Structure and Thermodynamic Insights on Acetylaminofluorene-Modified Deletion DNA Duplexes as Models for Frameshift Mutagenesis

Anusha Sandineni
University of Rhode Island, asandineni@my.uri.edu

Follow this and additional works at: http://digitalcommons.uri.edu/theses

Terms of Use
All rights reserved under copyright.

Recommended Citation
http://digitalcommons.uri.edu/theses/21

This Thesis is brought to you for free and open access by DigitalCommons@URI. It has been accepted for inclusion in Open Access Master's Theses by an authorized administrator of DigitalCommons@URI. For more information, please contact digitalcommons@etal.uri.edu.
STRUCTURE AND THERMODYNAMIC INSIGHTS ON
ACETYLAMINOFLUORENE-MODIFIED DELETION DNA DUPLEXES AS
MODELS FOR FRAMESHIFT MUTAGENESIS

BY
ANUSHA SANDINENI

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
BIOMEDICAL AND PHARMACEUTICAL SCIENCES

UNIVERSITY OF RHODE ISLAND
2013
2-Acetylaminofluorene (AAF) is a prototype aryl amine carcinogen that forms C8-substituted dG-AAF and dG-AF as the major DNA lesions. The bulky N-acetylated dG-AAF lesion can induce various frameshift mutations depending on the base sequence around the lesion. We hypothesized that the thermodynamic stability of bulged-out slipped mutagenic intermediates (SMIs) is directly related to deletion mutations. The objective of the present study was to probe the structural/conformational basis of various dG-AAF–induced SMIs formed during a translesion synthesis. We performed spectroscopic and thermodynamic studies of several AAF-modified 16-mer model DNA duplexes, including fully paired and -1, -2, and -3 deletion duplexes of the 5'-CTCTCGATG[AAF]CCATCAC-3' sequence and an additional -1 deletion duplex of the 5'-CTCTCGGC[AAF]CCATCAC-3' NarI sequence. Modified deletion duplexes existed in a mixture of external B and stacked S conformers, with the population of the S conformer being ‘GC’-1 (73%) > ‘AT’ -1 (72%) > full (60%) > -2 (55%) > -3 (37%). Thermodynamic stability was in the order of -1 deletion > -2 deletion > fully paired > -3 deletion duplexes. These results indicate that the stacked S-type conformer of SMIs was thermodynamically more stable than the conformationally flexible external B conformer. This order of lesion stability was in good agreement with the efficiencies of the frameshift mutations obtained previously by primer extension assays with the human DNA polymerase η [Schorr and Carell, ChemBioChem, 11, 2534-2537 (2010)]. Taken together, these results support a hypothesis that the conformational and thermodynamic stabilities of the SMIs are critical determinants for the induction of frameshift mutations.
ACKNOWLEDGEMENTS

I would like to thank Dr. Bongsup Cho, my major advisor, for providing me this opportunity to work in his group. He is a great teacher. I took his Drug metabolism and Medicinal Chemistry classes which were very interesting. He helped me in improving my presentation skills and always explained me clearly what I need to do when I have any problem in research and related issues. I was inspired by him both professionally and personally, he is an ideal mentor to me. I would also like to thank my Committee members Drs. King, Slitt, DeBoef, and Levine for their support.

I also like to thank my colleagues Dr. Patnaik for helping me with the synthesis of AAF carcinogen and HPLC training initially, Vipin Jain for guiding me with my course work and other experimental and technical problems in my research such as NMR experiments, Dr. Ganesan for providing literature which was immensely helpful for writing my Thesis. I would like to thank Lifang Xu for being my friend and giving a good company in most of the classes we had taken together, and helping me with research and Anshu Jain for helping me with MALDI-TOF experiments. I thank them for their friendships and their helpful discussions.

I would like to thank my uncle Dr. Srinivasa Rao Meneni and his family for helping me and taking care of me in the US. I thank my dear sister Dr. Tanuja Sandineni for being my ideal from my childhood. For sharing my problems and for being a good support when I am feeling low, I thank my dear friend Sushma Bollepally.
Finally, I’m indebted to my dearest parents for their unconditional love and continuous support. They helped me in choosing a right career path. They are always special in my life and I am inspired by them.
PREFACE

This thesis is written in Standard format.

Chapter 1 provides an introduction on aromatic amine-induced carcinogenesis and the work done on them in the past. Here I propose a hypothesis and specific plans to prove the hypothesis. I also briefly described the analytical techniques used in the present study.

Chapter 2 describes detailed experimental methods used in the present study.

Chapter 3 describes the results obtained from the proposed work.

Chapter 4 discusses the results in detail in terms of previous literature work and biological implications.
TABLE OF CONTENTS

Page

ABSTRACT.................................................................................................ii

ACKNOWLEDGEMENT.................................................................iii

PREFACE..........................................................................................v

TABLE OF CONTENTS...............................................................vi

LIST OF TABLES...........................................................................vii

LIST OF FIGURES..............................................................................viii

CHAPTER 1: INTRODUCTION..........................................................1

CHAPTER 2: EXPERIMENTAL PROCEDURES..............................13

CHAPTER 3: RESULTS...................................................................18

CHAPTER 4: DISCUSSION.................................................................26

BIBILOGRAPHY................................................................................58
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Thermal and thermodynamic parameters derived from UV-melting</td>
<td>30</td>
</tr>
<tr>
<td>curves</td>
<td></td>
</tr>
<tr>
<td>2. Thermal and thermodynamic parameters derived from Differential</td>
<td>31</td>
</tr>
<tr>
<td>Scanning Calorimetry (DSC)</td>
<td></td>
</tr>
<tr>
<td>3. Circular dichroism (CD) shifts in wavelength of FAAF modified</td>
<td>32</td>
</tr>
<tr>
<td>deletion duplexes with respective to full unmodified duplex</td>
<td></td>
</tr>
<tr>
<td>4. Circular dichroism (CD) shifts in wavelength of FAAF modified</td>
<td>33</td>
</tr>
<tr>
<td>duplexes with respective unmodified counterparts</td>
<td></td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Structures and template sequences used in this study. (a) Chemical structures of AAF, FAAF, AF, FAF adducts (b) 16-mer non NarI sequence and (c) 16-mer NarI sequences used in the present study</td>
<td>34</td>
</tr>
<tr>
<td>2. Mechanism for slippage induced -2 deletion on NarI sequence</td>
<td>35</td>
</tr>
<tr>
<td>3. Mechanisms for slippage induced (A) Single base frameshift mutagenesis (B) -3 frameshift mutagenesis</td>
<td>36</td>
</tr>
<tr>
<td>4. Model duplexes studied in the present study (a) Full duplex (b) -3 deletion duplex (c) -2 deletion duplex (d) ‘AT’ -1 deletion duplex (e) ‘GC’ -1 deletion duplex</td>
<td>37</td>
</tr>
<tr>
<td>5. Typical DSC melting profile for the two-state unfolding of a protein</td>
<td>38</td>
</tr>
<tr>
<td>6. Typical UV-melting profile for the two-state unfolding of a DNA helix</td>
<td>39</td>
</tr>
<tr>
<td>7. Temperature dependent $^{19}$F NMR spectra of FAAF modified G$_1$ in NarI 16-mer sequence</td>
<td>40</td>
</tr>
<tr>
<td>8. (a) HPLC chromatogram of a reaction mixture derived from treatment of the 16-mer sequence (5'-CTCTCG$_1$ATG$_2$CCATCAC-3') with an activated FAAF (N-acetoxy-N-2-(acetylamino)-7-fluorofluorene). Mono- (G$_1$, G$_2$) and di-FAAF adducts are eluted in the 15- 25 min range (b) On-line photodiode array UV spectra of unmodified, mono-, and di-FAAF adducts</td>
<td>41</td>
</tr>
<tr>
<td>9. MALDI-TOF mass spectra of peak 2 after 5'$\rightarrow$3'exonuclease digestion (BSP enzyme) in Reflectron mode at various time intervals. Inset provides theoretical MW of the corresponding fragments</td>
<td>42</td>
</tr>
<tr>
<td>10. MALDI-TOF mass spectra of peak 2 after 3'$\rightarrow$5'exonuclease digestion (SVP enzyme) in Reflectron mode at various time intervals. Inset provides theoretical MW of the corresponding fragments</td>
<td>43</td>
</tr>
<tr>
<td>11. MALDI-TOF mass spectra of peak 3 after 5'$\rightarrow$3'exonuclease digestion (BSP enzyme) in Reflectron mode at various time intervals. Inset provides theoretical MW of the corresponding fragments</td>
<td>44</td>
</tr>
</tbody>
</table>
12. MALDI-TOF mass spectra of peak 3 after 3’→5’ exonuclease digestion (SVP enzyme) in Reflectron mode at various time intervals. Inset provides theoretical MW of the corresponding fragments……………………………..45

13. ESI spectra of peak 2 obtained from FAAF-modification of 16-mer template (5’-CTCTCG1ATG2CCATCAC-3’) after (a) 5’-exonuclease digestion (b) 3’-exonuclease digestion……………………………………………………………..46

14. ESI spectra of peak 3 obtained from FAAF-modification of (5’-CTCTCG1ATG2CCATCAC-3’) 16-mer template after a) 5’-exonuclease digestion b) 3’-exonuclease digestion………………………………………………….47

15. UV-melting curves of FAAF modified duplexes (red) and their respective unmodified controls (blue) all at 10 μM in 0.2 M NaCl, 10 mM sodium phosphate, and 0.2 mM EDTA at pH 7. ……………………………………48

16. Differential scanning calorimetry (DSC) curves in 20 mM phosphate buffer containing 0.1 M NaCl at pH 7: FAAF-modified duplexes in red and unmodified controls in blue…………………………………………………………………..49

17. Comparative thermodynamic histograms of FAAF-modified duplexes with respective to their unmodified controls: ΔΔH=ΔH (modified duplex)-ΔH (control duplex), ΔΔS=ΔS (modified duplex) - ΔS (control duplex) and ΔΔG=ΔG (modified duplex) – ΔG (control duplex)………………………………………50

18. CD Spectral overlays of (a) fully-paired (b) -3 deletion (c) -2 deletion (d) ‘AT’ -1 deletion, and (e) ‘GC’-1 deletion duplexes at 30°C. FAAF modified duplexes (solid lines) and unmodified control duplexes (dotted lines)………51

19. CD Spectral overlays of FAAF-modified -3 deletion duplex (black) relative to the unmodified -3 deletion (black) and fully-paired unmodified duplexes (green)…………………………………………………………………………………52

20. Dynamic 19F NMR spectra of FAAF-modified duplexes in various sequence settings: (a) fully-paired, (b) -3 deletion, (c) -2 deletion, (d) ‘AT’ -1 deletion, and (e) ‘GC’-1 deletion duplexes. S, B, and W-conformers are labeled as red, purple, and green colors. Denatured FAAF-modified single strands (SS) are colored in cyan…………………………………………………………………..53

21. Line simulation of FAAF-modified 16-mer (5’-CTCTCG1ATG2CCATCAC-3’) (a) -3 deletion (b) -2 deletion (c) ‘AT’-1 deletion duplexes and FAAF modified NarI 16-mer (5’-CTCTCG1GC1GCATCAC-3’) (e) GC’-1 deletion duplex at 30 °C…………………………………………………………………..54
22. Imino proton region (11-15 ppm) of proton NMR of FAAF-modified 16-mer (5'-CTCTCGATG2CCATCAC-3') (a) Full duplex (b) -3 deletion (c) -2 deletion duplex (d) ‘AT’ -1 deletion (e) ‘GC’ -1 deletion duplex (upper panel) with respective to the unmodified duplexes (lower panel) at 5°C............55

23. Comparison of the dynamic imino proton spectra of FAAF-modified (a) fully-paired, (b) -3, (c) -2, and (d) ‘AT’ and (e) ‘GC’ -1 duplexes (right panel) relative to their respective controls (left panel) at 5, 20, 30, 40°C.................56

24. Comparison of the dynamic imino proton spectra of FAAF-modified (d) ‘AT’ and (e) ‘GC’ -1 duplexes (right panel) relative to their respective controls (left panel) at 5, 20, 30, 40°C.................................................................57
CHAPTER 1: INTRODUCTION

1.1. General Background

Genomic instability has been implicated in many human diseases including cancer and aging (1). DNA mutations may be due to damage from environmental factors, such as UV radiation or certain chemicals, as well as random mistakes during replication for cell division. Accumulation of DNA damage contributes to frameshift mutations which involve a gain or loss of one or more base pairs relative to the original sequence, thereby altering the informational content of the genome. The underlying molecular mechanisms of frameshift mutations are not well understood, however, those induced by aromatic amines have been studied extensively (2, 3).

Aromatic amines are an important class of chemical carcinogens that are present in the environment as byproducts of fossil fuel combustion, tobacco smoke, dyes and over-cooked meat (2, 4). Many of these compounds play a significant role in causing various human cancers related to breast, liver and bladder (5). For example, the human bladder carcinogen 4-aminobiphenyl (ABP) is produced in cigarette smoke, fossil fuel combustion, and textile industries. Similarly, 2-aminofluorene (AF) and 2-acetylaminofluorene (AAF) were originally developed as pesticides, but found to be potent liver carcinogen in rodents. Consequently, use of these compounds were banned; however, they have been used extensively to study the molecular mechanisms of mutagenesis and carcinogenesis (2, 6).

AAF is metabolically activated in the liver to produce a nitrenium ion, which directly reacts with cellular DNA, producing two common C8-substituted dG adducts i.e., dG-AAF and dG-AF (Figure 1a) (2, 7). During replication, the conformationally
flexible dG-AF adduct can accommodate either stacked (S) or B-type (B) conformation (8), whereas the bulky dG-AAF adopts the S or wedge (W)-conformation, which stalls DNA synthesis and thus facilitate frameshift mutations (9, 10). Primer extension studies revealed that a high fidelity polymerase blocks DNA synthesis at the site of dG-AAF in the template, however dG-AF in the similar sequence is bypassed easily (11).

1. 2. Translesion synthesis (TLS)

Translesion synthesis (TLS) occurs when a replicative polymerase was unable to copy a damaged DNA base during DNA replication. TLS can either be error-free or error-prone. A correct nucleotide is introduced opposite the lesion in error free bypass of lesion whereas in error prone lesion bypass an incorrect nucleotide is placed opposite the damaged base (12). Human DNA polymerase Pol η is one of those lesion bypass Y-family DNA polymerase. Unlike replicative polymerases, the larger open active site of this Pol η provides a room for bulky lesions such as thymine-thymine UV-induced dimer, thus promoting replication through bypass of the lesion. The activity of Pol η has been shown to be crucial in preventing the Xeroderma Pigmentosum (XPV) disease (13). However, replication by these bypass polymerases is of low fidelity and lacks high processivity (14, 15).

1. 3. Slippage mechanism of AAF-induced frameshift mutations

DNA is prone to constant attack from various agents that may alter and distort the DNA structure. These perturbations in DNA may act as a block to DNA polymerase, inducing frame shifts (16). Fuchs and coworkers observed that the majority of AAF-induced mutations in *E. coli* lacI gene are frameshift mutations. Examples include
additions and deletions of a single base pair, deletions of two adjacent base pairs, and a few three-base-pair deletions (17, 18). Molecular mechanisms causing these AAF-induced frameshift mutations have been explained using a slippage mechanism. Figure 2 explains the mechanism of slippage induced -2 deletion. In the first step, correct nucleotide dC (blue) is inserted opposite the modified guanine (red). A pause at the replication fork promotes a misalignment of the primer strand, resulting in looping out of two bases of from the template strand. Continued extension from the two nucleotide bulge in the template strand, leads to a newly synthesized strand two bases shorter than the template strand (19-21). Similarly, in -1 and -3 deletions polymerase promotes elongation of the primer by placing the correct nucleotide dC opposite the modified guanine: however, a significant proportion of the extending primer slips to pair with the dG next to the lesion in -1 deletion and two bases away from lesion in -3 deletion (shown in Figure 3 (A),(B) respectively) (22).

1.4. Previous studies on various Slipped mutagenic intermediates (SMI)

To understand the origin of adduct-induced mutations, it is necessary to evaluate the conformational differences produced by these adducts. Using high resolution $^1$H NMR, O’Handley et al (23) reported the presence of about 70% base displaced S-conformer in a AAF-modified 9-mer duplex in the CG*C sequence context. Cho and coworkers have used $^{19}$F NMR spectroscopy to study a fully paired 12-mer duplex in the TG*A context (24). The results showed the exclusive existence of the S-conformer with a possibility of the acetyl group in cis and trans equilibrium. Milhe et al. (25, 26) obtained $^1$H NMR data on the AAF modified -1 and -2 deletion duplexes in the CG*C sequence context. It was found that 70% of the AAF-population was external in the -1
deletion duplex, whereas in the -2 deletion duplex, 80% of the fluorene moiety was inserted into the bulged double helix. However, the NMR data was difficult to interpret because of conformational heterogeneity. Similar NMR studies were conducted on -1 and -2 deletion duplexes modified by the N-deacetylated dG-AF (27, 28) and anti-benzo[a]pyrene diol epoxide (BPDE) (29, 30). All these deletion duplexes except for the -1 AAF duplex, existed primarily in stacked (S) conformation, in which the aromatic carcinogen is inserted into the bulge (25). These modified duplexes showed consistent thermal stabilities (ΔTm = 11-15 °C) relative to controls, supporting well-stacked structures.

Schorr and Carell recently conducted primer extension studies with the human bypass polymerase η on variation of the NarI hot spot sequences (22). Structural instability ensued after addition of the correct dC opposite dG-AAF triggered a slippage of the dC primer end to form various bulges, depending on the base sequence around the lesion. For example, dG3-AAF in the NarI sequence (5’-G1G2CG3CC-3’) induced the -2 frameshift exclusively. Changing G2C to AT (5’-G1ATG3CC-3’) resulted in a -3 deletion mutation, presumably due to the presence of weak A:T base pairs around the lesion. Contiguous Gs around the lesion (5’-G1G2GG3GG-3’) yielded a strong -1 deletion mutation. These results suggest that the bulky dG-AAF:dC pair facilitates a slippage and the stability of the resulting looped-out SMI structure may be important for the manifestation of the frameshift mutations.

1. 5. Hypothesis and Experimental plans

We hypothesized that the thermodynamic stability of bulged-out slipped mutagenic intermediates (SMIs) is directly related to deletion mutations. The objective
of the present study was to probe the structural/conformational basis of various dG-AAF–induced SMIs formed during a translesion synthesis.

We performed spectroscopic and thermodynamic studies of several AAF-modified 16-mer model DNA duplexes, including fully paired and -1, -2, and -3 deletion duplexes of the 5’-CTCTCGATG[FAAF]CCATCAC-3’ sequence and an additional -1 deletion duplex of the 5’-CTCTCGGCG[FAAF]CCATCAC-3’ NarI sequence. The model 16-mer template used in the present study is patterned after the sequence used by Schorr and Carell in their primer kinetic studies using polymerase η, except that G2 was adducted with the fluorinated N-acetylaminofluorene (FAAF) (Figure 1b). The utility of fluorine tagged aromatic amines as an effective structure probe has been shown (24, 31). The modified template was annealed with various primers to create four model duplexes: full duplex (16/16-mer), -3 deletion (16/13-mer), -2 deletion (16/14-mer) and -1 deletion (16/15-mer). Using these duplexes, we conducted systematic spectroscopic (19F NMR, ICD) and thermodynamic (UV/melting, DSC) studies. The results show that the thermal and thermodynamic stabilities of bulky adduct-induced SMIs are key factors for determining the propensity to form different frameshift mutations.

1.6. Experimental Techniques used in this Thesis

**Duplex Melting:**

Thermal denaturation of a DNA duplex provides information about the stability of the DNA. Such information was obtained by differential scanning calorimetry (DSC) and UV-thermo melting, which are complementary to each other. DSC is a micro calorimetry technique that directly provides enthalpy and entropy values. On
the other hand, in UV melting these parameters are obtained using van’t Hoff analysis of melting curves. Melting of a DNA duplex produces a cooperative transition of a double strand to single strands. $T_m$ is the temperature at which half of the mixture exists in double stranded and other half in single stranded, thus serving as a good indicator for duplex stability. $\Delta G$ is a free energy change which represents the thermodynamic stability of a DNA duplex relative to the single strands; it depends upon environmental changes such as temperature, pressure and concentration. $\Delta H$ is an enthalpy change which deals with chemical interactions, which in turn are dependent on forces such as hydrogen bonding, van der Waal’s, dispersion forces. For example, decreased enthalpy causes destabilization in DNA because of de stacking of bases. $\Delta S$ is an entropy change associated with the conformational restrictions originated in a duplex compared to single strands (32). Adduct formation usually leads to thermal and thermodynamic destabilization (33). It is important to study the thermodynamic process as it plays an important role in understanding the mechanisms of DNA replication and repair (34). Studying the differences between thermal and thermodynamic parameters of adducted DNA duplexes relative to unmodified duplexes (eg: $\Delta \Delta G = \Delta G_{\text{modified}} - \Delta G_{\text{unmodified}}$) unfolds the secondary structural details of DNA obtained after carcinogen modification.

**1.6.1. Differential Scanning Calorimetry (DSC)**

DSC is a thermodynamic apparatus used for monitoring the changes in heat capacity ($\Delta C_p$) of a protein or DNA in aqueous solution as a function of temperature (°C). This technique monitors the thermally induced conformational changes and
stability of bio molecules, such as folding, unfolding, or melting of a DNA or protein (35).

DSC instrument consists of two cells: sample (buffer with sample of interest) and reference (only buffer). Both cells are supplied with energy simultaneously resulting in a transition endothermic peak. A typical DSC melting profile is shown in Figure 5. The midpoint of the curve determines $T_m$, which determines the stability of complex. Higher $T_m$ means a greater thermal stability. Pre and post transition baseline difference is used to determine the heat capacity, $\Delta C_p$. $\Delta H$ is enthalpy determined by calculating the areas under the curve which indicates total heat energy absorption by the sample (36, 37). The $\Delta H$ value is model independent unlike UV-melting which depends on two state model (32). The remaining thermodynamic parameters free energy ($\Delta G$) and entropy ($\Delta S$) are calculated indirectly based on the enthalpy ($\Delta H$) and heat capacity ($\Delta C_p$) values. A high concentration of sample (protein 1 mg/ml) is usually used to determine the thermodynamic parameters. This may cause problems due to accumulation of denatured sample or self-association. Because of these difficulties it is necessary to interpret the accurate concentration in DSC studies (38). In the present study we used 0.1 mM concentration of DNA duplex.

1.6.2. UV-melting experiments

UV-visible spectrometry is used for determining the stability of a DNA duplex. Absorbance at 260 nm is measured as a function of temperatures. As the temperature increases, a transition from double stranded DNA to single stranded DNAs occurs because of de stacking of DNA base pairs and uncoupling of Watson-Crick base pairs.
Absorbance increases (hyperchromicity) because of unstacked base pairs. On the other hand, duplex formation occurs on cooling, thereby decreasing absorbance (39).

A typical UV-melting profile of a double helical DNA duplex is shown in Figure 6. \( T_m \) is calculated when the UV absorbance reaches half of its final value. Enthalpy \( (\Delta H^0) \) and entropy \( (\Delta S^0) \) of melting transition is calculated by employing two different methods: fits of individual curves or plot of \( T_m^{-1} \) vs. logarithm of duplex concentration \((\ln C_t)\) in the equation \( T_m^{-1} = \frac{R}{\Delta H^0} \ln C_t/4 + \frac{\Delta S^0}{\Delta H^0} \), in which \( T_m \) is in Kelvin, \( C_t \) is concentration of the sample and \( R \) is universal gas constant. \( \Delta G^0 \) is determined through standard thermodynamic equation \( \Delta G^0 = \Delta H^0 - T\Delta S^0 \) (33). UV melting requires a small amount of samples and the variation in absorbance provides details about the transitions associated with the molecule (32). Although useful in predicting duplex stability, the UV method cannot provide structural information and is not suitable for longer (14 bases or more) oligonucleotide sequences (40). In addition, use of the two-state model in the presence of any intermediate conformation may produce inaccurate results (41).

1.6.3. Circular Dichroism (CD)

Circular dichroism spectroscopy (CD) is a spectroscopic technique for measuring the differences in absorption of a right and left handed circular polarized light of a chiral molecule. CD is sensitive in changes in the secondary structures of proteins and DNA in solution. It is a reliable technique for observing the structural alterations of macromolecule when there is change in temperature and pH or interaction with other molecules (42). Asymmetric nature of sugars in DNA backbone and its helical structures produce CD signals (43). CD spectra are usually obtained as molar
ellipticity (deg.cm$^2$/dmol$^{-1}$) vs. wavelength (nm). Typically, CD of DNA is sensitive between 200-320 nm. CD requires a small amount of samples, which could be recovered for other measurements. However, CD produces excess noise when high concentration buffers are used or when there is any particulate matters in the solution. In addition, CD cannot give detail structural information as NMR and X-ray techniques (42).

A typical B-DNA displays CD with an intense positive and negative S-shape band around 280 and 245 nm, respectively (44). Previously, Cho and coworkers carried out CD experiments on various aromatic amine-DNA adducts such as AF and AAF modified DNA duplexes. Their CD results provided insights on adduct-induced heterogeneity, distortion and disruption of Watson-Crick base pairs of DNA duplex. They observed induced circular dichroism (ICD), a phenomenon arising from the interaction of a carcinogen with DNA or protein. For example AF modified DNA duplexes showed ICD in the 290-350 nm range, which is a valuable conformational marker. Similarly, AAF modified DNA duplexes exhibited ICD around 290-320 nm (33, 45, 46). CD spectral shape explains difference in stacking of base pairs, bending and ICD of the DNA duplex after interaction with carcinogen molecule (47). An increased absorbance at 275 nm represents increased stacking between DNA base pairs and carcinogen. The wavelength shifts at 275 nm represents adduct induced DNA bending (46). Consequently, we used CD in the present study to probe the conformational changes in AAF-modified bulged duplexes. The obtained CD spectra were correlated with $^{19}$F NMR data to provide detail insights on AAF modified bulged duplexes.
1.6.4. Dynamic $^{19}$F NMR Spectroscopy

$^1$H NMR was used as an extremely valuable tool to explore and understand the solution structures of various carcinogen-DNA adducts. While the $^1$H NMR provides information about a major single conformation, the technique is not suitable for studying multiple conformations. Consequently, we employed $^{19}$F NMR to study multi-conformeric adducts. The advantage of $^{19}$F over $^1$H is its low background noise in analyzing the complex spectra. In addition, fluorine signals are easy to detect and their chemical shifts and relaxation times are highly sensitive to the local environment through hydrogen bonding, electrostatic and van der Waal’s interaction. As such, $^{19}$F makes an ideal probe in analyzing adduct-induced conformational heterogeneities (8).

In the present study, we incorporated a single fluorine probe at the longest axis of the aminofluorene moiety, i.e., at position 7. The model showed no significant differences in carcinogenicity, metabolic and conformational profiles compared to the non-fluoro parent compound. $^{19}$F acts as a conformational probe for determining the position of a lesion in various DNA duplexes (24).

Temperature dependent $^{19}$F NMR was used to study the dynamic interconversions among different conformers, adduct-induced heterogeneity and kinetic and thermodynamics parameters. At low temperature the exchange between fluorine signals was slow. As the temperature increases, the exchange became fast, so fluorine signals are broadened and move close together. At intermediate temperature the signals merge and form one coalescent signal. The $^{19}$F chemical shifts vs. temperature experiments known as dynamic NMR provide insights into both dynamic and transition of double strand helix to single strand. We have conducted a number of
experiments on 12-mer DNA duplexes modified by various carcinogens such as AF, AAF and ABP (31, 46). Usually $^{19}$F-NMR spectra are obtained in a pH 7.0 aqueous NMR buffer in 10% D$_2$O/90% H$_2$O buffer. In the present study $^{19}$F NMR was carried out to probe the conformational heterogeneity induced by FAAF-modified bulged duplexes. Assignment of $^{19}$F NMR signals is often difficult due to lack of related signals (i.e., NOE or scalar couplings, etc.). However, general dynamic, chemical shift information, and magnetic anisotropy could be obtained as reported previously. For example, dynamic $^{19}$F NMR spectra of FAAF modified guanine (G$_1$) in NarI 16mer sequence is shown in Figure 7. The B, S, W conformations were assigned accordingly based on chemical shifts, isotope effect, and ring current effects at 5°C (46, 47).

1.6.5. MALDI-TOF MS

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) is a soft ionization technique. Here mass spectra are acquired by laser ionization of a sample, followed by accelerating the ions to reach a TOF detector. The ionization process takes place by embedding the sample in an appropriate matrix. Ideally the matrix is of low molecular weight molecule and absorbs laser light effectively in the UV region thereby intensifying the ionization and vaporization of sample. With the aid of electric and magnetic field, sample with the acquired charge is accelerated in the drift tube. Mass to charge (m/z) ratio of the sample is calculated based on time required to reach the detector from source. The molecule with a higher mass takes longer time to reach the detector. Hence, based on time of flight an m/z ratio of a specific compound can be found (48).
The ability to determine accurate masses in a broad mass range makes MALDI-TOF ideal for characterizing macromolecules. The technique is also applied in characterizations of DNA and protein adducts (49, 50), sequencing (51, 52), identifying DNA-protein interactions (53), in sequencing using Sanger method (54). It is also easy to operate with high sensitivity and is predominantly flexible towards typical adulterants such as salts and buffers in samples. Although useful, it is not productive in analyzing long oligonucleotides with greater than 50 nucleotide bases due to decreased signal sensitivity and consistency. In addition, oligonucleotide samples must be desalted offline before applying to MALDI-TOF. This is contrast to Electrospray Ionization technique (ESI-MS), which can be readily coupled to chromatography apparatus for purification (55). MALDI mass spectra of oligonucleotide are easy to analyze as they generate single charged ions. However, ESI-MS produces many peaks because of multiple charged ions (56). In the present study, MALDI-TOF in conjunction with exonuclease digestion was used effectively to map FAAF modified oligonucleotides.
CHAPTER 2: EXPERIMENTAL PROCEDURES

2. 1. Materials

Oligodeoxynucleotide sequences used in the present study (shown in Figure 4) were purchased from Eurofins MWG Operon (Huntsville, AL, USA). Phosphodiesterases (3', 5') employed in mass spectrometry were purchased from Worthington Biochemical (Lakewood, NJ) and HPLC grade solvents from Fisher Inc. (Pittsburgh, PA, USA).

2. 2. Synthesis and purification of FAAF-modified oligonucleotide

To prepare a FAAF-modified 16-mer template, 0.10 mM of crude Oligodeoxynucleotide was dissolved in 1-2 ml of sodium citrate buffer (pH 6.0) with 4-5 mg of synthesized carcinogen (N-acetoxy-N-2-(acetylamino)-7-fluorofluorene) mixed in 0.5 ml of absolute ethanol and incubated at 37°C about 15-20 minutes in a water bath shaker. The reaction was periodically monitored to check the modification of Guanines in template on reversed phase HPLC. Figure 8a shows a typical HPLC chromatographic profile on a reverse phase column before purification. Unreacted oligo appeared at 11.7 min (peak 1) and modified oligos (peak 2,3,4) in the 15 - 25 min range. Peaks 2 and 3 were confirmed as G₁ and G₂-mono-FAAF-adduct, respectively, on the basis of the UV absorption intensity in the 300-350 nm region (Figure 8b) and nuclease digestion mass spectrometry analyses. Peak 3 eluting at 25 min was assigned as a di-FAAF-modified oligo (Figure 8b, see below, Results). The HPLC (Hitachi) used for the separation and purification of adducts consists of L2450 (DAD) diode array detector and L2130 Pump. A C₁₈ column from Phenomenex with dimensions (150×10mm) and 5.0 µm particle size was utilized. A linear gradient
method with 3-15% Acetonitrile for 15 min and 15-30% for 25 min was used for separating and purifying the ODNs and rate of flow was 2.0 ml/min throughout the run. The desired G₂-modified oligo (5’-CTCTCG₁ATG₂[FAAF]CCATCAC-3’) was annealed with an appropriate complementary sequence to form model duplexes at the various sequence settings (Figure 4). An identical set of unmodified control duplexes were similarly prepared. We reported previously the preparation and characterization of the other modified oligo (5’-CTCTCG₁GCG₂[FAAF] CCATCAC-3’) used for preparation of the ‘GC’ -1 deletion duplex (47).

2. 3. Characterization of FAAF-modified oligonucleotide

After collection of FAAF-modified oligonucleotides using HPLC system, the obtained modified oligonucleotides were characterized either by matrix assisted laser desorption ionization-time of flight (MALDI-TOF) or Electrospray ionization (ESI) LC/MS detection. Before characterizing the modified oligonucleotide the template sequence was digested using phosphodiesterase enzyme (both 5’ and 3’).

2. 3. 1. Enzymatic digestion/MALDI-TOF

By using equal ratio of 3-Hydroxypicolinic acid (3-HPA) and di-ammonium hydrogen citrate (DHAC), matrix solution for MALDI experiments was prepared. 1 µL of analyte was mixed with the matrix. From the above prepared solution 1 µL was taken and applied on the MALDI plate and allowed to dry for further analysis.

SVP and BSP enzymes chop nucleotide from the 3’- or 5’- end, respectively. For the SVP digest, 10× 10⁻² units enzyme was added to solution with 1 µL of the oligonucleotide sample (100-200 pmol), 6 µL of 0.1M ammonium citrate, and 6 µL of deionized water. In the BSP digestion, 1× 10⁻² units of BSP enzyme was added to
solution containing oligonucleotide sample (100-200 pmol) and 7µL of distilled water. For SVP digestion, the digest solution was incubated at 37°C, however in BSP digestion the solution was placed at room temperature. The digest solution containing sample in it (1 µL) was removed at standard time interval and spotted on the MALDI plate along with the matrix solution (1: 1 ratio of 3-HPA and DHAC) (57). All spectra were acquired using Reflectron mode through Shimadzu Axima Performance mass spectrometer. Along with MALDI-TOF we also used Electrospray ionization and quadruple time-of-flight (TOF) mass spectrometry to map the position of FAAF modification on template as described previously (47, 58).

2.4. UV-melting

Thermo melting data was acquired by employing a Cary100 Bio UV/VIS spectrophotometer holding 6×6 cell block and 1cm path length. Spectrophotometer has a peltier temperature controller to control the sample cell temperature. Oligo samples of 0.5-10 µM concentration range were prepared in buffer solution consisting of 0.2 M NaCl, 10 mM sodium phosphate and 0.2 mM EDTA at pH 7.0. Melting curves were obtained by recording the absorbance at 260 nm as a function of regulated temperature. An alternate series of heating (from 15-85°C) and cooling cycles (from 85-15°C) at 1°C/min speed was repeated for five times. Acquired data was then exported to Meltwin software (version 3.5) to obtain the thermodynamic parameters (33).

2.5. Differential Scanning Calorimetry (DSC)

DSC measurements of all five FAAF-modified duplexes were obtained by a Nano-DSC from TA Instruments (Lindon, UT, USA). 0.1 mM concentration of duplex
was dissolved in a DSC buffer (pH 7.0) consisting of 0.02 M sodium phosphate, 100 mM Sodium chloride. The sample and reference cells were loaded with the sample solution and buffer solution (control) respectively using a pipette.

In the present study at varying temperatures (15°C to 90°C) the samples were scanned against buffer at a speed of 0.75°C/min and repeated for another few scans. Raw data were collected as microwatt vs. temperature. To obtain base-line corrected $C_p^{ex}$ vs. temperature curves a buffer to buffer scan (control) was deducted from sample scan and regulated for speed of heat. $T_m$ is obtained at the midpoint of the curve. Enthalpy (ΔH) is found through integration of area under the transition curve that determines the total amount of heat absorbed by the sample after appropriate baseline correction that influences the transition. By using procedures described by Chakrabarti et al. (59) other thermodynamic parameters free energy (ΔG) and entropy (ΔS) were calculated.

2.6. Circular Dichroism (CD)

To measure CD spectra of the model duplexes, Jasco J-810 Spectropolarimeter was employed. It has a peltier system to regulate the temperature of the cell. Oligonucleotide template strand with a concentration of 6.5 µM was hybridized with equimolar concentration of primer strand and the samples were dissolved in 300 µl of a pH 7.0 buffer (0.2 M NaCl, 10 mM sodium phosphate, 0.2 mM EDTA) and placed in a 1.0 mm path length cell. To ensure correct formation of a duplex, the samples were heated and then cooled at 85°C and 15°C respectively for 10 min. Samples were scanned at a speed of 50 nm/min at 30°C from 200 to 400 nm with data points every
0.2 nm and 2 sec response time. After averaging 10 accumulations of the spectra, data was smoothed using 17-point adaptive smoothing algorithm.

2.7 Dynamic $^{19}$F Nuclear Magnetic Resonance ($^{19}$F-NMR) Spectroscopy

Approximately 0.2 mM of the G$_2$-FAAF-modified 16-mer ODN template (5′-CTCTCGATG[FAAF] CCATCAC-3′) were annealed with an equal concentration of the respective complementary sequence to produce fully paired, -3 deletion, -2 deletion, -1 deletion duplexes (Figure 4 a-d). Similarly, a FAAF-modified 16-mer NarI Sequence (5′-CTCTCGGGCG[FAAF]CCATCAC-3′) was annealed with an equimolar amount of the respective complementary sequence to produce ‘GC’-1 deletion duplexes (Figure 4e). Obtained model duplexes were mixed in 0.3 ml of NMR buffer (pH 7.0) consisting of 10% D$_2$O/90% H$_2$O, 0.1 M sodium chloride, 0.01 M sodium phosphate and 0.1M EDTA. By using sample filter of 0.45 μm membrane, samples were filtered into a Shigemi tube. Varian NMR spectrometer functioning at 500.0 and 476.5 MHz with HFC probe was used to acquire $^1$H- and $^{19}$F-NMR data respectively. Parameters for acquiring the data were obtained according to the previous study (60). By employing Wet-1D sequence imino proton spectra were recorded at 5-40°C in reference to DSS. Proton decoupled mode was used to acquire $^{19}$F NMR data with reference to CFCl$_3$ by allotting external C$_6$F$_6$ in C$_6$D$_6$ at -164.9 ppm. 1-D and 2-D NMR spectra were acquired between 5 and 78°C by slowly raising the temperature with increments of 5–10°C. BRUKER-VT unit consisting of liquid N$_2$ is used to regulate the temperatures. Computer line shape simulations were performed using WINDNMR-Pro software (61).
CHAPTER 3: RESULTS

3.1. Model Systems

The model 16-mer template (5'-CTCTCG₁ATG₂CCATCAC-3’) is patterned after the sequence used by Schorr and Carell in their primer kinetic studies (22) using polymerase $\eta$, except that G₂ was adducted with the fluorinated N-acetylamino fluorene (FAAF) (Figure 1a). The utility of fluorine tagged aromatic amine DNA adducts as an effective structure probe was mentioned previously (24, 31, 61, 62). The modified template was annealed with various primers to create four model duplexes: full duplex (16/16-mer), -3 deletion (16/13-mer), -2 deletion (16/14-mer) and -1 deletion (16/15-mer) (Figure 4 a-d). We also studied the ‘GC’ version of the 16/15-mer -1 deletion duplex, in which the two A:T base pairs between G₁ and G₂ are replaced with two G:C base pairs (5'-CTCTCG₁GCG₂CCATCAC-3’) (Figure 4-e). The same 16-mer NarI sequence in the fully paired duplex was used previously as a substrate for structural and nucleotide excision repair studies (47).

3.2. Characterization of FAAF-modified oligonucleotide templates

Figure 8a shows the HPLC profile of a reaction mixture after 20 min of reaction between a 16-mer oligo (5'-CTCTCG₁ATG₂CCATCAC-3’) and the activated FAAF carcinogen (see Experimental Procedures). Unreacted oligo appeared at 11.7 min (peak 1) and three modified oligos (peaks 2-4) followed in the 15 ~ 25 min time period. The shoulder intensity in the 290-320 nm of peak 4 (diadduct) was twice as the two early eluting ones, which confirms peaks 2 and 3 as mono-FAAF modified (Figure 8b) (33).
Figure 11 shows the MALDI-TOF spectra of the 5ꞌ→3ꞌ exonuclease digestion fragments of peak 3 at different time intervals. The ions observed at m/z 5,016 at 0 time interval (Figure 11-a) represents the molecular weight of the FAAF modified 16-mer oligo template (5ꞌ-CTCTCG₁ATG₂CCATCAC-3ꞌ) before digestion. An increase in incubation time leads to the digestion of subsequent 5ꞌ-unmodified bases. The 5ꞌ-digestions were significantly slowed down at m/z 2,897 and 2,593 for the 5ꞌ-T(FAAF)G₂CCATCAC-3ꞌ and 5ꞌ-(FAAF)G₂CCATCAC-3ꞌ fragment, respectively (see insets for theoretical MW values). Figure 12 shows the results for the 3ꞌ→5ꞌ exonuclease digestion. As expected, the 3ꞌ-digestion was significantly slowed down at m/z 2,929 after 30 s, which is consistent with the 5ꞌ-CTCTCG₁ATG₂(FAAF)-3ꞌ fragment. The 5ꞌ→3ꞌ and 3ꞌ→5ꞌ digests of peak 3 were also subjected to ESI-QTOF. The ions observed at m/z 863 and m/z 1,295 in Figure 14a are the (M-3H)³⁻ and (M-2H)²⁻ ions from the fragment 5ꞌ-G₂(FAAF)CCATCAC-3ꞌ. Similarly, the ions at m/z 975 and m/z 1,463 in the 3→5ꞌ digest correspond to the (M-3H)³⁻ and (M-2H)²⁻ ions formed from the fragment 5ꞌ-CTCTCG₁TAG₂(FAAF)-3ꞌ (Figure 14b). Taken together, these results confirm peak 3 as G₂-FAAF modified 16-mer template.

The early eluting mono adduct peak 2 was characterized similarly. As expected, the 5ꞌ→3ꞌ digestion (Figure 9) was slowed down significantly at 60 min at m/z 3,540, representing an ion formed from the 5ꞌ-G₁(FAAF)ATG₂CCATCAC-3ꞌ fragment. By contrast, the 3ꞌ digestion (Figure 10) was very efficient, but halted at m/z 1,982, consistent with the 5ꞌ-CTCTCG₁(FAAF)-3ꞌ fragment ion. The ESI spectra of the 5ꞌ→3ꞌ digest exhibited the (M-3H)³⁻ of the 5ꞌ-G₁(FAAF)ATG₂CCATCAC-3ꞌ fragment at m/z 1,178 ion is the (M-3H)³⁻ (Figure 13a). The 3ꞌ→5ꞌ digest shows ions at m/z 990, which
corresponds to the \((\text{M-2H})^{2-}\) ion formed from the fragment \(5'\text{-CTCTCG}_{1}(\text{FAAF})-3'\) (Figure 13b). These results confirm peak 2 as FAAF-G\(_1\).

3.3. UV-melting

Figure 15 shows the UV-melting profiles of FAAF-modified duplexes (red) relative to their respective unmodified controls (blue) at 10 \(\mu\)M. All of the duplexes showed monophasic and sigmoidal curves with a strong linear correlation \((R^2 > 0.9)\) between \(T_m^{-1}\) and \(\ln C_t\), confirming their typical cooperative helix-coil melting transition. Thermal and thermodynamic parameters calculated from these curves are tabulated in Table 1. It is clear that the full and -3 deletion duplexes were destabilized thermally \((\Delta T_m)\) and thermodynamically \((\Delta\Delta G_{37^\circ C})\) upon FAAF modification: full duplex \((-7.7 \, ^\circ C, 4.5 \, \text{kcal/mol}) > -3\) deletion \((-5.7 \, ^\circ C, 2.0 \, \text{kcal/mol})\). By contrast, -2 and -1 deletion duplexes displayed a strong stabilization: ‘AT’ -1 deletion \((13.0 \, ^\circ C, -8.5 \, \text{kcal/mol})\) > -2 deletion \((9.5 \, ^\circ C, -7.2 \, \text{kcal/mol})\). The ‘GC’ -1 deletion duplex displayed thermal and thermodynamic stabilization \((13.3 \, ^\circ C, -7.1 \, \text{kcal/mol})\) similar to that of the ‘AT’ counterpart.

3.4. Differential Scanning Calorimetry (DSC)

Figure 16 shows DSC plots of excess heat capacity \(C_p^{ex}\) vs. temperature for all five FAAF-modified duplexes relative to their unmodified controls. These curves are transformed into the corresponding thermodynamic histograms (Figure 17), and the results are tabulated in Table 2. The median peak of each bell curve indicates \(T_m\), at which half of the duplexes melt, whereas the areas under the curve represent the enthalpy \((\Delta H)\) of duplex formation \((37)\). Consistent with the UV melting data, the ‘AT’-1 deletion duplex was most stabilized \((\Delta\Delta G_{37^\circ C} = -6.3 \, \text{kcal/mol}, \Delta T_m = 15.2 \, ^\circ C)\),
followed by -2 deletion duplex ($\Delta \Delta G_{37^\circ C} = -3.5$ kcal/mol, $\Delta T_m = 10.8 \, ^\circ C$). As expected, the fully paired duplex was most destabilized ($\Delta \Delta G_{37^\circ C} = 3.5$ kcal/mol, $\Delta T_m = -8.0 \, ^\circ C$) followed by -3 deletion duplex ($\Delta \Delta G_{37^\circ C} = 0.8$ kcal/mol, $\Delta T_m = -3.1 \, ^\circ C$).

Unlike fully-paired duplexes where stacking usually promotes entropy, the mostly stacked (72-73% S) modified -1 deletion duplexes exhibited decreased entropy relative to the the control deletion duplexes (‘AT’; $\Delta \Delta S = -50.8$ eu and ‘GC’ $\Delta \Delta G_{37^\circ C} = -31.0$ eu). The decreased entropy, however, was compensated by large enthalpies to produce a net gain in overall free energy (‘AT’; $\Delta \Delta H = -6.3$ kcal/mol and ‘GC’ $\Delta \Delta G_{37^\circ C} = -3.5$ kcal/mol). The -3 deletion duplex (52% B-type), however, displayed a slight entropy gain ($\Delta \Delta S = 8.4$ eu), but was compensated ($\Delta \Delta H = 3.5$ kcal/mol) to produce a small loss of free energy ($\Delta \Delta G_{37^\circ C} = 0.8$ kcal/mol). Largely stacked (60%) fully-paired duplexes are known to have disrupted Watson-Crick base pairing at the lesion site, thus resulting in enthalpy reduction ($\Delta \Delta H = 13.9$ Kcal/mol). Here too enthalpy-entropy compensation afforded a loss of free energy ($\Delta \Delta G_{37^\circ C} = 3.5$ kcal/mol) (63).

3.5. **Induced Circular Dichroism (ICD).**

Figure 18 shows CD spectral overalys of all five FAAF-modified duplexes (solid lines) relative to their respective unmodified controls (dotted lines) at 30°C. All duplexes exhibited a CD pattern characteristic for a typical B-form DNA duplex with a (+)275nm/(-)250nm S-shape curve. All modified duplexes displayed small negative ICD ellipticity around the 290-320 nm range (45, 46, 63).

A slight increase in the positive intensity (hyperchromicity) at 275 nm was noted for -2 and -1 (both ‘AT’ and ‘GC’) deletion duplexes. This is likely due to lesion-
induced duplex stability caused by increased stacking of the planar aromatic carcinogen in the bulge (see Thermodynamics above). The opposite (hypochromicity) was observed for the fully-paired and -3 deletion duplexes with the effect much greater for the latter. Again, the trend is in agreement with the differences in conformational population (inserted S-type vs. external B-type) as well as the thermodynamic results.

The modified duplexes also displayed significant blue shifts relative to their respective unmodified controls, signifying adduct-induced DNA bending (47). Initially, we probed the bending of unmodified deletion duplexes relative to the fully-paired unmodified counterparts. No wavelength shift was observed except for -3 deletion duplex which exhibited 5 nm of blue shift (Table 3). Similarly, we compared modified deletion duplexes relative to their unmodified counterparts (Figure 18). All except for -3 deletion exhibited significant blue shifts upon FAAF-modification: full duplex ($\Delta G^{*,G} = 5$ nm) > ‘AT’-1 deletion ~ ‘GC’ -1 deletion ($\Delta G^{*,G} = 4$ nm) > -2 deletion ($\Delta G^{*,G} = 3$ nm). We have shown previously that FAAF in the CG*C context exists mostly (61%) in the stacked conformation (47). Hence, the greater blue shift observed for a fully-paired duplex could be attributed to major structural disturbance at the lesion site. The significant blue shifts in deletion duplexes could also indicate a stacked-induced bending of DNA, i.e., ‘AT’ and ‘GC’ -1 (72-73% S) followed by -2 (55%) deletion duplexes. An exception was the FAAF-modified -3 deletion duplex, which displayed a shift to longer wavelengths ($\Delta G_{G,G^*} = 3$ nm). This may be due to its relatively high B-type conformation (52%) compared to other deletion duplexes.
When compared to the fully-paired unmodified duplex, however, the -3 deletion duplex displayed a blue shift of 3 nm (Figure 19).

3.6. Dynamic $^{19}$F NMR: conformational heterogeneity

Figure 20 shows dynamic $^{19}$F NMR spectra of all five FAAF-modified duplexes. While signal patterns vary, all exhibited sharp single signals around -115 ppm at coalescence temperatures, signifying duplex melting, $i.e.$, 70, 60, 65, 75, 78°C for full, -3, -2, ‘AT’ -1, ‘GC’ -1 deletion duplexes, respectively.

Our previous $^{19}$F NMR studies of fully paired FAAF-duplexes in various flanking sequence contexts (TG*A, CG*C, CG*G, GG*C) have revealed a mixture of B-, S- and W-conformations in the -115.0 ~ -115.5 ppm, -115.5 ~ -117.0 ppm and -116.5 ~ -118.0 ppm ranges, respectively (46, 47). Consequently, we assigned the -115.6 and -117.9 ppm signals at 5°C of the fully paired duplex as the S- and W-conformations, respectively (Figure 20a). A similar strategy was used to assign the signals in the $^{19}$F NMR spectra of deletion duplexes, $i.e.$, the signals (~ -115 ppm) near the coalescence temperatures arise from non-stacked external binding B-type conformers and $^{19}$F shielding is a hallmark for the van der Waals interactions of the carcinogen moiety within the bulge duplexes ($e.g.$, S-type conformation).

The ‘GC’ -1 deletion duplex displayed two major signals at 10°C (Figure 20e). The signal at -116.0 ppm could be assigned as the B-type conformer (purple) because of its close proximity to the coalescent signal (single strand). The major signal at -116.5 ppm was resistant to melting and its sharp signal persisted even at 70°C! This must be an inserted S-type conformation (red), undergoing a melting transition in the 70-75°C
range. Moreover, the unusually high coalescence melting (78°C) is in excellent agreement with greater thermal and thermodynamic stability (Table 1 and 2).

A similar $^{19}$F NMR pattern was obtained for the ‘AT’ -1 deletion duplex, i.e., B (-115.5 ppm, purple) and S (-116.6 ppm, red) at 10°C (Figure 20d). A strong signal at -116.2 ppm (marked in asterisk) is located in between S and B and coalesced to S around 40°C. This signal could be an intermediate conformer between B and S conformers (see below), but is unique to the ‘AT’ -1 duplex. Other than that, the two -1 duplexes exhibited similar dynamic patterns. The greater conformational stability (NMR) of the ‘GC’ over ‘AT’ -1 deletion duplex is in good agreement with the thermodynamic (DSC) results above and probably due to the better hydrogen-bonding capability of the 5’-G:C over A:T.

The -3 and -2 deletion duplexes also displayed two major $^{19}$F signals in the -114 ~ -115 and -114.5 ~ -115.5 ppm range, respectively (Figure 20-b,c). However, these signals are collectively shifted downfield by about 1 ppm relative to those of the -1 duplexes, suggesting an altered electronic environment, i.e., an open and flexible bulge. Again, the downfield and upfield signals could be assigned as B and S conformations, respectively, based on the $^{19}$F shielding and ring current effect. Additional signals marked with asterisk could be intermediate conformers.

Figure 21 compares $^{19}$F NMR spectra of all the modified deletion duplexes at 30°C. The percent population ratios were calculated by line simulations (red). The population of S conformers was in order of ‘GC’-1 (73%) ~‘AT’ -1 (72%) > full (60%) > -2 (55%) > -3 (37%) for the deletion duplexes. The S/W population ratio for the full duplex was 60:40. Lesion-induced conformational change in the bulge
structures is also evident from their imino proton spectra at 5°C (Figure 22). In all cases, a mixture of broad imino signals was observed for the Watson-Crick hydrogen bonds (12~14 ppm) as well as those at and near the lesion site (10 ~ 12 ppm). As shown in Figure 23, unlike the fully-paired duplex the high field imino proton signals of the -3 deletion duplex disappeared rapidly as the temperatures increase, indicating lesion-induced conformational flexibility at and near the lesion site. In contrast, the -1 deletion duplexes, particularly ‘GC’ -1, showed strong persistence of the high field imino protons even at higher temperatures (Figure 24). These imino proton results indicate a tightly packed structure at the lesion site, consistent with the thermodynamic stabilization and increased stacking with strongly positive CD around 270 nm.
CHAPTER 4: DISCUSSION

The solution structures of arylamine-modified DNA duplexes, including AF- and AAF, have been studied in various sequence settings (8, 10, 64). In the fully paired duplexes, these lesions exist largely in a mixture of the major groove anti-glycosidic B and the intercalated syn-glycosidic S conformers (8, 10, 61, 64). Additionally, the bulky N-acetylated AAF adopts a W conformer, in which the carcinogen is placed in the narrow minor groove (46). The relative population balance of these conformers is governed primarily by the nature of the base sequences surrounding the lesion. The resulting conformational heterogeneity has been shown to be responsible for different mutational and repair outcomes. Milhe and coworkers conducted \(^{1}\)H NMR study on AAF modified 11/10-mer -1 deletion duplex in the CG*C context (25). They found that 70% of the aminofluorene moiety was located externally (e.g., B-type) and conformational heterogeneity prohibited further analysis of the remaining conformers. This is clearly contrasted to the present result, which showed 73% of S conformation for the 16/15-mer ‘GC’-1 deletion duplex. However, a AAF on the 12/10-mer -2 deletion duplex exhibited S-type conformation (26), in good agreement with our result in the present study. Similar NMR studies on -1 and -2 deletion duplexes modified by the N-deacetylated dG-AF (27, 28) showed exclusively S-type conformations. All these cases, greater thermal stabilities (\(\Delta T_m= 11-15^\circ\)) have been observed (10, 30).

We hypothesized that the dG-AAF-modified deletion duplexes exist as a combination of the external-binding B, the inserted S, and some intermediate conformations between them. Our working hypothesis was that dG-AAF induces
various sequence-dependent slippages during replication, and that the conformational and thermodynamic stabilities of the resulting SMIs are key factors for determining the different types of frameshift mutations.

Our dynamic \(^{19}\)F-NMR results (Figure 20) showed that dG-FAAF existed in varying ratios of B- and S-types, along with a few intermediate conformers. A greater population of the S-type conformation was observed for the ‘GC’ (73%) and ‘AT’ (72%) -1 duplexes as compared to the -2 (55%) and -3 (37%) deletion duplexes. This result was probably due to the favorable stacking of the intercalated aminofluorene moiety in the bulges. The increased positive CD at 275 nm was caused by an increase of \(\pi-\pi\) stacking interactions. A blue CD shift indicated distortion and bending of the lesion-induced DNA (Figure 18).

Fully paired arylamine-modified duplexes are known to destabilize thermally and thermodynamically relative to the unmodified controls (10). The dG-FAAF produced a mixture of complex S/B/W-conformers. The bulky N-acetyl group was apparently responsible for producing up to 40% W-conformer in the fully paired 16-mer duplex (Figure 20-a) (46, 47). The destabilizing effect of the dG-FAAF lesion was related to the difference in the conformational populations. Thus, the 60% S-conformeric full duplex disrupted the Watson-Crick base pairs at the lesion site, resulting in thermal (\(\Delta T_m = -8.0^\circ\text{C}\)) and enthalpic destabilizations (\(\Delta\Delta H = 13.9\) kcal/mol). Our previous conformational and thermodynamic studies of the NarI 16-mer full duplex also revealed 61% S conformations, which resulted in thermal (\(\Delta T_m = -8.3^\circ\text{C}\)) and enthalpic destabilizations (\(\Delta\Delta H = 24.7\) kcal/mol).
In contrast to the fully-paired duplex, the S-conformeric bulged structures resulted in thermal and thermodynamic stabilizations. For example, a dramatic thermal stabilization has been observed for several bulge duplexes modified by the bulky aryl amine and benzo[a]pyrene carcinogens (10, 30). Modeling studies suggested that the syn-glycosidic conformer SMI was more stable than the anti-external conformation. This syn-SMI was stabilized by favorable interactions between the carcinogen and flanking base pairs inside the bulge pocket. The highly stacked nature of the ‘GC’ (73%) and ‘AT’ (72%) -1 deletion duplexes resulted in enthalpic stabilization (ΔΔH = -15.2 kcal/mol and ΔΔH = -21.9 kcal/mol, respectively). As expected, the 55% S conformeric -2 deletion duplex displayed intermediate thermal (ΔTm = 10.8°) and enthalpic stabilizations (ΔΔH = -9.6 kcal/mol). A similar, but truncated (12/10mer), NarI -2 duplex displayed thermal (ΔTm = 11.7°) and enthalpic stabilizations (ΔΔH = -5.3 kcal/mol). As expected, the highly flexible B-conformeric (52%) -3 deletion duplex exhibited enthalpic destabilization (ΔΔH = 3.5 kcal/mol).

Schorr and Carell (22) conducted primer extension studies with the human bypass polymerase η on several dG-AAF–modified NarI sequences. The addition of the correct dC opposite the lesion during replication triggered the dC primer end to slip and resulted in the formation of various SMIs, depending on the base sequence context around the lesion. The dC located 3’ to the AAF lesion was found to be critical for the misalignment/shifting process. Strong -2 deletion was observed in the NarI 5’-CGGCG[AAF]CC---3’ or 5’---CGGCG[AAF]CT---3’ sequences. When all of the dC bases around the lesion were replaced by dG (i.e., 5’---CGGGG[AAF]GG---3’), only the -1 slippage was obtained with a high yield. A rare -3 deletion occurred when the
sequence was modified appropriately to exclude 3 bases. In other words, the efficiency of the frame shifting process was in order of -1 > -2 > -3 deletions.

In full agreement with the Pol η primer extension assay data, our $^{19}$F NMR results showed that the AAF-modified opposite -1, -2, and -3 deletion duplexes had 72%, 55%, and 37% intercalated bulge conformations. The highly stacked S-conformeric -1 and -2 deletion duplexes displayed tight compactness of the lesion in the bulge, and consequently displayed greater thermal and thermodynamic stabilities. Previous modeling studies found that the AAF-modified SMI was quite stable relative to its normally extended counterpart in the presence of the syn-carcinogen conformation (65). In the -3 deletion duplex, however, the lesion was not stacked well, resulting in greater conformational flexibility. These results indicate that the optimum space required to incorporate the AAF lesion is the -1 followed by the -2 bulge.

In conclusion, a frameshift is triggered by the inclusion of a correct dC opposite the bulky AAF lesion at the replication fork, which can be accommodated by either the anti- or syn-glycosidic conformation. The structural and conformational instabilities of such an accommodation facilitate the development of various misalignments, which depend on the nature of the base sequences that are in close proximity to the modified dG. The structural data in the present study show that the thermal and thermodynamic stabilities of bulky adduct-induced SMIs are key factors for determining the propensity to form different frameshift mutations.
<table>
<thead>
<tr>
<th></th>
<th>-ΔH Kcal/mol</th>
<th>-ΔS eu</th>
<th>-ΔG37°C Kcal/mol</th>
<th>T_{m}^{b} °C</th>
<th>ΔΔH^{c} Kcal/mol</th>
<th>ΔΔS^{d} eu</th>
<th>ΔΔG^{e} 37°C Kcal/mol</th>
<th>ΔT_{m}^{f} °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full duplex</td>
<td>122.3</td>
<td>101.0</td>
<td>333.2</td>
<td>19.0</td>
<td>14.5</td>
<td>62.0</td>
<td>55.3</td>
<td>21.3</td>
</tr>
<tr>
<td>-3 deletion</td>
<td>99.0</td>
<td>88.7</td>
<td>276.6</td>
<td>13.3</td>
<td>11.5</td>
<td>52.2</td>
<td>46.5</td>
<td>10.3</td>
</tr>
<tr>
<td>‘AT’ -1 deletion</td>
<td>100.2</td>
<td>122.5</td>
<td>280.0</td>
<td>13.3</td>
<td>15.7</td>
<td>47.0</td>
<td>56.5</td>
<td>-21.8</td>
</tr>
<tr>
<td>‘GC’ -1 deletion</td>
<td>103.1</td>
<td>123.7</td>
<td>282.9</td>
<td>15.4</td>
<td>17.5</td>
<td>53.7</td>
<td>67.0</td>
<td>-20.6</td>
</tr>
</tbody>
</table>

Blue: Unmodified, Red: Modified

a) The average standard deviation for -ΔΔG, -ΔH and T_{m}^{b} are ±0.4, ±3.0 and ±0.4 respectively.

b) T_{m} value is the temperature at half the peak area.

c) ΔΔH = ΔH(modified duplex) - ΔH(Control)

d) ΔΔS = ΔS(modified duplex) - ΔS(Control)

e) ΔΔG = ΔG(modified duplex) - ΔG(Control)

f) ΔT_{m} = T_{m}(modified duplex) - T_{m}(Control)
Table 2: Thermal and thermodynamic parameters derived from Differential Scanning Calorimetry (DSC)

<table>
<thead>
<tr>
<th></th>
<th>$-\Delta H$ Kcal/mol</th>
<th>$-\Delta S$ eu</th>
<th>$-\Delta G$EC Kcal/mol</th>
<th>$T_m^b$ °C</th>
<th>$\Delta \Delta H$ Kcal/mol</th>
<th>$\Delta \Delta S$ d eu</th>
<th>$\Delta \Delta G$EC Kcal/mol</th>
<th>$\Delta T_m^f$ °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full duplex</td>
<td>115.7</td>
<td>101.8</td>
<td>320.8</td>
<td>287.0</td>
<td>16.3</td>
<td>12.8</td>
<td>66.8</td>
<td>58.8</td>
</tr>
<tr>
<td>-3 deletion</td>
<td>76.9</td>
<td>73.4</td>
<td>214.6</td>
<td>206.2</td>
<td>10.3</td>
<td>9.5</td>
<td>54.9</td>
<td>51.8</td>
</tr>
<tr>
<td>-2 deletion</td>
<td>99.6</td>
<td>121.5</td>
<td>254.1</td>
<td>274.1</td>
<td>9.5</td>
<td>13.0</td>
<td>49.4</td>
<td>60.2</td>
</tr>
<tr>
<td>‘AT’ -1 deletion</td>
<td>90.3</td>
<td>105.5</td>
<td>250.0</td>
<td>281.0</td>
<td>12.8</td>
<td>17.6</td>
<td>61.7</td>
<td>77.7</td>
</tr>
<tr>
<td>‘GC’ -1 deletion</td>
<td>88.3</td>
<td>97.9</td>
<td>254.1</td>
<td>274.1</td>
<td>9.5</td>
<td>13.0</td>
<td>54.0</td>
<td>69.2</td>
</tr>
</tbody>
</table>

Blue: Unmodified, Red: Modified

a) The average standard deviation for $-\Delta G$, $-\Delta H$ and $T_m^b$ are $\pm 0.4$, $\pm 3.0$ and $\pm 0.4$ respectively.

b) $T_m$ value is the temperature at half the peak area.

c) $\Delta \Delta H = \Delta H$(modified duplex) - $\Delta H$(Control)

d) $\Delta \Delta S = \Delta S$(modified duplex) - $\Delta S$(Control)

e) $\Delta \Delta G = \Delta G$(modified duplex) - $\Delta G$(Control)

f) $\Delta T_m = T_m$(modified duplex) - $T_m$(Control)
<table>
<thead>
<tr>
<th>Type of Shift</th>
<th>Wavelength (nm)</th>
<th>Shift in wavelength (nm) ($\Delta G - G^*$)</th>
<th>Type of Shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>275 nm</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-3 deletion</td>
<td>270 nm</td>
<td>5 nm</td>
<td>Blue shift</td>
</tr>
<tr>
<td>-2 deletion</td>
<td>270 nm</td>
<td>0 nm</td>
<td>no shift</td>
</tr>
<tr>
<td>'AT'-1 deletion</td>
<td>270 nm</td>
<td>0 nm</td>
<td>no shift</td>
</tr>
</tbody>
</table>

G: Full Unmodified duplex wavelength

G*: Unmodified duplex wavelength
Table 4: Circular dichroism (CD) shifts in wavelength of FAAF modified duplexes with respective unmodified counterparts

<table>
<thead>
<tr>
<th>Type of Shift</th>
<th>Unmodified (G) Wavelength (nm)</th>
<th>FAAF modified (G*) Wavelength (nm)</th>
<th>Shift in wavelength (nm) ($\Delta_{G,G^*}$)</th>
<th>Type of Shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full duplex</td>
<td>275 nm</td>
<td>270 nm</td>
<td>5 nm</td>
<td>Blue shift</td>
</tr>
<tr>
<td>-3 deletion</td>
<td>270 nm</td>
<td>273 nm</td>
<td>-3 nm</td>
<td>Red shift</td>
</tr>
<tr>
<td>-2 deletion</td>
<td>275 nm</td>
<td>272 nm</td>
<td>3 nm</td>
<td>Blue shift</td>
</tr>
<tr>
<td>'AT'-1 deletion</td>
<td>275 nm</td>
<td>271 nm</td>
<td>4 nm</td>
<td>Blue shift</td>
</tr>
<tr>
<td>'GC'-1 deletion</td>
<td>275 nm</td>
<td>271 nm</td>
<td>4 nm</td>
<td>Blue shift</td>
</tr>
</tbody>
</table>

G: Unmodified duplexes

G*: FAAF modified duplexes
Figure 1: Structures and template sequences used in this study. (a) Chemical structures of AAF, FAAF, AF, FAF adducts (b) 16-mer non NarI sequence and (c) 16-mer NarI sequences used in the present study.
Figure 2: Mechanism for slippage induced -2 deletion on NarI sequence [Roy, D., Hingerty, B. E., Shapiro, R., and Broyde, S. Chem Res Toxicol 11, 1301-1311 (1998)]
Figure 3: Mechanisms for slippage induced (A) Single base frameshift mutagenesis (B) -3 frameshift mutagenesis.

Figure 3(A) adapted from [Roy, D., Hingerty, B. E., Shapiro, R., and Broyde, S. Chem Res Toxicol 11, 1301-1311, (1998)] and 3(B) from [S. Schorr and T. Carell; ChemBioChem 11, 2534-2537 (2010)]
Figure 4: Model duplexes studied in the present study (a) Full duplex (b) -3 deletion duplex (c) -2 deletion duplex (d) ‘AT’ -1 deletion duplex (e) ‘GC’ -1 deletion duplex
**Figure 5:** A Typical DSC melting profile for the two-state unfolding of a protein.

Figure 6: Typical UV-melting profile for the two-state unfolding of a DNA helix
Figure 7: Temperature dependent $^{19}$F NMR spectra of FAAF modified G$_1$ in NarI 16-mer sequence [Jain, V., Hilton, B., Patnaik, S., Zou, Y., Chiarelli, M. P., and Cho, B. P. *Nucleic Acids Res* 40, 3939-395 (2012)]
Figure 8: (a) HPLC chromatogram of a reaction mixture derived from treatment of the 16-mer sequence (5'-CTCTCG1ATG2CCATCAC-3') with an activated FAAF (N-acetoxy-N-2-(acetylamino)-7-fluorofluorene). Mono- (G1, G2) and di-FAAF adducts are eluted in the 15-25 min range (b) On-line photodiode array UV spectra of unmodified, mono-, and di-FAAF adducts.
Figure 9: MALDI-TOF mass spectra of peak 2 after 5’→3’ exonuclease digestion (BSP enzyme) in Reflectron mode at various time intervals. Insets provide theoretical MW of the corresponding fragments.
Figure 10: MALDI-TOF mass spectra of peak 2 after 3’→5’ exonuclease digestion (SVP enzyme) in Reflectron mode at various time intervals. Insets provide theoretical MW of the corresponding fragments.
**Figure 11:** MALDI-TOF mass spectra of peak 3 after 5’→3’ exonuclease digestion (BSP enzyme) in Reflectron mode at various time intervals. Insets provide theoretical MW of the corresponding fragments.
Figure 12: MALDI-TOF mass spectra of peak 3 after 3'→5'exonuclease digestion (SVP enzyme) in Reflectron mode at various time intervals. Insets provide theoretical MW of the corresponding fragments.
Figure 13: ESI spectra of peak 2 obtained from FAAF-modification of 16-mer template (5'-CTCTCG\textsubscript{1}ATG\textsubscript{2}CCATCAC-3') after (a) 5'-exonuclease digestion (b) 3'-exonuclease digestion
Figure 14: ESI spectra of peak 3 obtained from FAAF-modification of (5’-CTCTCG₁ATG₂CCATCAC-3’) 16-mer template after a) 5’-exonuclease digestion b) 3’-exonuclease digestion.
Figure 15: UV-melting curves of FAAF modified duplexes (red) and their respective unmodified controls (blue) all at 10 μM in 0.2 M NaCl, 10 mM sodium phosphate, and 0.2 mM EDTA at pH 7.
Figure 16: Differential scanning calorimetry (DSC) curves in 20 mM phosphate buffer containing 0.1 M NaCl at pH 7: FAAF-modified duplexes in red and unmodified controls in blue.
Figure 17: Comparative thermodynamic histograms of FAAF-modified duplexes with respective to their unmodified controls: $\Delta \Delta H = \Delta H$ (modified duplex) - $\Delta H$ (control duplex), $\Delta \Delta S = \Delta S$ (modified duplex) - $\Delta S$ (control duplex) and $\Delta \Delta G = \Delta G$ (modified duplex) – $\Delta G$ (control duplex).
Figure 18: CD Spectral overlays of (a) fully-paired (b) -3 deletion (c) -2 deletion (d) ‘AT’ -1 deletion, and (e) ‘GC’ -1 deletion duplexes at 30 °C. FAAF modified duplexes (solid lines) and unmodified control duplexes (dotted lines)
Figure 19: CD Spectral overlays of FAAF-modified -3 deletion duplex (black) relative to the unmodified -3 deletion (black) and fully-paired unmodified duplexes (green).
Figure 20: Dynamic $^{19}$F NMR spectra of FAAF-modified duplexes in various sequence settings: (a) fully-paired (b) -3 deletion (c) -2 deletion (d) ‘AT’ -1 deletion and (e) ‘GC’-1 deletion duplexes. S, B, and W-conformers are labeled as red, purple, and green colors. Denatured FAAF-modified single strands (SS) are colored in cyan.
Figure 21: Line simulation of FAAF-modified 16-mer (5'-CTCTCG1ATG2CCATCAC-3') (a) -3 deletion (b) -2 deletion (c) ‘AT’-1 deletion duplexes and FAAF modified NarI 16-mer (5'-CTCTCG1GC3CCATCAC-3') (e) ‘GC’-1 deletion duplex at 30 °C.
Figure 22: Imino proton region (11-15 ppm) of proton NMR of FAAF-modified 16-mer (5'-CTCTCG_{1}ATG_{2}CCATCAC-3') (a) Full duplex (b) -3 deletion (c) -2 deletion (d) 'AT' -1 deletion (e) 'GC' -1 deletion duplex (upper panel) with respective to the unmodified duplexes (lower panel) at 5°C.
Figure 23: Comparison of the dynamic imino proton spectra of FAAF-modified (a) fully-paired, (b) -3, and (c) -2 (right panel) relative to their respective controls (left panel) at 5, 20, 30, 40°C
Figure 24: Comparison of the dynamic imino proton spectra of FAAF-modified (d) ‘AT’ and (e) ‘GC’ -1 duplexes (right panel) relative to their respective controls (left panel) at 5, 20, 30, 40°C.


65. Roy, D., Hingerty, B. E., Shapiro, R., and Broyde, S. (1998) A slipped replication intermediate model is stabilized by the syn orientation of N-2-aminofluorene- and N-