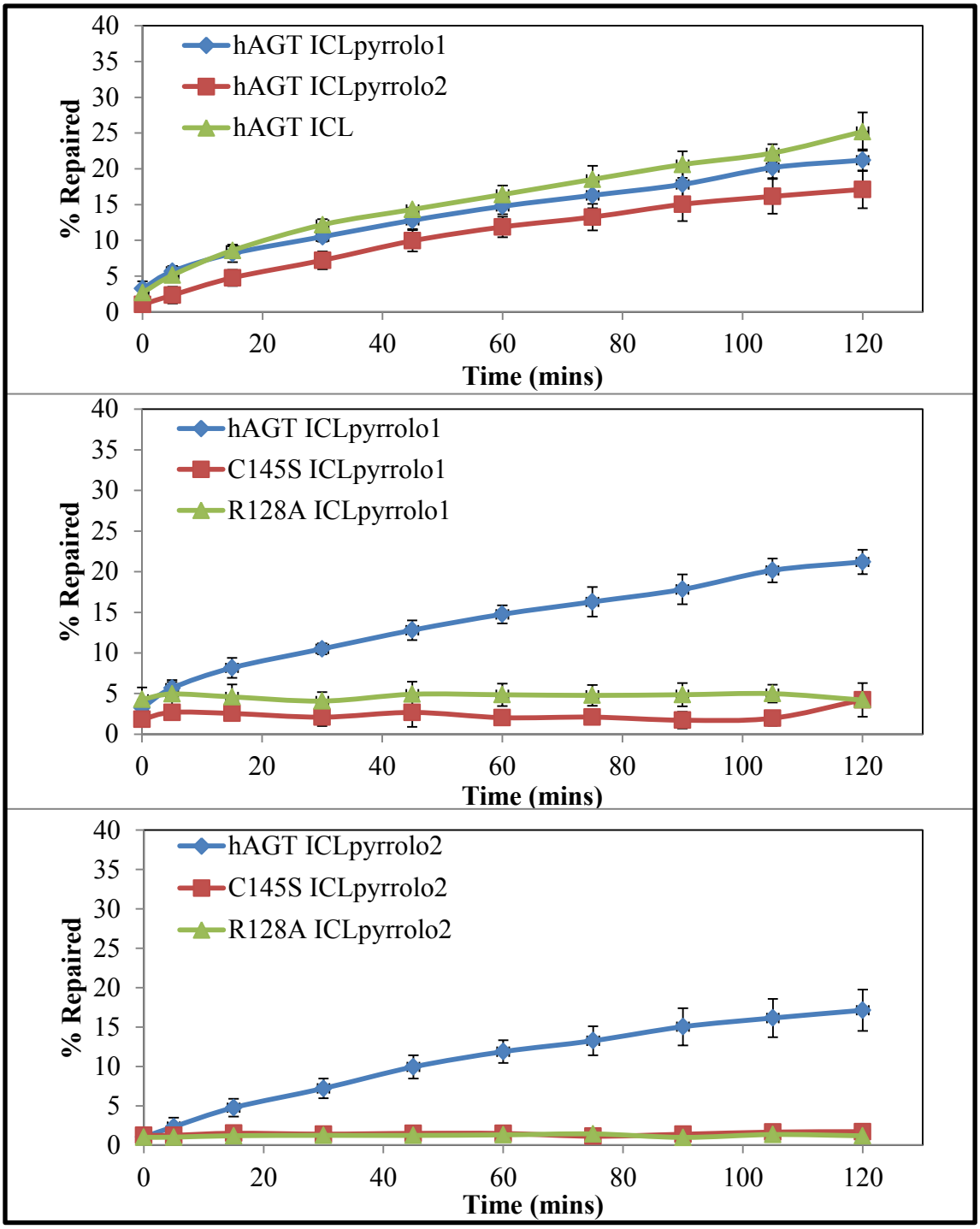


Figure 33 Radioactivity repair assays done in triplicate comparing AGT wild type (6 mM) with all ICL DNAs (0.13 mM, top graph) and AGT variants repair of ICLpyrrolo1 (middle graph) and ICLpyrrolo2 (bottom graph).

3.4.2 Binding studies

All proteins unable to repair ICL DNA were used for binding assays. The binding affinity was most important as a determinant factor to choose an AGT variant for fluorescent binding studies. **Table 6** shows results of binding of C145S, R128A and Ada-C to unmodified DNA, ICL, ICLpyrrolo1 and ICLpyrrolo2. As observed in **Figure 34 top**, increasing concentration of protein results in increased binding of DNA and migration of bands to the upper band position. Constants of dissociation (K_d) were calculated using the equations as shown in **section 2.25**.

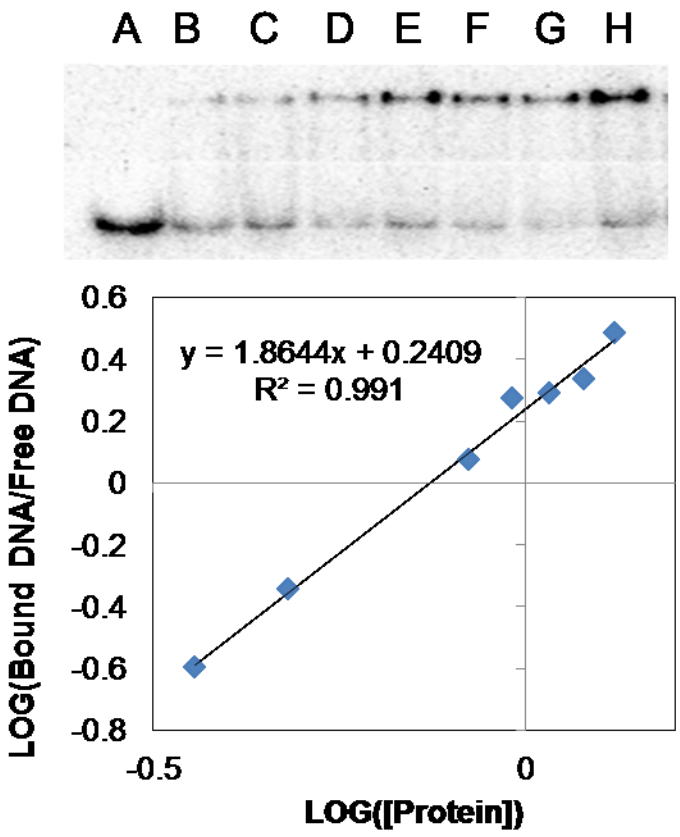


Figure 34 An example of binding data of C145S protein of various concentrations with 0.5 nM of DNA CJW322/323, with the polyacrylamide gel on top and the Hill plot analysis of the raw data on bottom. Lanes A-H of the gel contained DNA and designates the following: (A) DNA only, (B) 0.36 µM C145S, (C) 0.48 µM C145S, (D) 0.84 µM C145S, (E) 0.96 µM C145S, (F) 1.08 µM C145S, (G) 1.20 µM C145S and (H) 1.32 µM C145S.

C145S showed the strongest binding with K_d values for ICL and non-ICL containing DNA within the same order of magnitude. We would have thought that ICL DNA would be bound more strongly because, as a repair protein, we would assume AGT would preferentially bind damaged DNA to repair it. C145S binds more efficiently than R128A (ICL DNA only) and Ada-C. R128A binds very poorly to unmodified DNA but not with ICL DNA. Whereas Ada-C binds most inefficiently to both ICL and non-ICL DNA, with a slight increase in binding strength for ICL DNA over unmodified DNA. The presence of pyrrolo-dC did not result in significant differences in AGT binding.

Binding data suggests that the R128 residue is not essential for detecting damage but is required to associate with unmodified DNA, hinting at the possible mechanism of damage detection. Given R128A's binding preferences it is plausible that damage detection is completed by detecting intrahelical and capturing extrahelical modifications as Duguid had postulated and not by flipping out each base. However, more mutants such as modification of tyrosine 114 are required to provide more information. It must be noted that pyrrolo-dC did not significantly affect AGT variants binding. Furthermore, when comparing to simpler oligonucleotides it is possible that larger DNA substrates will show similar results and smaller dissociation constants, whereas shorter DNA substrates will shower a greater variance in K_d values between ICL and non-ICL containing DNA. Furthermore, the strongest binding protein C145S showed overall dissociation constants (K_d) in the high nM range, where R128A and Ada-C binds at least 10 times less strongly.

DNA	Overall protein dissociation constants K_d (μM)		
	C145S	R128A	Ada-C
Unmodified DNA	0.578±0.005	ND*	22.72±6.96
ICL	0.249±0.012	2.924±0.575	10.53±1.25
ICLpyrrolo1	0.141±0.014	4.442±0.846	4.268±0.883
ICLpyrrolo2	0.163±0.014	7.322±0.008	7.554±0.010

Table 6 Apparent binding dissociation constants of AGT C145S, R128A and *E. coli* homolog Ada-C. ND* no definitive K_d value could be determined up to 100μM.

3.4.3 Fluorescence binding studies

Fluorescence binding experiments were conducted using pyrrolo-dC containing DNA to attempt to elucidate AGT's association with ICL DNA. During optimization it was determined that dsDNA with and without pyrrolo-dC do not fluoresce, with the buffer having no significant fluorescence profile. Upon the addition of protein there is a significant increase in fluorescence, where increasing fluorescence can often be observed with increasing protein concentration depending on the presence of ICL and the position of pyrrolo-dC (**Figure 35 bottom** shows an example of spectra). In addition, protein samples without DNA appear to fluoresce and overshadow pyrrolo-dC fluorescence, which was unexpected (**Figure 35 top**). Therefore, all fluorescence spectra were corrected for protein fluorescence by subtracting fluorescence as observed with protein and buffer only (**Figure 35 top**) with each protein sample with DNA. It must be noted, however, that increasing protein concentration tended to increase background fluorescence, which could most possibly be due to the presence of oxidized DTT, as proteins do not tend to fluoresce with excitation at 346nm. We were able to reduce background fluorescence by using a more concentrated protein stock solution, yet for

concentration above 9μM another reducing agent, such as TECEP, could be investigated as solubility becomes an issue for highly concentrated protein stock solutions.

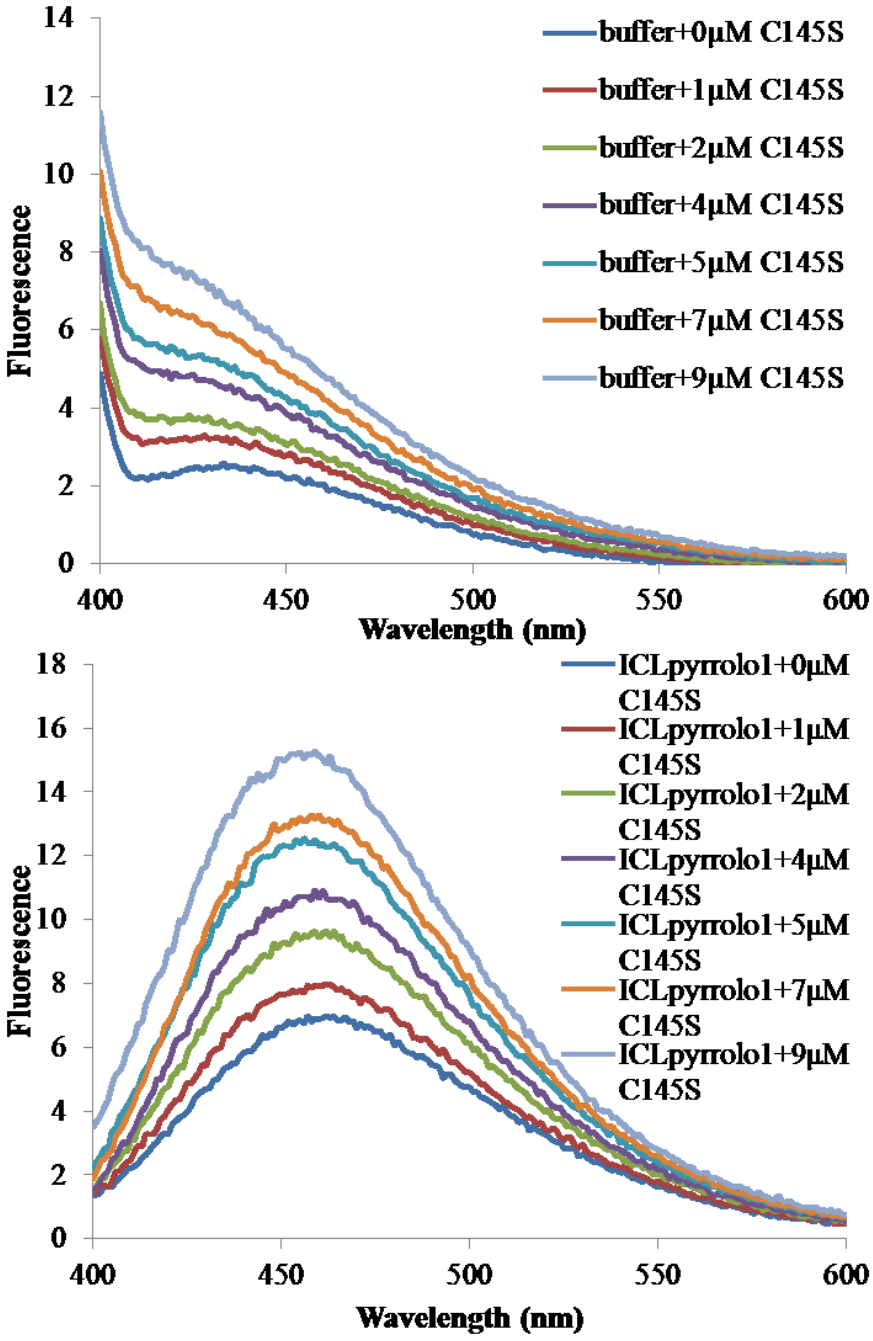


Figure 35 Fluorescence spectra of buffer with varying concentrations of C145S protein (top) and corrected fluorescence spectra of 391 nM ICLpyrrolo1 DNA with varying amounts of C145S protein (bottom).

Fluorescence binding studies were conducted with C145S with ICL and non-ICL containing DNA. We observed that pyrrolo-dC exhibited greater fluorescence when located at the 5'-end (pyrrolo2 position) rather than a central location (pyrrolo1 position) for DNA duplexes lacking and containing the ICL (see **Figure 36**). Although the pyrrolo1 position shows lower overall fluorescence, there is a constant increase as protein concentration increases (**Figure 35 bottom and 36 top**). The difference between ICL and non-ICL containing DNA for pyrrolo1 is minimal with both demonstrating similar trends. Furthermore, the pyrrolo2 position may show overall greater fluorescence but the least difference in fluorescence as protein concentration increases (**Figure 36 bottom**). In addition, unmodified DNA spectra appear to have increased fluorescence as compared to ICL containing DNA for pyrrolo2.

The fluorescence spectra of pyrrolo1 and pyrrolo2 presents an issue as we can no longer specify whether the difference in fluorescence between the two positions is due to the presence of the crosslink or not. The pyrrolo1 position may fluoresce less due to increased stability in the center of a double helix and the presence of the crosslink. While, greater fluorescence at the 5'-end suggests that pyrrolo-dC is in a more dynamic environment where the double helix can open more readily (being near the end of the helix), which is possibly independent of significant changes in AGT concentration. Furthermore, the difference observed between ICL and non-ICL DNA for the pyrrolo2 position (**Figure 36 bottom**) may be due to the presence of single stranded DNA in the control sample (for increased fluorescence) and/or the presence of ICL (for decreased fluorescence for ICLpyrrolo2).

While we did not observe a maximum fluorescence for the amounts of protein used, it is possible that dissociation constants could be determined. However, more work has to be done as this was completed once, a maximum fluorescence was not reached and a more pure substrate must be used in conjunction with other protein variants to observe consistency.

Further optimization of the fluorescence assay coupled with the use of a greater variety of AGT variants may aid in elucidating AGT association with ICL DNA. Specifically C145S binding curves could be used as a standard with which to compare all other protein variants. Alternatively, fluorescence binding experiments can be used to possibly determine dissociation constants, providing a safer non-radioactive option. Unfortunately based on the fluorescence data obtained we cannot definitively determine the extent of DNA double helix melting due to AGT bound to ICL damage.

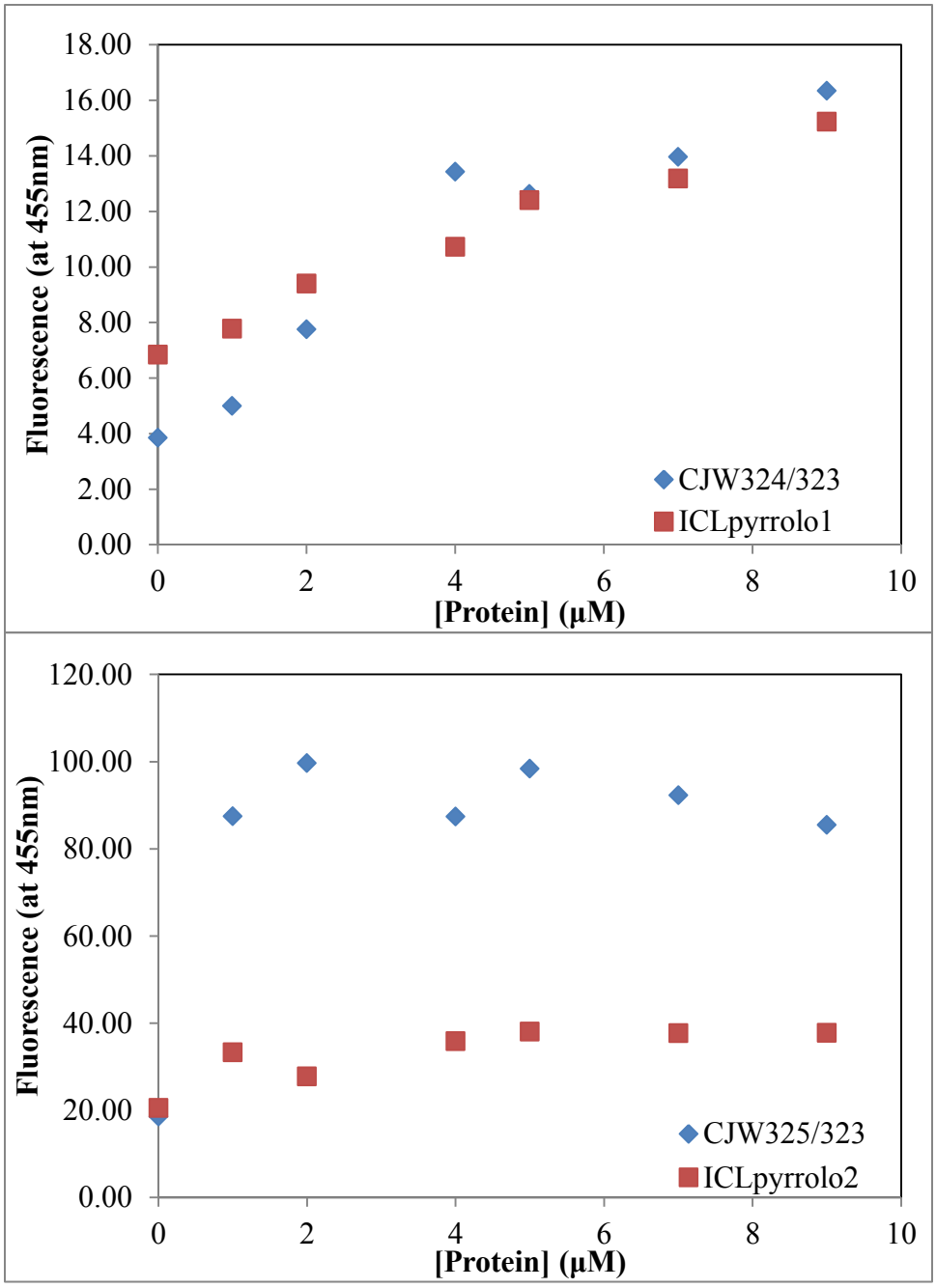


Figure 36 Comparison of fluorescence intensities at 455nm as C145S protein concentration increases with 391 nM DNA. ICLpyrrolo1 is compared to its pyrrolo-dC containing control (top) and ICLpyrrolo2 compared to its pyrrolo-dC containing control (bottom).

3.4.4 Summary

Repair assays revealed the apparent activity of C145S, which was removed upon the addition of EDTA. Furthermore, the R128A variant proved that the arginine "finger" plays an essential role in ICL DNA repair. Binding assays illustrated C145S's strong binding affinity to both unmodified DNA and ICL DNA with R128A and Ada-C binding less strongly. In addition, R128A's very weak binding to unmodified DNA was a surprise as it was able to bind ICL DNA suggesting R128A may be required to associate with unmodified DNA but that it does not play an essential role in damage detection.

Fluorescence studies did not specifically demonstrate the extent to which AGT melts the double helix upon binding ICL DNA over non-ICL containing DNA. In addition, pyrrolo appears to fluoresce more when positioned near the 5'-end (pyrrolo2) than when centrally positioned (pyrrolo1) for both ICL and non-ICL DNA. This results in lower overall fluorescence at the pyrrolo1 position. Future work with protein variants may reveal important residue for binding and damage detection, using C145S binding curves as a standard.

Conclusion

We attempted to understand the influence of AGT binding and effect on the structure of the ICL DNA duplex by having synthesized an O^6-2'-deoxyguanosine-heptylene-O^6-2'-deoxyguanosine (O^6-dG-alkylene-O^6-dG) ICL in a 5'-GNC-3' motif that mimics the lesion formed by hepsulfam. To better understand the effect of AGT binding, we also incorporated pyrrolo-dC at various positions.

Several AGT variants (AGT wild type, C145S mutant, R128A mutant and Ada-C) were expressed and prepared to high purity, identified by MS-Q-TOF and biophysically characterized, showing that they were stable above 37^{o}C and properly folded.

Synthesis of the ICL DNA (tri-partite synthesis) proved difficult with the crosslink coupling yield being poor and the residue (Lev-dC) properly capping the 5'-ends from undesired extension. Most significantly purification revealed the presence of two major products (top and bottom bands). Standard characterization methods proved futile in identifying the desired product, as all ICL DNA containing pyrrolo-dC showed similar DNA composition, overlapping CD data (all B-form) and thermal denaturation studies did not provide any conclusive results. Surprisingly, pyrrolo-dC seemed to reduce stability induced by incorporation of O^6-methyl-dG directly opposed by increasing thermal stability by several degrees.

As we could not identify the desired product of the tri-partite project, we prepared ICL DNA (bi-partite synthesis) with a presence of minor impurities, due to poor 5'-amidite coupling, which could not be removed without significant loss of product. ICL

compositional analysis confirmed the identity of the desired products and the presence of both crosslink and pyrrolo-dC. ICL incorporation appeared to increase thermal stability by 5^{o}C while not causing substantial distortion of the double helix. Importantly, the fluorophore did not seem to decrease thermal stability, nor cause distortion of the B-form DNA and did not significantly affect AGT's ability to repair and bind ICL DNA. Repair assays revealed that the R128A variant was unable to repair ICL DNA and the apparent activity of C145S, which was removed in the presence of EDTA. Binding assays illustrated C145S's strong binding affinity with R128A and Ada-C binding less strongly. Surprisingly R128A bound very weakly to unmodified DNA while it was able to bind ICL DNA suggesting that it may not play an essential role in damage detection. Fluorescence studies with C145S were unable to solve the ICL-AGT association dilemma, however could possibly be used to provide approximate dissociation constants. Fluorescence appeared to be quenched more for centrally located pyrrolo-dC's as compared to those nearer to the 5'-end. The fluorescence data obtained provides a base with which to further develop the fluorescence assay to possibly elucidate protein mutant binding to various oligonucleotides.

Future work

Development of new AGT variants, such as mutagenesis of the residue tyrosine 114, may reveal residues with novel functions and aid in solving the damage detection and ICL-AGT association dilemmas. In addition, an in-depth search into the effects of zinc and other divalent cation on AGT activity may reveal the specific changes in stability and structure as well as explaining what caused C145S's apparent activity. Furthermore, we would like to improve upon the tri-partite synthesis, as it allows for a more diverse base sequence in a non-symmetrical fashion. In addition, further refinement of the synthesis of the bi-partite synthesis ICL DNA should be completed in order to reduce oligonucleotide by-products, and hence identify ICL DNA using Orbitrap MS, which can be used to improve repair and binding studies. Most importantly, we would like to see further optimization of the fluorescence experiments in tandem with development of AGT variants and new more complex DNA substrates. This development could help to understand the quenching effect due to pyrrolo's position, and more significantly, elucidate AGT's method of damage detection and ICL repair. Furthermore, identification of possible binding partners to AGT and the function of the N-terminal could possibly reveal protein partners required in signalling and/or regulation of AGT.

References

1. Lindahl T. (1993) “Instability and decay of the primary structure of DNA.” Nature **362:709-715.**

2. Noll DM, Mason TM, Miller PS. (2006). “Formation and repair of interstrand cross-links in DNA.” Chem Rev. **106:277-301.**

3. Friedberg EC. (1995). “Out of the shadows and into the light: the emergence of DNA repair.” Trends Biochem Sci. **20:381.**

4. Lawley PD, Brookes P. (1968). “Cytotoxicity of alkylating agents towards sensitive and resistant strains of Escherichia coli in relation to extent and mode of alkylation of cellular macromolecules and repair of alkylation lesions in deoxyribonucleic acids.” Biochem J. **109(3):433-447.**

5. Rupert CS. (1962). “Photoenzymatic repair of ultraviolet damage in DNA. II. Formation of enzyme-substrate complex.” J Gen Physiol. **45:725-741.**

6. Bender K, Federwisch M, Loggen U, Nehls P, Rajewsky MF. (1996). “Binding and repair of O^6-ethylguanine in double-stranded oligodeoxynucleotides by recombinant human O^6-alkylguanine-DNA alkyltransferase do not exhibit significant dependence on sequence context.” Nucleic Acids Res. **24:2087-2094.**

7. Brent TP, Houghton PJ, Houghton JA. (1985). "O^6-Alkylguanine-DNA alkyltransferase activity correlates with the therapeutic response of human rhabdomyosarcoma xenografts to 1-(2-chloroethyl)-3-(trans-4-methylcyclohexyl)-1-nitrosourea." Proc Natl Acad Sci USA. **82: 2985-2989.**

8. Daniels DS, Mol CD, Arvai AS, Kanugula S, Pegg AE, Tainer JA. (2000). "Active and alkylated human AGT structures: a novel zinc site, inhibitor and extrahelical base binding." EMBO **19: 1719–1730**

9. Daniels DS, Woo TT, Luu KX, Noll DM, Clarke ND, Pegg AE, Tainer JA. (2004). "DNA binding and nucleotide flipping by the human DNA repair protein AGT." Nat Struct Mol Biol. **11: 714-720.**

10. Dolan ME, Moschel RC, Pegg AE. (1990). "Depletion of mammalian O^6-alkylguanine-DNA alkyltransferase activity by O^6-benzylguanine provides a means to evaluate the role of this protein in protection against carcinogenic and therapeutic alkylating agents." Proc. Natl. Acad. Sci. **87: 5368-5372.**

11. Drabløs F, Feyzi E, Aas PA, Vaagbø CB, Kavli B, Bratlie MS, Peña-Diaz J, Otterlei M, Slupphaug G, Krokan HE. (2004). "Alkylation damage in DNA and RNA-repair mechanisms and medical significance." DNA Repair **3:1389-1407.**

12. Duguid EM, Mishina Y., He C. (2003). "How do DNA repair proteins locate potential base lesions? A chemical crosslinking method to investigate O^6-alkylguanine-DNA alkyltransferases." Chem Biol. **10: 827-835.**

13. Duguid EM, Rice PA, He C. (2005). "The structure of the human AGT protein bound to DNA and its implications for damage detection." J Mol Biol. **350: 657-666.**

14. Fang Q, Kanugula S., Pegg AE. (2005). "Function of domains of human O^6-alkylguanine-DNA alkyltransferase." Biochemistry. **44: 15396-15405.**

15. Fang Q, Noronha A., Murphy SP, Wilds CJ, Tubbs JL, Tainer JA, Chowdhury G, Guengerich FP, Pegg AE. (2008). "Repair of O^6-G-alkyl-O^6-G interstrand cross-links by human O^6-alkylguanine-DNA alkyltransferase." Biochemistry **47: 10892-10903.**

16. Gerson SL, Willson JK. (1995). "O^6-alkylguanine-DNA alkyltransferase. A target for the modulation of drug resistance." Hematol Oncol Clin North Am. **9:431-450.**

17. Gerson SL. (2002). "Clinical relevance of MGMT in treatment of cancer." J. Clin. Oncol. **20:2388-2399.**

18. Goodtzova K, Kanugula S, Edara S, Pauly GT, Moschel RC, Pegg AE. (1997). "Repair of O6-benzylguanine by the Escherichia coli Ada and Ogt and the human *O*^{6}-alkylguanine-DNA alkyltransferases." J Biol Chem. **272:8332-8339.**

19. Graves RJ, Li BF, Swann PF. (1989). "Repair of *O*^{6}-methylguanine, *O*^{6}-ethylguanine, *O*^{6}-isopropylguanine and *O*^{4}-methylthymine in synthetic oligodeoxynucleotides by *Escherichia coli ada* gene O^6-alkylguanine-DNA-alkyltransferase." Carcinogenesis. **10:661-666.**

20. Guengerich FP, Fang Q, Liu L, Hachley DL, Pegg AE. (2003). "O^6-alkylguanine-DNA alkyltransferase: low pK_a and high reactivity of cysteine 145." Biochemistry **42:10965-10970.**

21. Horiguchi M, Kim J, Matsunaga N, Kaji H, Egawa T, Makino K, Koyanagi S, Ohdo S. (2010). "Glucocorticoid-dependent expression of *O*^{6}-methylguanine-

DNA methyltransferase gene modulates dacarbazine-induced hepatotoxicity in mice." J. Pharmacol. Exp. Ther. **333:782-787.**

22. Hu J, Ma A, Dinner AR. (2008). "A two-step nucleotide-flipping mechanism enables kinetic discrimination of DNA lesions by AGT." Proc. Natl. Acad. Sci. U.S.A. **105:4615-4620.**

23. Juillerat A, Gronemeyer T, Keppler A, Gendreizig S, Pick H, Vogel H, Johnsson K. (2003). "Directed evolution of O^6-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo." Chem Biol. **10:313-317.**

24. Kaina B, Fritz G, Mitra S, Coquerelle T. (1991). "Transfection and expression of human O6-methylguanine-DNA methyltransferase (MGMT) cDNA in Chinese hamster cells: the role of MGMT in protection against the genotoxic effects of alkylating agents.." Carcinogenesis. **12: 1857-1867.**

25. Kanugula S, Goodtzova K, Pegg AE. (1998). "Probing of conformational changes in human O^6-alkylguanine-DNA alkyl transferase protein in its alkylated and DNA-bound states by limited proteolysis." Biochem J. **329: 545-550.**

26. Kanugula S, Pegg AE. (2001). "Novel DNA repair alkyltransferase from *Caenorhabditis elegans*." Environ Mol Mutagen. **38:235-243.**

27. Kanugula S, Pauly GT, Moschel RC, Pegg AE. (2005). "A bifunctional DNA repair protein from Ferroplasma acidarmanus exhibits O^6-alkylguanine-DNA

alkyltransferase and endonuclease V activities." Proc Natl Acad Sci USA. **102:3617-3622.**

28. Kleibl K. (2002). "Molecular mechanisms of adaptive response to alkylating agents in Escherichia coli and some remarks on O^6-methylguanine DNA-methyltransferase in other organisms." Mutat Res. **512:67-84.**

29. Klein S, Oesch F. (1992). "Assay for O^6-alkylguanine-DNA-alkyltransferase using oligonucleotides containing O^6-methylguanine in a BamHI recognition site as substrate." Anal Biochem. **205:294-299.**

30. Kotandeniya D, Murphy D, Seneviratne U, Guza R, Pegg A, Kanugula S, Tretyakova N. (2011). "Mass spectrometry based approach to study the kinetics of O^6-alkylguanine DNA alkyltransferase-mediated repair of O^6-pyridyloxobutyl-2'-deoxyguanosine adducts in DNA." Chem Res Toxicol. **24:1966-1975.**

31. Mattern J, Eichhorn U., Kaina B, Volm M. (1998). "O^6-methylguanine-DNA methyltransferase activity and sensitivity to cyclophosphamide and cisplatin in human lung tumor xenografts." Int J Cancer. **77: 919-922.**

32. McElhinney RS, McMurry TB, Margison GP. (2003). "O^6-alkylguanine-DNA alkyltransferase inactivation in cancer chemotherapy." Mini Rev Med Chem. **3:471-485.**

33. McManus FP, Fang Q, Booth JD, Noronha AM, Pegg AE, Wilds CJ. (2010). "Synthesis and characterization of an O^6-2'-deoxyguanosine-alkyl-O^6-2'-

deoxyguanosine interstrand cross-link in a 5'-GNC motif and repair by human O^6-alkylguanine-DNA alkyltransferase." Org Biomol Chem. **8: 4414-4426.**

34. Mishina Y, Duguid EM, He C. (2006) "Direct reversal of DNA alkylation damage." Chem Rev. **106:215-232.**

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

36. Pegg AE, Perry W, Bennett RA. (1981). "Partial hepatectomy increases the ability of rat liver extracts to catalyze removal of O^6-methylguanine from alkylated DNA." Biochem. J. **197:195-201.**

37. Pegg AE. (2000). "Repair of O6-alkylguanine by alkyltransferases." Mutat Res. **462:83-100.**

38. Pegg AE, Goodtzova K, Loktionova NA, Kanugula S, Pauly GT, Moschel RC. (2001). "Inactivation of human O^6-alkylguanine-DNA alkyltransferase by modified oligodeoxyribonucleotides containing O^6-benzylguanine." J Pharmacol Exp Ther. **296:958-965.**

39. Pegg AE. (2011). "Multifaceted roles of alkyltransferase and related proteins in DNA repair, DNA damage, resistance to chemotherapy, and research tools." Chem Res Toxicol. **24:618-639.**

40. Philip PA, Souliotis VL, Harris AL, Salisbury A, Tates AD, Mitchell K, van Delft JH, Ganesan TS, Kyrtopoulos SA. (1996). "Methyl DNA adducts, DNA repair,

and hypoxanthine-guanine phosphoribosyl transferase mutations in peripheral white blood cells from patients with malignant melanoma treated with dacarbazine and hydroxyurea." Clin Cancer Res. **2:303-310.**

41. Reagan MS, Pittenger C, Siede W, Friedberg EC. (1995) "Characterization of a mutant strain of *Saccharomyces cerevisiae* with a deletion of the RAD27 gene, a structural homolog of the RAD2 nucleotide excision repair gene." J Bacteriol. **177:364-371.**

42. Schoonhoven NM, Murphy S., O'Flaherty DK, Noronha AM, Kornblatt MJ, Wilds CJ. (2008). "Synthesis, biophysical and repair studies of O^6-2'-deoxyguanosine adducts by *Escherichia coli* OGT." Nucleic Acids Symp Ser (Oxf). **52: 449-450.**

43. Scicchitano D, Jones RA, Kuzmich S, Gaffney B, Lasko DD, Essigmann JM, Pegg AE. (1986). "Repair of oligodeoxynucleotides containing O^6-methylguanine by O^6-alkylguanine-DNA-alkyltransferase." Carcinogenesis. **7:1383-1386.**

44. Srivenugopal KS, Yuan XH, Friedman HS, Ali-Osman F. (1996). "Ubiquitination-dependent proteolysis of O^6-methylguanine-DNA methyltransferase in human and murine tumour cells following inactivation with O^6-benzylguanine or 1,3-bis(2-chloroethyl)-1-nitrosurea." Biochemistry **35: 1328-1334.**

45. Takahashi M, Sakumi K, Sekiguchi M. (1990). "Interaction of Ada protein with DNA examined by fluorescence anisotropy of the protein." Biochemistry **29: 3431-3436.**

46. Toorchen D, Topal MD. (1983). "Mechanisms of chemical mutagenesis and carcinogenesis: effects on DNA replication of methylation at the O^6-guanine position of dGTP." Carcinogenesis. **4:1591-1597.**

47. Toorchen D, Lindamood C 3rd, Swenberg JA, Topal MD. (1984). "O^6-methylguanine-DNA transmethylase converts O^6-methylguanine thymine base pairs to guanine thymine base pairs in DNA." Carcinogenesis. **5:1733-1735.**

48. Tubbs JL, Pegg AE, Tainer JA. (2007). "DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by O^6-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy." DNA Repair (Amst). **6:1100-1115.**

49. Vaughan P, Lindahl T, Sedgwick B. (1993). "Induction of the adaptive response of Escherichia coli to alkylation damage by the environmental mutagen, methyl chloride." Mutat Res. **293:249-257.**

50. Verdemato PE, Brannigan JA, Damblon C, Zuccotto F, Moody PC, Lian LY. (2000). "DNA-binding mechanism of the Escherichia coli Ada O^6-alkylguanine-DNA alkyltransferase." Nucleic Acids Res. **28:3710-3718.**

51. Vora RA, Pegg AE, Ealick SE. (1998). "A new model for how O^6-methylguanine-DNA methyltransferase binds DNA." Proteins **32:3-6.**

52. Wilkinson MC, Potter PM, Cawkwell L, Georgiadis P, Patel D, Swann PF, Margison GP. (1989). "Purification of the *E. coli ogt* gene product to homogeneity and its rate of action on O^6-methylguanine, O^6-ethylguanine and O^4-

methylthymine in dodecadeoxyribonucleotides." Nucleic Acids Res. **17:8475-8484.**

53. Wilson BD, Strauss M, Stickells BJ, Hoal-van Helden EG, van Helden P. (1994). "An assay for O^6-alkylguanine-DNA alkyltransferase based on restriction endonuclease inhibition and magnetic bead separation of products." Carcinogenesis. **15:2143-2148.**

54. Xu-Welliver M, Pegg AE. (2002). "Degradation of the alkylated form of the DNA repair protein O^6-alkylguanine-DNA alkyltransferase." Carcinogenesis. **23:823-830.**

55. Zang H, Fang Q., Pegg AE, Guengerich FP. (2005). "Kinetic analysis of steps in the repair of damaged DNA by human O^6-alkylguanine-DNA alkyltransferase." J Biol Chem. **280: 30873-30881.**

56. Kozekov ID, Nechev LV, Moseley MS, Harris CM, Rizzo CJ, Stone MP, Harris TM. (2003). "DNA interchain cross-links formed by acrolein and crotonaldehyde." J Am Chem Soc. **125:50-61.**

57. Sancar A. (2003). "Structure and function of DNA photolyase and chryptochrome blue-light photoreceptors." Chem Rev. **103:2203-2237.**

58. Rydberg B, Lindahl T. (1982). "Non-enzymatic methylation of DNA by the intracellular methyl group donor *S*-adenosyl-L-methionine is a potentially mutagenic reaction." EMBO J. **1:211-216.**

59. Taverna P, Sedgwick B. (1996). "Generation of an endogenous DNA-methylating agent by nitrosation in *Escherichia coli*." J Bacteriol. **178:5105-5111.**

60. Friedman JI, Jiang YL, Miller PS, Stivers JT. (2011). "Unique dynamic properties of DNA duplexes containing interstrand cross-links." Biochemistry **50:882-890.**

61. Jamieson ER, Lippard SJ. (1999). "Structure, Recognition, and Processing of Cisplatin-DNA Adducts." Chem Rev. **99:2467-2498.**

62. Wang D, Hara R, Singh G, Sancar A, Lippard SJ. (2003). "Nucleotide excision repair from site-specifically platinum-modified nucleosomes." Biochemistry. **42:6747-6753.**

63. Dooley PA, Tsarouhtsis D, Korbel GA, Nechev LV, Shearer J, Zegar IS, Harris CM, Stone MP, Harris TM. (2001). "Structural studies of an oligodeoxynucleotide containing a trimethylene interstrand cross-link in a 5'-(CpG) motif: model of a malondialdehyde cross-link." J Am Chem Soc. **123:1730-1739.**

64. Dooley PA, Zhang M, Korbel GA, Nechev LV, Harris CM, Stone MP, Harris TM. (2003). "NMR determination of the conformation of a trimethylene interstrand cross-link in an oligodeoxynucleotide duplex containing a 5'-d(GpC) motif." J Am Chem Soc. **125:62-72.**

65. Guengerich FP. (2006). "Interactions of Carcinogen-Bound DNA with Individual DNA polymerases." Chem Rev. **106:420-452.**

66. Huang H, Kim HY, Kozekov ID, Cho YJ, Wang H, Kozekova A, Harris TM, Rizzo CJ, Stone MP. (2009). "Stereospecific formation of the (R)-γ-hydroxytrimethylene interstrand N^2-dG: N^2-dG cross-link arising from the γ -OH-1,N^2-propano-2'-deoxyguanosine adduct in the 5'-CpG-3' DNA sequence." J Am Chem Soc. **131:8416-8424.**

67. Stevens K, Madder A. (2009). "Furan-modified oligonucleotides for fast, high-yielding and site-selective DNA inter-strand cross-linking with non-modified complements." Nucleic Acids Res. **37:1555-1565.**

68. Streeper RT, Cotter RJ, Colvin ME, Hilton J, Colvin OM. (1995). "Molecular pharmacology of hepsulfam, NSC 3296801: identification of alkylated nucleosides, alkylation site, and site of DNA cross-linking." Cancer Res. **55:1491-1498.**

69. Wilds CJ, Booth JD, Noronha AM. (2006). "Synthesis of oligonucleotides containing an O^6-G-alkyl-O^6-G interstrand crosslink." Elsevier Tetrahedron Letters **47: 9125-9128.**

70. Blanpain C, Mohrin M, Sotiropoulou PA, Passegué E. (2011). "DNA damage response in tissue-specific and cancer stem cells." Cell Stem Cell. **8:16-29.**

71. Burns JA, Dreij K, Cartularo L, Scicchitano DA. (2010). "*O*^{6}-methylguanine induces altered proteins at the level of transcription in human cells." Nucleic Acids Res. **38:8178-8187.**

72. Nagel G, Brenner W, Johnsson K, Kaina B. (2003). "DNA repair protein O^6-methylguanine-DNA methyltransferase in testis and testicular tumors as determined by a novel nonradioactive assay." Anal Biochem. **321:38-43.**

73. Teo AK, Oh HK, Ali RB, Li BF. (2001). "The modified human DNA repair enzyme O^6-methylguanine-DNA methyltransferase is a negative regulator of estrogen receptor-mediated transcription upon alkylation DNA damage." Mol Cell Biol. **21:7105-7114.**

74. Bhavsar YP, Reilly SM, Wadkins RM. (2011). "Evaluation of fluorescent analogs of deoxycytidine for monitoring DNA transitions from duplex to functional structures." J Nucleic Acids. **2011:986820.**

75. Hardman SJ, Thompson KC. (2006). "Influence of base stacking and hydrogen bonding on the fluorescence of 2-aminopurine and pyrrolocytosine in nucleic acids." Biochemistry **45:9145-9155.**

76. Hudson RH, Choghamarani AG. (2007). "The 6-methoxymethyl derivative of pyrrolo-dC for selective fluorometric detection of guanosine-containing sequences." Nucleosides, Nucleotides, Nucleic Acids **26:533-537.**

77. Li T, Fu R, Park HG. (2010). "Pyrrolo-dC based fluorescent aptasensors for the molecular recognition of targets." Chem Commun (Camb.) **46:3271-3273.**

78. Martí AA, Jockusch S, Li Z, Ju J, Turro NJ. (2006). "Molecular beacons with intrinsically fluorescent nucleotides." Nucleic Acids Res. **34:e50.**

79. Ming X, Ding P, Leonard P, Budow S, Seela F. (2012) “Parallel-stranded DNA: enhancing duplex stability by the 'G-clamp' and a pyrrolo-dC derivative.” Org Biomol Chem. **10:1861-1869.**

80. Moser AM, Patel M, Yoo H, Balis FM, Hawkins ME. (2000). “Real-time fluorescence assay for O^6-alkylguanine-DNA alkyltransferase.” Anal Biochem. **281:216-222.**

81. Park KS, Lee JY, Park HG. (2012). “Mismatched pyrrolo-dC-modified duplex DNA as a novel probe for sensitive detection of silver ions.” Chem Commun (Camb.) **48:4549-4551.**

82. Sandin P, Börjesson K, Li H, Mårtensson J, Brown T, Wilhelmsson LM, Albinsson B. (2008). “Characterization and use of an unprecedentedly bright and structurally non-perturbing fluorescent DNA base analogue.” Nucleic Acids Res. **36: 157-167.**

83. Thompson KC, Miyake N. (2005). “Properties of a new fluorescent cytosine analogue, pyrrolocytosine.” J Phys Chem B. **109:6012-6019.**

84. Wilhelmsson LM. (2010). “Fluorescent nucleic acid base analogues.” Q Rev Biophys. **43:159-183.**

85. Zhang H, Wang M, Gao Q, Qi H, Zhang C. (2011). “Fluorescent detection of single nucleotide polymorphism utilizing a hairpin DNA containing a nucleotide base analog pyrrolo-deoxycytidine as a fluorescent probe.” Talanta **84:771-776.**

86. Zhang X, Wadkins RM. (2009). "DNA hairpins containing the cytidine analog pyrrolo-dC: structural, thermodynamic, and spectroscopic studies." Biophys J. **96: 1884-1891.**

87. Creighton, TE (editor); Pace NC, Scholtz JM. (1997). "*Protein structure: a practical approach.* Ch. 12: Measuring the conformational stability of a protein." 2nd ed. Oxford: IRL Press at Oxford University Press.

88. Damha MJ, Giannaris PA, Zabarylo SV. (1990). "An improved procedure for derivatization of controlled-pore glass beads for solid-phase oligonucleotide synthesis." Nucleic Acids Res. **18:3813-3821.**

89. Trempe JF, Wilds CJ, Denisov AY, Pon RT, Damha MJ, Gehring K. (2001). "NMR solution structure of an oligonucleotide hairpin with a 2'F-ANA/RNA stem: implications for RNase H specificity toward DNA/RNA hybrid duplexes." J Am Chem Soc. **123:4896-4903.**

90. Wilds CJ, Damha MJ. (2000). "2'-Deoxy-2'-fluoro-beta-D-arabinonucleosides and oligonucleotides (2'F-ANA): synthesis and physicochemical studies." Nucleic Acids Res. **28:3625-3635.**

91. Bishop GR, Chaires JB. (2003). "Characterization of DNA structures by circular dichroism." Curr Protoc Nucleic Acid Chem. **Chapter 7: Unit 7.11.**

92. Garbett NC, Ragazzon PA, Chaires JB. (2007). "Circular dichroism to determine binding mode and affinity of ligand-DNA interactions." Nature Protoc. **2:3166-3172.**

93. Gray DM, Ratliff RL, Vaughan MR. (1992). "Circular dichroism spectroscopy of DNA." Methods Enzymol. **211:398-406.**

94. Smith, CWJ (editor); Black DL *et al*. (2002). "RNA: Protein interactions: a practical approach. p.134 *ii*: *If intermediates are not visible, determine the overall K_d and the number of binding sites*." Oxford University Press.

95. Rasimas JJ, Pegg AE and Fried MG. (2003). "DNA-binding mechanism of O^6-alkylguanine-DNA-alkyltransferase." J. Biol. Chem. **278, 7973-7980.**