STRUCTURAL STUDIES OF FLUXIONAL LESIONS IN DEOXYRIBONCULEIC ACID

By

KYLE LAMAR BROWN

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry

August, 2008

Nashville, Tennessee

Approved:

Michael P. Stone

Brian O. Bachmann

Terry P. Lybrand

Andrzej M. Krezel

F. Peter Guengerich

To my parents,

Joe and Barbara

ACKNOWLEDGEMENTS

I would like to thank Vanderbilt University for the opportunity to study at this institution. I thank the Vanderbilt Department of Chemistry, Center in Molecular Toxicology, Center for Structural Biology, and the Vanderbilt-Ingram Cancer Center for contributions to this work. I especially thank the National Institutes of Health for funding.

I would like to thank my advisor, Dr. Michael Stone, for providing a stable work environment to learn and grow while pursuing this degree. In addition I thank members of my Ph.D. committee: Dr. Brian Bachmann, Dr. Fred Guengerich, Dr. Andrzej Krezel, and Dr. Terry Lybrand. I have benefited from the opportunity to interact and learn from these individuals through classroom settings, seminars, and conversation. I offer a special thanks to Drs. Tom and Connie Harris. Although retired, Tom and Connie Harris have shown a special interest in my development as a scientist and for that I will always be grateful. What I have learned from all these people will continue to influence and shape my life long into the future.

A multidisciplinary approach to science often requires the involvement of various collaborators; the present work is no different. This research would not have been possible without the contributions of Dr. Ashis Basu and Dr. Carmelo Rizzo. Dr. Basu and members of his lab were instrumental in the preparation of thymine glycol samples. Dr. Rizzo has offered sage advice concerning my studies. I especially thank the members of Dr. Rizzo's lab, particularly Dr. Ivan Kozekov and Mrs. Pam Tamura for their

instruction in various analytical techniques and Dr. Plamen Christov for his preparation of methyl formamidopyrimidine samples.

Technical expertise was essential during many aspects of this research. I thank Dr. Nick Ulyanov for his assistance with molecular dynamic calculations and strategy. Also, I thank Dr. Jarrod Smith for his instruction in computational techniques in regard to structural biology. Likewise, Dr. Jonathan Sheehan and Dr. Eric Dawson were always helpful with molecular modeling techniques. I thank Dr. Prasad Polavarapu, Dr. Andes Hess, and Dr. Joel Tellinghuisen for their assistance with quantum theory and calculations. Thanks to Dr. Carlos de los Santos for help with thermodynamic measurements. I appreciate the assistance of Ellen Rochelle and Roy Hoffman with computer related problems. Dr. Joel Harp was helpful with crystallography. I also thank Dr. Don Stec and Dr. Jaison Jacob for their assistance with NMR spectroscopy. I would be remiss not to thank the enormous contributions of Dr. Markus Voehler. Without Markus' technical expertise and friendship this research would not have been possible.

Members of the Stone lab have been supportive of this research over the years. I especially thank the following individuals: Keith Merrit, Young Jin Cho, Ewa Kowal, Hai Huang, Feng Wang, Wen Xu, Ganesh Shanmugam, Surajit Banerjee, Tandace Scholdberg, Sarah Musser, Travis Nielsen, Hye-Young Kim, Mary Kerske, Amy Williams, and Eric Chapdelaine. In addition to these members, I have had the privilege of working with many helpful graduate and undergraduate rotation students: Dan Brown, Travis Adams, Shane Magee, Jerry Yen, Jamil Stetler, Stephen Dean, Nick Demuro, and Yury Kuzmichev.

In closing, I want to thank friends and family for all the love, encouragement, and support they have provide to me, I love you all. A special thanks goes to parents, Joe and Barbara, a guy could not ask for better parents.

Above all, I want to thank God for "miracling" me through this entire process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

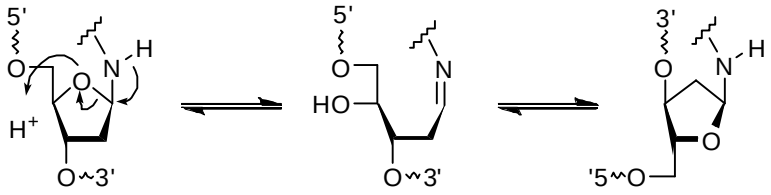
LIST OF SCHEMES

CHAPTER I

INTRODUCTION

Qualitative changes in the informational macromolecule, deoxyribonucleic acid (DNA), are involved in the aging process, neurological degradation, and carcinogenesis [1-10]. Reactive oxygen species and multiple endogenous and exogenous chemical agents are known to produce structural changes in nucleic acids [11-13]. The biological consequences of modified DNA occur with genomic replication and transcription and interaction with repair enzymes [14, 15]. A DNA lesion may be innocuous or it may be repaired prior to deleterious results, but it could also be mutagenic or toxic [14, 16-18]. The consequences of these interactions are a direct result of altered shape, stability, or chemistry of modified DNA [19]. The analysis of damaged DNA structure and stability provides a valuable tool for the elucidation of a lesion's biological consequences [20, 21].

Anomers are a specific type of epimer found in cyclic saccharides. Specifically, they are diastereomers differing in their configuration at the hemiacetal carbon [22, 23]. Canonical nucleic acids contain β ribose or deoxyribose anomers while other biological molecules (e.g. carbohydrates, polysaccharides) may contain both α and β anomers [24]. However, α anomers maybe incorporated into a DNA chain or rearrange from a preexisting β anomer [25-29]. For anomerization to occur the sugar ring must open. In saccharides this is contingent on conversion of the hemiacetal to a hemiketal [22]. In nucleic acids, the C1' hydroxyl group is replaced by a purine / pyrimidine base, thus a hemiketal is not possible; these cyclic bases protect the glycosidic nitrogen from



Scheme 1-1: Acid catalyzed anomerization in oligonucleotides

reduction. However, if the glycosidic nitrogen is not involved in the aromatic ring system, a transient iminium bond may form with the glycosidic carbon (Scheme 1-1) [25, 29-31]. This provides the possibility of deoxyribose ring opening and subsequent epimerization. Anomerization in nucleic acids and subsequent biological interactions are poorly understood.

The formamidopyrimidines (FAPYs) are examples of DNA lesions that are of increasing chemical and biological interest [26, 27, 29, 32-36]. FAPYs are linked to their sugar moiety via an amino group (Figure 1-1 A). This allows for the possibility of anomerization by way of the previously discussed mechanism. Although not immediately obvious, thymine glycols also share the potential for anomerization by way

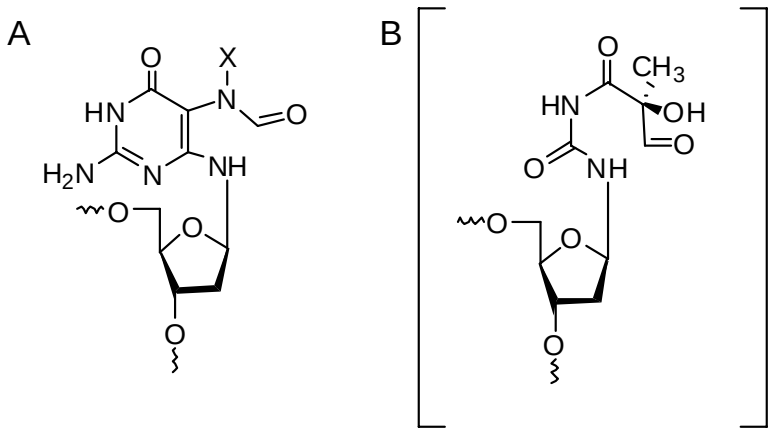
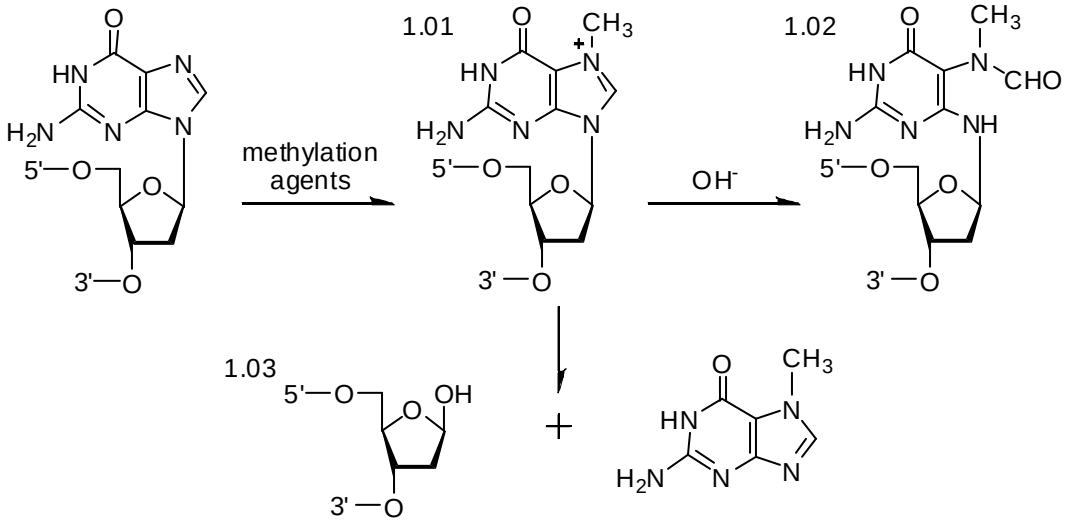


Figure 1-1: A) Guanine formamidopyrimidine [X = CH_3 or AFB_1] B) Proposed carbamoyl-2-hydroxy-2-methyl-3-oxopropanamide intermediate of *cis*-5*R*,6*S*-thymine glycol to *trans*-5*R*,6*R*-thymine glycol epimerization.

of a carbamoyl-2-hydroxy-2-methyl-3-oxopropanamide intermediate (Figure 1-1 B) [37]. The α anomers have been demonstrated to be a block to replication [28, 38, 39], however, little else is known regarding anomer equilibrium in DNA.

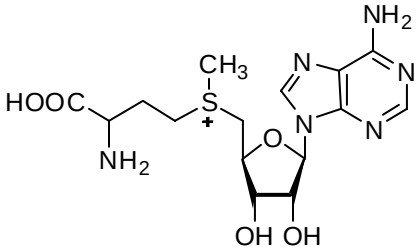
Methyl Formamidopyrimidine

DNA alkylation by methylating agents can generate many products. The major product of guanine methylation is generally N7-methyl-2'-deoxyguanosine (**1.01**)(7-Me-dGuo) [40]. The positively charged 7-Me-dGuo (**1.01**) can be converted to either an apurinic site **(1.03)** or an imidazole ring opened form, 2,6-diamino-4-hydroxy-5-(N-methylformamido)-pyrimidine (Me-dGuo-FAPY) **(1.02)** [41, 42] (Scheme 1-2). The principal source of endogenous methylation is S-adenosyl-L-methionine (SAM) (Figure 1-2). SAM regulates gene expression by methylation of the DNA bases cytosine and adenine DNA [43-45]. Another endogenous methyl source is N-methylnitrosourea; a byproduct of pyrimidine biosynthesis [46, 47]. Methylation of guanine may also occur

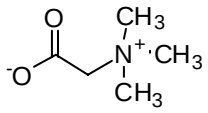


Scheme 1-2: Formation of FAPY-N7-Me-dGuo

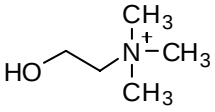
by reaction with betaine, choline, and N-nitroso compounds [48]. In addition, the cationic 7-Me-dGuo **(1.01)** is produced by reaction with exogenous methyl sources.



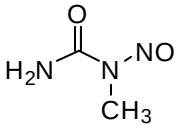
S-adenosyl-L-methionine (SAM)



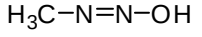
Betaine



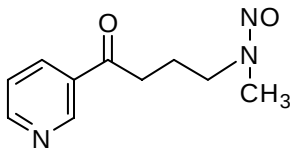
Choline



N-Methylnitrosourea



Methyldiazohydroxide



4-(methylnitrosamine)-1-(3-pyridyl)-1-butanone (NNK)

Figure 1-2: Methylating agents

Examples included 4-(methylnitrosamine)-1-(3-pyridyl)-1-butanone (NNK), the most potent carcinogen in tobacco [49, 50] and methyldiazohydroxide, a product of metabolic activation of N-nitrosamines [51, 52].

Conversion of the cationic 7-Me-dGuo **(1.01)** to an AP site **(1.03)** or Me-dGuo-FAPY **(1.02)** is influenced by multiple factors. Depurination is favored at neutral to acidic pH and alkali conditions catalyze ring opening of the imidazole of 7-Me-dGuo **(1.01)** [53]. Depurination occurs faster in single strand DNA than double strand DNA. In ribonucleic acids (RNA) depurination is slow and the rate of imidazole ring opening is three fold higher than in DNA [54, 55]. Slow depurination is attributed to destabilization of a depurination transition state by the electron withdrawing 2'-hydroxyl group of RNA [56]. Under physiological conditions, depurination is favored over formation of Me-dGuo-FAPY in DNA **(1.02)** [57, 58]. Nevertheless, once formed, Me-dGuo-FAPYs are stable under physiological conditions [59, 60]. In

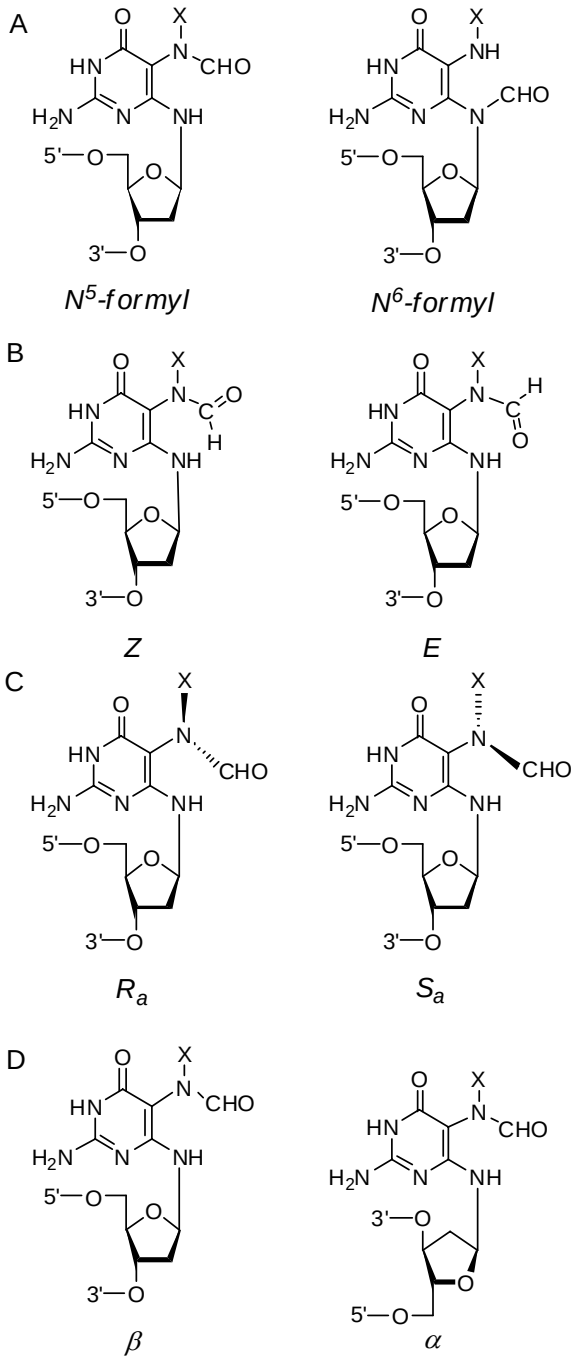
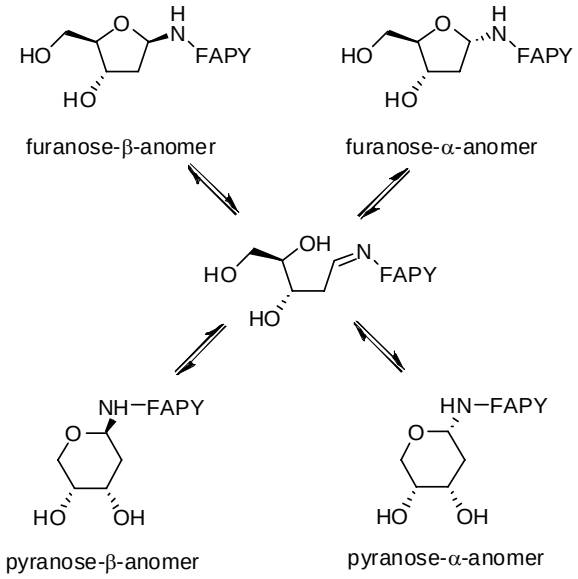


Figure 1-3: FAPY Isomers A) regioisomers B) geometrical isomers C) atropisomers D) anomers

animal models Me-dGuo-FAPY **(1.02)** was found in the liver and bladder epithelial DNA of rats treated with N-methylnitrosourea, N,N-dimethylnitrosamine, and 1,2-dimethylhydrazine [61, 62].

Formamidopyrimidines exist as a complex mixture of isomers. The regioisomers having the formyl group on N^6 are possible but have never been observed [63] (Figure 1-3 A). Rotation of the formyl moiety about the N^5-CHO bond results in geometrical isomers (Figure 1-3 B). It is uncertain if geometrical isomers are chromatographically separable. However, rotation on the millisecond time scale gives rise to chemical exchange cross peaks in NMR spectroscopy [31]. Rotation about the C5-N^5 bond produces the chromatographically separable atropisomers with AFB_1-FAPY nucleosides (Figure 1-3 C) [31]. This slow rotation is attributed to steric interference with the N^6 amine proton [25, 42, 64]. Deoxyribose ring opening of FAPY lesions may

produce a mixture of α and β anomers [29] (Figure 1-4 D; Scheme 1-3). In FAPY nucleosides furanose ring expansion may occur when the ring closure reaction targets the 5'-hydroxyl group producing α and β pyranose FAPYs (Scheme 1-3) [65]. The pyranose form is more stable [25, 66]; however, in DNA the 5'-hydroxyl group is involved in the phosphodiester bond, prohibiting pyranose formation [67]. Production of the chromatographically separable furanose anomers is possible in oligonucleotides [25, 26, 59].



Scheme 1-3: Formamidopyrimidine anomerization and deoxyribose ring expansion

Repair of Me-dGuo-FAPY **(1.02)** is efficient and sequence dependent [57, 68, 69]. Both Me-dGuo-FAPY **(1.02)** and 8-oxodeoxyguanosine (8-oxoG) are repaired by human 8-oxoG DNA glycosylase (hOGG1) and FAPY DNA glycosylase (Fpg) in human and *E. coli* cells respectively [70]. Unlike 8-oxoG, repair of Me-dGuo-FAPY **(1.02)** was not base-pair dependent. However, repair of Me-dGuo-FAPY **(1.02)** randomly incorporated into M13 DNA varied with sequence. It has been suggested that Fpg repair of Me-dGuo-FAPY **(1.02)** is most efficient in dG-rich sequences [71]. Asagoshi et al. [70] reported that Me-dGuo-FAPY **(1.02)** flanked by 5'-dG was excised more efficiently than that flanked by 3'-dG. The influences of 5'- and 3'-flanking dG were observed for both Fpg and hOGG1. Additionally, 8-oxoG with a 3'-dA is not repaired efficiently by Fpg; this is not necessarily true for Me-dGuo-FAPY **(1.02)**.

Repair studies have been conducted on FAPYs of unknown isomeric composition; differential isomeric repair has not been addressed [57, 68-71].

Me-dGuo-FAPY **(1.02)** is toxic in mutagenesis and primer extension assays. Me-dGuo-FAPY **(1.02)** has been described as a "fairly strong but not absolute" [34] block to replication in *Escherichia coli* with little mutagenic activity [32, 41]. The *in vitro* replication properties ascribed to Me-dGuo-FAPY **(1.02)** do not address potential differences in isomeric forms. Recently, a series of primer extension assays using a variety of prokaryotic and eukaryotic DNA polymerases were performed to assess Me-dGuo-FAPY **(1.02)** insertion and extension properties [56]. Me-dGuo-FAPY was determined to be highly miscoding with little primer extension. Specific mutational properties of formyl and C5-N^5 rotamers are unknown at this time. However, in other cases α anomers have been demonstrated to produce a lethal block to replication [28, 38, 39]. The possibility of α anomers producing the toxicity associated with Me-dGuo-FAPY **(1.02)** lesions has been recognized, but subsequently dismissed due to a lack of evidence supporting the presence of the α anomer of Me-dGuo-FAPY in DNA [34].

Aflatoxin B_1 Formamidopyrimidine

Aflatoxins are mycotoxins produced during the growth phase of *Aspergillus flavus* and other related fungi which have the propensity to contaminate improperly stored food [72-76]. Aflatoxins were first characterized in the 1960s after the death of more than 100,000 turkey poults (turkey X disease) following consumption of mold-contaminated peanut meal [77, 78]. Exposure to aflatoxin is particularly high in underdeveloped countries with a humid climate [76, 79, 80]. These mycotoxins are

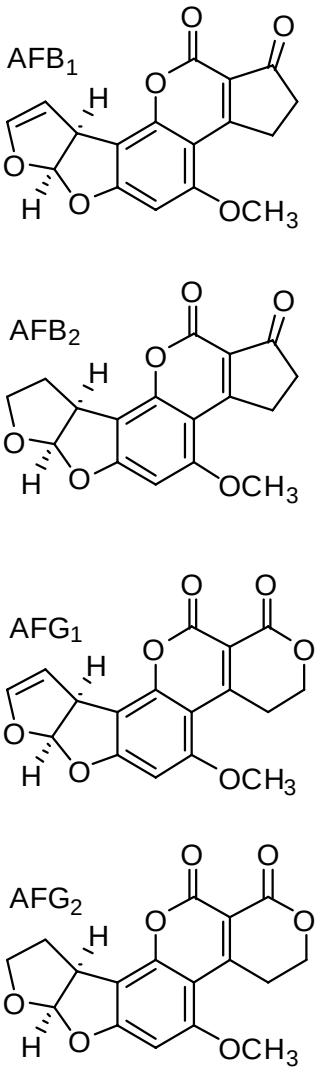


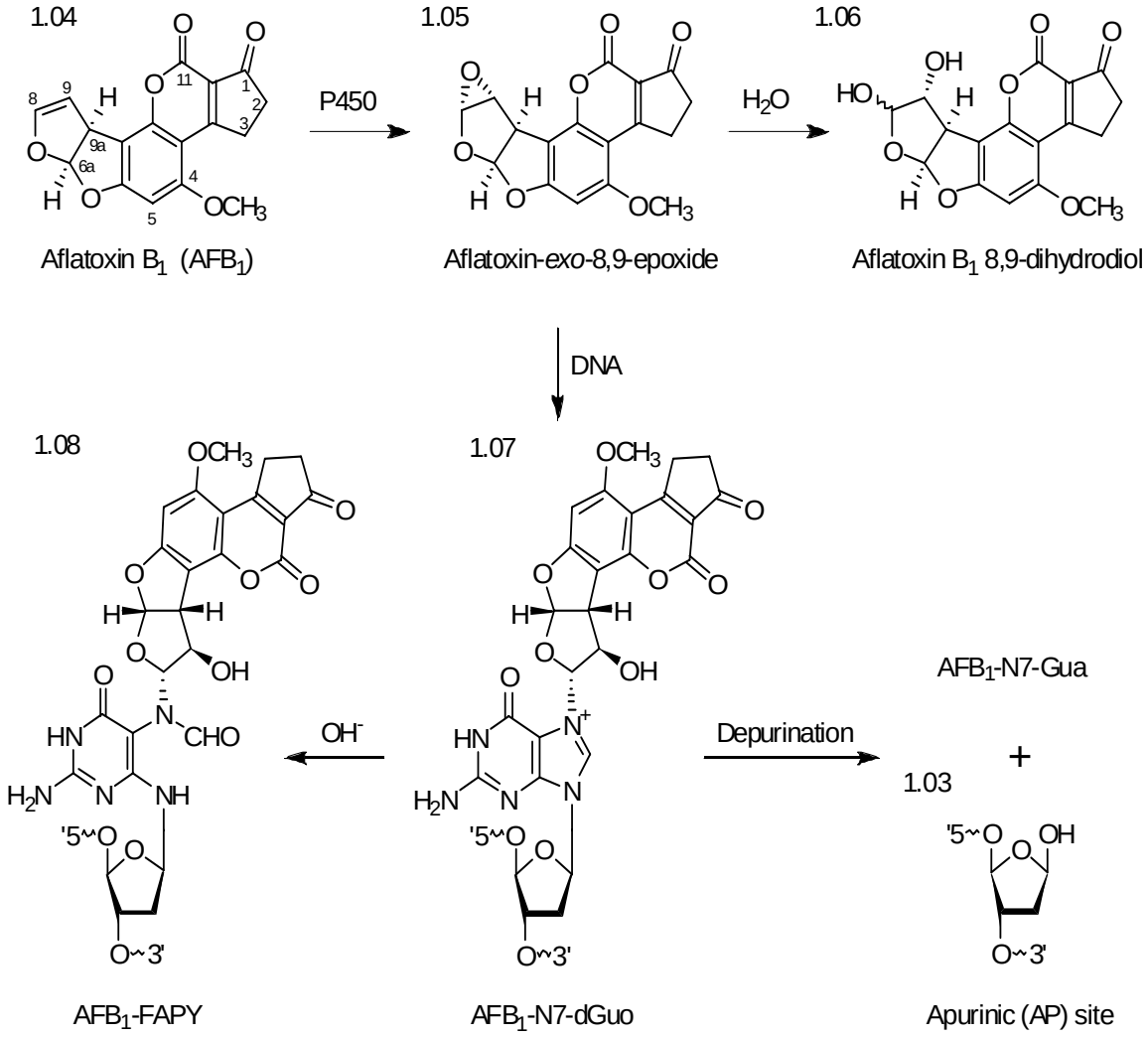
Figure 1-4: Aflatoxins

responsible for significant crop damage, even in developed countries like the United States [81].

Over a dozen different aflatoxin compounds have been described [82]. However, there are four major aflatoxins designated B_1, B_2, G_1, and G_2 (Figure 1-4) [83]. The aflatoxin naming convention is based on blue/green UV fluorescence (i.e. B/G) and relative chromatographic mobility. Aflatoxin B_1 (AFB_1) is the most toxic and most extensively studied constituting the majority aflatoxin literature to date; AFB_1 is the principal aflatoxin produced by most toxigenic strains [80] and it is the most potent natural carcinogen known to man [72, 84].

Aflatoxin B_1 **(1.04)** is mutagenic and cytotoxic in both bacteria and mammals [72, 73]. At one point AFB_1 was thought to be a direct-acting carcinogen [72]. Subsequent investigation has demonstrated that bioactivation is essential in AFB_1-related toxicosis and mutagenesis [85]. AFB_1 is metabolized in humans to its genotoxic form by the liver enzyme cytochrome P450. This reaction is facilitated principally by cytochrome P450 3A4 [86-88] and to a lesser extent P450 1A2 [88]. The AFB_1 *exo*-8,9-epoxide **(1.05)** is the electrophilic species that covalently modifies DNA with high affinity [88-90]. AFB_1 *exo*-8,9-epoxide reacts exclusively or almost exclusively at the N7 position of guanine (Scheme 1-4); limited evidence has been reported for trace reactions at other sites in guanine, adenine, and

cytosine [91-95]. The kinetics of this reaction have been analyzed in detail [92, 96]. The isomeric AFB_1 *endo*-8,9-epoxide [97] does not damage DNA because the stereochemistry of the epoxide does not permit the requisite S_N2 reaction [92, 98, 99]. Nevertheless, rapid hydrolysis of both epoxides [100] produces a dihydrodiol/dialdehyde mixture [101] that is believed to be responsible for AFB_1 cytotoxicity [102, 103]. The epoxidation of AFB_1 is mimicked *in vitro* by reaction with dimethyldioxirane and m-chloroperbenzoic acid [104, 105].



Scheme 1-4: AFB_1 induced DNA damage *in vivo*

Numerous observations have contributed to our current understanding of how AFB_1 forms lesions in DNA. To begin, the epoxide will not react with guanine in Z- or A-form DNA [106]. Because AFB_1-*endo*-8,9-epoxide does not react with DNA, it is not mutagenic [92]. In reaction studies of isomeric oligodeoxyribonucleotides having CpG and GpC sequences, the epoxide reacted with different stoichiometries, suggesting that reaction only occurs when the AFB_1 is intercalated on the 5' side of guanine [98, 107]. The epoxides of AFG_1 react with guanine in DNA less efficiently than the epoxide of AFB_1; this is attributed to better intercalation of AFB_1 than AFG_1 hence better reaction [108, 109]. Reaction with single stranded DNA is less efficient than with duplex [110]; this is consistent with the behavior of other intercalators [111].

Although unstable, AFB_1 *exo*-8,9-epoxide **(1.05)** reacts with DNA with yields up to 98% [96, 99]. The epoxide has been demonstrated to intercalate precovalently into a DNA duplex on the 5' side of guanine [98, 99]. A proton field peripheral to the DNA is postulated to facilitate hydrolysis and conjugation of AFB_1 *exo*-8,9-epoxide [96, 112, 113]. Covalent modification involves a facile S_N2 reaction to yield *trans*-8,9-dihydro-8-(N7-deoxyguanosyl)-9-hydroxyaflatoxin B_1 (AFB_1-N7-dGuo) **(1.07)** [92, 99, 105, 106, 114]. AFB_1-N7-dGuo increases the T_m of the duplex but the positively charged imidazole ring causes the adduct to be labile resulting in depurination leaving apurinic (AP) sites **(1.03)**. Alternatively, the imidazole ring of AFB_1-N7-dGuo **(1.07)** can open to form the more chemically and biologically stable *trans*-8,9-dihydro-8-(2,6-diamino-4-oxo-3,4-dihydropyrimid-5-yl formamido)-9-hydroxyaflatoxin B_1 (AFB_1-FAPY) **(1.08)** [72, 115-117]. (Scheme 1-4)

The influence of varying DNA sequences on AFB_1 reactivity has been studied extensively [118-128]. Benasutti *et al.* [119] conducted a systematic survey where it was determined that AFB_1 reactivity toward guanine was highly dependent on both the 3' and 5' neighboring bases. In the 5' position reactivity favored G > C > A > T and in the 3' position reactivity favored G > T > C > A. Thus, a central guanine flanked by 3' guanine and 5' guanine was the most reactive; a central guanine flanked by a 3' adenine and a 5' thymine was least reactive. The *sup*F gene of the pS189 shuttle vector as observed in the human xeroderma pigmentosum (XP) cell line possesses seven sequences predicted to have high relative reactivities by Benasutti's rules [120]. Levy *et al* concluded that not all sites incurred damage consistent with their predicted levels of reactivity [120]. Furthermore, some mutational hotspots were predicted to have weak reactivity. These observations are consistent with other AFB_1 induced mutation studies [121, 129-131]. This would suggest that, although AFB_1 mutagenesis is affected by sequence, additional factors influence modification and processing [73].

Chronic consumption of AFB_1 contaminated food has been correlated with increased incidence of hepatocellular carcinoma (HCC) [72, 80]. HCC is the fifth most commonly occurring cancer in the world and the third greatest cause of cancer mortality [132]. Liver tumors believed to result from AFB_1 exposure have a common mutational hotspot, a G→T transversion at the 3rd position of codon 249 of the p53 tumor suppressor gene (AG<u>G</u>: targeted G underlined) [133]. This mutation will produce an Arg→Ser alteration. Codon 249 represents one of four Arg residues in the highly conserved DNA binding motif of p53 [134]. This mutation has been found in approximately 50% of HCC tumors in global regions having a high potential for AFB_1 exposure. Over 2000 HCC

samples from around the world have been assayed for this hotspot mutation. In Qidong and Tongan (China), India, Southern Africa, The Gambia, Senegal, and Mozambique, regions where AFB_1 exposure is high, approximately 44% of examined tumors were positive for G→T transversions at the 3rd position of codon 249 of the p53 gene. In contrast, in the US, Japan, Europe, and Australia, regions where AFB_1 exposure is low, less than 1% of examined tumors had this mutation. These data suggest this particular p53 mutation is unique to AFB_1 induced liver tumors [73].

Carcinogenicity is considerably increased in individuals who test positive for hepatitis B virus (HBV), a common occurrence in regions where aflatoxin exposure is high [135-143]. One study suggested that patients who test positive for AFB_1-N7-Gua antigen are three times more likely to develop HCC, patients testing positive for HBV surface antigen are seven times more likely to develop HCC, and when patients test positive for both, they are sixty times more likely to develop HCC [141]. There are several proposed explanations for this apparent synergism. Some have suggested that the HBx, a gene product of HBV, binds and inactivates p53 [144, 145]. HBx may also inhibit p53 induced apoptosis [146]. Others have not observed direct interaction between HBx and p53, but alteration in localization, phosphorylation, or transcription of wild-type p53 during HBx expression has been observed [147-151]. Regardless, a definitive explanation of this synergism has yet to be established.

Transformation of normal cells to malignant cells may also occur by activation of cellular *ras* genes by single base mutations [152-154]. Activation of c-Ki-*ras* genes in rat liver tumors was evidenced by mutations in codon 12 [155, 156]. The first and second positions of codon 12 incur mutations in rats while trout tumors showed mutations in the

second position of codons 12 and 13 [157]. In humans, mutations at the first and second positions of codon 12 of the Ha-*ras* proto-oncogene have been reported [158]. Of likely significance, codons 12 and 13 of c-Ki-*ras* (GCAGGA) bear a striking resemblance to codon 249 of p53 (AGGC) [73].

The lesions AFB_1-N7-dGuo **(1.07)**, AFB_1-FAPY **(1.08)**, and AP site **(1.03)** are all candidate precursors to aflatoxin induced mutation. A common G→T transversion has been reported in all of the following experimental systems: in human liver tumors [140, 159, 160]; in cultured human hepatocytes [161]; in the human HPRT gene [162]; in the human Ha-*ras* proto-oncogene [163]; in human cells transfected with an AFB_1-modified pS189 shuttle vector [21]; in an intra-sanguinous host-mediated assay [164]; in the *lacI* gene of transgenic C57BL/6 mice and F344 rats exposed to AFB_1 [165]; in transgenic C57BL/6N mice [166]; in the *ras* gene of rainbow trout [167]; in the *lacI* gene of SOS-induced *E. coli* containing the *mucAB* mutagenesis enhancing operon [168]. These data do not differentiate between AFB_1-FAPY, AFB_1-N7-dGuo, and AP site induced mutation. It could be argued that aflatoxin induced AP sites are the prime source of G→T transversions.

In the 1980s, techniques were developed to assess the mutational specificity and quantitative mutagenicity of individual lesions resulting from damaged DNA [21]. This process involves the incorporation of a known DNA lesion into the genome of a virus or a plasmid by recombinant DNA techniques. The adduct-containing vector is transfected into bacterial or mammalian cell lines. At this stage, the lesion encounters the replication and repair mechanisms of the host cell. The progeny of the modified vector are then quantitatively and qualitatively analyzed.

This site specific mutagenesis approach was used to study the known DNA lesions resulting from AFB_1 exposure (i.e. AP, AFB_1-FAPY, and AFB_1-N7-dGuo). These experiments were conducted in *E. coli* expressing the SOS induced UmuDC and MucAB error prone polymerases using single stranded bacteriophage M13 genomes [162, 169]. (Note: As of 2005, structural studies of AFB_1 adducts were conducted in double stranded DNA [98, 115, 170-173]) AFB_1-N7-dGuo primarily caused primarily G→T transversions at a rate of 2-6%. A significant fraction (13%) of the mutations occurred 5' to the modified base. This mutational asymmetry has been attributed to the 5' location of the AFB_1 moiety [162]; it has been speculated that in this location the adduct may act as a pseudo base leading to miscoding [73]. The bulky benzo[a]pyrene also has the ability to induce asymmetric mutations [174]. The AFB_1-N7-dGuo derived transversions were highly dependant on MucAB and UmuDC unlike transversions derived from AP sites. Vectors containing apurinic lesions gave rise only to mutations targeted to the damaged site; these lesions are not able to induce non-targeted mutations. The pattern of non-targeted mutations and error-prone polymerase dependent transcription matched the mutational pattern observed in cells treated with aflatoxin disproving the idea that AP sites are responsible for AFB_1 induced mutations [162].

AFB_1-FAPY exists as a pair of chromatographically separable isomers originally designated "major" and "minor" based on their equilibrium populations [175, 176]. The major species was a block to replication while the minor species caused G→T transversions at a rate of 36%. AFB_1-FAPY is detected at near maximal levels in rat DNA weeks after AFB_1 exposure, underscoring its high persistence *in vivo* [117, 177, 178]. On the basis of that fact it was concluded that AFB_1-FAPY may be responsible for

the majority of mutations observed in human HCC as a consequence of its longevity [169, 178-182].

Although AFB_1-FAPY is a substrate for both nucleotide excision repair (NER) and base excision repair (BER), it is preferentially repaired by NER [183]. Many bulky adducts, like AFB_1, are substrates for nucleotide excision repair (NER) [120, 184-188]. AFB_1-FAPY and AFB_1-N7-dGuo are both efficiently repaired in NER-proficient cells [183]. Similarly, in *E. coli*, AFB_1-FAPY and AFB_1-N7-dGuo are excised by uvrABC with equal efficiency [189]. Interestingly, the rate of incision can vary as much as 15-fold depending on sequence with 3'-dA being the most resistant [189]. Little else is known about sequence dependent repair of AFB_1 lesions. Considering that AFB_1-FAPY possesses sequence specific structural properties [115], sequence related differences in repair are not surprising. AFB_1 lesions are also substrates for base excision repair (BER) [190]. The Fpg protein (MutM), a BER glycosylase, repairs oxidative damage to DNA [191]. MutM removes AFB_1-FAPY, but not AFB_1-N7-dGuo *in vitro,* demonstrating a preference between the two lesions [190]. Further investigation revealed that AFB_1-FAPY was a poor substrate for both MutM and also its human functional analog hOGG1; these results were compared to canonical substrates for these proteins, 8-oxoguanine and Me-dGuo-FAPY which are excised more efficiently [70, 183]. Nevertheless, in mammalian cells AFB_1-FAPY persists for days to weeks whereas AFB_1-N7-dGuo is expeditiously removed, suggesting a differential repair mechanism or spontaneous depurination of the cationic adduct [117, 177, 192].

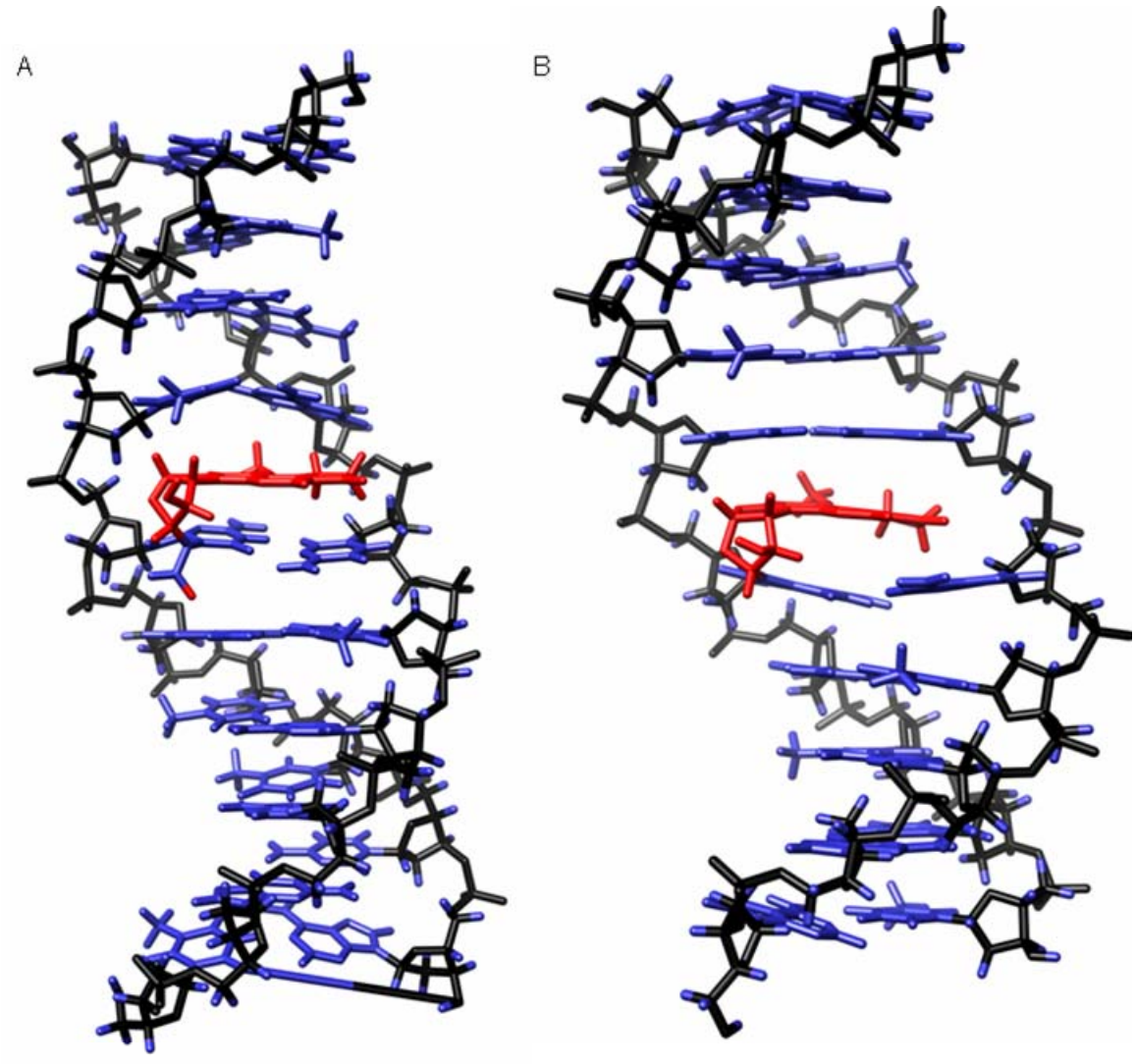


Figure 1-5: Previously published NMR solution structures of AFB_1 modified DNA. A) 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-5' (X = AFB_1-FAPY) B) 5'-$A^1C^2A^3T^4C^5X^6A^7T^8C^9T^{10}$-3'•5'-$A^{11}G^{12}A^{13}T^{14}C^{15}G^{16}A^{17}T^{18}G^{19}T^{20}$-5' (X = AFB_1-N7-dGuo)

The chemistry of the AFB_1-FAPY adduct is poorly understood. Qualitative identification of AFB_1-FAPY isomers has lead to several ideas concerning their identities. NMR spectroscopy of AFB_1-FAPY nucleoside indicates that it exists as four species that can be chromatographically separated into two species before re-equilibration; one HPLC fraction was larger than the other (hence "major" and "minor" respectively), each fraction contained two inseparable species [175, 176]. Initially,

"major" and "minor" were proposed to be N^5 and N^6 formyl regioisomers (Figure 1-3 A) [176]. This was subsequently disproven by incorporation of an isotopically labeled ^{15}N at position N7 of guanine [31, 63]. Upon AFB_1 adduction and base catalyzed rupture of the imidazole ring, ^1H NMR studies revealed that the ^{15}N was located solely at the N^5 position. Heteronuclear ^1H-^{15}N coupling was observed between N^5 and the CHO proton for both "major" and "minor" isomers. No coupling between CHO and N^6 was observed. This places the formyl group on N^5 for both isomers, dispelling the regioisomer isomer theory.

The quest for AFB_1-FAPY isomer clarification led to the use of AFB_1 modified nucleosides. When studying a mixture, nucleosides are simpler systems than oligonucleotides and therefore more tractable to NMR analysis. In AFB_1-FAPY nucleosides, exactly like Me-dGuo-FAPY, possible isomers were determined to be: rotamers of the formyl moiety (*E* and *Z* geometrical isomer), rotamers of the C5-N^5 bond (R_a and S_a atropisomers), and epimers of the anomeric carbon (α and β anomers) [31] (Figure 1-3). In AFB_1-FAPY nucleosides, the R_a atropisomer was favored over S_a by a factor of 7:1; the *Z* geometrical isomer was favored over *E* 3:1 [31].

Successful analysis of biomolecules by NMR necessitates that the analyte comprise a commanding majority of the mixture. In 1998 a detailed solution structure of the duplex DNA 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-5' (X = AFB_1-FAPY) was reported [115, 172] (Figure 1-5). The structure is an R_a atropisomer, *E* geometrical isomer, and β anomer. Logically, it was concluded that the solution structure in dsDNA represents the "major" FAPY isomer.

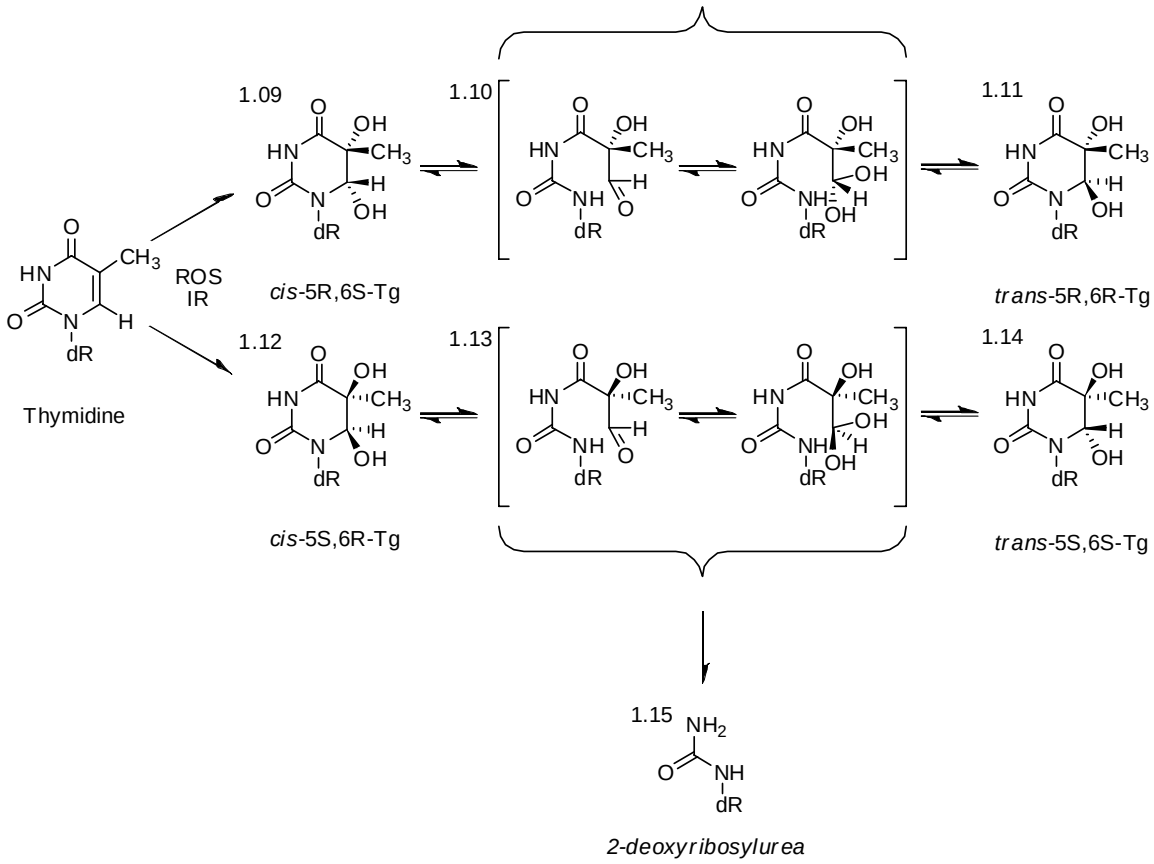
The conclusion that the solution structure FAPY species equates to the "major" form analyzed in mutagenesis experiments is flawed. The solution structures of AFB_1-FAPY and AFB_1-N7-dGuo are similar in many ways (Figure 1-5). The AFB_1 is located on the 5' side of the modified base, AFB_1 spans the helix and is fully intercalated. Ignoring the obvious chemical differences between the two adducts, there are subtle difference in the structures. The AFB_1-N7-dGuo helix is slightly bent at the lesion site and the AFB_1 moiety cannot stack parallel to neighboring bases. In the AFB_1-FAPY structure, the helix is unwound 10 degrees more than in the AFB_1-N7-dGuo structure and there is a hydrogen bond between the CHO oxygen and the 3' adenine N^6 amine. These differences are minor and one might predict a similar biological response to both lesions. However, AFB_1-N7-dGuo produced G→T transversions in *E. coli* and AFB_1-FAPY "major" was a block to replication. This radically different mutational spectrum was difficult to reconcile in light of the NMR solution structures.

One of the primary focuses of this dissertation is the elucidation of structures of the various AFB_1-FAPY species in DNA and the impact of these structures on replication bypass. Was the previously proposed assignment of FAPY "major" and "minor" as atropisomers correct [115, 116, 169]? Reactivity of the epoxide of AFB_1 with guanine is influenced by sequence context; could sequence context have also produced a structural effect? Most importantly, the difference in mutation patterns between "major" and "minor" posed an interesting structural question: What kind of isomers could be responsible for such a different mutational effect?

Thymine glycol

Thymine bases are oxidized by ionizing radiation or reactive oxygen species. In the case of ionizing radiation, the hydroxyl radical is the principal reactive species responsible for DNA damage [12, 13, 193]. The hydroxyl radical is also an *in vivo* product of aerobic metabolism [12, 194]. The DNA damage resulting from ionizing radiation and cellular oxidation is qualitatively similar [195]. Thymine bases are the most susceptible to oxidation [196, 197]. Many products may result, i.e., thymine peroxide, thymine hyperperoxide, and others; however, the major stable product *in vivo* and *in vitro* is thymine glycol (Tg) [197-200]. Oxidative free radical attack on the C5-C6 ethylenic bond produces a pair of *cis* thymine glycol **(1.09,1.12)** enantiomers, each in equilibrium with its *trans* epimer **(1.11,1.14)** (Scheme 1-5) [37, 201-205]. This equilibrium favors the *cis* isomers **(1.09,1.12)** approximately 4:1 [37, 206]. Loss of aromaticity prohibits planer conformation for thymine glycol bases [207]. An axial Tg CH_3 results in a ^{5}E conformation for the 5*R* isomer for example, while an equatorial Tg CH_3 produces a ^{6}E conformation. Therefore, each Tg isomer [5R,6S **(1.09)**; 5R,6R **(1.11)**; 5S,6S **(1.12)**; 5S,6R **(1.14)**] pucker conformation can be generalized to either ^{5}E or ^{6}E.

Thymine glycol lesions are common in cells of all organisms [208]. Thymine glycol and thymidine glycol have been detected in the urine of humans and laboratory animals fed a normal diet [198, 209]. It has been estimated that the average human cell will repair over 400 of these lesions per day [198, 210]. The presence of thymine glycol in bodily fluids of non-irradiated subjects is attributed to enzymatic removal from cellular DNA where it is produced as a consequence of normal oxidative stress [198, 209].



Scheme 1-5: Formation of cis-5*R*,6*S*-thymine glycol and epimer equilibration

Thymine glycol is relatively stable, however, in alkaline environments, the thymine glycol ring may be hydrolytically fractured during epimerization **(1.10, 1.13)** leaving a urea residue **(1.15)** composed of the N1-C2-N3 pyrimidine atoms bound to the deoxyribose (Scheme 1-5) [211, 212].

Thymine glycol is a strong block to DNA replication. Multiple studies have shown that thymine glycol inhibits DNA synthesis *in vivo* in most sequences [211, 213-215]. However, polymerase read-through has been observed in certain contexts [215]. The presence of a pyrimidine 5' to Tg enhances the probability of translesion synthesis more than a 5' purine [215-217]. The 3' base determines the extent of replication block [218]. Site-specific incorporation of thymine glycol into the M13mp19 bacteriophage

and subsequent transfection into *E. coli* produced targeted T→C transitions at a frequency of 0.3%. This was attributed to formation of a Tg•G wobble base pair that was predicted by molecular modeling [219]. Unfortunately, mutagenesis experiments were conducted on a mixture of thymine glycol lesions. More recently, it has been demonstrated that the bypass of Tg lesions by Y-family DNA polymerases is stereospecific, with pol η bypassing the 5*R* **(1.09)** epimers more efficiently [220] and pol κ bypassing the 5*S* epimers **(1.12)** more efficiently [221].

Repair of thymine glycols is highly dependant on isomeric configuration and on the complementary base. Under normal circumstances the lesion is paired with adenine. Therefore, most Tg lesions are located opposite adenines in DNA. However, Tg opposite guanine may occur when a 5-methyl cytosine is oxidized to 5-methycytosine glycol that undergoes facile hydrolytic deamination to Tg [222, 223]. This is significant considering > 70% of cytosines in a CpG context are believed to be methylated in mammalian cells [222]. Studies by Teebor and co-workers have shown the repair of these lesions by DNA N-glycosylases/AP lyases to be dependent on both absolute configuration of the lesion and the opposing base [224]. Using *E. coli* endonuclease III (Nth), *cis*-5*S*,6*R*-Tg **(1.12)** opposite adenine was repaired with greater efficiency than its diastereomer **(1.09)**. Studies with human endonuclease III (hNth), an endonuclease that repairs pyrimidine lesions arising from oxidative damage [225, 226], indicated a 13:1 preference for excising the 5*R* epimers **(1.09, 1.11)** vs. the 5*S* epimers **(1.12, 1.14)**. Additionally, the 5*R*,6*S* **(1.09)** isomer placed opposite adenine was repaired more efficiently than when it is opposite a guanine. This endonuclease was virtually inactive against the 5*S*,6*R* **(1.12)** diastereomer regardless of the opposing base [227, 228]. The human endonuclease-like

protein hNEIL1 glycosylase [229] showed a 1.5:1 preference for excising the 5*R* epimers **(1.09, 1.11)** vs. the 5*S* epimers **(1.12, 1.14)** [228], but seemed to be more efficient when Tg was opposite a guanine [224, 230]. Similar observations have been made for prokaryotic, yeast, and murine glycosylases [230].

The current structural knowledge of thymine glycol in DNA is limited and contradictory in some instances. Predictions from molecular modeling studies suggested that 5*R*,6*S*-Tg **(1.09)** could be extrahelical as a result of loss of planarity of the Tg base [198, 219]. This was postulated to be principally a consequence of steric clash with the Tg CH_3. An equatorial configuration of 5*R*,6*S*-Tg **(1.09)** CH_3 was predicted to be more sterically favorable than axial (Figure 1-6) [231]. The degree of extrahelicity is believed to be sequence dependent with 3' pyrimidines producing the most perturbation [219]. However, 3' purines are predicted to stabilize Tg by a 3'-N7 → Tg-OH^6 hydrogen

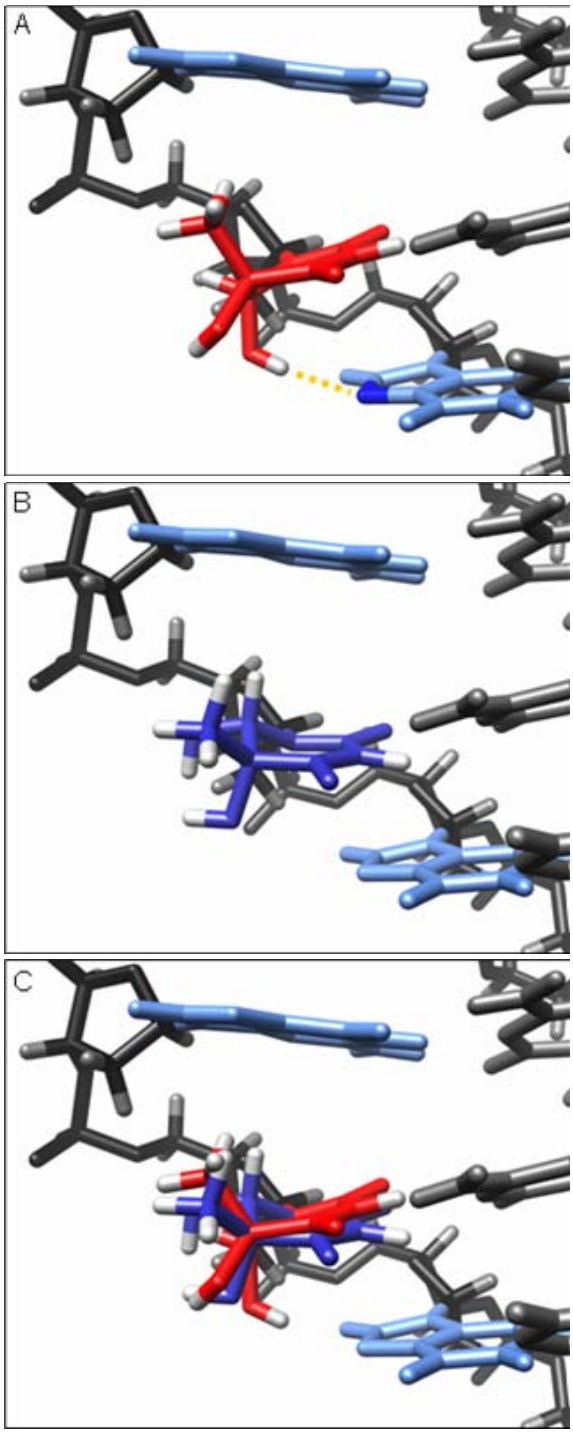


Figure 1-6: Model structures of *cis*-5*R*,6*S*-Tg flanked by guanines. A) Tg CH_3 axial. B) Tg CH_3 equatorial. C) Superposition of pucker conformations.

bond [231] (Figure 1-6 A). NMR data have been obtained in 5'-GTgC-3' and 5'-ATgA-3' sequence contexts where Tg = *cis*-5*R*,6*S* **(1.09)**. The 5'-GTgC-3' structure was disordered at the lesion site and opposing base [232]. The 5'-ATgA-3' structure indicated the Tg residue was approximately half extrahelical [233]. This study did not address intra-strand lesion hydrogen bonding. In addition, these NMR studies did not address the possibility of *cis* and *trans* geometrical isomers or α and β anomers present in the analyte [232, 233]. More recently, a binary primer-template complex, containing a site-specifically Tg-5*R* **(1.09)** modified template, was crystallized with the replicative RB69 DNA polymerase [234]. The resulting structure, representing the situation immediately following incorporation of dATP opposite Tg, revealed the presence of the *cis*-5*R*,6*S* Tg epimer **(1.09)** at the active site. The *cis*-5*R*,6*S* Tg epimer was intrahelical and formed a Watson-Crick base pair with the dA at the primer 3'-terminus [234]. This confirmed modeling studies that predicted *cis*-5*R*,6*S* Tg **(1.09)** would successfully pair with dA [231]. Moreover, in the crystal structure with the RB69 polymerase, the Tg methyl group was in the axial conformation, hindering stacking of the adjacent 5´-template guanine [234]. These structural results provided a possible rationale for earlier observations that extension past the 5*R*-Tg lesion by the Klenow fragment of *E. coli* DNA polymerase I or T4 DNA polymerase was prohibited [214].

Many questions remain concerning thymine glycol processing. First, how can an extrahelical lesion's complementary base modulate its repair? Second, is it possible that the toxic properties ascribed to thymine glycol maybe a consequence of inversion of the anomeric carbon resulting in α anomers? Finally, thymine glycol is reported to exist as a mixture of *cis – trans* epimers in nucleosides, is this equilibrium present in

oligonucleotides? NMR structural studies have been undertaken with a Tg modified oligonucleotide duplex to see whether more modern NMR techniques would reveal more structural details in the region of the lesion than had been observed in earlier investigations.

CHAPTER II

MATERIALS AND METHODS

Biological Hazards

Aflatoxin B_1 and many of its derivatives are potently carcinogenic. Great care should be exercised to avoid personnel exposure. Crystalline material presents an inhalation hazard because the crystals develop electrostatic charge causing aerosols to form. Manipulations should be carried out in a well-ventilated hood with suitable containment procedures. Aflatoxins can be destroyed by oxidation with NaOCl.

Materials

Unmodified oligodeoxynucleotides were purchased from the Midland Certified Reagent Co. (Midland, TX). Samples were purified by reverse phase HPLC chromatography and analyzed by mass spectrometry by Midland technicians. Oligodeoxynucleotide concentrations were measured by UV absorbance at 260 nm. AFB_1 was purchased from Sigma-Aldrich Chemicals (Milwaukee, WI).

Dimethyldioxirane Synthesis

A closed distillation apparatus was flushed with nitrogen. Once the cold finger was chilled, 15 mL of reagent grade H_2O, 16 g $NaHCO_3$, and 15 mL of acetone were added to a stirred reaction vessel. A slight vacuum was applied as nitrogen flowed

through the system (ca. 180 Torr, water aspirator). To this, 30 g of OXONE ® (monopersulfate compound, CAS 70693-62-8) was added via a solid addition funnel and 10 mL of H_2O and 10 mL of acetone via a liquid addition funnel. OXONE was added in approximately 2 g increments as water/acetone dripped into the reaction vessel from the liquid addition funnel. The distillate was collected for approximately 45 min. The yellowish dimethyldioxirane solution was quickly transferred to polypropylene containers containing anhydrous $MgSO_4$. The solution was then refrigerated (-20 °C) and allowed to dry for approximately 24 h. This method produced approximately 5 mL of 0.03 to 0.07 M dimethyldioxirane [235]. The product was assayed for concentration and water content by ^{1}H NMR (solvent = acetone). The concentration was determined by comparing the height of the dimethyldioxirane methyl proton peak (δ 1.65) to that of the ^{13}C satellite peak to the right of the acetone signal [236]. Successful drying of the dimethyldioxirane was established by magnitude of residual water peak (δ 2.8).

Aflatoxin B_1 Adduct Synthesis

Dimethyldioxirane was used to epoxidize AFB_1 [105]. One mg of AFB_1 was dissolved in 0.5 mL of anhydrous CH_2Cl_2 (3.2 mM). Two equivalents of dimethyldioxirane were then added to the AFB_1 solution. An amber glass reaction vial was used to block ambient light and avoid potential photo products. The reaction was allowed to proceed for approximately 15 min. The resulting epoxide was dried by gently blowing nitrogen over the solution, care being taken to avoid formation of airborne crystalline AFB_1 epoxide. The completeness of the reaction and possible presence of dihydrodiol, which is formed by reaction of the epoxide with adventitious moisture, was

established by ^{1}H NMR [105]. Of particular value for this determination are the signals for the H6a protons; unmodified AFB_1 appears at δ 6.9, epoxide δ 6.1, and dihydrodial at δ 6.6 [105].

Typically, oligonucleotides with one guanine were used for the adduction reactions. Efficiency of the reaction was improved by using double stranded DNA. This was achieved in one of three ways depending on sequence; use of a self complementary sequence, use of a complementary strand that did not contain guanine, or blocking intercalation sites by triplex formation [107, 237-239]. Typically, the oligonucleotide was dissolved in 100 μL of 20 mM sodium phosphate buffer (pH 6.5, 100 mM NaCl) and cooled to ~ 5 °C. The appropriate amount of epoxide was dissolved in 100 μL of extra dry CH_2Cl_2 (< 0.003% H_2O). The epoxide was then added in two separate additions. A 5:1 epoxide to oligonucleotide ratio was optimal in most cases. The sample was mixed at 5 °C for 15 min. The organic layer was removed and oxidized with NaOCl.

Table 2-1: Oligonucleotides used in structural analyses

	Sequence	*Size*	*T_m of Unmodified Duplex(100mM NaCl) [°C]*	*Reference*
1	5'-CCT CTT CXA ACT C-3'	13	44.6	[169]
2	5'-CTA TXA TTC A-3'	10	22.2	[115]
3	5'-CTA TXT TTC A-3'	10	23.2	
4	5'-CTT TXA ACC C-3'	10	31.7	
5	5'-CTT CXA ACT C-3'	10	31.4	
6	5'-CTXA-3'	4	<5	[240]
7	5'-GAG GAX GCC CTT-3'	12	66.7	[237, 241]
8	5'-TCA TTX AAT CCT TCC CCC-3'	18	56.6	[242-244]
9	5'-AGA GTC GAC-3'	9	29.3	
10	5'-GTG CGT GTT TGT-3'	12	45.4	[237, 241]

Cationic AFB_1-N7-dGuo was purified from unreacted and complementary DNA by HPLC. Upon isolation, AFB_1-N7-dGuo modified DNA was lyophilized. AFB_1-N7-dGuo samples were placed in 500 μL of sodium carbonate solution (pH 10, 100 mM, 37 °C) to facilitate imidazole ring opening producing AFB_1-FAPY. The conversion from

AFB_1-N7-dGuo to AFB_1-FAPY was monitored by HPLC (λ 260 and 360 nm). AFB_1-FAPY was successfully incorporated into sequences 1-8 of Table 2-1.

AFB_1-FAPY modified oligonucleotides equilibrate to a mixture of α and β anomers. To study individual anomers, HPLC purification was conducted with an ammonium formate mobile phase adjusted to pH 8.0 with NaOH. Lyophilization of ammonium formate caused the pH to drop <6.0. To correct for this event, samples were saturated with an excess of $Na_2HPO_4 \bullet 7H_2O$. Lyophilized samples were quickly desalted prior to NMR by use of a C-18 Sep-Pak (Waters Corporation, Milford MA). Sep-Pak activation was achieved by washing with 5 mL of acetonitrile. The Sep-Pak was then rinsed with 10 mL of reagent grade water. The dissolved sample was then passed over the Sep-Pak five times to ensure complete sample binding to C-18 media. The bound sample was washed with 5 mL water to remove salt. Elution was achieved with 3 mL of a 50/50 water/acetonitrile solution. The eluted sample was collected in a polypropylene receptacle containing buffer for NMR studies (20 mM sodium phosphate buffer, pH 8.7-9.0, 100 mM NaCl, 10 mM NaN_3, and 50 μM Na_2EDTA). Where appropriate, the eluted anomer was combined in a 1:1 molar ratio with its appropriate complement for duplex studies. CGE was used to estimate excess ssDNA; samples containing more than10% excess were repurified by HPLC.

Thymine glycol and urea deoxyriboside samples were prepared in the lab of Dr. Ashis Basu (University of Connecticut, Storrs). The *cis-5R*-thymine glycol phosphoramidites were incorporated into the sequences 5'-GTGCG<u>X</u>GTTTGT-3' and 5'-AGAG<u>X</u>CGAC-3' where X represents the site of modification [245] (Table 2-1,

sequences 9-10). Thymine glycol and AFB_1 modified samples were analyzed by MALDI-TOF and CGE to confirm sample constitution and homogeneity.

Matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF)

Oligodeoxynucleotides were analyzed using a Voyager-DE (PerSeptive Biosystems, Inc., Foster City, CA) MALDI-TOF spectrometer. Samples were suspended in a matrix consisting of 0.5 M 3-hydroxypicolinic acid in 1:1 CH_3CN:H_2O and spotted onto sample plates. Mass spectra were recorded in the reflector mode using laser attenuation of 3030. The accelerating voltage was 20 kV, with a grid voltage of 85.00%, guide wire voltage of 0.050%, and a delay of 100 ns. The spectra were an averaged from 256 scans.

Capillary Gel Electrophoresis (CGE)

CGE data was obtained using a Beckman Coulter P/ACE MDQ Capillary Electrophoresis System. An injection voltage of 10 kV for 2 s was used; in some cases it was advantageous to pressurize the loop (20 psi) to avoid air in the capillary. A 30 cm capillary, packed with ssDNA 100-R gel (Beckman Coulter Cat #477489) was used at a separation voltage of 9.0 kV for 55 min with a Tris-borate-urea buffer (tris 44%, boric acid 56%, 7 M urea). Approximately 0.1 OD of oligonucleotide and 1 μL of 1% Orange G reference marker (Beckman Coulter Cat #S906682) were placed in the injection capsule (total volume 40 μL). UV absorbance was recorded at 254 nm or 340 nm.

Thermodynamic Measurements

UV-absorption thermal denaturation experiments were conducted on a CARY 4E UV-Vis spectrophotometer (Varian, Inc. Palo Alto, CA). Data was collected and analyzed using the Cary WinUV Thermal application (v. 2.0). Samples were suspended in 10 mM phosphate buffer, containing 500 mM NaCl, and 10 mM Na_2EDTA (pH 7.0). For thymine glycol sample analysis, four different concentrations of 1 mL of 0.1, 0.5, 0.8, and 1.0 μM solutions in 1 cm capped cuvettes were placed in the spectrometer's multicell temperature regulation block along with a blank. For analysis of AFB_1-FAPY samples, one concentration of approximately 1.0 μM was used. During the serial analysis of the samples, temperature was increased at a rate of 0.3 ºC /min from 5 to 80 ºC and absorbance was measured at 260 nm. Before each 5 - 80 ºC sweep, temperature was allowed to equilibrate for 5 min at 5 ºC or 80 ºC. Absorbance versus temperature profiles were analyzed by first derivative and hyperchromicity to determine T_m. Hybridization versus temperature plots, α curves, and van't Hoff plots were produced where appropriate [246].

Thermodynamic Data Analysis

Thermodynamic parameters were generated by analysis of absorbance versus temperature profiles within in the Meltwin (v. 3.5 McDowell) and Thermal (v. 2.0 Cary) applications. A Marquardt-Levenburg algorithm was used for individual curve fitting. Thermodynamic values (ΔH^o and ΔS^o) were obtained by two methods [247-249]. The analysis of thermodynamic properties by two different methods is necessary to ascertain the presence of a two state transition of duplex denaturation. Previous studies have

defined a one step transition as agreement within 10% of enthalpy values calculated by two methods [250]. In method 1, parameters were extracted from individual melting curves. In method 2, linear regression analysis of van't Hoff plots (SigmaPlot v. 9.0) produced straight lines where the slope equals $R/\Delta H^o$ and the intercept equals $\Delta S^o/\Delta H^o$ by the following equation ($R = 1.9872$ cal K^{-1} mol^{-1}; C = molarity):

$$\frac{1}{T_m} = \frac{R}{\Delta H^o} \ln\left(\frac{C}{4}\right) + \frac{\Delta S^o}{\Delta H^o} \qquad (2\text{-}1)$$

Gibbs free energy (ΔG^o) was then determined at 25 ºC and 37 ºC. Error values were represented as standard deviation. Theory and analysis of the thermodynamic properties of DNA have been discussed in detail elsewhere [246, 251].

Electronic Circular Dichroism (ECD)

CD spectra were collected on a Jasco J-720 spectropolarimeter (Japan Spectroscopic Co., LTD, Tokyo, Japan) at 25 °C. Duplex and single stranded samples (4.0×10^{-6} M) were analyzed in a 1 cm optical cell. Data from 210 to 400 nm were signal averaged over 10 scans with a sensitivity of 50 mdeg and resolution of 1 nm (scan speed = 50 nm/min).

Quantum Mechanical Calculations

Partial charges for the non-standard residues (i.e. thymine glycol and AFB_1-FAPY) were generated using the program GAUSSIAN 03 [252]. Geometry optimization

and frequency calculations were performed using the B3LYP density functional (DFT) method with the 6-31G* basis set. Additionally, thymine glycol energies were recorded at the 6-31G**, 6-31+G*, and 6-311++G** levels of theory. Potential points were written out with a density of 6 points per unit area in the electrostatic potential (ESP) fit. Gaussian ESP output was converted to restrained electrostatic potential (RESP) charges using the program ANTECHAMBER [253].

In the case of thymine glycol, four sets of GAUSSIAN calculations were performed with each basis set. Coordinates were chosen for *cis* 6S-Tg (CH_3 axial), *cis* 6S-Tg (CH_3 equatorial), *trans* 6R-Tg (CH_3 axial), and *trans* 6R-Tg (CH_3 equatorial). Quantum calculations were expedited by replacing the sugar moiety with a CH_3 group at the N1 position of the modified base.

The AFB_1-FAPY residue is not directly tractable to quantum calculations due to its comparatively large size (35 heavy atoms). Therefore, RESP charges were calculated for AFB_1 and the modified base + sugar separately. For this reason, unique library input files for AMBER were prepared for the AFB_1 adduct and modified base – deoxyribose.

The resultant optimized structures and charges were used as parameters for rMD calculations. Frequency analysis was used to test for convergence. Diagonalization of the Hessian matrix will produce positive eigenvalues when a structure is at a minimum [254-257].

Nuclear Magnetic Resonance Spectroscopy (NMR)

DNA samples were prepared at concentrations of approximately 1.5 mM and placed in 5 mm NMR tubes for analyses. Samples prepared for the observation of non-

exchangeable protons were dissolved in 500 μL of 20 mM sodium phosphate buffer, 100 mM NaCl, 10 mM NaN_3, and 50 μM Na_2EDTA at pH 7.0. Aflatoxin samples were studied at a pH of 8.7 to 9.0. Samples were lyophilized from 99.9% D_2O three times and suspended in 500 μL of 99.996 % D_2O. For the observation of exchangeable protons, samples were suspended in 500 μL of 9:1 H_2O:D_2O. ^{1}H and ^{31}P spectra were referenced to internal 3-(trimethylsilyl)propionic-2,2,3,3-d_4 acid, sodium salt (3-TMSP) and external 85% H_3PO_4 (capillary in H_2O) respectively.

Bruker Avance spectrometers operating at ^{1}H frequencies of 800, 600, and 500 MHz were used for data acquisition. Carrier frequencies were set at 4.7 ppm for ^{1}H, 80 ppm for ^{13}C, and 0 ppm for ^{31}P. Spectrometers were equipped with 5 mm, triple resonance, CP-TCI-Z cryoprobes and quadruple resonance QXI-XYZ probes. Thymine glycol modified 5'-GTGCGXGTTTGT-3'•5'-ACAAACACGCAC-3' duplexes were analyzed at 30 ± 0.5 °C. Single stranded and AFB_1-FAPY modified samples were analyzed at 7 ± 0.5 ºC. The observation of exchangeable protons of all duplex samples was conducted at 7 ± 0.5 ºC. Data was processed with XWINNMR (v 3.5 patch level 6, Bruker Inc., Karlsruhe, Germany), TOPSPIN (v 2.0.b.6, Bruker, Karlsruhe, Germany), and/or NMRPipe [258].

Nuclear Overhauser Effect Spectroscopy (NOESY)

Phase-sensitive ^{1}H-^{1}H NOESY [259, 260] experiments used in the assignment of exchangeable protons were conducted at ^{1}H frequencies of 500, 600, and 800 MHz. Water signal suppression was achieved using the 3919 watergate pulse; the binomial null was set at 11 ppm [261]. A relaxation delay of 1.0 s and sweep width of 20 ppm was used for these experiments. NOESY spectra used for the assignment of nonexchangeable

protons were collected at ^1H frequencies of 600 and 800 MHz with a sweep width of 10 ppm. Water suppression for samples dissolved in D_2O was achieved with a square presaturation pulse. Distance restraints were derived from four successively collected NOESY experiments at mixing times of 250, 200, 150, and 80 ms. The assignment of ssDNA samples required mixing times of 250 or 300 ms. A relaxation delay of 2.0 s was used for nonexchangeable NOESY experiments. States-TPPI quadrature detection was used for all NOESY experiments. Thymine glycol related NOESY experiments were recorded with 2048 real data points in d2 and 1024 real data points in d1. The indirect dimension was zero filled to obtain an overall matrix size of 2048 X 2048 real points. AFB_1 and ssDNA related NOESY experiments were recorded with 2048 real data points in d2 and 512 real data points in d1. In this case d1 was zero filled and linear predicted to obtain an overall matrix size of 2048 x 2048 real points. A skewed sinebell-squared apodization function with a 90º phase shift was applied in both dimensions.

Correlated Spectroscopy (COSY)

COSY experiments were conducted in double quantum filtered (DQF) [262] and magnitude modes at ^1H frequencies of 500, 600, and 800 MHz. Spectra were recorded with a sweep width of 10 ppm and a relaxation delay of 2.0 s. Water suppression was achieved with a presaturation pulse. States-TPPI quadrature detection was used. Experiments were recorded with 2048 real data points in d2 and 512 real data points in d1. The indirect dimension was zero filled and linear predicted to obtain an overall matrix size of 2048 x 2048 real points. A skewed sinebell-squared apodization function with a 180° phase shift was applied in both dimensions. Coupling constants were fit using constrained multiplet evaluation (ACME) [263] of a phase sensitive COSY

recorded at ^{1}H of 800 MHz. In this case, spectra were obtained with 2K x 1K (real data points) and a relaxation delay of 12 s; no residual water presaturation was applied. States quadrature detection was used. The indirect dimension was zero filled to obtain an overall matrix size of 2048 x 2048 real points; no linear prediction was used. Remaining acquisition and processing parameters were identical to those previously stated.

Heteronuclear ^{1}H–^{31}P COSY [264] experiments were performed at a ^{1}H frequency of 600 MHz on Bruker Avance spectrometer with a QXI-XYZ probe. Spectra were recorded with a sweep width of 10 ppm in d2 and 5 ppm in d1. Water suppression was achieved with a Dante pulse sequence [265]. States quadrature detection and a relaxation delay of 2.0 s were used. Experiments were recorded with 1024 real data points in d2 and 256 real data points in d1. Zero filling produced an overall matrix size of 2048 x 512 real points. A skewed sinebell-squared apodization function with a 90º phase shift was applied in both dimensions. Pulses were optimized for $^{3}J_{PH}$ coupling constants (12 Hz).

Total Correlated Spectroscopy (TOCSY)

Two dimensional ^{1}H-^{1}H TOCSY [266] data was collected at a ^{1}H frequency of 500 MHz for single stranded DNA samples. Spectra were recorded with a sweep width of 10.5 ppm and a relaxation delay of 1.5 s. Water suppression was achieved with a square wave presaturation pulse. States-TPPI quadrature detection was used with a TOCSY mixing times of 80 and 120 ms. Trim pulse duration was set to 2.5 ms. Experiments were recorded with 2048 real data points in d2 and 512 real data points in d1. The indirect dimension was zero filled to obtain an overall matrix size of 2048 x 1024 real points. A skewed sinebell-squared apodization function with a 90º phase shift was applied in both dimensions.

Heteronuclear Single Quantum Correlation Spectroscopy (HSQC)

Multiplicity edited ^{1}H-^{13}C heteronuclear single quantum correlation spectroscopy (HSQC) [267] experiments were used to ascertain ^{13}C resonance assignments. Spectra were recorded at ^{1}H frequencies of 500 and 600 MHz for single stranded DNA samples and 800 MHz for duplex samples. Spectra were recorded with a ^{1}H sweep width of 10.5 ppm and a ^{13}C sweep width of 180 ppm. A relaxation delay of 21.8 s was used. Echo-Antiecho quadrature detection was used. An 80 μs pulse was used for garp proton decoupling. Pulses were optimized for $^{1}J_{CH}$ coupling constants (140 Hz). Experiments were recorded with 1024 real data points in d2 and 256 real data points in d1. The indirect dimension was zero filled and linear predicted to obtain an overall matrix size of 1024 x 1024 real points. A skewed sinebell-squared apodization function with a 90º phase shift was applied in both dimensions.

Heteronuclear Multiple Bond Correlation Spectroscopy (HMBC)

Heteronuclear ^{1}H-^{13}C multiple bond correlation spectroscopy (HMBC) experiments were used to determine heteronuclear zero and double quantum coherence correlation. Spectra were recorded at ^{1}H frequency of 500 MHz for single stranded DNA samples. Spectra were recorded with a ^{1}H sweep width of 10.5 ppm and a ^{13}C sweep width of 180 ppm. A relaxation delay of 1.5 s was used. Pulses were optimized for $^{1}J_{CH}$ coupling constants (9 Hz). Experiments were recorded with 1024 real data points in d2 and 512 real data points in d1. The indirect dimension was zero filled and linear predicted to obtain an overall matrix size of 1024 x 1024 real points. A skewed sinebell-squared apodization function with a 180º phase shift was applied in both dimensions.

T_1 Measurement (Inversion Recovery)

T_1 inversion recovery experiments [268] were collected at a ^1H frequency of 800 MHz and processed by Fourier transform [269]. Samples were dissolved in D_2O buffer. Recovery times of 0.01, 0.05, 0.1, 0.25, 0.5, 1, 2, 3, 4, 5, 7, 10 s were used for the inversion recovery profile for a total of 12 points for T_1 fitting. Spectra were recorded with a ^1H sweep width of 10.5 ppm in d1. Data was collected in a pseudo 2D array of 16K real data points in d2 and 12 real data points in d1. No acquisition mode was selected.

T_2 Measurement (Carr-Purcell-Meiboom-Gill)

T_2 experiments were collected using the Carr-Purcell-Meiboom-Gill (CPMG) method [268] at a ^1H frequency of 800 MHz. Samples were dissolved in D_2O buffer. The CPMG spin echo time increment profile consisted of values 4, 6, 8, 10, 12, 14, 16, 18, and 20 s for a total of 9 points for T_2 fitting. Spectra were recorded with a ^1H sweep width of 10.5 ppm in d1. Data was collected in a pseudo 2D array of 16K real data points in d2 and 9 real data points in d1. No acquisition mode was selected.

Starting Structures

Starting structures were produced using the MKDNA program written by Jarrod Smith (Department of Biochemistry, Vanderbilt University) using the nucleic acid builder (NAB) suite [270]. Input parameters consisted of a mixture of values for inclination, twist, rise, and x offset. These values were selected to sample both A and B form DNA. 32 starting structures were generated for thymine glycol. Sixteen of the

coordinates had the Tg methyl in the axial orientation while the other sixteen had the equatorial orientation. Global parameters for the core sixteen structures are listed in table 2-2. NAB generated structures were then modified in xLEaP and energy minimized for 1000 steps (100 steepest descent) in SANDER while holding base heavy atoms static [253]. This was necessary to alleviate NAB artifacts, e.g. unnatural backbone bond lengths. Five structures were generated for use in AFB_1-FAPY rMD calculations. In this instance, an extra A:T basepair was placed 5' to the modified base and subsequently removed to provide space for the AFB_1 residue. Insertion of the FAPY residue was accomplished using InsightII (Accelyris, San Diego, CA) and later energy minimized in SANDER.

Table 2-2: Input parameters for NAB

Structure number	*X offset*	*inclination*	*twist*	*Rise*
Tg 1	6.96	22.0	32.7	2.56
Tg 2	2.25	22.0	32.7	2.56
Tg 3	6.96	22.0	36.0	2.56
Tg 4	2.25	22.0	36.0	2.56
Tg 5	6.96	22.0	32.7	3.38
Tg 6	6.96	-5.30	31.7	3.12
Tg 7	6.96	2.90	31.7	3.12
Tg 8	6.96	-5.30	31.7	3.52
Tg 9	6.96	2.90	31.7	3.52
Tg 10	2.25	-5.30	31.7	3.12
Tg 11	2.25	2.90	31.7	3.12
Tg 12	2.25	-5.30	31.7	3.52
Tg 13	2.25	2.90	31.7	3.52
Tg 14	2.25	19.0	33.6	2.30
Tg 15	6.90	19.0	33.6	2.30
Tg 16	2.25	6.00	36.0	3.40
AFB 1	2.25	-1.20	35.9	2.56
AFB 2	2.25	3.02	35.9	2.56
AFB 3	6.96	3.02	35.9	2.56
AFB 4	2.25	3.02	33.6	2.56
AFB 5	2.25	6.00	36.0	3.40

Distance Restraints

NOE intensities were determined from integration of peak volumes by a Gaussian fit function or box summation using the program SPARKY (v. 3.11) [271] for four different mixing times. Intensities measured from interconverting thymine glycol samples were scaled by (100/80) for intensities between variable and constant resonances, and by (100/80) squared for thymine glycol intra-residue signals. Experimentally determined intensities were combined with intensities generated from complete relaxation of a B form starting structure of DNA to produce a hybrid intensity matrix [272, 273]. The program MARDIGRAS [274-276], using the RANDMARDI function, was used to iteratively refine the hybrid intensity matrix and to optimize the agreement between the experimental and calculated intensities. Calculations at mixing times of 250, 200, 150, 80 ms were each run at isotropic correlation times of 2, 3, 4, and 5 ns with a canonical B form starting structure, producing 16 sets of distances. Each set was a product of 50 RANDMARDI iterations. A refined structure resulting from initial rounds of simulated annealing was used as a starting structure for the second round of MARDIGRAS calculations. This produced improved distance sets used to subsequently refine structures that fit more closely to experimental data. Results were indexed to short- and long-range distances common to both A and B form DNA as a means to assess the accuracy of resultant distance sets. Data sets that best matched canonical A-form and B-form distances were averaged and used in subsequent restrained molecular dynamics calculations.

The distance restraint sets were divided into five categories indicative of the confidence level of the experimental data. Class 1, 2, and 3 distances were calculated

from completely resolved, no spin diffusion; slightly resolved, no spin diffusion; and medially resolved; possible spin diffusion, respectively. Class 1, 2, and 3 peaks were at least 0.4 ppm away from diagonal or water resonances. Class 4 and 5 distances were calculated from cross peaks that were highly overlapped and spin diffused.

Additional empirical distance restraints were used to define Watson-Crick hydrogen bonding and counter ion backbone proximity. Generic base pair distances were assigned as follows: For G•C base pairs; G H1 → C N3 = 1.84 – 2.04 Å; G H22 → C O2 = 1.75 – 1.95 Å; G N1 → C N3 = 2.85 – 3.05 Å; G O6 → C H42 = 1.80 – 2.00 Å; G O6 → C N4 = 2.81 – 3.01 Å; for A•T base pairs; A N1 → T H3 = 1.71 – 1.91; A N1 → T N3 = 2.72 – 2.92; A H61 → T O4 = 1.84 – 2.04 Å. During explicit solvent calculations loose distance restraints (3.0 Å - 8.0 Å lower and upper bounds) were used to constrain sodium counter ions near two phosphate groups of the helix so that their position was localized to the backbone phosphates.

Torsion Angle Restraints

Deoxyribose pseudorotation was determined by fitting ^{3}J ^{1}H coupling constants for deoxyribose protons by either amplitude constrained multiplet evaluation (ACME) of COSY data [263] or by direct measurement of coupling constants form DQF-COSY experiments. Electronegitivity of substituent (EOS) Karplus curves were generated and converted to phase angle space assuming a maximum pucker amplitude (Φ) of 44 [277, 278]. $J_{1'2'}$, $J_{1'2''}$, and $J_{1'3'}$ were fit to the curve to determine phase angle ranges (ρ) for deoxyribose rings. The sugar pseudorotation and amplitude ranges were converted to upper and lower bound restraints for the five dihedral angles ν_0 to ν_4. Deoxyribose

pucker conformations were confirmed to be N-type or S-type by an approximate measurement of the mole fraction of sugar pucker in the S configuration (X_S) determined from the sum of $J_{1'2'}$ and $J_{1'2''}$ [278]. Residues that had less than 50% X_S or indicated potential for C3' endo configuration from EOS Karplus curves were constrained such that they were allowed to explore both N and S conformations during rMD calculations ($\rho = 0 – 210$). Residues that has X_S greater than 50% were restrained such that $\rho = 125 – 210$.

Backbone torsion angles were restrained with experimental data where available. Remaining angles were limited based on known A form / B form ranges. The ε dihedral angles (C4'-C3'-O3'-P) were determined by using coupling constants determined from constant time NOESY or selective $J_{H3'(i)\text{-}P(i\text{-}1)}$ COSY experiments [279, 280] and fit by a Karplus relationship [281]. Remaining backbone torsion angles were limited in range based on previously determined backbone relationships [282].

Restrained Molecular Dynamics Calculations

Restrained molecular dynamics were conducted using the AMBER 9 suite [283]. Initial calculations involved a simulated annealing strategy for multiple starting trajectories. Convergent structures were analyzed for the presence of conformational families. Representative structures were then placed in truncated octahedron water box for further equilibrium refinement.

Simulated Annealing

Multiple starting structures were generated using the program NAB (nucleic acid builder) [270]. An RMSD of 4.3 Å was calculated between all 32 starting trajectories

used for Tg rMD calculations. The starting structures used in AFB_1-FAPY rMD calculations had a RMSD of 3.4 Å. Coordinate and topology files were generated with xLEaP [283] using ff99 parameters [284]. The restraint energy function included terms that defined distance and dihedral restraints as square-well potentials [285]. The generalized Born solvent model was used with a salt concentration of 0.2 mM [286, 287]. 20 ps of simulated annealing molecular dynamics were carried out with a non-bonded interaction cutoff at 15 Å. Structure coordinates were recorded every 1 ps. During the calculations, the system was heated to 600 K and returned to 0 K; temperature was maintained using the Berendsen thermostat algorithm [288]. In brief, the simulated annealing protocol utilized a starting temperature of 600 K. From 5 to 18 ps the temperature was decreased to 100 K. The temperature was decreased to 0 K during the final 2 ps of the simulations. A time constant of 0.4 ps was used for heat bath coupling during the first 0.5 ps of the simulation. This was changed to 4 ps from 0.5 to 18 ps. From 18 to 19 ps, the coupling was increased to1 ps. Coupling was further increased during the final 19 to 20 ps of the simulation with a constant of 0.05 ps. During the first 3 ns of the simulation the relative weights of the NMR restraint energy terms was ramped up from 0.1 to 1; this was maintained for the remaining 17 ns of the simulation. Force constants were set as follows: Watson-Crick hydrogen bonding = 32 kcal mol^{-1} $Å^{-1}$; Torsion angles (backbone and sugar) = 5 kcal mol^{-1} $Å^{-1}$; Class 1 and 2 distance restraints = 20 kcal mol^{-1} $Å^{-1}$; Class 3 distance restraints = 18 kcal mol^{-1} $Å^{-1}$; Class 3 distance restraints = 18 kcal mol^{-1} $Å^{-1}$; Class 4 distance restraints = 16 kcal mol^{-1} $Å^{-1}$; Class 5 distance restraints = 10 kcal mol^{-1} $Å^{-1}$.

Explicit Solvent

A representative structure from simulated annealing calculations was charge neutralized by the addition of sodium counter ions and placed in a truncated octahedral TIP3P water box with periodic boundaries at a distance of 8 Å from the solute. The solvated system was then equilibrated using standard protocols. Briefly, after the initial minimization, the solvated system was minimized for 1000 steps (500 steepest descent / 500 conjugate gradient) at constant volume with the solute held fixed by positional restraints, thus allowing the solvent and counter ions to equilibrate. Next the system was minimized for 2500 steps with no positional restraints at a constant volume allowing the solute to equilibrate to its solvent. The system was subsequently heated to 300 K over 100 ps; NMR restraints and empirical restraints were slowly built up during the heating period. Production calculations for rMD were run for a period of 10 ns. Throughout the equilibration and production periods, the temperature was held constant using the Langevin thermostat [289, 290] with a collision frequency of 1 ps^{-1} and electrostatic interactions were treated with the particle mesh Ewald (PME) method [291]. A 15 Å cutoff for non-bonded interactions was used and bond lengths involving hydrogen were held fixed using the SHAKE algorithm [292]. Assessment of the accuracy of calculated MD structures was achieved by complete relaxation matrix analysis (CORMA). RMSD values of heavy atoms between each final structure were compared using SUPPOSE. Hydrogen bonding occupancies, ring puckers, average structures, and other dynamic properties were extracted from molecular dynamics trajectories with PTRAJ. Both SUPPOSE and PTRAJ are distributed with AMBER [283].

Helicoidal Analysis

Helicoidal analysis was performed using CURVES [293] on a average of ten structures taken from the final 100 ps of rMD calculations in explicit solvent of duplex samples. Backbone torsion angles for an average single strand tetramer sample were measured manually using Insight II (Accelyris, San Diego, CA) from graphical display.

Complete Relaxation Matrix Analysis (CORMA)

Theoretical NOE intensities were calculated from rMD structures by complete relaxation matrix analysis (CORMA) [272]. Calculations were performed on an ensemble of 10 structures of equal probability (0.1 %). Theoretical NOE intensities were supplemented with NOE intensities taken from cross peak volumes from spectra with a mixing time of 250 ms. An isotropic correlation (τ_c) of 3 ns was used in all analyses. Input intensities were normalized will all proton intensities, not just highly spin diffused geminal and vicinal protons. Random noise was added to individual intensities to simulate spectral baseline error. The noise value was obtained by taking half the intensity of the weakest input peak. The cutoff level for intensity precision was set to 0.001. Methyl free rotation was simulated using a 3-site jump model

A sixth root residual ($R_x^{1/6}$) was calculated for each structural ensemble to measure how accurately structures fit experimental NOESY data. The sixth root residual is calculated using the following relationship [294]:

$$R_x^{1/6} = \sum \left| I_o^{1/6} - I_c^{1/6} \right| \Big/ \sum I_o^{1/6} \qquad (2\text{-}2)$$

The experimentally observed NOE intensity (I_o) is compared to a theoretically calculated NOE intensity (I_c) generated from refined input structures. The sixth root residual reflects the accuracy between calculated and observed NOE intensities.

Crystallography

Primer sequences and conditions used for crystallographic structure determination are similar to those previously described [242-244]. Two binary complexes were prepared. The first was comprised of Dpo4 and 5'-TCA TTX AAT CCT TCC CCC-3' (X = FAPY-AFB_1) annealed with a 13mer complement 5'-GGG GGA AGG ATTC-3'. The second was comprised of Dpo4 and 5'-TCA TTX AAT CCT TCC CCC-3' (X = FAPY-AFB_1) annealed with a 12mer complement 5'-GGG GGA AGG ATT-3'. Polymerase and DNA complexes were combined in a 1:1.2 molar ratio and incubated in 20mM Tris-HCl buffer, pH 7.5 for 20 minutes at room temperature. Drops contained a final protein concentration of 5 mg/mL. Well solutions contained 9% polyethylene glycol 3350, 20 mM Tris-HCl, pH 7.5, and 100 mM $Ca(CH_3CO_2)_2$.

CHAPTER III

METHYL FORMAMIDOPYRIMIDINE LESION EQUILIBRIA

Introduction

This chapter addresses the complex equilibria of Me-dGuo-FAPY in DNA. The base-catalyzed imidazole ring opening of Me-dGuo-FAPY produces a complex mixture of Me-dGuo-FAPY isomers. This mixture may include rotamers of the formamide (CHO *E* and *Z* geometrical isomers), rotamers of the C5-N^5 bond (R_a and S_a atropisomers), and/or C1' epimers (α and β anomers) (Figure 1-3) [25, 26, 31, 67, 295-297]. The equilibrium state of this complex mixture is poorly understood. The potential of different conformational and configurational isomers of specific DNA lesions to produce differing biological effects is a concept of increasing importance [31, 224, 227]. The elucidation of rotamer equilibria in oligonucleotides by NMR requires clear, unambiguous resonance assignments. However, HPLC analysis is tractable to the study of chromatographically separable acid-catalyzed anomerization products [25, 31]. This was used to provide insight into the roles of anomers in Me-dGuo-FAPY equilibria. Because α anomers have been demonstrated to produce a block to replication, it seems possible this isomer may contribute significantly to the cytotoxicity of Me-dGuo-FAPY [28, 38, 39]. The idea of anomers contributing to the toxicity of Me-dGuo-FAPY has been recognized, but subsequently dismissed due to a lack of evidence supporting the presence of α–Me-dGuo-FAPY in oligonucleotides [34].

Results

Synthesis of all Me-dGuo-FAPY-modified samples was conducted by Dr. Plamen Christov (Department of Chemistry, Vanderbilt University). Me-dGuo-FAPY was site specifically placed in 5'-CTATXATTCA-3'•3'-GATACTAAGT-5' and 5'-AXC-3' where X represents the site of modification. Samples were purified by HPLC. Subsequent HPLC and NMR analysis of both the 10mer and trimer indicated multiple isomeric forms of Me-dGuo-FAPY were present.

High Performance Liquid Chromatography (HPLC)

Reverse-phase HPLC analyses of the modified 10mer indicated at least five Me-dGuo-FAPY isomers were present (Figure 3-1). The possibility of additional isomers cannot be excluded due to the broadness of the HPLC peaks. The 10mer duplex was purified by HPLC, under basic conditions (mobile phase = HCO_2NH_4 @ pH 8.0) into six fractions labeled A-F. Peak A is the complementary strand and is not relevant to these discussions; thus B through F are single strand oligonucleotides containing the Me-dGuo-FAPY lesion. Peak B re-equilibrated independent of pH to a mixture of peaks B, C, & D.

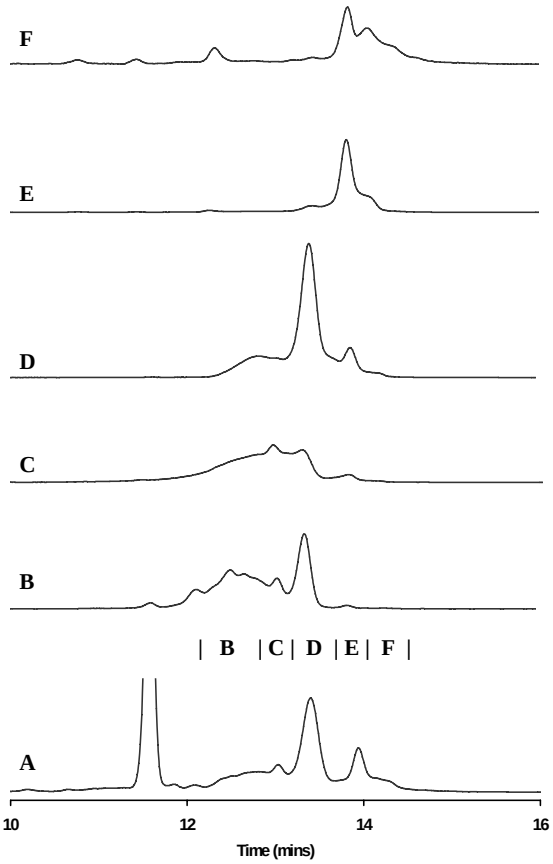


Figure 3-1: HPLC analysis of Me-dGuo-FAPY modified 5'-CTATXATTCA-3'•3'-GATACTAAGT-5' was indicative of multiple isomeric species in dsDNA and ssDNA. Aliquots B through F were held at pH 8.0 and reanalyzed.

Peak C re-equilibrated independent of pH to predominantly B and C. Peak D re-equilibrated to a mixture of peaks B, C and D; E was observed, but had not reached equilibrium. Peak E equilibrated with F but not with peaks B, C, D, or E. Peak F re-equilibrated with peak E. Upon lowering the pH to 6.0, samples B-F readily equilibrated within 2 hrs. Presence of the complementary strand did not measurably affect equilibria between isomers.

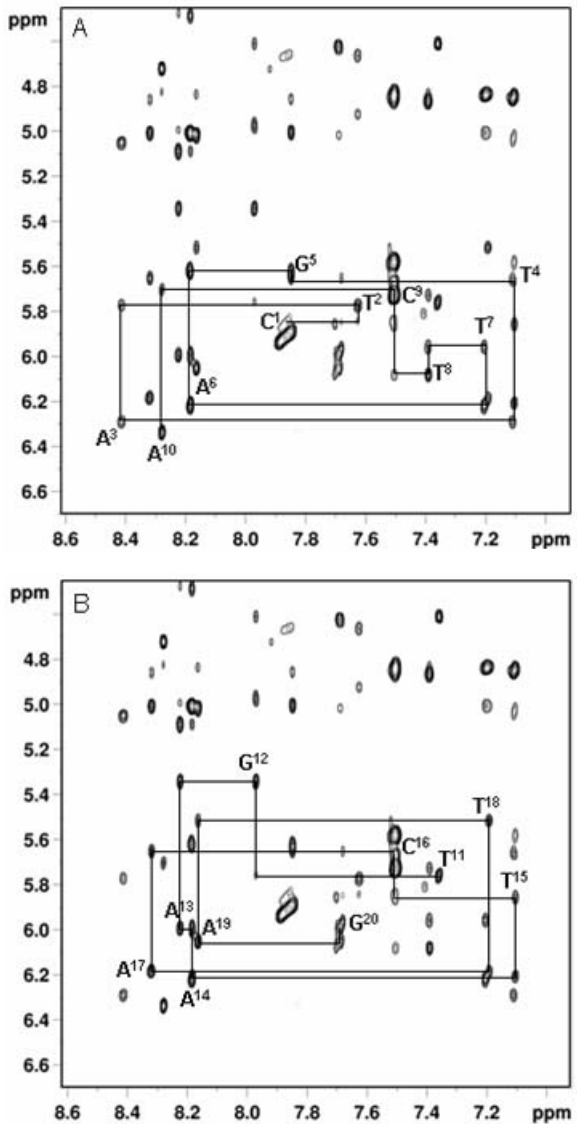


Figure 3-2: Sequential assignments of NOESY H8/H6 to H1' protons for the unmodified 5'-$C^1T^2A^3T^4G^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-5' duplex. Intra residue aromatic to H1' cross peaks are labeled. Assignments for the primary strand are in panel A while the complementary strand assignments are in panel B

<u>Nuclear Magnetic Resonance Spectroscopy (NMR) of Modified 10mer Duplex</u>

The spectroscopic partial assignment of non-exchangeable NMR resonances was possible by comparison of the Me-dGuo-FAPY modified 10mer duplex to an unmodified duplex (Figure 3-2). It was determined that there was a break in intrastrand NOE connectivity in both the modified and complementary strands of the modified 10mer (Figures 3-3). In the modified strand, the break occurred between cross peaks T^4 H6→X^5 H1' and A^6 H1'→A^6 H8. In the complementary strand, connectivity was broken between the T^{15} H1'→T^{15} H6

cross peak and the A^{17} H1'→A^{17} H8 cross peak. Multiple resonances were observed for both T^4 H6→X^5 H1' and A^6 H1'→A^6 H8 cross peaks of the modified strand. In the complementary strand, multiple resonances were observed between T^{15} H6→T^{15} H1' and A^{13} H1'→A^{13} H8 (Figure 3-3 D). Chemical exchange cross peaks associated with rotameric transitions have been observed in nucleoside AFB_1-FAPY [31]. The existence of chemical exchange cross peaks could not be conclusively identified in the congested 10mer NOESY spectra without assignment of X^6 CHO and X^6 CH_3 resonances.

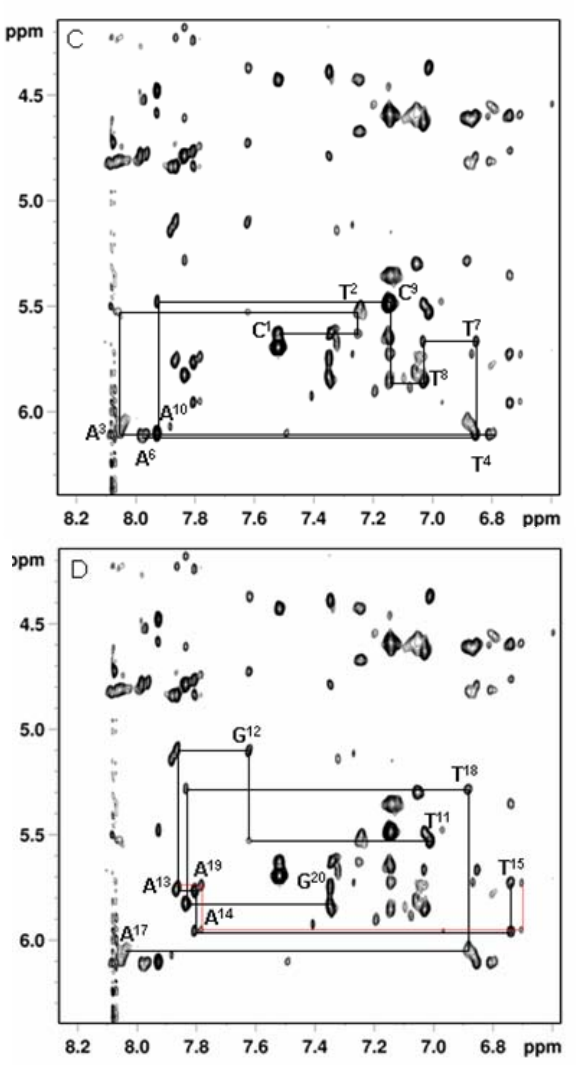


Figure 3-3: Sequential assignments of NOESY H8/H6 to H1' protons for the Me-dGuo-FAPY modified 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-5' duplex. Intra residue aromatic to H1' cross peaks are labeled. Assignments for the primary strand are in panel C while the complementary strand assignments are in panel D. Connectivity was broken in both strands. Peak splitting was apparent in residues A^{13} through T^{15} (red line, panel D).

Nuclear Magnetic Resonance Spectroscopy (NMR) of Modified Trimer

During the synthesis of Me-dGuo-FAPY phosphoramidites, there was concern that Me-dGuo-FAPY pyranoside was a reaction byproduct (Scheme 1-3). A combination of ^{1}H-^{31}P COSY, ^{1}H-^{1}H TOCSY, and ^{1}H-^{13}C multiplicity edited HSQC experiments were used to address this concern. To reduce the complexity of the NMR spectra, the 5'-$A^1X^2C^3$-3' trimer sample was analyzed. The disorder of the trimer

prohibited the evolution of inter-residue NOEs. The Me-dGuo-FAPY CHO and CH_3 resonances were tentatively assigned based on ^1H intensity, chemical shift, and process of elimination (Figure 3-4). In both Me-dGuo-FAPY pyranose and furanose samples a methyl singlet was identified at 1.71 ppm. Both samples had CHO proton resonances at 8.30 and 7.46 ppm.

Sugar proton resonance assignments were based on connectivity and chemical shift obtained from TOCSY spectra. Multiple conformations were observed as evidenced by multiple spin systems of the same residue. For example, 4 sets of peaks associated with X^2 of the furanose trimer were recorded (Appendix A). Deoxyribose ^{13}C resonance assignments were correlated from proton assignments by HSQC. Digital resolution of ^{13}C resonances was calculated to be approximately 0.5 ppm.

Inter-residue scalar couplings, $^3J_{P\text{-}H3'}$ and $^3J_{P\text{-}H5''}$, provided inter-residue connectivity for some, but not all spin systems. In the pyranose sample, H3' (i) to H4' (i+1) connectivity was observed; X^2 C4' was confirmed to be a methine carbon from multiplicity edited HSQC data (Figure 3-5). The pyranose X^2 C4' chemical shift was comparable to literature values of other carbohydrate pyranose C4'

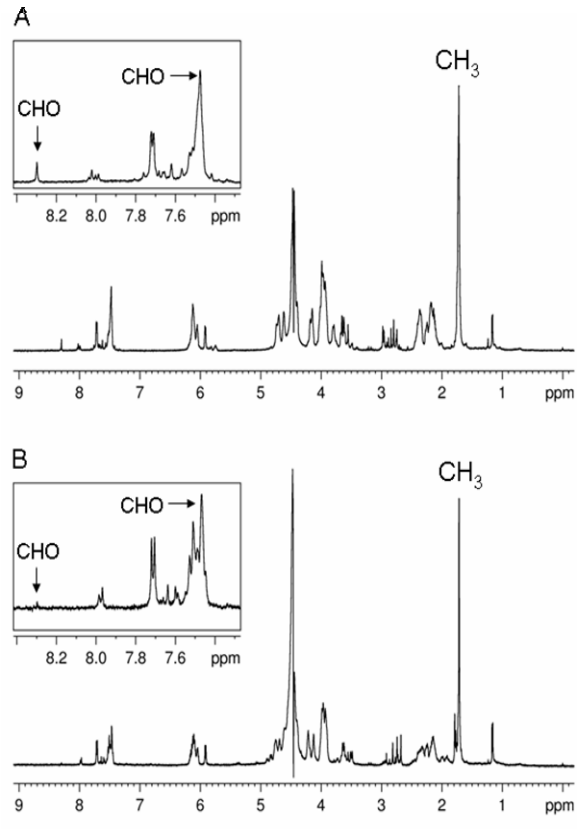


Figure 3-4: ^1H-NMR analysis of the Me-dGuo-FAPY modified trimer 5'- A^1X^2C -3' pyranose (panel A) and furanose (panel B) indicated multiple CHO resonances (insert) but a single methyl resonance was identified.

resonances [298, 299]. A furanose trimer sample exhibited canonical H3' (i) to H5'' (i+1) ribose connectivity (Figure 3-6). HSQC data confirmed the furanose X^2 C5' to be a methylene carbon.

Discussion

NMR Analysis of Me-dGuo-FAPY Trimer

The study of Me-dGuo-FAPY in oligonucleotides necessitated the site specific incorporation of a Me-dGuo-FAPY phosphoramidite. The Me-dGuo-FAPY in DNA existed as a complex mixture of isomers in equilibrium, with the exception of one isomer that could be separated from the mix and remained chromatographically pure [56]. It was proposed that the non-equilibrating species was a Me-dGuo-FAPY pyranoside. During deprotection of the 5'-hydroxy group of Me-dGuo-FAPY phosphoramidite, the deoxyribose can undergo ring expansion to a pyranose under acidic conditions (Scheme 1-3). Shortening the deprotection time reduced the formation of the chromatographically pure isomer. This result supported the

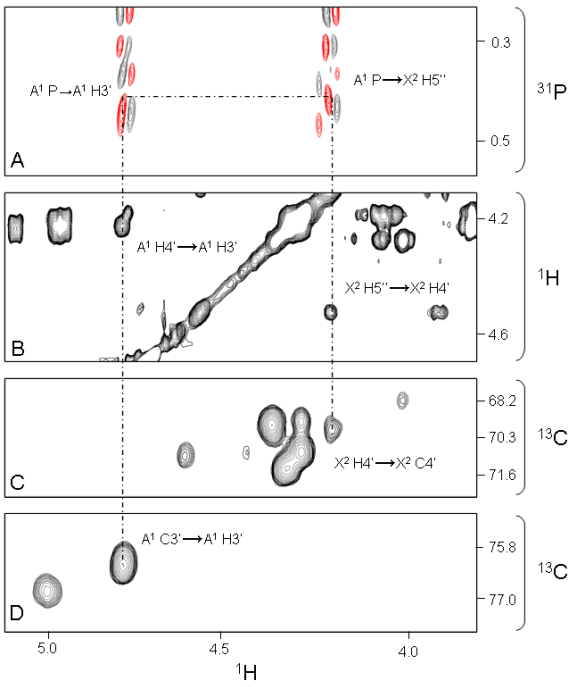


Figure 3-5: NMR analysis of the pyranose Me-dGuo-FAPY in 5'- $A^1X^2C^3$-3'. Tile plots of ^{31}P-^{1}H COSY (A), ^{1}H-^{1}H TOCSY (B), and ^{1}H-^{13}C HSQC (C & D) data indicate that connectivity can be traced from H3' of adenosine to H4' of the pyranose of Me-dGuo-FAPY. The multiplicity edited HSQC confirms the pyranose C4' is a methine carbon (negative phase peaks = red, positive peaks = black).

pyranoside hypothesis [56].

The existence of the Me-dGuo-FAPY pyranoside was confirmed by NMR analysis. A combination of ^{1}H-^{31}P COSY, ^{1}H-^{1}H TOCSY, and ^{1}H-^{13}C multiplicity-edited HSQC experiments were used to address the question of pyranose-furanose identity. To simplify NMR analysis, the 5'-AXC-3' trimer was synthesized and purified into pyranose and furanose mixtures. Identification of the A^{1} H3' resonance allowed for assignment of the X^{2} H5" or X^{2} H4' resonance (furanose/pyranose respectively) through ^{1}H-^{31}P scalar coupling. In the pyranose sample, H3' (i) to H4' (i+1) connectivity was observed; X^{2} C4' was confirmed to be a methine carbon from multiplicity edited HSQC data (Figure 3-5). The furanose trimer sample exhibited typical H3'(i) to H5" (i+1) deoxyribose connectivity (Figure 3-6). HSQC data confirmed the furanose X^{2}:C5' to be a methylene carbon.

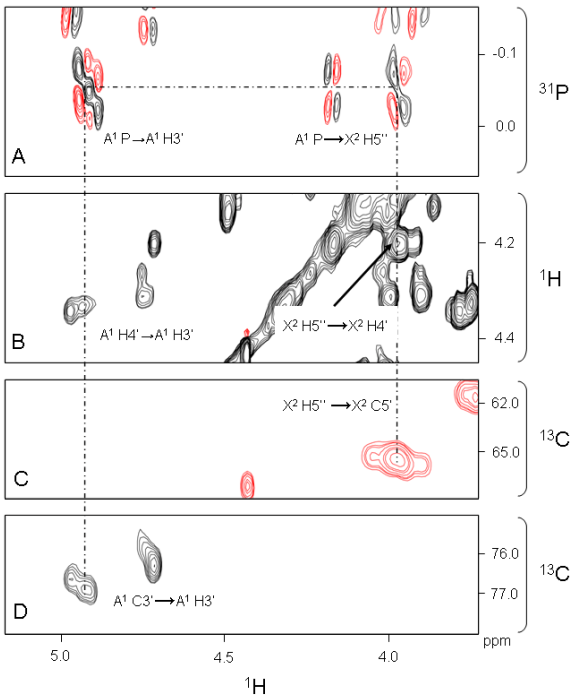


Figure 3-6: NMR analysis of the furanose Me-dGuo-FAPY in 5'- A^{1}X^{2}C^{3}-3'. Tile plots of ^{31}P-^{1}H COSY (A), ^{1}H-^{1}H TOCSY (B), and ^{1}H-^{13}C HSQC (C & D) data indicate that connectivity can be traced from H3' of adenosine to H5'' of the furanose of Me-dGuo-FAPY. The multiplicity edited HSQC confirms the furanose C5" is a methylene carbon (negative phase peaks = red, positive peaks = black).

Chemical shift analysis further supported the pyranose–furanose assignments. Surprisingly, the chemical shifts of the C5' of pyranose and furanose samples were not significantly different (66.1 vs. 65.2 ppm, respectively) (Appendix A; Table A-1 & A-2). However, there was a pronounced difference between the pyranose and furanose C4'

chemical shifts (68.2 vs. 82.0 ppm). Shielding of the C4' pyranose resonance is consistent with published carbohydrate chemical shifts [298, 299].

Assignments were made for some of the Me-dGuo-FAPY CHO and CH_3 resonances in the trimer. NMR analysis of AFB_1-FAPY nucleosides indicate that *E* and *Z* formyl rotamer resonances were observed for each C5-N^7 (R_a / S_a) atropisomer [31]. Therefore, based on the AFB_1-FAPY nucleoside results, it was anticipated that there should be four identifiable Me-dGuo-FAPY CHO resonances; two were assigned (8.30 & 7.46 ppm). In addition, it was anticipated that there should be two Me-dGuo-FAPY CH_3 resonances; one was assigned (1.71 ppm). This result does not exclude the existence of other potential rotamers. It is possible that remaining resonances were overlapped in the spectrum or were below NMR detectable limits.

NMR Analysis of Me-dGuo-FAPY 10mer

In addition to the 5'-AXC-3' trimer, Me-dGuo-FAPY was placed in the 5'-CTATXATTCA-3'•3'-GATACTAAGT-5' duplex for NMR studies. Partial assignment of non-exchangeable NMR resonances was possible by comparison of the modified duplex to an unmodified sample (Figures 3-2 & 3-3). There was a break in intra-strand connectivity in both the modified and complementary strands of the modified duplex (Figures 3-3). In the modified strand, the break occurred between T^4 H6→X^5 H1' and A^6 H1'→A^6 H8 cross peaks. In the complementary strand, connectivity was broken between the T^{15} H1'→T^{15} H6 cross peak and A^{17} H1'→A^{17} H8 cross peak. Loss of connectivity in the modified strand was attributed to conformational flexibility.

Although disrupted connectivity in the complementary strand was anticipated, the degree of disruption was surprising. Biological effects of Me-dGuo-FAPY have been compared to those of thymine glycol [34, 300]. Thymine glycol does not induce a break in connectivity in its complementary strand (Chapter V) [233]. Perturbation at C^{16}, due to increased conformational flexibility, may result from disrupted base pairing. Nevertheless, it is striking that the extent of disturbance in the complementary strand results in a separate NOE connectivity from T^{15} H6→T^{15} H1' and A^{13} H1'→A^{13} H8 (Figure 3-3, panel D). It is not immediately obvious how the Me-dGuo-FAPY lesion can induce dramatic disruption in the complementary strand. It is possible that the disruption is exaggerated by the instability of this sequence. Thermal melting values of Me-dGuo-FAPY modified 5'-CTATXATTCA-3'•5'-TGAATCATAG-5' were below 5 °C [56]; NMR experiments were conducted at 7 °C.

Indications of multiple isomeric species in the sample were observed in the form of multiple resonances near the site of modification. Multiple resonances were observed at both T^4 H6→X^5 H1' and A^6 H1'→A^6 H8 cross peaks of the modified strand. Chemical exchange cross peaks associated with rotameric transitions have been reported in nucleoside FAPYs [31]. The existence of such interactions could not be confirmed in duplex samples without unambiguous assignment of X^6 CHO and X^6 CH_3 proton resonances.

HPLC Analysis of Me-dGuo-FAPY 10mer

Although NMR analysis of the 10mer duplex was disappointing with respect to isomer elucidation, HPLC provided observations concerning α and β anomer populations. HPLC analysis was indicative of at least five furanose isomers in this

sequence (Figure 3-1). Although several of the collected peaks equilibrated after HPLC separation independent of pH, equilibration of peaks E and D was slow at basic pH. The interconversion of anomers is an acid catalyzed event [25, 29, 31, 296, 301]. Rotamer interconversion is independent of pH. Therefore, it was concluded that peaks E and D are anomers. Peak B was a rotamer of D; peak F was a rotamer of E. Further isomeric characterization was not possible.

There are reported differences in Me-dGuo-FAPY rotamer populations in free base vs. poly(dGC) [42]. Dr. Plamen Christov's (Department of Chemistry, Vanderbilt University) initial Me-dGuo-FAPY studies were in the 5'-CTTXTT-3' hexamer [56]. HPLC chromatograms of Me-dGuo-FAPY in 5'-CTATXATTCA-3' vs. 5'- CTTXTT -3' are significantly different with respect to the relative areas of the major components. In the 5'-CTTXTT-3' context, the relative abundance of the four principal isomers were: peak 1, 20%; peak 2, 45%; peak 3, 5%; and peak 4, 30% [56]. Precise HPLC peak integration of 5'-CTATXATTCA-3' was precluded by peak broadening. A cursory comparison of the two sequence chromatograms is indicative of incomparable elution profiles. Oligonucleotide length could cause the HPLC elution profile to vary between samples. However, relative isomer abundance is expected to remain similar. This was not observed, suggesting a sequence dependent effect on relative isomer population.

In AFB_1-FAPY modified oligonucleotides the R_a atropisomers and Z geometrical isomers are stabilized by a CHO hydrogen bond with the N^6 amino group of the 3' adenine (Chapter VI) [115]. AFB_1-FAPY and Me-dGuo-FAPY were studied in the same 10mer sequence; therefore, Me-dGuo-FAPY may exhibit a similar hydrogen bonding motif. More studies are necessary to test this hypothesis. Nevertheless, this could

explain sequence-dependent equilibria positions; a hydrogen bond with the 3' amino group cannot form in 5'- CTTXTT -3'.

Biological Significance

The equilibrium position in site-specific mutagenesis studies is unknown; thus, the contributions of specific isomers to the mutational spectrum could not be determined. Me-dGuo-FAPY was concluded to be highly miscoding at the lesion site, but prohibited read-through in primer extension assays with various polymerases [56]. The α anomers are an important component of the Me-dGuo-FAPY mixtures and are blocks to replication [28, 38, 39]. This work demonstrates that both α and β anomers are present in 5'-CTATXATTCA-3'. It seems possible that the reported cytotoxicity of these lesions is not exclusively a product of the FAPY component. It is possible that the α anomer of the FAPY is responsible for cytotoxicity ascribed to FAPY. Ide and coworkers dismissed this possibility [34]. There were two reasons for their conclusion. First, they did not think that anomers were possible with Me-dGuo-FAPY in their studies based on their understanding of anomer formation. The present work contradicts their conclusion; anomeric isomers are possible, at least in 5'-CTATXATTCA-3'. Unfortunately, Ide et al. did not study Me-dGuo-FAPY in the 5'-TXA-3' context. Secondly, Ide et al. reports that αdA blocks replication by Pol I Klenow fragment at the insertion step [38, 39]. Klenow fragment may not necessarily reflect replication properties of other polymerases. Me-dGuo-FAPY blocked replication in the extension step. Ide argues this difference in bypass does not support the idea of anomers [34]. Although logical, perhaps a direct

comparison of αdA and α-Me-dGuo-FAPY mutagenicity is not appropriate given the obvious differences in lesion structure.

Structural studies of oligonucleotides containing the Me-dGuo-FAPY lesion should be feasible by NMR using a thermally stable DNA sequence. Although it is unknown if the formamide is stabilized by a hydrogen bond to 3' adenine, it would be reasonable to initially study Me-dGuo-FAPY in this context to maintain as simple a system as possible. In addition, purification of the modified oligonucleotides and subsequent NMR analysis should be conducted at basic pH to slow anomeric conversion. All things considered, this may provide a stable platform for structural studies of Me-dGuo-FAPY and possibly for unsubstituted FAPY lesions as well. Modifications to the sequence would then reveal context specific properties directing preferred isomer conformations.

CHAPTER IV

AFB_1-FAPY LESION EQUILIBRIA

Introduction

Site-specific mutagenesis studies of AFB_1-FAPY and AFB_1-N7-dGuo were conducted in *E. coli* using single stranded viral genomes [162, 169]. Structural studies of both adducts have been conducted in double stranded DNA [98, 115, 170-173]. AFB_1-N7-dGuo caused primarily G→T transversions at a rate of 2-6% [162]. AFB_1-FAPY was reported to exist as a pair of chromatographically separable isomers designated "major" and "minor" based on their equilibrium populations [169]. The "major" isomer was a block to replication while the "minor" isomer caused G→T transversions at a rate of 36%. Qualitative identification of these isomers has lead to varying ideas concerning their identities. It was postulated that structural studies had been conducted on the "major" FAPY isomer [73, 116]. This chapter addresses the characterization of AFB_1-FAPY "major" and "minor" isomers [73] as α and β anomers respectively. Equilibrium of α and β AFB_1-FAPY isomers is dependant on DNA environments, i.e. the α-AFB_1-FAPY isomer is preferred 2:1 in ssDNA while β-AFB_1-FAPY is the only anomer normally seen in dsDNA [31]. The R_a atropisomer is exclusively observed in both ssDNA and dsDNA. Equilibrium of geometrical isomers of the formamide is sequence dependent, the *E* isomer being preferred in a 5'-TXA-3' sequence. The *Z* isomer is preferred in AFB_1-FAPY nucleosides 3:1 over *E* [31]. It follows that structural studies of

duplex DNAs [98, 115, 170-173] were conducted on β-AFB_1-FAPY while mutagenesis experiments were conducted on a mixture of α and β AFB_1-FAPY anomers [162, 169]. The α anomer was a block to replication while the β was highly mutagenic in *E. coli*.

Results

AFB_1-FAPY Oligonucleotide Analysis

AFB_1 epoxide was adducted to unmodified DNA as described in Chapter II. Separation of AFB_1-N7-dGuo modified DNA from unreacted and scaffold oligonucleotides was achieved by HPLC with UV monitoring at 260 and 360 nm; AFB_1-N7-dGuo modified DNA registers at both wavelengths. Confirmation of the identity of the oligonucleotides was obtained by MALDI-TOF mass spectrometry. Absolute mass values depend on the specific AFB_1 modified sequence (Table 2-1). AFB_1-N7-dGuo modified DNA has an additional mass of 312 amu; AFB_1-FAPY modified oligonucleotides have additional mass of

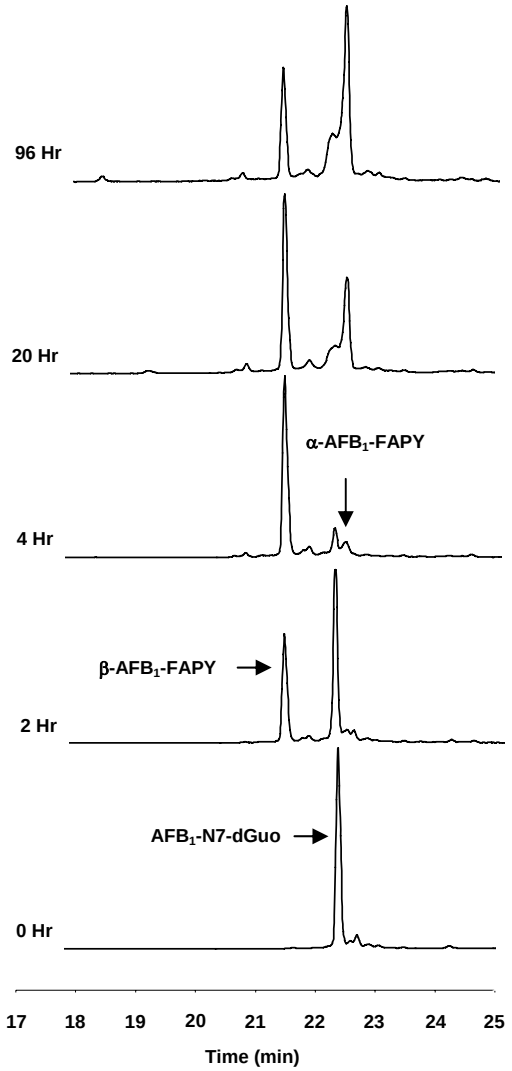


Figure 4-1: Base catalyzed conversion of AFB_1-N7-dGuo modified 5'-CTATXATTCA-3' to β-AFB_1-FAPY and α-AFB_1-FAPY was monitored by HPLC (λ 360 nm, 25 °C).

329 atomic mass units.

Base catalysis of imidazole ring opening and isomer equilibration of AFB_1-N7-dGuo oligonucleotides were monitored by HPLC (Figure 4-1). Initially, AFB_1-N7-dGuo modified 5'-CTATXATTCA-3' eluted at 22.4 min. Formation of the first AFB_1-FAPY isomer was observed at 21.4 min. As the concentration of AFB_1-N7-dGuo was depleted (4 hours), the conversion of the first AFB_1-FAPY isomer to the second AFB_1-FAPY isomer was evidenced by a peak at 22.6 min. Eventually, the AFB_1-FAPY isomers equilibrated to a 2:1 mixture (96 h). The second AFB_1-FAPY isomer had a shoulder peak at 22.4 min. Isolation of the shoulder peak at 22.4 min or the parent peak at 22.6 min indicated that the species equilibrated independent of pH in less than 20 min.

Characterization of AFB_1-FAPY Isomers

Insight into isomer equilibrium was obtained by HPLC. The fractions eluting at 21.4 and 22.6 min for AFB_1-FAPY modified 5'-CTATXATTCA-3' were brought to 90% purity or better by HPLC. Samples (20 μM) in sodium phosphate buffer (0.2 M) at pH 6.0, 7.0, and 8.0 were prepared in triplicate and allowed to equilibrate at room temperature. The equilibration was monitored by HPLC; 0.5 nmol injections gave a signal to noise ratio of 20:1 or better. A 4:1 equilibrium ratio

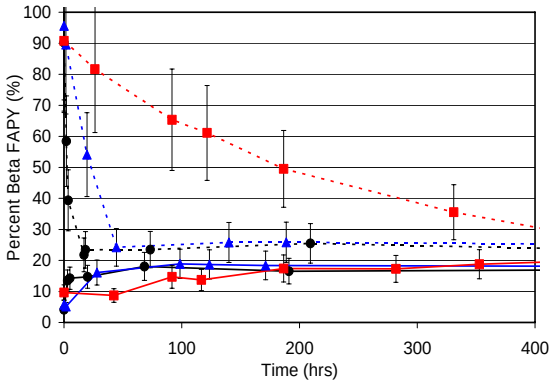


Figure 4-2: Anomeric interconversion of AFB_1-FAPY modified 5'-CTATXATTCA-3'. Purified β-AFB_1-FAPY (broken lines) and α-AFB_1-FAPY (solid lines) were placed in buffers of varying pH (pH 6 = ●, pH 7 = ▲, pH 9 = ■) and equilibration was monitored as a function of time. (Reprinted with permission from J. Am. Chem. Soc. Vol. 128, pgs 15188-99, 2006, Copyright 2006 American Chemical Society)

was reached in less than 24 h for samples at pH 6.0 and 7.0 (Figure 4-2). Alkaline samples did not reach equilibrium for over 16 days. The observed pH dependent equilibration of AFB_1-FAPY isomers is consistent with an acid catalyzed anomer equilibration mechanism (Scheme 1-1).

Evidence that the major FAPY species in single-stranded DNA was the α anomer was obtained by 2D NMR with the key results being obtained by ^{1}H-NOESY experiments (Figure 4-3). The major AFB_1-FAPY isomer in ssDNA was prepared in 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' and kept at pH 9.0, 7 °C during NMR data collection. Assignment of X^5 H2' and H2" resonances was determined by measurement of cross peak intensity with X^5 H3' (2.93 and 2.79 ppm respectively). Relative intensities were determined by integration of 1D projections taken from the 2D matrix at

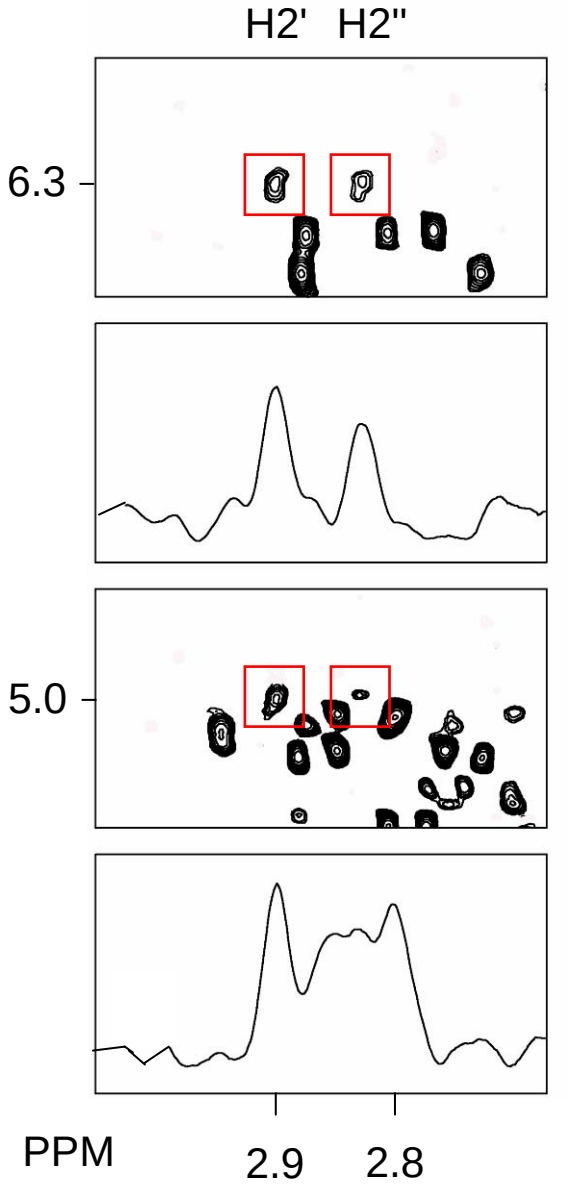


Figure 4-3: Analysis of ^{1}H-^{1}H NOESY intensities indicate the deoxyribose of the AFB_1-FAPY was in the α orientation at C1' for 5'-CTATXATTCA-3'•5'-TGAATCATAG-3'. Results indicate that H1' was spatially closer to H2' than H2". Assignments of H2' and H2" was determined by relative proximity to H3'.

the H3' resonance (5.03 ppm). It is of interest to note that in canonical DNA the H2" resonance is typically downfield of H2'. Subsequent X^5 H1' orientation was determined by cross peak analysis with X^5 H2' and H2". X^5 H1' had a more intense cross peak with H2' placing H1' in the α anomeric configuration. In a β anomer, H1' has a more intense cross peak with H2" (For example, see Figure 5-10).

Anomer Equilibrium

For single stranded oligonucleotides, the α / β equilibrium ratio was both sequence and concentration dependent. This ratio was monitored by serial dilution of the modified oligonucleotides 5'-CTATXATTCA-3' and 5'-CCTCTTCXAACTC-3'. In 5'-CTATXATTCA-3' the β-AFB_1-FAPY was favored at lower concentrations (Figure 4-4). In 5'-CCTCTTCXAACTC-3' the α-AFB_1-FAPY configuration was favored at low concentrations. At concentrations greater than 0.5 mM, α was favored 2:1 over β in ssDNA of both sequences. Mixtures of AFB_1-FAPY modified 5'-CTATXATTCA-3' and 5'-CTATXTTTCA-3' were analyzed at 6 hrs into FAPY equilibration. The 5'-TXA-3' context produced HPLC traces consistent with previous data (Figures 4-1 and 4-5). The 5'-TXT-3' trace was indicative of more than 3 isomeric species (Figure 4-5). Analysis of UV traces of the 5'-TXT-3' peaks at t_R 19.2 and 19.8 min offered limited insight to the identity of these isomers.

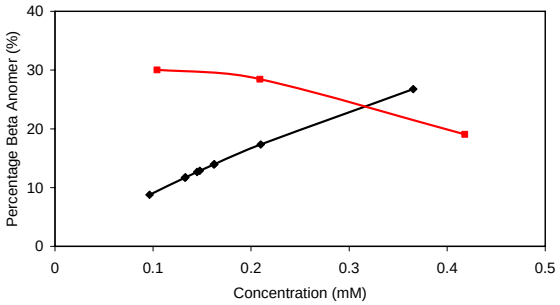


Figure 4-4: Concentration effects on equilibrium of β-AFB_1-FAPY and α-AFB_1-FAPY were measured by HPLC. The oligonucleotides 5'-CTATXATTCA-3' (■) and 5'-CCTCTTCXAACTC-3' (■) were compared. The β anomer was favored at higher concentrations in 5'-CCTCTTCXAACTC-3'. The β anomer was favored at lower concentrations in 5'-CTATXATTCA-3'.

The α / β equilibrium ratio is dependant on single strand and duplex environments. Approximately 0.5 mM AFB_1-FAPY modified 5'-CCTCTTCXAACTC-3' equilibrated to a 4:1 mixture of α and β anomers representing the single strand equilibrium. (Figure 4-6) An excess of complementary strand (t_R 18.5 min) was added to the mixture. This shifted the equilibrium to > 95% β-AFB_1-FAPY in less than 2 h.

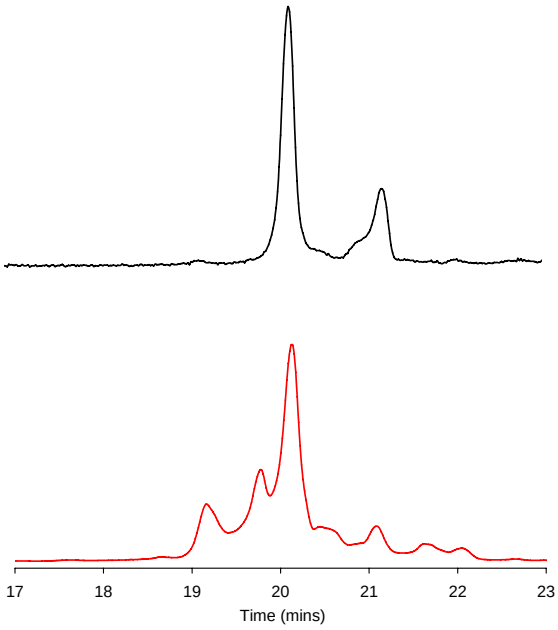


Figure 4-5: HPLC elution profiles of AFB_1-FAPY modified oligonucleotides (λ 360 nm) indicated a sequence dependent effect on equilibrium. Analysis of 5'-CTATXATTCA-3' (—) was indicative of three chromatographically identifiable isomers. Analysis of 5'-CTATXTTTCA-3' (—) was indicative of seven or more species. (0.8 mM, 25 °C)

The chemical stability of AFB_1-FAPY is greater than that of AFB_1-N7-dGuo in that it does not depurinate as easily. However, AFB_1-FAPY does depurinate and sequence has an effect on the rate. The stability of AFB_1-FAPY in the sequences 5'-CTATXATTCA-3' and 5'-CTATXTTTCA-3' was monitored by HPLC as a function of time (Figure 4-7). The sum of the integrals of α-AFB_1-FAPY and β-AFB_1-FAPY were compared to that of the depurinated oligonucleotides (18.2 and 18.5 min; λ =260 nm). The data were fit by linear least squares analysis in Excel 2003 (Microsoft). During the 3 hr time course of the experiment, the 5'-TXT-3' sequence depurinated at a rate nearly three fold faster than 5'-TXA-3'.

Discussion

Assignment of α and β AFB_1-FAPY anomers was achieved by analysis of HPLC and NMR results. Initial work to assign AFB_1-FAPY isomers was conducted at the nucleoside level [31]. It was demonstrated that AFB_1-FAPY exists as a mixture of not just one set of isomers, but three sets, i.e. geometrical, atropisomers, and anomeric (Figure 1-3). This made the identification of the preferred AFB_1-FAPY isomers in DNA more important.

Anomeric Assignment

Evaluation of AFB_1-FAPY equilibria in oligonucleotides by HPLC unraveled the isomer conundrum. An essential aspect of AFB_1-FAPY equilibration is a pH dependent interconversion mechanism (Figure 4-2). Purified samples of the principal HPLC fraction were placed in buffers of varying pH. Samples at neutral and acidic conditions reached equilibrium in < 3 h. Samples at a basic pH did not reach equilibrium for days. This provided direct evidence of an acid-catalyzed isomerization; anomerization is an acid catalyzed process (Scheme 1-1). Rotamer interconversion is not known to be pH dependent.

Supporting evidence for the assignment of AFB_1-FAPY anomers was obtained by ^{1}H-NMR. By manipulation of the pH, it was possible to retard conversion long enough to study the individual anomers by NMR. Expeditious sample preparation and data collection was a crucial element for acquisition of useful data. The α-AFB_1-FAPY adduct was studied in 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' at pH 9; this provided direct comparison to the β-AFB_1-FAPY structural studies carried out earlier which

showed that at pH 7 the FAPY adduct was the β anomer in dsDNA [115]. Assignments of X^6 H2' and X^6 H2" were determined by relative NOE cross peak intensities with X^6 H3' (Figure 4-3). The α vs. β orientation of X^6 H1' was then determined by measuring relative intensities to X^6 H2' and X^6 H2". It was of interest to note the chemical shift flip of X^6 H2' and X^6 H2". Typically, the H2" resonance is downfield of H2'. This chemical shift inversion is consistent with studies of α-deoxyadenosine in duplex DNA [28]. Deoxyribose rings pucker and the distances between deoxyribose protons are expected to change as a consequence of repuckering. This is expected to have a minimal effect on the accuracy of the assignment of anomeric protons; with the exception of the H1'→H4' distance, other deoxyribose proton distances vary 0.24 angstrom or less regardless of pucker [277, 278]

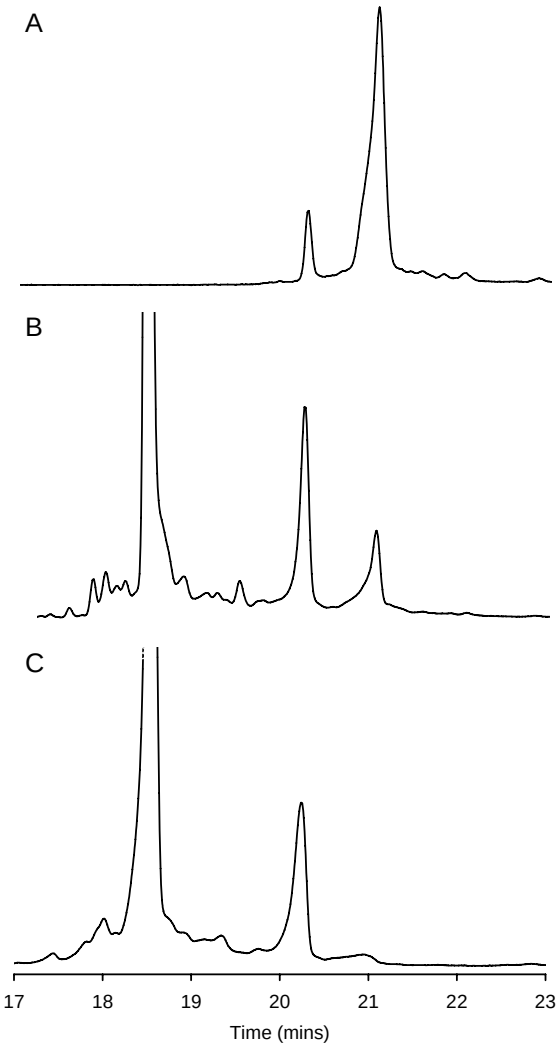


Figure 4-6: The α and β anomer equilibrium of AFB_1-FAPY modified 5'-CCTCTTCXAACTC-3' was dependent on single strand vs. duplex environments. HPLC analysis (λ 260 nm) indicated that an equilibrated mixture of single strand AFB_1-FAPY favors the α form (21.2 min) more than β (20.5 min) approximately 4:1 (panel A). Within 30 min of introducing an excess of complementary strand (18.5 min) to the mixture, the equilibrium began to favor the β anomer (20.5 min) (panel B). After 2 h the equilibrium favored the β anomer almost exclusively (panel C). (Reprinted with permission from J. Am. Chem. Soc. Vol. 128, pgs 15188-99, 2006, Copyright 2006 American Chemical Society)

Sequence Effects on AFB_1-FAPY Equilibria

It has been demonstrated that DNA sequence can affect AFB_1 behavior [118-128]. These data support the idea that the AFB_1-FAPY adduct is influenced by sequence and double strand vs. single strand environments. An AFB_1-FAPY modified, single strand oligonucleotide, 5'-CCTCTTCXAACTC-3', was allowed to reach equilibrium (approximately 2:1 α:β). An excess of complementary strand was added to the mix forming a duplex. Over the course of a few hours the equilibrium completely shifted toward the β anomer (Figure 4-6). This experiment was first conducted by Zengwu Deng (Department of Chemistry, Vanderbilt University) in 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' [110]. By repeating this experiment in 5'-CCTCTTCXAACTC-3', a direct comparison to mutagenesis results was possible [169]. In both sequences, α-AFB_1-FAPY ("major") was preferred in ssDNA and β-AFB_1-FAPY ("minor") was preferred in dsDNA. Mutagenesis experiments were conducted in single stranded vectors while structural studies were in a double stranded environment [115, 169]. It is concluded that the highly mutagenic "minor" species used by Smela et al. [169] in mutagenesis experiments corresponds to the published solution structure in duplex DNA of Mao et al. (β-AFB_1-FAPY) [115]. The blocking "major" species is thus assigned as the α anomer.

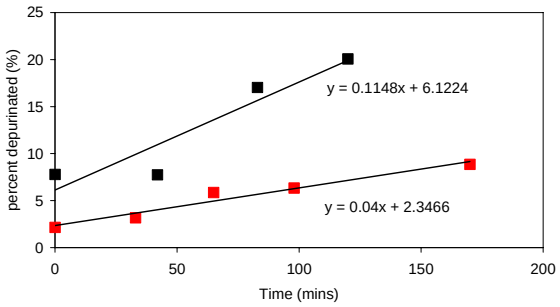


Figure 4-7: Depurination of AFB_1-FAPY modified oligonucleotides was monitored by HPLC as a function of time. The 10mer 5'-CTATXTTTCA-3' (■) released AFB_1-FAPY at a rate 3 fold greater than that of 5'-CTATXATTCA-3' (■).

Mutagenicity of β-AFB_1-FAPY was inferred from a mixture of anomers [169]. Although β-AFB_1-FAPY produced 32% G→T transversions in *E. coli* expressing the error prone polymerases MucAB and UmuDC, it should be considered that this rate may be underestimated. In light of the present AFB_1-FAPY equilibrium data, the progeny from the β-AFB_1-FAPY modified vector may be underestimated; it is likely that mutagenesis experiments designed to be conducted on mostly β-AFB_1-FAPY in fact had a high percentage of α-AFB_1-FAPY.

Effects associated with increasing concentration are related to self-association of the modified oligonucleotides (Figure 4-4). In the 5'-CCTCTTC<u>X</u>AACTC-3' sequence the core six bases are self-complementary (-TTC<u>X</u>AA-). Thus, shift of the α / β ratio toward β-AFB_1-FAPY in this sequence at increasing concentrations can be explained by formation of a self-complementary Watson-Crick duplex around the adduct site. The AFB_1 moiety is expected to compete for the same intercalation site, therefore, it is unclear how both AFB_1 adducts are accommodated. The preference for α-AFB_1-FAPY in the 5'-CTAT<u>X</u>ATTCA-3' sequence at increasing concentration is not easily explained. The driving force behind this concentration dependent equilibration event is not based on Watson-Crick base pair formation. The self-association of 5'-CTAT<u>X</u>ATTCA-3' may be more complicated than a simple bimolecular process.

Placement of AFB_1-FAPY in a 5'-T<u>X</u>T-3' sequence allowed for the equilibration of rotameric and anomeric isomers (Figure 4-5). In the published structure of β-AFB_1-FAPY, the formyl group was stabilized by an inter-residue hydrogen bond with the 3' adenosine N^6 amine (X^5 CHO→A^6 H62) [115]. This effectively stabilized the *E* formyl

geometrical isomer in dsDNA; the *Z* geometrical isomer was favored in nucleoside by a factor of 3:1 [31]. The *Z* isomer was not reported in the duplex sample; the exclusive observation of *E* in DNA can be explained by hydrogen bond stabilization. This X^5 CHO→A^6 H62 hydrogen bond is also present in varying degrees in α-AFB_1-FAPY structures (Chapter IV). In the 5'-T<u>X</u>T-3' sequence, a corresponding hydrogen bond with 3'-thymine is not possible. The observation of multiple rotameric forms in 5'-T<u>X</u>T-3' is attributed to loss of the stability provided by this hydrogen bond. The rate of depurination is noticeably increased in a 5'-T<u>X</u>T-3' sequence (Figure 4-7), therefore, a 3'-dA must have an effect on chemical stability. The reduced conformational stability of 5'-T<u>X</u>T-3' may result in a loss of resonance allowing for increased depurination. It is unknown if a 3'-cytosine could form a X^n:CHO→C^{n+1}:H42 hydrogen bond; it seems plausible that a 3'-dG could stabilize the *Z* formyl geometrical isomer by means of a X^n:CHO→G^{n+1}:O6 hydrogen bond. If this hydrogen bond is responsible for the reduced depurination of 5'-T<u>X</u>A-3', 5'-T<u>X</u>T-3' should be the least stable of possible 3' neighbors whereas a 3' guanine and cytosine are predicted to be more stable than 5'-T<u>X</u>T-3' but less stable than 5'-T<u>X</u>A-3'. Position 2 of codon 249 (p53) is flanked by a 3'-dG while position 3 has a 3'-dC. Position 3 is the hallmark mutational hotspot associated with AFB_1 exposure. It is uncertain if this is a result of site reactivity, poor site-specific repair, or perhaps a tertiary structural effect. Nevertheless, it seems plausible that inter-residue hydrogen bonding is a contributor to sequence related effects of all FAPY lesions.

Biochemical Implications

The identities of the AFB_1-FAPY isomers are of considerable interest to the structural biology community. Many DNA adducts exist as mixtures of isomers. In the case of AFB_1-FAPY, two isomers of the lesion produce uniquely different biological effects. The β-AFB_1-FAPY lesion produces G→T transversions like AFB_1-N7-dGuo [73, 162, 169]. However, α-AFB_1-FAPY is a toxic lesion; this is consistent with the biological processing of other α anomers [28, 38, 39, 297].

In AFB_1-FAPY modified nucleosides, both geometrical and atropisomers were present at equilibrium [31]. In this work, it is demonstrated that 5'-CTATXATTCA-3' and 5'-CCTCTTCXAACTC-3' modified oligonucleotides exist as equilibrium mixtures involving α and β anomers. However, when the 3' deoxyadenosine is replaced with a deoxythymidine, more than four isomers are observed. It seems logical that these additional isomers are rotamers. An understanding of AFB_1-FAPY equilibria is necessary to interpret the divergent responses to the principal FAPY isomers in replication experiments. In earlier studies of AFB_1-FAPY and related lesions in oligonucleotides it has been suggested that the principal equilibrating species are atropisomers [116, 169, 302, 303] (Figure 1-3 C). In light of present results, it is concluded that the principal AFB_1-FAPY isomers are anomers (Figure 1-3 D). This is consistent with mutagenesis results where the α-AFB_1-FAPY lesion ("major") blocked replication [169]; α anomers are known to be cytotoxic [28, 38, 39]. It should also be noted that the historical terms "major" and "minor" have been applied to different isomers in different publications [115, 169]. This work allows for the retirement of the antiquated monikers "major" and "minor" in favor of more descriptive α and β terminology.

Effects of duplex vs. single-stranded environments on isomer equilibration are an essential element to AFB_1-FAPY equilibrium. Site-specific mutagenesis experiments were performed using single stranded viral genomes [162, 169]. The highly mutagenic "minor" form, now identified as β-AFB_1-FAPY, was initially believed to be less favored than α-AFB_1-FAPY. The β anomer may be considerably more dangerous than initially projected. It seems logical that at the replication fork, conversion of β-AFB_1-FAPY to α-AFB_1-FAPY may be relatively slow compared to polymerase read-through. At pH 7.0 conversion from β to α took hours (Figure 4-2). Therefore, in genomic DNA, β-AFB_1-FAPY would be favored and very little of α-AFB_1-FAPY would be predicted to form during replication. However, it is unknown whether β to α conversion is accelerated as a result of conformational effects when bound by a replication enzyme. Detailed crystallographic analysis of AFB_1-FAPY modified oligonucleotides in complex with replicative and bypass enzymes will be necessary to address this question.

CHAPTER V

THYMINE GLYCOL LESION EQUILIBRIA

Introduction

This chapter addresses the complex equilibria of thymine glycol (Tg) in DNA. Thymine glycol (*cis*-5*R*) exists primarily as a β anomer. However, *cis*-5*R*,6*S*-Tg was in equilibrium with *trans*-5*R*,6*R*-Tg in dsDNA. The *cis*-*trans* equilibrium was affected by the complementary base.

The most common thymine oxidation product in DNA, 5,6-dihydroxy-dihydro-2′-thymidine, known as thymine glycol (Tg), is formed by exposure to ionizing radiation, as well as oxidizing agents (Scheme 1-5) [197, 199, 200]. The C5 and C6 atoms in Tg are chiral and thus Tg exists in DNA as two diastereomeric pairs of epimers, the 5*S cis*, *trans* pair (5*S*,6*R*; 5*S,*6*S*) and the 5*R cis*, *trans* pair (5*R*,6*S*; 5*R*,6*R*). The 5*R* pair is more abundant and more stable [37]. In both cases, the *cis* epimers predominate [37, 206].

Surprisingly, previous solution structural studies of the 5*R*-Tg adducts did not report that they exist in DNA as equilibrating diastereomeric 6*S*/6*R cis*, *trans* epimers. The 5*R*-Tg adduct was previously examined in the 5′-ATgA-3′ sequence, paired opposite dA [232, 233]. These NMR studies concluded that Tg induced a localized structural perturbation, and that Tg was extrahelical [233]. The structure of the 5*R*-Tg adduct in the 5'-GTgC-3' sequence was reported to be disordered [232]. These solution structures differed from recent results utilizing a binary primer-template complex of the replicative

RB69 DNA polymerase, in which the template adducted with the 5*R*-Tg lesion was crystallized. The analysis of this binary complex, representing the situation immediately following incorporation of dATP opposite Tg, revealed the presence of the *cis*-5*R*,6*S*-Tg epimer, which was intrahelical and formed a Watson-Crick base pair with the dA at the primer 3'-terminus [234]. This was consistent with modeling studies in which Tg-5*R* was predicted to successfully pair with dA [231]. Moreover, for the *cis*-5*R*,6*S*-Tg epimer, the Tg methyl group was in the axial conformation, hindering stacking of the adjacent 5′-template guanine, providing a rationale for the observation that extension past this lesion is not observed [214].

The present work describes the 5*R*-Tg adduct, incorporated site-specifically into 5′-$G^{1}T^{2}G^{3}\underline{C^{4}G^{5}Tg^{6}}G^{7}T^{8}T^{9}T^{10}G^{11}T^{12}$-3′ (codon 273 of p53 underlined) and annealed with 5′-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}\underline{A}^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3′ and 5′-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}\underline{G}^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3′ producing $Tg^{6}\bullet A^{19}$ and $Tg^{6}\bullet G^{19}$ base pairs. In contrast to previous NMR studies, [11,12] it was demonstrated that in solution in duplex DNA, the *cis*-5*R*,6*S* and *trans*-5*R*,6*R* epimers are both present at significant levels opposite adenine, with the *cis*-5*R*,6*S* epimer predominating. It was shown that for the 5*R*-Tg adduct, the predominant deoxyribose C1′ epimer in double-stranded DNA is the β anomer. This is significant because proposed *cis*- *trans* intermediates present the possibility of a transient glycosyl iminium bond formation and subsequent anomerization similar to that found in FAPY DNA lesions (Figure 1-1) [25-27, 31]. This data suggests that the significant levels of the *trans*-5*R*,6*R* epimer cannot be ignored with respect to the biological processing of these adducts, corroborating studies showing that the repair of

Tg adducts by DNA N-glycosylases/AP lyases is modulated by the *cis-trans* epimerization [224].

Results

Oligodeoxynucleotide Analysis

The dodecamer 5′-GTGCGTgGTTTGT-3′, Tg = 5*R*-Tg, was synthesized as reported [224]. Mass spectrometric analysis yielded the anticipated molecular ion peak with mass 3732 amu. Capillary gel electrophoretic analysis showed that the modified oligodeoxynucleotide eluted as a single peak. Enzymatic hydrolysis of the oligodeoxynucleotide to deoxynucleosides, followed by C-18 HPLC chromatography, also yielded a single peak (Figure 5-1). The UV trace of the Tg mononucleoside matched that of previous reports [198]. Additional peaks corresponding to dA, dC, dG, and dT were observed in the anticipated intensity ratios. Thus, the Tg-adducted oligodeoxynucleotide consisted of a single chromatographically separable species. The modified single strand

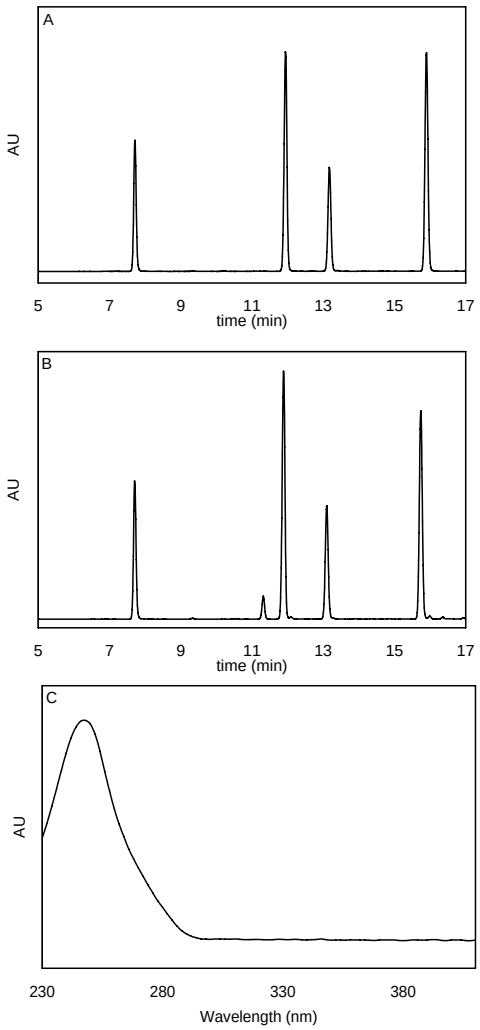


Figure 5-1: HPLC analysis of enzyme digest products of control 12mer duplex (panel A) and Tg modified 12mer duplex (Tg•A) (panel B). Cytidine eluted first at 7.7 mins followed by guanosine (11.9 min), deoxythymidine (13.2 min), and adenosine (15.9 min). The thymine glycol nucleoside eluted at 11.2 min (panel B). The UV trace of Tg nucleoside is show in panel C.

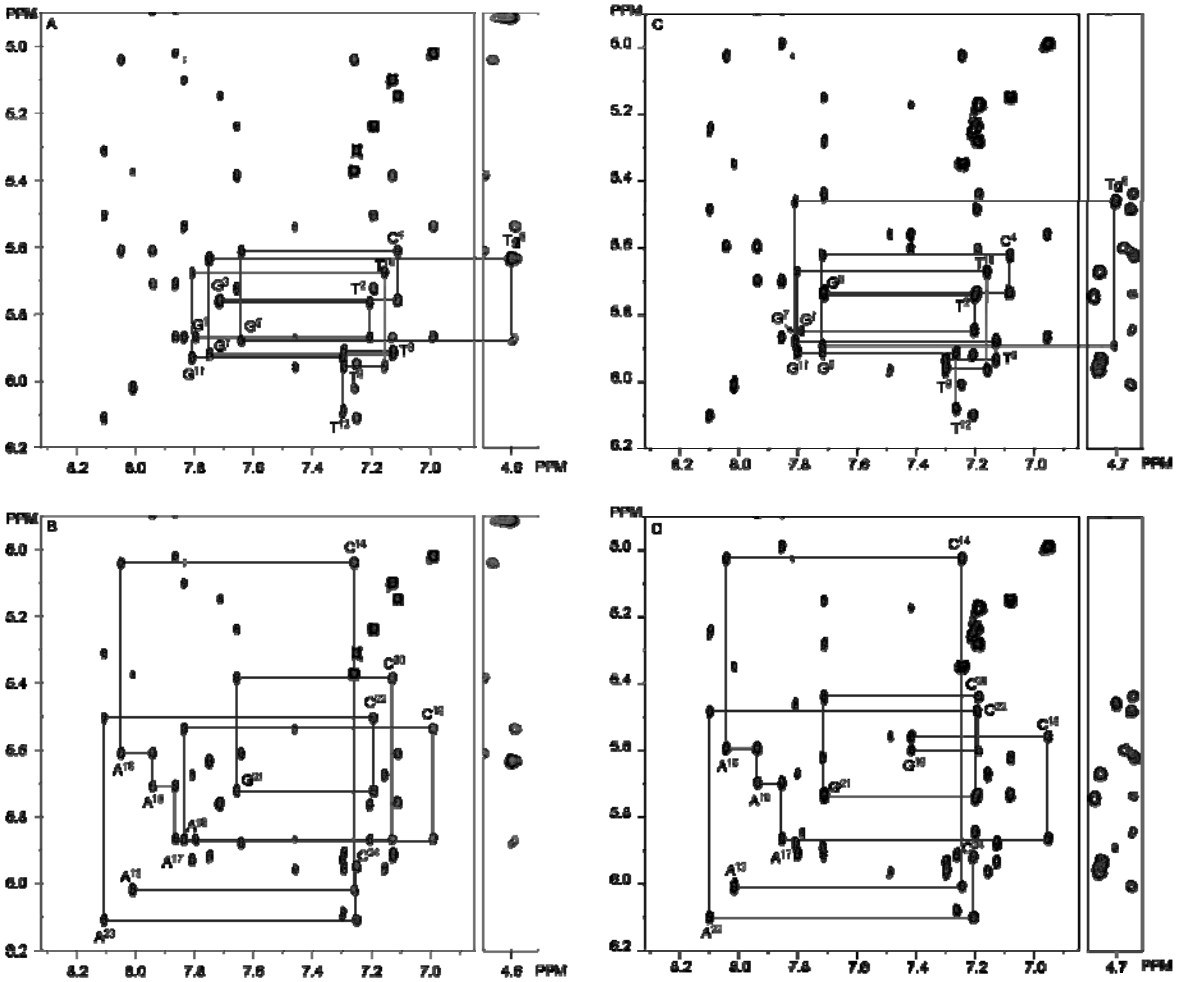


Figure 5-2: Sequential assignments of NOESY H8/H6 to H1' protons for the $Tg^{6}•A^{19}$ and $Tg^{6}•G^{19}$ duplexes. Connectivity for the Tg modified strand of $Tg^{6}•A^{19}$ (panel A) and the adenine containing complementary strand (panel B) are unbroken. Connectivity for the Tg modified strand of $Tg^{6}•G^{19}$ (panel C) and the adenine containing complementary strand (panel D) are also unbroken. Intra-residue aromatic to H1' cross peaks are labeled.

oligonucleotide was incubated for 12 days at pH 6.0. The sample was monitored by HPLC. Over this period of time, only a single chromatographic peak was observed.

NMR Spectroscopy of Non-exchangeable DNA Protons

Resonances for the 5′-GTGCGTgGTTTGT-3′•5′-ACAAACACGCAC-3′ duplex (Tg•A) were assigned using standard strategies [304, 305]. Figure 5-2 shows an expansion of the NOESY spectrum including the NOEs between purine H8 and

pyrimidine H6 protons and the corresponding deoxyribose H1´ protons. In the modified strand (Figure 5-2, panel A), the sequential NOE connectivity was not interrupted at the lesion site. Also, there was no break in connectivity for the complementary strand (Figure 5-2, panel B). The 5´-GTGCGTgGTTTGT-3´•5´-ACAAACGCGCAC-3´ duplex (Tg•G) also indicated no break in connectivity at Tg of the modified strand (Figure 5-2, panel C). Expanded NOESY plots of H8/H6 to H1' for Tg•G indicated complete intra-strand connectivity for the complementary strand (Figure 5-2, panel D).

The deoxyribose protons were assigned from a combination of COSY and NOESY data. With the exception of several of the H4' protons, and the stereotopic assignments of the H5' and H5" sugar protons, assignments were made unequivocally. NOE intensities were used to assign the deoxyribose H2' and H2" resonances based on the fact that H2'→H3' cross peak was anticipated to have a greater intensity than the H2"→H3' cross peak. These distances are expected to vary approximately 0.24 Å or less regardless of pucker [277, 278]. In general, canonical B-DNA distances, chemical shift, and scalar coupling, between the H4', H5', and H5" protons were used to tentatively assign the H5' and H5" deoxyribose protons. The assignments of the non-exchangeable protons are provided in Appendix A.

NMR Spectroscopy of Exchangeable DNA Protons

The assignments of the Watson-Crick imino and amino protons of the modified oligodeoxynucleotides were made using standard methods [306] (Figure 5-3 & 5-4). Regarding the Tg•A duplex, the G^5 H1 imine resonance arising from the 5' neighbor base pair G^5•C^{20} was broad at 5 ºC and disappeared when the temperature was increased

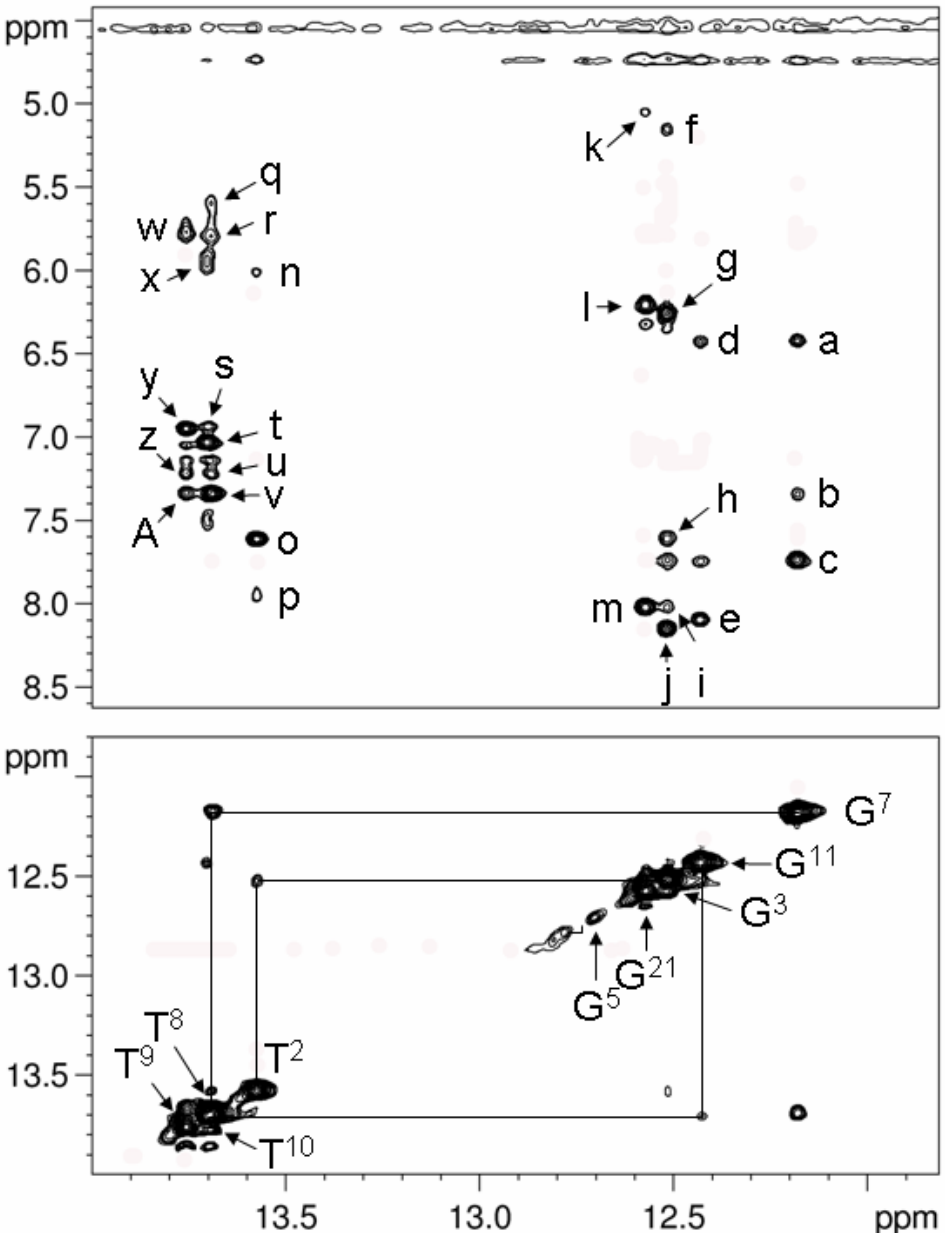


Figure 5-3: The upper panel is an expanded plot showing the sequential NOE connectivity for the amino protons to amino protons for the Tg•A duplex. The cross peaks are assigned as (a) G^{7}:H22→G^{7}:H1 (b) A^{17}:H2→G^{7}:H1 (c) C^{18}:H42→G^{7}:H1 (d) G^{11}:H22→G^{11}:H1 (e) C^{14}:H42→G^{11}:H1 (f) C^{22}:H5→G^{3}:H1 (g) G^{3}:H22→G^{3}:H1 (h) A^{23}:H2→G^{3}:H1 (i) C^{4}:H42→G^{3}:H1 (j) C^{22}:H42→G^{3}:H1 (k) C^{4}:H5→G^{21}:H1 (l) G^{21}:H22→G^{21}:H1 (m) C^{4}:H42→G^{21}:H1 (n) A^{19}:H62→T^{2}:H3 (o) A^{23}:H2→T^{2}:H3 (p) A^{23}:H61→T^{2}:H3 (q) A^{17}:H62→T^{8}:H3 (r) A^{16}:H62→T^{8}:H3 (s) A^{16}:H2→T^{10}:H3 (t) A^{15}:H2→T^{10}:H3 (u) A^{16}:H61→T^{8}:H3 (v) A^{17}:H2→T^{8}:H3 (w) A^{16}:H62→T^{9}:H3 (x) A^{15}:H62→T^{10}:H3 (y) A^{16}:H2→T^{9}:H3 (z) A^{16}:H61→T^{9}:H3 (A) A^{17}:H2→T^{9}:H3. The lower panel is an expanded plot showing the sequential NOE connectivity for the imino protons. The data was collected at 800 MHz at 250 ms mixing time and a temperature of 7 °C.

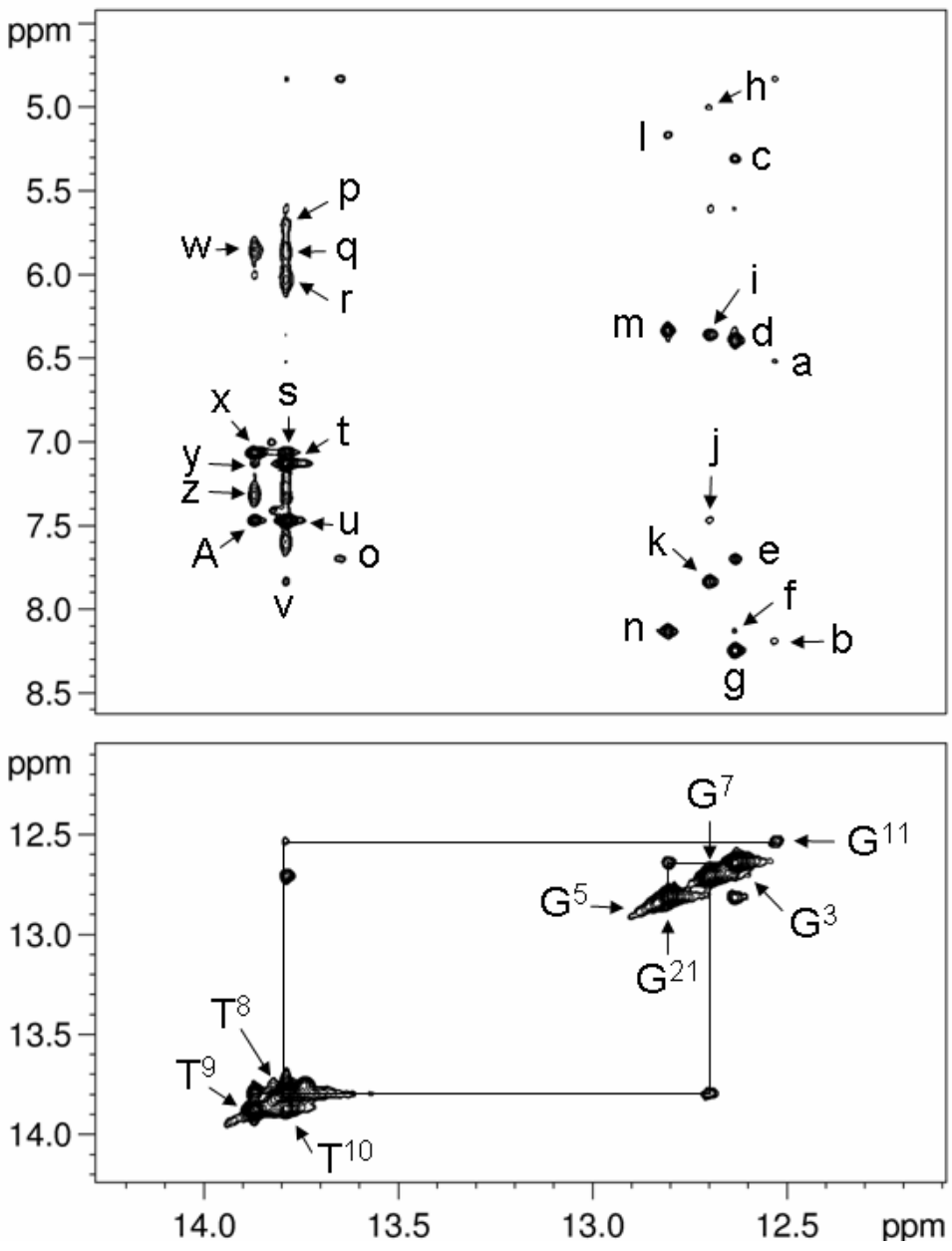


Figure 5-4: The upper panel is an expanded plot showing the sequential NOE connectivity for the amino protons to amino protons for the Tg•G duplex. The cross peaks are assigned as (a) G^{11}:H22→G^{11}:H1 (b) C^{14}:H62→G^{11}:H1 (c) C^{22}:H5→G^{3}:H1 (d) G^{3}:H22→G^{3}:H1 (e) A^{23}:H2→G^{3}:H1 (f) C^{4}:H42→G^{3}:H1 (g) C^{22}:H42→G^{3}:H1 (h) C^{18}:H5→G^{7}:H1 (i) G^{7}:H22→G^{7}:H1 (j) A^{17}:H2→G^{7}:H1 (k) C^{18}:H42→G^{7}:H1 (l) C^{4}:H5→G^{21}:H1 (m) G^{21}:H22→G^{21}:H1 (n) C^{4}:H42→G^{21}:H1 (o) A^{23}:H2→T^{2}:H3 (p) A^{17}:H62→T^{8}:H3 (q) A^{16}:H62→T^{8}:H3 (r) A^{15}:H62→T^{10}:H3 (s) A^{16}:H2→T^{10}:H3 (t) A^{15}:H2→T^{10}:H3 (u) A^{17}:H2→T^{8}:H3 (v) C^{18}:H42→T^{8}:H3 (w) A^{16}:H62→T^{9}:H3 (x) A^{16}:H2→T^{9}:H3 (y) A^{15}:H2→T^{9}:H3 (z) A^{16}:H61→T^{9}:H3 (A) A^{17}:H2→T^{9}:H3. The lower panel is an expanded plot showing the sequential NOE connectivity for the imino protons. The data was collected at 800 MHz at 250 ms mixing time and a temperature of 7 °C.

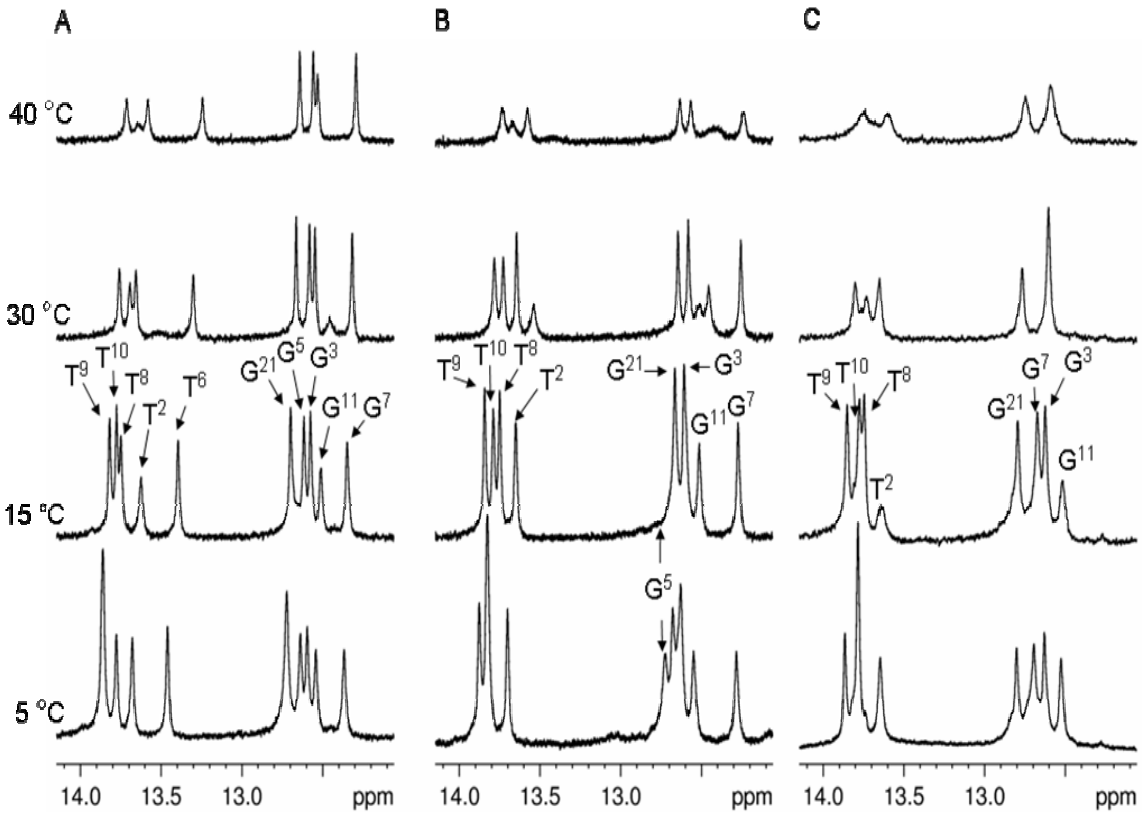


Figure 5-5: Temperature dependent analysis of imino protons of T•A (panel A), Tg•A (panel B), and Tg•G (panel C) duplexes as monitored by ^{1}H NMR. Melting of the G^{5}•C^{20} base pair neighbor of Tg^{6} occurs at a lower temperature compared to the unmodified sample. G^{7} H1 is deshielded in Tg•G (panel C) compared to Tg•A (panel B).

to 15 ºC (Figure 5-5). In contrast, for the unmodified sample, the G^{5} H1 imine resonance was sharp and was observed at temperatures as high as 40 ºC. In Tg•A, there was no cross peak between the broad G^{5} H1 resonance and G^{21} H1 (Figure 5-3). The Tg^{6} H3 amine resonance, anticipated to resonate in the amino proton region of the ^{1}H NMR spectrum, was not identified. The A^{19} H61 and H62 resonances were not observed. This was attributed to solvent exchange. The imino resonances for base pairs T^{2}•A^{23}, G^{3}•C^{22}, C^{4}•G^{21}, G^{7}•C^{18}, T^{8}•A^{17}, T^{9}•A^{16}, T^{10}•A^{15}, and G^{11}•C^{14} were observed (Figure 5-3). The imino resonances for the terminal base pairs G^{1}•C^{24} and T^{12}•A^{13} were not observed, this was attributed to exchange broadening with water.

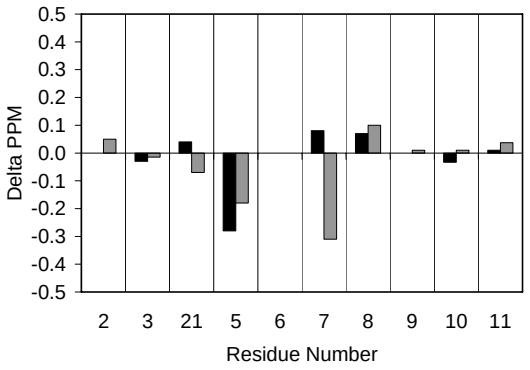


Figure 5-6: Chemical shift perturbation of amino protons of Tg•A and Tg•G samples relative to T•A are indicative of change in chemical environment. In the Tg•A duplex (black bars) the H1 of G^5 was deshielded 0.3 ppm relative to G^5 H1 of T•A; G5 H1 was deshielded in 0.2 ppm in Tg•G (grey bars). The G^7 H1 was affected differently in each duplex. In Tg•A, G^7 H1 was shielded 0.1 ppm whereas in Tg•G it was deshielded 0.3 ppm. Tg^6 H3 was not detected in either duplex.

The assignments of the Watson-Crick imino and amino protons of the Tg•G duplex are seen in Figure 5-4. The G^5 H1 imine resonance arising from the 5′ neighbor base pairs G^5•C^{20} were broad at 5 ºC (Figure 5-5). Much like the Tg•A duplex, in Tg•G the Tg^6 H3 amine resonance was not identified. Again, this was attributed to solvent exchange. The imino resonances for base pairs T^2•A^{23}, G^3•C^{22}, C^4•G^{21}, T^8•A^{17}, T^9•A^{16}, T^{10}•A^{15}, and G^{11}•C^{14} were observed (Figure 5-4). The imino resonances for the terminal base pairs G^1•C^{24} and T^{12}•A^{13} were not observed.

Chemical shifts of modified duplex imino resonances were compared to the T•A control (Figure 5-6). In Tg•A and Tg•G the Tg^6 H3 was not observed. However, the G^7 H1 resonance was perturbed significantly in both modified duplexes. In Tg•A, G^7 H1 was shielded 0.1 ppm. In the Tg•G duplex, G^7 H1 was deshielded 0.3 ppm.

NMR Spectroscopy of Tg Protons

Analysis of NOESY data for the Tg•A duplex obtained at multiple mixing times showed the presence of intense cross peaks arising from chemical exchange between two chemical species for both the Tg^6 CH_3 and Tg^6 H6 protons (Figure 5-7). The integrated volumes of the two exchange cross peaks were consistent at multiple mixing times. Integration of the two Tg^6 CH_3 resonances indicated that the two species were present at

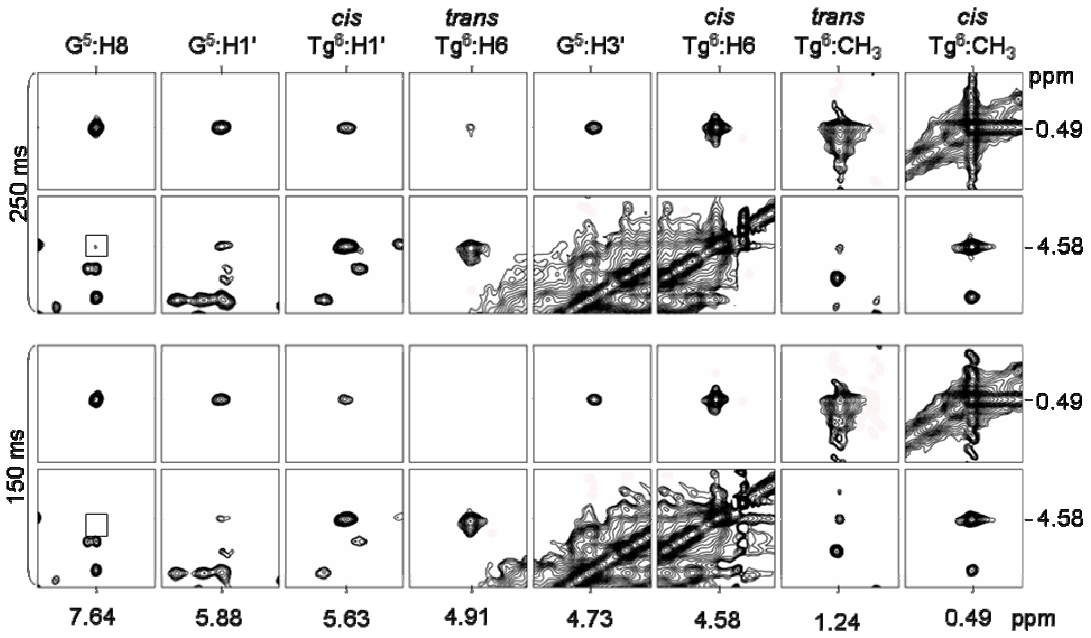


Figure 5-7: ^{1}H-^{1}H NOESY tile plot of Tg•A at mixing times of 250 and 150 milliseconds. Cross peaks between *cis* CH_3 / H6 and G^5, the 5' neighbor, decreased as a function of mixing time. Chemical exchange cross peaks between the *cis* and *trans* epimers of Tg did not significantly decrease in intensity as a function of mixing time.

an equilibrium ratio of 7:3 (Figure 5-8). For the major species, the Tg^6 CH_3 protons exhibited a chemical shift of 0.49 ppm, while the Tg^6 H6 proton resonated at 4.58 ppm. These chemical shifts were consistent with previously reported values [233]. A total of 23 NOE cross peaks were assigned between Tg^6 CH_3 and H6 in major species and DNA (7 for Tg^6 H6 and 16 for Tg^6 CH_3) in the Tg•A duplex. The chemical shift values for the Tg^6 CH_3 and H6 resonances arising from the minor species were significantly downfield relative to those from the major species, located at 1.24 ppm and 4.91 ppm, respectively. The Tg^6 CH_3 resonance for the minor species was overlapped with the T^2 CH_3 resonance. For the minor species, there was only one NOE cross peak, observed between Tg^6 H6 and Tg^6 H2'. This cross peak was observed in both *cis* and *trans* samples. Integration of the cross peak was consistent with a 7:3 equilibrium. A single set of chemical shifts were

observed for G^5 and G^7; the mixture of species at Tg^6 did not extend to the neighboring nucleotides.

Analysis of NOESY data for the Tg•G duplex obtained at multiple mixing times did not exhibit chemical exchange cross peaks for either the Tg^6 CH_3 and Tg^6 H6 protons (Figure 5-9). This indicated that one major chemical species was observable on the millisecond timescale in the Tg•G sample. The Tg^6 CH_3 protons exhibited a chemical shift of 0.91 ppm, while the Tg^6 H6 proton resonated at 4.70 ppm. A total of 20 NOE cross peaks were assigned between Tg^6 CH_3 and H6 in major species and DNA (9 for Tg^6 H6 and 11 for Tg^6 CH_3). A single set of chemical shifts were observed for G^5 and G^7; the mixture of species at Tg^6 did not extend to the neighboring nucleotides.

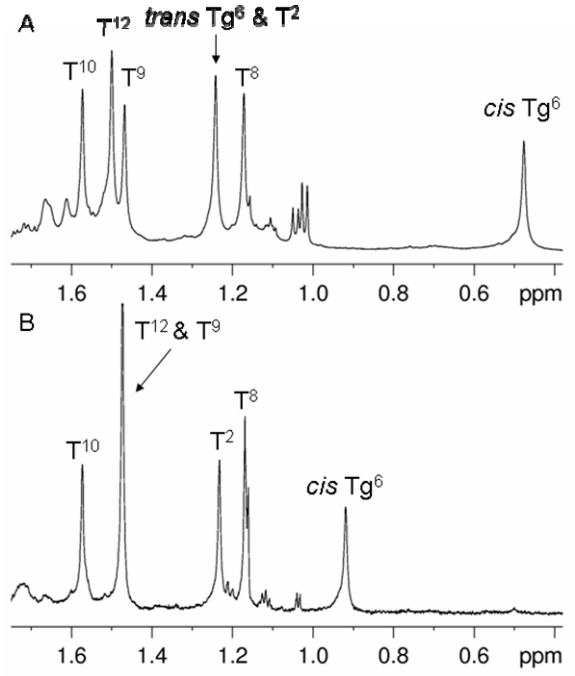


Figure 5-8: 1D ^{1}H NMR of thymine and thymine glycol methyl groups. In the Tg•A duplex (panel A) the *trans* Tg^6 CH_3 and T^2 CH_3 overlapped. Integration and intensity of *cis* and *trans* Tg methyl peaks were used, in addition to cross peak integration, to determine relative equilibrium of the Tg lesion in Tg•A. A single Tg^6 CH_3 peak was observed in the Tg•G duplex (panel B).

<u>NMR Spectroscopic Assignment of Tg^6 Isomers</u>

The two Tg species observed in the Tg^6•A^{19} duplex were assigned as arising from slow configurational exchange (NMR time scale) between the *cis*-5*R*,6*S* and *trans*-5*R*,6*R* epimers. Seven NOESY cross peaks were observed between the Tg^6 H6 resonance of the major species and surrounding protons (Table 5-1). Their intensities at mixing times of 80 and 250

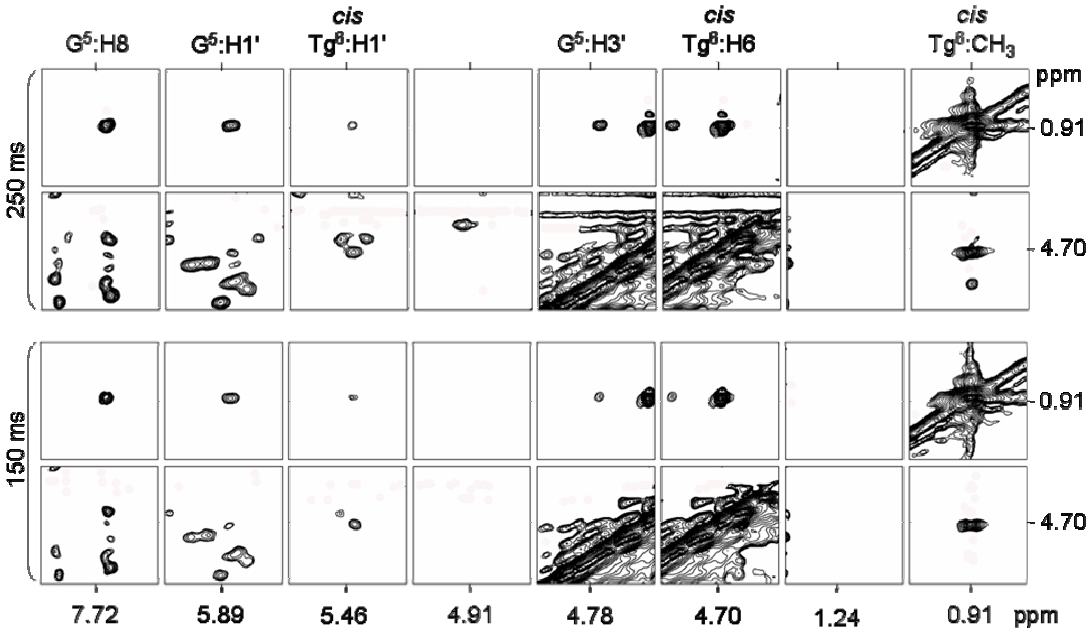


Figure 5-9: ^{1}H-^{1}H NOESY tile plot of Tg•G at mixing times of 250 and 150 milliseconds. No chemical exchange cross peaks were observed at resonances (1.24 and 4.91 ppm). Cross peaks with G^5 decrease in intensity as a function of mixing time.

ms were compared with the corresponding distances predicted for each epimer on the basis of molecular modeling. The spectral overlap of Tg^6 H3' and Tg^6 H6 resonances (< 0.01 ppm) in the major species hindered the assessment of the Tg^6 H3'→Tg^6 H1' and Tg^6 H6→Tg^6 H1' cross peaks. For the *cis*-5*R*,6*S* configuration, Tg^6 H6 and Tg^6 CH_3 are spatially proximate, yielding a strong Tg^6 H6→Tg^6 CH_3 NOE even at the short mixing time of 150 ms (Figure 5-7). Likewise, the G^5 H1'→Tg^6 H6 and G^5 H8→Tg^6 H6 NOE were diagnostic of the *cis*-5*R*,6*S* configuration. On this basis, the major species, present at ~ 70% population, was assigned as the *cis* epimer.

Table 5-1: NOESY cross peaks from the major Tg H6 species to neighboring protons.

Major Tg^6 H6 cross peak	Trans Model Distance (Å)	Cis Model Distance (Å)	NOESY Intensity (80 ms)	NOESY Intensity (250 ms)
G^5 H2"	5.7	2.2	w	m
Tg^6 H5'	4.5	3.3	s	s
G^5 H1'	3.9	2.3	nd	m
G^5 H8	7.1	6.1	nd	w
G^5 H4'	6.8	5.4	nd	w

There was a single isomeric species present in the Tg•G duplex as observed by NMR. Its configuration was determined to be *cis* by the same distance filtering methodology previously described. In the Tg•G sample Tg^6 H3' and Tg^6 H6 resonances were dispersed (4.53 and 4.70 ppm respectively).

Anomeric Configuration of the Deoxyribose Sugar

The possibility that Tg^6 could undergo configurational exchange between α and β deoxyribose anomers was considered. The NOESY data for the Tg•A deoxyribose ring showed that the intensity of the Tg^6 H1'→Tg^6 H2" NOE was greater than the Tg^6 H1'→Tg^6 H2' NOE, which placed H1' in the β configuration (Figure 5-10). Anomeric analysis of Tg•G NOE cross peak volumes was precluded as a result of resonance overlap of Tg^6 H2' and H2".

It was considered that anomeric equilibration may be dependant on single strand verses double strand environments in Tg samples like AFB_1-FAPY. The modified 5′-GTGCGTgGTTTGT-3′ was purified by HPLC, placed in sodium phosphate buffer (pH 6.0) and monitored for 3 weeks (25 °C). During the course of this experiment, a single

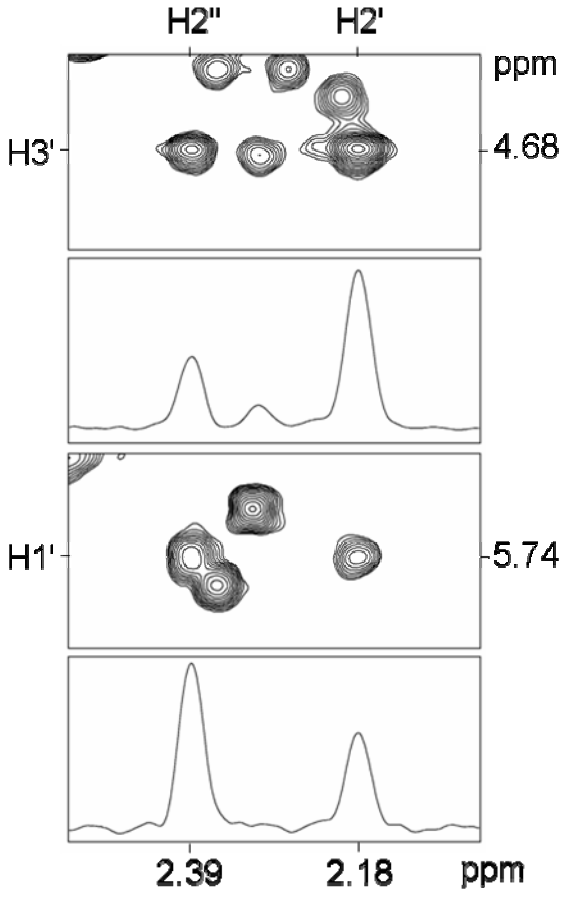


Figure 5-10: NOE intensities indicated the deoxyribose of the Tg^6 is in the β orientation.

chromatographically separable oligonucleotide was observed.

Potential Energy Minimization

As a result of the loss of aromaticity of the thymine base, Tg puckers. Thus, for either the *cis*-5*R*,6*S* or *trans*-5*R*,6*R* configurations, the Tg CH_3 may be in either axial or equatorial conformations. Density functional theory (DFT) and second order Møller-Plesset (MP2) quantum calculations using GAUSSIAN 03 [252] with the 6-31G*, 6-31G**, and 6-311++G** basis sets (Table 5-2) indicated that the *cis*-5*R*,6*S* configuration was predicted to be of slightly lower energy than the *trans*-5*R*,6*R* configuration (~1 kcal/mol), which was consistent with the present experimental observations, as well as previous experimental observations [3-5]. The calculations also predicted that for the *cis*-5*R*,6*S* configuration, the Tg CH_3 group favored the axial conformation by ~ 5 kcal/mol, whereas for the *trans*-5*R*,6*R* configuration, the Tg CH_3 group favored the equatorial conformation.

Table 5-2: Sum of electronic and zero-point energies of thymine glycol bases. †‡

cis-5R,6S-Tg	*B3LYP/6-31G**	*B3LYP/6-31G***	*B3LYP/6-31++G***	*MP2/6-31++G***
CH_3 Axial	-404,683.2973	-404,698.5177	-404,817.4954	-403,849.9161
CH_3 Equatorial	-404,679.2802	-404,679.2802	-404,812.9824	-403,844.9469
trans-5*R*,6*R*-Tg				
CH_3 Axial	-404,678.4983	-404,693.7951	-404,812.3523	-403,845.2241
CH_3 Equatorial	-404,682.0525	-404,697.2935	-404,816.5090	-403,847.9367

†Reported energies are in units of kcal/mol.
‡DFT calculation energies are zero-point energy corrected.

Discussion

Despite the fact that it has been recognized that Tg exists in DNA as two diastereomeric pairs of epimers, 5*S*,6*R*, 5*S*,6*S* and 5*R*,6*S*, 5*R*,6*R* [37, 206], previous

solution structure studies of the 5*R*-Tg adduct (no structural studies of the 5*S*-Tg adduct have been conducted) overlooked the fact that the adduct exists in DNA as an equilibrating pair of diastereomeric 5*R* *cis*, *trans* epimers [232, 233]. The present NMR data reveal that in the Tg•A duplex oligodeoxynucleotide the 5*R*-Tg adduct actually exists as a 7:3 equilibrium mixture of *cis* and *trans* epimers, which equilibrate in slow exchange on the NMR timescale. This ratio is comparable to the 87% *cis* to 13% *trans* ratio of epimers reported at the nucleoside level [37]. Thus, we conclude that significant levels of the *trans*-5*R*,6*R* epimer maybe present in both duplex DNA and nucleosides, and that in duplex DNA base paired opposite adenine, the 5*R*-Tg adduct should be considered to exist as an equilibrium mixture of the two epimers. The presence of *cis* and *trans*-Tg^6 seems to have little effect on neighboring bases as is evidenced by the presence of a single set of resonances for G^5 and G^6.

However, Tg does have a negative effect on base pairing. This is evidenced by the fact that the base pair G^5•C^{18} melts approximately 35 °C lower in Tg modified DNA compared to unmodified DNA (Figure 5-5). In addition, NMR analysis of both the exchangeable Tg^6 N3H and A^{19} H61 & H62 amine resonances suggest that these protons are in exchange with solvent. This is indicative of reduced base pairing between Tg^6 and A^{19}. Taken together, this reduced stability is expected to significantly contribute to the observed reduction in duplex T_m (Δ 13 °C) (Chapter VII).

Similarly, in the Tg•G duplex oligodeoxynucleotide, neighboring bases produced a single set of NMR resonances for G^5 and G^6. However, the adjacent base pair does have an effect on the *cis*-*trans* equilibrium. In Tg•G the *cis*-5R,6S isomer was the only

observed Tg epimer. The T_m of Tg•G was also reduce by 13 °C, however, this sample was less thermodynamically favorable than Tg•A (Chapter VII).

An explanation for the significant equilibrium differences observed in the Tg•A and Tg•G duplexes must reside in altered base pairing schemes. The failure to observe hydrogen bonding between the Tg•A and Tg•G base pairs may be a result of increased dynamics at the lesion site. A complete lack of interstrand hydrogen bonding would be expected to produce disordered structures. Ordered structures are observed suggesting some form of stabilization, yet experimental evidence supporting hydrogen bonding was not observed. Restrained molecular dynamics calculation of the Tg•A duplex predicts hydrogen bonding is possible between the base pairs during a 5 ns trajectory. Deficiencies in hydrogen bonding in Tg•A and Tg•G are a consequence of steric clash between Tg^6 CH_3 and G^5. A possible explanation for the observed *cis-trans* equilibrium in the Tg•A sample may be reduced steric clash between G^5 and a *trans* Tg-*5R,6R* lesion where the CH_3 is in an equatorial conformation. Clark et al. have shown that an equatorial CH_3 produces less steric interaction with the 5' neighbor. Therefore, in the Tg•A duplex, a thermodynamic drive to conserve Watson-Crick hydrogen bonding may facilitate epimerization to the *trans* epimer on a limited basis. This conclusion draws support from DFT calculations predicting that *trans* 5*R*-Tg with CH_3 in an equatorial position is the next most energetically favorable Tg configuration second only to *cis* 5*R*-Tg with CH_3 in the axial arrangement. In the Tg•G duplex there is no possibility for Watson-Crick hydrogen bonding although orientation similarly to a wobble G•T pair is possible. This would shift Tg^6 toward the major groove, as compared to its orientation when placed opposite A^{19}, and would enable the Tg^6 CH_3 group to be maintained in the

energetically more favorable axial conformation, consistent with the observation that in the Tg•G pair the *cis*-5*R*,6*S* epimer is favored.

Biological Significance

It is estimated that human cells repair hundreds of Tg adducts per day [198]. The Tg adduct is a substrate for base excision repair, both in *E. coli* and in mammalian cells [307]. In *E. coli*, repair of Tg is initiated by endonuclease III (Nth) [308] and endonuclease VIII (Nei) [309]. Yeast [23], mammalian [24,25], and human orthologs [26-28] of Nth have been characterized. Likewise, human orthologs of Nei have been characterized [29, 30]. The base excision repair of Tg lesions is dependant both on their stereochemical configurations (5*R*,6*S* vs. 5*S*,6*R*), and the identity of the complementary base [224]. Using Nth, the *cis*-5*S*,6*R* adduct placed complementary to adenine was repaired with greater efficiency than was the *cis*-5*R*,6*S* adduct. With human hNth, the *cis*-5*R*,6*S* Tg adduct placed complementary to adenine was repaired more efficiently than when it was complementary to guanine. However, hNth was inactive against the *cis*-5*S*,6*R* adduct, regardless of the opposing base. The human endonuclease-like protein (hNeill) did not differentiate between the stereoisomers, but was more efficient when Tg was opposite a guanine [7, 10].

Significantly, the Tg adduct is also a substrate for nucleotide excision repair (NER) proteins. Both randomly-introduced Tg adducts and abasic sites were substrates for the UvrABC NER enzymes of *E. coli* [310]. Subsequently, it was determined that the Tg adduct was excised from DNA *in vitro* by human NER enzymes [311]. However, DNA containing dihydrothymine, a lesion with a similar structure to thymine glycol, was

not incised [312]. The structure of Tg adducts in duplex DNA and structure-activity relationships with regard to their repair are of considerable interest.

If not repaired, the 5*R*-Tg adduct is lethal to cells [34-38]. This is attributed to the observation that the lesion is a strong block to DNA replication for both replicative and repair polymerases [15, 39-44]. However, pyrimidines 5′ to template Tg allow residual polymerase read-through more often than do purines [216]. In any case, polymerase blockage is characterized by the termination of primer extension following the incorporation of dATP opposite Tg; *i.e.*, it is caused by an inability to extend beyond the Tg•A pair, rather than failure to insert dATP at template Tg. Aller et al [234], utilizing a binary complex of a *cis*-5*R*,6*S* Tg-adducted template:primer with the replicative RB69 DNA polymerase, concluded that the template *cis*-5*R*,6*S* Tg was intrahelical and formed a Watson-Crick base pair with the incorporated dA. However, the axial conformation of the *cis*-5*R*,6*S* Tg CH_3 group hindered stacking of the 5' neighbor template guanine, presumably hindering incorporation of the next incoming nucleotide into the growing primer strand, and providing a rationale as to why primer extension past the lesion was prohibited even though DNA polymerases readily incorporated dATP across from the 5*R*-Tg lesion. These structural studies corroborated the modeling work of Clark et al. [231] who reported that the Tg•A base pair was stable, suggesting that Tg retains the ability to direct the insertion of the correct nucleotide during DNA synthesis, whereas interactions with the 5′-neighboring base pair were destabilized.

Based on previous solution structures of Tg, the lesion was reported to be "extrahelical" in dsDNA. What is unclear is how repair enzymes can differentiate an extrahelical lesion based on the opposing base. An extrahelical lesion is expected to have

little if any interaction with its opposite base. Here we show that the opposing base pair has a significant effect on *cis*-*trans* equilibrium. Therefore, an interstrand interaction between Tg and dA or dG must exist. This interaction may alter other properties of the helix, such as the degree of Tg extrahelicity, base pair stability, base pair geometry, or lesion dynamics. It seems logical that a repair enzyme may recognize these resultant effects.

Despite the fact that Tg was a strong block to replication in *E. coli*, it was weakly mutagenic, causing < 0.5% T→C transitions [219]. It was concluded that the *cis*-5*R*,6*S* Tg adduct was displaced laterally toward the major groove as compared to an unmodified thymine, perhaps increasing the likelihood of G•Tg wobble pairing, potentially explaining the observed T→C transitions [219]. The degree of extrahelicity was predicted to be sequence-dependent, being modulated by the identity of the 3′-neighbor nucleotide [219]. An interesting possibility is that the *trans*-5*R*,6*R* lesion is responsible for the < 0.5% mutagenicity reported in *E. coli*. Specifically, for the *trans*-5*R*,6*R* epimer, the equatorial conformation of the Tg CH_3 group is predicted to be more energetically favorable (Table 3-2). The equatorial orientation of the Tg CH_3 group in the *trans*-5*R*,6*R* epimer may allow stacking of the 5' neighbor template guanine, presumably facilitating incorporation of the next incoming nucleotide into the growing primer strand.

CHAPTER VI

NMR STRUCTURAL STUDIES OF THE α-AFB_1-FAPY LESION IN DNA

Introduction

This chapter addresses the solution structures of α-AFB_1-FAPY in 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' and 5'-CTXA-3'. Conclusions drawn from these structures are specific to the localized 5'-TXA-3' sequence context. Under basic conditions (pH 8.0) and low temperature (7 °C), conversion of α to β is sufficiently slow to permit expeditious collection of NMR data. The β anomer of AFB_1-FAPY was previously studied in 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' [115]. Use of identical duplex sequences allowed direct comparison of α and β anomer structure alterations. It seemed plausible that placing α-AFB_1-FAPY in an unfavored duplex environment may not accurately reflect the anomer's structure in a single strand context. Therefore, α-AFB_1-FAPY was also analyzed in a single stranded tetramer. Comparison of the α anomer in ssDNA and dsDNA along with the previously published β anomer in dsDNA leads to the conclusion that favorable stacking interactions are an important factor in anomeric equilibrium. In addition, distortion of the phosphodiester backbone of α-AFB_1-FAPY modified DNA may contribute to associated cellular toxicity.

Results

HPLC Analysis

Anomeric purity of AFB_1-FAPY modified samples was determined by HPLC prior to spectroscopic analysis. Alkaline conditions (pH 8.0-8.5) sufficiently retard anomeric interconversion to permit for expeditious analysis of α-AFB_1-FAPY modified oligonucleotides (Figure 6-1). The α anomers remained at 90% purity or better after 3 days at low temperatures (~ 5 °C). The β anomers produced a single HPLC peak at 19.5 min. However, α-AFB_1-FAPY produced a major peak at 20.1 min preceded by a shoulder peak at 20.0 min. Attempts to isolate the principal α-AFB_1-FAPY peak (20.1 min) and its shoulder (20.0 min) were unsuccessful as the two re-equilibrated in less than 30 minutes independent of pH.

UV Melting

The denaturation of native, α-AFB_1-FAPY, and β-AFB_1-FAPY modified duplexes were monitored by temperature dependent UV hyperchromicity (pH 8.0). The unmodified 5'-CTATGATTCA-3'•5'-TGAATCATAG-3' produced a single transition; T_m was determined to be 36

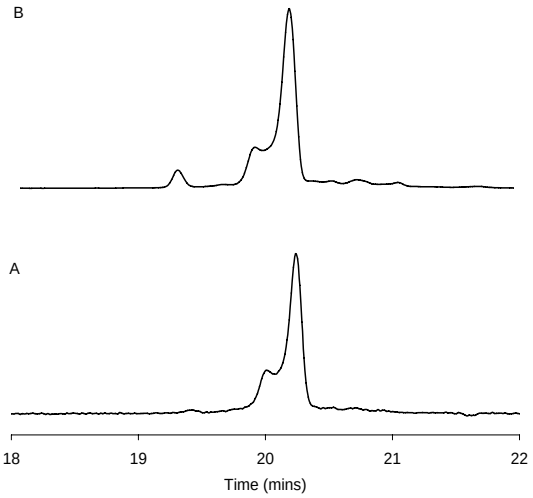


Figure 6-1: Anomeric purity of α-AFB_1-FAPY modified 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' was determined by HPLC (λ 360 nm) pre-NMR (panel A) and post-NMR (panel B). NMR experiments began with > 95% α-AFB_1-FAPY (20.1 min) (panel A). Purity was approximately 90% after 3 days (panel B). (β-AFB_1-FAPY (19.5 min)

°C by first derivative calculation (Figure 6-2, panel A). The α-AFB_1-FAPY and β-AFB_1-FAPY modified 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' was purified to > 90 % by HPLC prior to UV analysis. The β-AFB_1-FAPY modified duplex (Figure 6-2, panel B) produced a single transition at 50 °C during the heating stage. During the cooling stage, a biphasic transition was observed at 50 °C and 23 °C. The α-AFB_1-FAPY modified duplex (Figure 6-2, panel C) produced a biphasic transition at 50 °C and 23 °C during both heating and cooling stages.

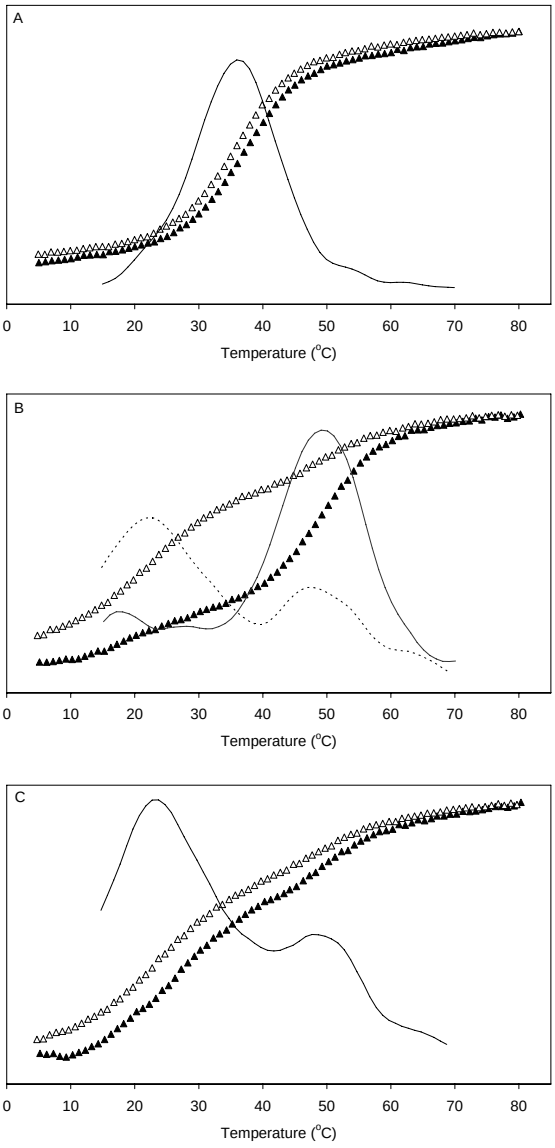


Figure 6-2: UV thermal melting analysis of AFB_1-FAPY modified oligonucleotides. The T_m of 5'-CTATGATTCA-3'•5'-TGAATCATAG-3' was determined to be 36 °C by first derivative analysis (panel A). The β-AFB_1-FAPY modified duplex (panel B) produced a single transition at 50 °C during the heating stage (5→ 80 °C = ▲). A biphasic transition was observed at 50 °C and 23 °C for the cooling stage (80→ 5 °C = ▲). The α-AFB_1-FAPY modified duplex (panel C) produced a biphasic transition at 50 °C and 23 °C during both stages.

Electronic Circular Dichroism (ECD) Spectroscopy

As a complement to NMR spectroscopy, ECD spectra were obtained to more fully analyze AFB_1-FAPY induced structural effects. Tertiary structural effects of AFB_1-FAPY anomers were analyzed by ECD in ssDNA and

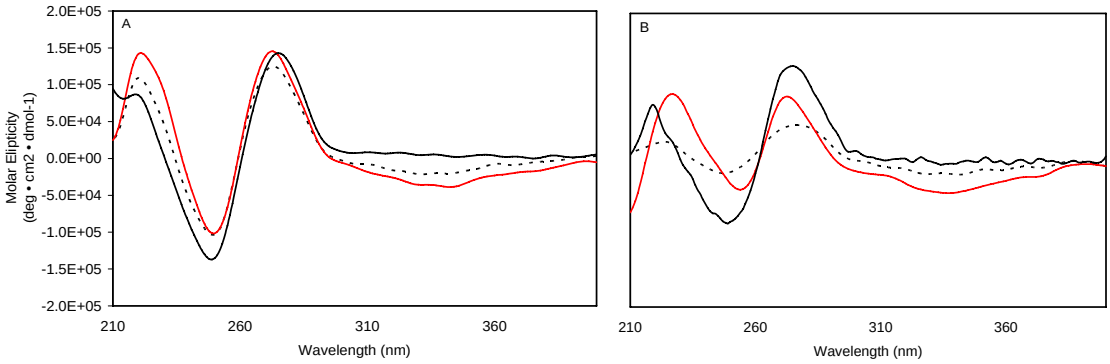


Figure 6-3: Electronic circular dichroism of AFB_1-FAPY anomers in oligonucleotides. The duplex 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' (panel A) was compared to 5'-CTATXATTCA-3' (panel B). Unmodified oligonucleotides (solid black line) were compared to β-AFB_1-FAPY (broken black line) and α-AFB_1-FAPY (solid red line).

dsDNA (Figure 6-3). The 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' duplex containing anomers of AFB_1-FAPY were compared to the native duplex (Figure 6-3, panel A). There were differences in molar ellipticity at λ 215, 255, 275, and 350 nm for both anomers between ssDNA and dsDNA. Bathochromic shifts of approximately 5 nm were observed in ssDNA samples near λ 215 nm (5'-CTATXATTCA-3') relative to the control (Figure 6-3, panel B). Values of molar ellipticity were divergent from native

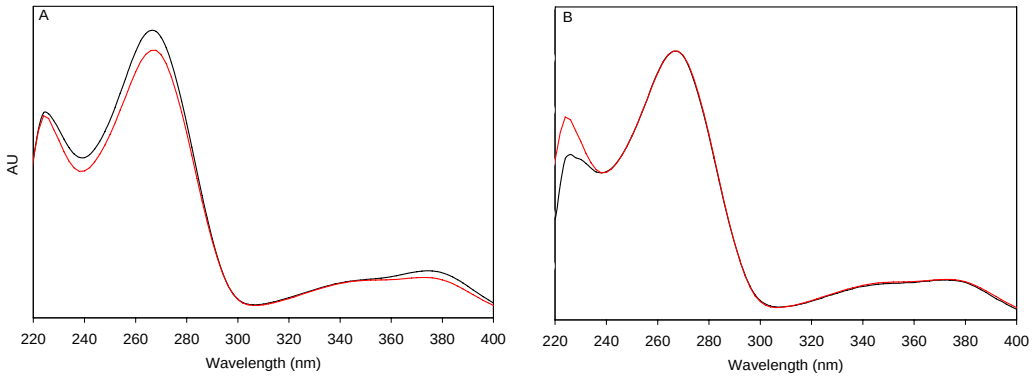


Figure 6-4: UV spectra of AFB_1-FAPY modified 5'-CTATXATTCA-3'. The spectra of the α anomer (—) was compared to that of the β anomer (—) in panel A. The spectra of the α anomer (—) was compared to that of the chromatographically inseparable preceding α anomer shoulder (—) in panel B.

ssDNA samples between λ 210 and 310 nm. However, AFB_1 induced ellipticity (~λ 310 – 360 nm) was consistently weak and negative in both ssDNA and dsDNA.

UV Spectroscopy

Ultraviolet spectra of AFB_1-FAPY modified 5'-CTATXATTCA-3' were extracted from HPLC diode array data obtained during isomer purification. The β anomer (19.5 min) exhibited hyperchromicity between λ 240 and 280 nm and AFB_1 induced hyperchromicity at 380 nm when compared to the principal α anomer (20.1 min) (Figure 6-4, panel A). The spectrum of the major (20.1 min) isomer of the α anomer was compared to that of the minor isomer (20.0 min) (Figure 6-4, panel B). The UV traces were nearly identical with minor differences in the far UV range (~ λ 225 nm).

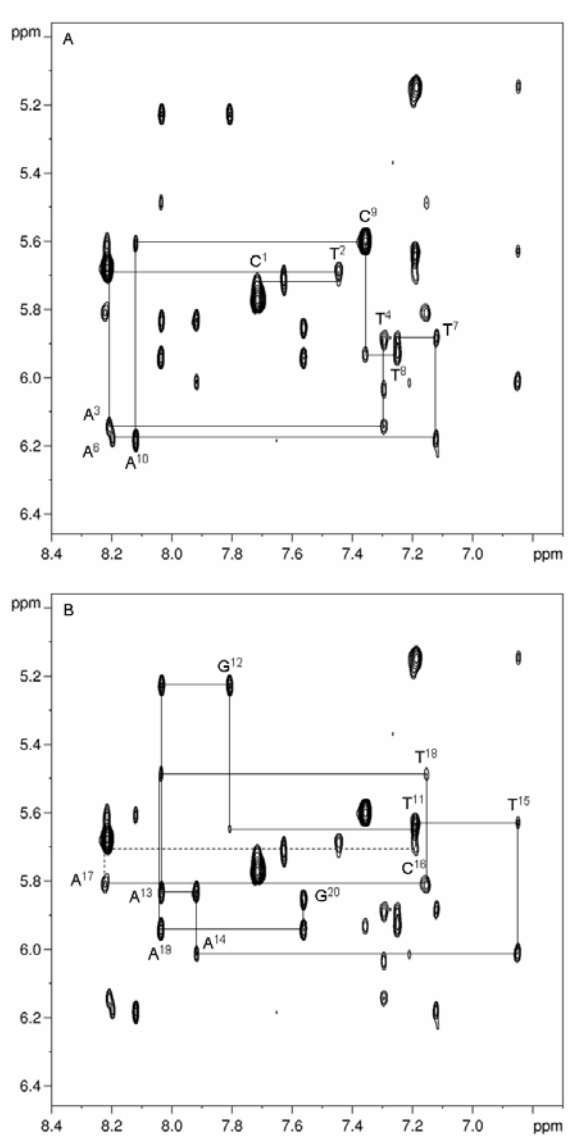


Figure 6-5: Sequential assignments of NOESY H8/H6 to H1' protons for the α-AFB_1-FAPY modified duplex 5'-CTATXATTCA-3'•5'-TGAATCATAG-3'. Connectivity for the modified strand (panel A) was broken between T^4 H1' and A^6 H8. Connectivity for the complementary strand (panel B) appeared to be broken between C^{16} H6 and A^{17} H1'. Intra-residue aromatic to H1' cross peaks are labeled.

NMR Spectroscopy of Non-exchangeable DNA Protons

Duplex DNA resonances were assigned using standard strategies [304, 305, 313]. Figure 6-5 shows an expansion of the NOESY spectrum in the region of the NOEs between purine H8 and pyrimidine

H6 protons and the deoxyribose H1′ protons for the 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-3' duplex. Sequential NOE connectivity was observed except for a break in the modified strand between T^4 H6 and A^6 H1' and for the complementary strand between C^{16} H1' and A^{17} H8. Multiple resonances were observed for the aromatic signals of residues T^{15} and A^{14} and some deoxyribose resonances of C^{16} (Appendix A; Table A-5). A total of 197 non-exchangeable resonances were assigned for the duplex. Chemical shifts were compared to those of an unmodified duplex (Figure 6-6). As expected, chemical shift perturbation was observed at the X^5 lesion site. However, disturbance was also observed at residues T^4, A^6, T^7, A^{14}, T^{15}, C^{16}, A^{17}, and T^{18}.

Single strand tetramer resonances were assigned based on chemical shift, pyrimidine vicinal coupling, and process of elimination. Intra-residue dipolar interactions were observed and used for resonance assignment (Figure 6-7). High mobility of the tetramer prohibited the evolution of most inter-residue dipolar interactions traditionally used in dsDNA resonance assignment. The 5'-$C^1T^2G^3A^4$-3' sequence was selected to minimize resonance overlap. Deoxyribose proton spin systems were correlated with pyrimidine bases by unambiguous assignment of cytosine H5→H6 and thymine CH_3→H6 cross peaks. The adenine deoxyribose ring system was assigned based on correlation with the characteristically downfield shifted adenine H8 resonance. Assignment of the G^3 and X^3 resonances were based on process of elimination and previously published literature values [240]. A total of 33 non-exchangeable resonances were assigned for the tetramer. Chemical shifts of residues 2, 3, and 4 of the tetramer were compared to residues 4, 5 and 6 of the modified duplex (Figure 6-6, panel F).

Tetramer resonances were shielded an average of 0.5 ppm relative to those of the modified duplex.

NMR Spectroscopy of Exchangeable DNA Protons

Fast exchange with solvent prohibited the assignment of tetramer imino and amino proton resonances. The assignments of Watson-Crick hydrogen-bonded imino and amino protons of the duplex oligodeoxynucleotides were made using standard methods

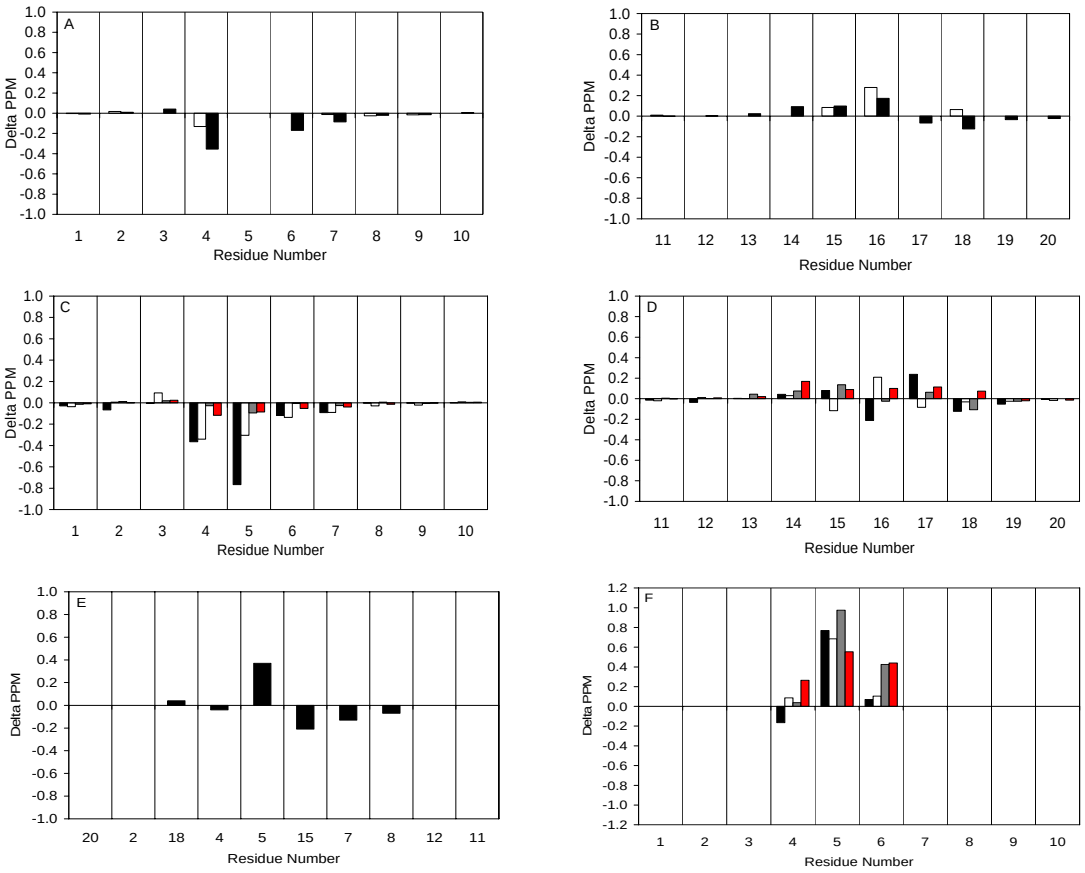


Figure 6-6: Chemical shift perturbation of α-AFB$_1$-FAPY modified 5'-C^{1}T^{2}A^{3}T^{4}X^{5}A^{6}T^{7}T^{8}C^{9}A^{10}-3'•5'-T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}-3' relative to unmodified oligonucleotide. The aromatic H6/H8 (■) and pyrimidine H5/CH$_3$ (□) resonance perturbations for the modified strand and complementary strand are listed in panel A and B respectively. Chemical shift perturbation for the deoxyribose resonances H1' (■), H2' (□), H2" (■), and H3' (■) for the modified strand and complementary strand are listed in panel C and D respectively. The exchangeable N1H/N3H (■) resonance differences are reported in panel E. Comparison of duplex and single strand tetramer deoxyribose resonances shifts H1' (■), H2' (□), H2" (■), and H3' (■) are reported in panel F. The resonances for residue 3 were not compared, the tetramer had a cytosine in place of an adenine in this position.

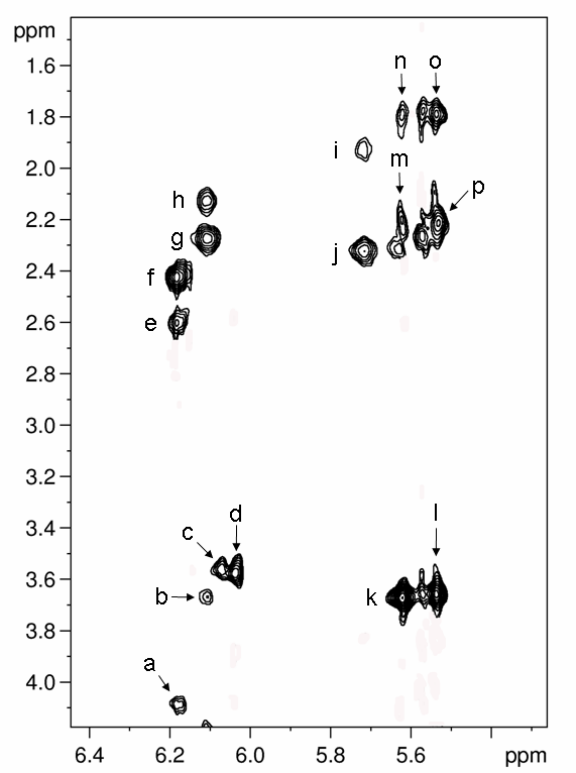


Figure 6-7: Select NOE assignments of AFB_1 inter residue and deoxyribose intra residue cross peaks for the α-AFB_1-FAPY modified 5'-CTXA-3'. The cross peaks are assigned as (a) A^4 H4'→A^4 H1' (b) AFB_1 OCH_3→T^2 H1' (c) AFB_1 H9a→ AFB_1 H9 (d) AFB_1 H9a→ AFB_1 H6a (e) A^4 H2'→ A^4 H1' (f) A^4 H2"→ A^4 H1' (g) T^2 H2"→T^2 H1' (h) T^2 H2'→T^2 H1' (i) C^1 H2'→C^1 H1' (j) C^1 H2"→C^1 H1' (k) AFB_1 OCH_3→ AFB_1 H5 (l) X^3 H5"→ X^3 H1' (m) X^3 H2'→ AFB_1 H5 (n) X^3 H2"→ AFB_1 H5 (o) X^3 H2"→ X^3 H1' (p) X^3 H2'→ X^3 H1'

(Figure 6-8) [306]. There was a break in sequential connectivity of the imino resonances X^5 H3 and T^4 H3 in the duplex sample. Both X^5 H3 and T^4 H3 had a cross peak with AFB_1 -OCH_3. The strong cross peaks X^5 H3→ C^{16} H41 and X^5 H3→ C^{16} H42 indicated that Watson-Crick hydrogen bonding between X^5 and C^{16} was intact. Similarly, imino resonances for the base pairs A^3•T^{18}, T^4•A^{17}, X^5•C^{16}, A^6•T^{15}, T^7•A^{14}, T^8•A^{13}, and C^9•G^{12} were observed suggesting Watson-Crick hydrogen bonding was preserved. Imino resonances for the base pairs C^1•G^{20}, T^2•A^{19}, and A^{10}•T^{11} were not observed; this is attributed to line broadening due to exchange with water. A total of 25 exchangeable resonances were assigned for the duplex. Imino chemical shifts were compared to those of unmodified duplex (Figure 6-6, panel E). The X^5 H3 resonance was shielded 0.4 ppm while H3 of T^{15}, T^7, and T^8 was deshielded an average of 0.15 ppm.

NMR Spectroscopy of Aflatoxin B_1 Protons

The AFB_1 H5, H6a, H8, H9, H9a, and -OCH3 resonances were assigned from NOE connectivities, chemical shift, and literature values [115] (Figure 6-9). AFB_1 H6a and H9a were identified from both COSY and NOESY experiments. H8 and H9 were assigned based on NOEs to H6a or H9a, and between themselves. A strong NOE from AFB_1 H5 to AFB_1 -OCH3 revealed that the latter resonance was at δ 3.51 ppm for the duplex and δ 3.76 ppm for the single strand tetramer, while AFB_1 H5 was at δ 5.7 ppm for both ssDNA and dsDNA. The X^5 CHO resonance was observed at δ 8.3 ppm for both duplex and tetramer samples. The duplex assignment of X^5 CHO was supported by NOEs to X^5 H1', and AFB_1 H8a; in the tetramer X^5 CHO had a cross peak only with AFB_1 H8a. Table 6-1 lists the assignments of the AFB_1 protons of both samples.

Table 6-1: Observed NOEs between AFB_1-FAPY lesion and DNA protons.

AFB_1-FAPY	*ds DNA*	*ss DNA*
H2α	C^{16} H1', C^{16} H3', C^{16} H2", C^{16} H4', C^{16} H5, C^{16} H6, A^{17} H2,	
H2β	A^{17} H2,	
H3α	A^{17} H3',	
H3β	A^{17} H2,	
H5	A^{17} H2, T^4 H6, T^4 H4',	T^4 H2', T^4 H2"
H6a	T^4 CH_3, T^4 H6, X^5 H8,	
H8a	X^5 H8, X^5 CHO,	X^3 CHO
H9a	T^4 CH_3, T^4 H6, T^4 H3', X^5 H8,	
H9		
-OCH_3	A^{17} H2, T^4 H1', T^4 H2', T^4 H2", T^4 H5', T^4 H3, X^5 H3,	T^4 H1'

NMR Spectroscopy of NOEs from Aflatoxin B_1 to DNA

In dsDNA, 30 inter-residue NOEs from AFB_1 to DNA were assigned (Table 6-1) but only 4 AFB_1 inter-residue NOEs were observed for ssDNA. The protons of the two AFB_1 furanose moieties exhibited NOEs to major groove and imino protons in dsDNA; most were to the 5'-neighboring base-pair T^4•A^{17} (Figure 6-9). Thus, both H6a and H9a, which are located on the same face of the AFB_1 residue, produced NOEs to T^4 H6 and CH_3. Similar NOEs were observed for AFB_1 H8 and H9. In ssDNA, AFB_1 inter-residue NOEs comprised exclusively H5 and $-OCH_3$ to T^2 deoxyribose protons H1', H2', and H2". In dsDNA AFB_1 H5 and $-OCH_3$ resonances produced NOEs to various minor groove and imino protons. These were primarily in the 5' direction to base pair T^4•A^{17}, and the modified nucleotide X^5. They included NOEs between AFB_1 -OCH_3 and T^4 H1', H2', H2", H5', H3, X^5 H3, and A^{17} H2. The cyclopentenone ring H2α exhibited NOEs to H1', H3', H2", H4', H5, of C^{16}, and H2 of A^{17} in the complementary

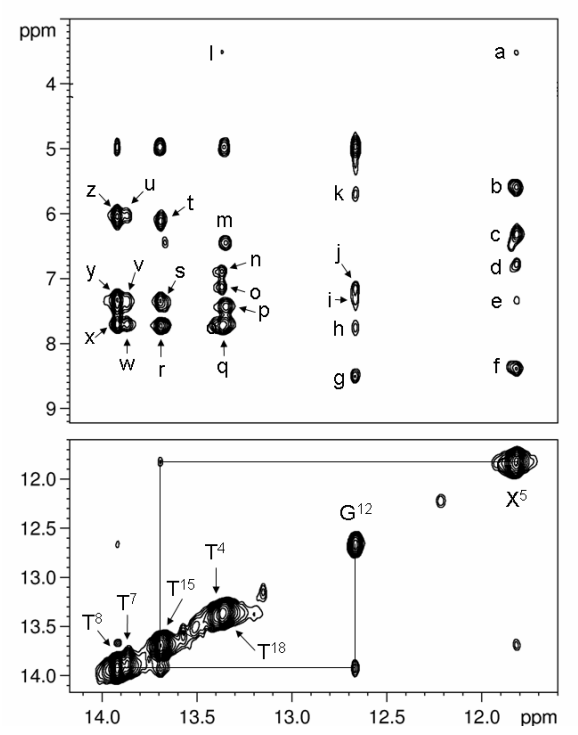


Figure 6-8: The upper panel is an expanded plot showing the sequential NOE connectivity for the amino protons to amino protons for the α-AFB_1-FAPY modified duplex. The cross peaks were assigned as (a) AFB_1 $OCH_3 \rightarrow X^5$ H3 (b) C^{16} H41$\rightarrow X^5$ H3 (c) X^5 H21$\rightarrow X^5$ H3 (d) X^5 H22$\rightarrow X^5$ H3 (e) A^6 H61$\rightarrow X^5$ H3 (f) C^{16} H42$\rightarrow X^5$ H3 (g) C^9 H42$\rightarrow G^{12}$ H1 (h) C^1 H42$\rightarrow G^{20}$ H1 (i) A^{13} H61$\rightarrow G^{12}$ H1 (j) C^9 H41$\rightarrow G^{12}$ H1 (k) C^9 H5$\rightarrow G^{12}$ H1 (l) AFB_1 $OCH_3 \rightarrow T^4$ H3 (m) A^3 H61$\rightarrow T^{18}$ H3 (n) A^{17} H62$\rightarrow T^4$ H3 (o) A^{17} H61$\rightarrow T^4$ H3 (p) A^3 H2$\rightarrow T^{18}$ H3 (q) A^{17} H2$\rightarrow T^4$ H3 (r) A^6 H2$\rightarrow T^{15}$ H3 (s) A^{14} H61$\rightarrow T^{15}$ H3 (t) A^6 H62$\rightarrow T^{15}$ H3 (u) A^{14} H62$\rightarrow T^7$ H3 (v) A^{14} H61$\rightarrow T^7$ H3 (w) A^{14} H2$\rightarrow T^7$ H3 (x) A^{13} H2$\rightarrow T^8$ H3 (y) A^{13} H61$\rightarrow T^8$ H3 (z) A^{13} H62$\rightarrow T^8$ H3. The lower panel is an expanded plot showing the sequential NOE connectivity for the imino protons. The data was collected at 500 MHz at 250 ms mixing time and a temperature of 5 °C.

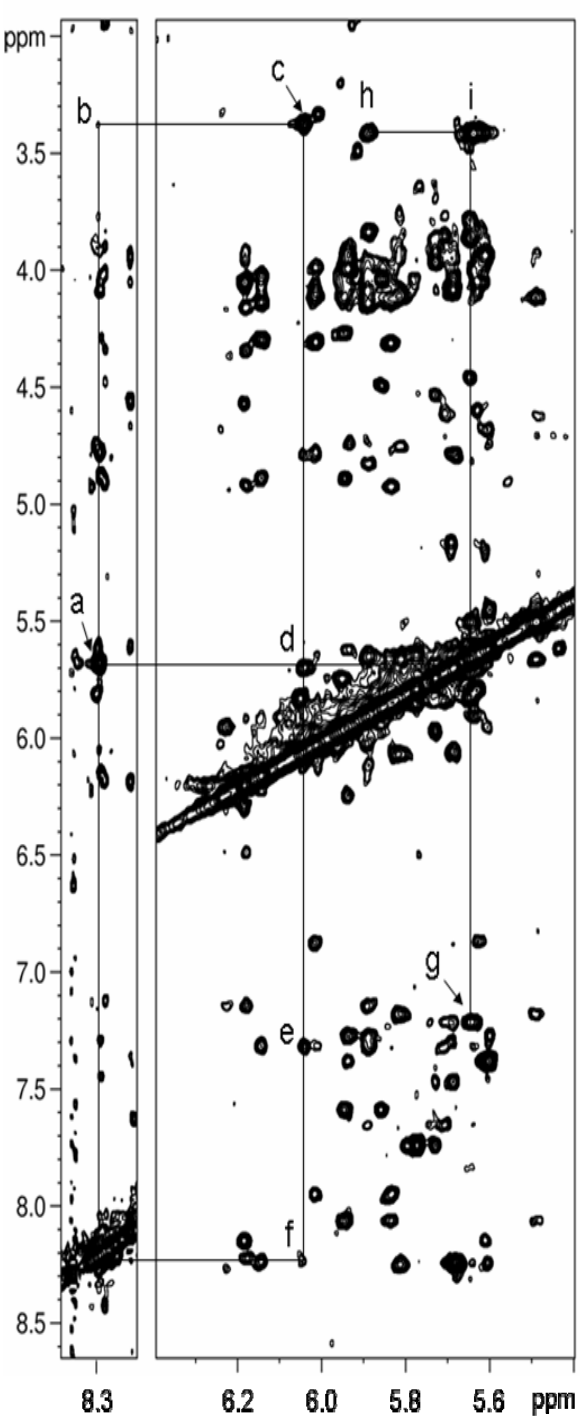


Figure 6-9: Select NOE assignments of the α-AFB_1-FAPY modified duplex. The cross peaks are assigned as (a) AFB_1 H8a →X^5 CHO (b) AFB_1 H9a →X^5 CHO (c) AFB_1 H9a→AFB_1 H6a (d) AFB_1 H6a→AFB_1 H8a (e) T^4 H6→AFB_1 H6a (f) X^5 CHO→AFB_1 H6a (g) A^{17} H2→AFB_1 H5 (h) AFB_1 OCH_3→T^4 H1' (i) AFB_1 OCH_3→ AFB_1 H5

strand. Cross peaks for H2β, H3α, and H3β could not be conclusively identified due to spectral overlap. Inter-residue NOEs between AFB_1 and the 3'-neighbor A^6•T^{15} base pair in dsDNA or A^4 in ssDNA were not observed.

NMR Spectroscopy of Anomeric Configuration

The anomeric configuration of AFB_1-FAPY modified oligonucleotides was established by NOEs within the deoxyribose of the modified residue. Assignment of H2' and H2" resonances was based on relative peak intensity to H3'. Configuration of H1' was subsequently determined by cross peak intensity to H2' and H2". The intensity of the X^5 H1′→X^5 H2′′ NOE was less than the X^5 H1′→X^5 H2′ NOE in dsDNA, which placed H1' in the α configuration (Figure 6-10, panel A). The same conclusion was reached for the tetramer (Figure 6-10, panel B).

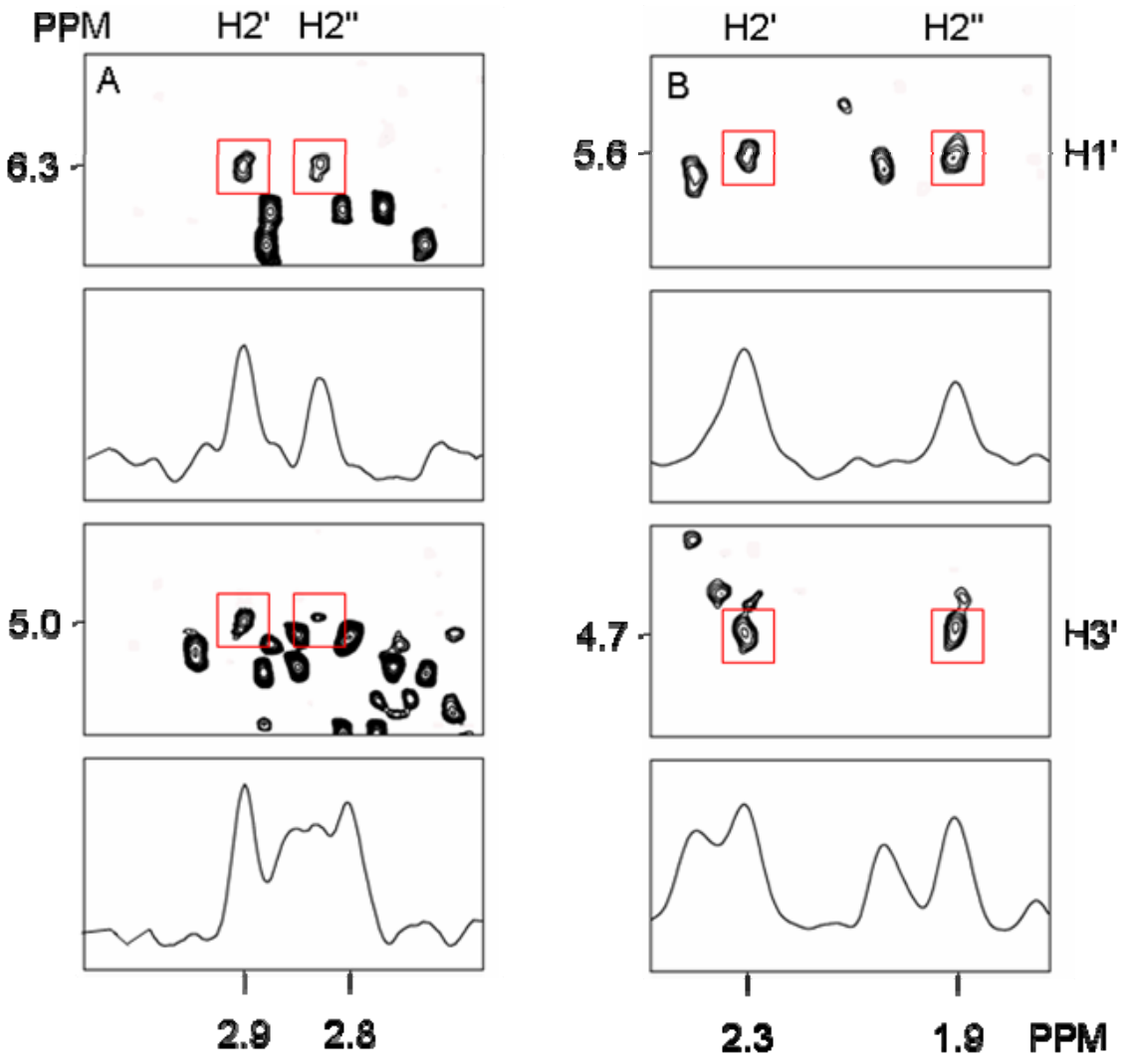


Figure 6-10: Analysis of NOE intensities indicate the deoxyribose of the AFB_1-FAPY was in the α orientation at C1'. This observation was true for 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' (Panel A) and 5'-CTXA-3' (Panel B).

NMR Spectroscopy of Formyl Proton Resonance (CHO)

A single formyl proton resonance was observed for α-AFB_1-FAPY in the tetramer and duplex samples (δ 8.3 ppm) in spectra acquired at 5 °C. 1D ^{1}H NMR experiments were conducted on the modified tetramer at variable temperatures (Figure 6-11). Analysis of downfield resonances at 30 °C revealed a sharp proton resonance at δ 7.5 ppm that produces a weak NOE to AFB_1 H8. This δ 7.5 ppm resonance was assigned as

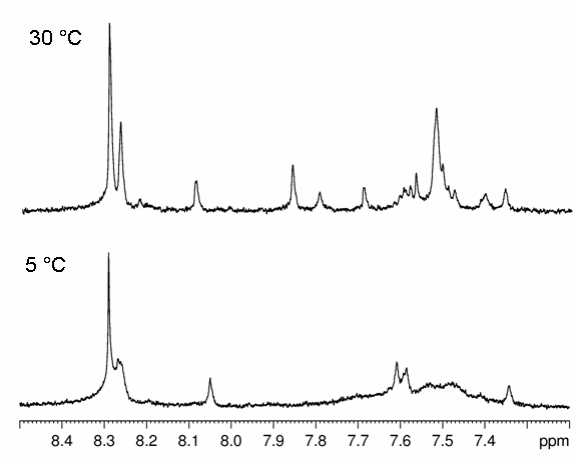


Figure 6-11: Variable temperature ^{1}H NMR of α-AFB$_1$-FAPY modified tetramer aromatic

the formyl proton of the *Z* geometrical isomer. No chemical exchange cross peak was observed between formyl resonances in 2D NOE experiments at 30 °C.

NMR Spectroscopy of ^{31}P

Previous reports indicate a significant deshielding of one of the phosphorus resonances of α deoxyadenosine in duplex DNA [28]. The phosphorus resonance of the α-AFB$_1$-FAPY 5' phosphate was not significantly disrupted relative to other phosphorus resonances in dsDNA. Perturbation relative to an unmodified duplex was inconclusive as a result of ambiguous ^{31}P resonance assignments in the control sample. However, the ^{31}P resonance associated with the phosphate group located between C^{16} and A^{17} was deshielded relative to other ^{31}P resonances in the duplex (Figure 6-12). Phosphorus spectra of α-AFB$_1$-FAPY in ssDNA showed no significant shielding or deshielding and were indicative or rapidly equilibrating conformers.

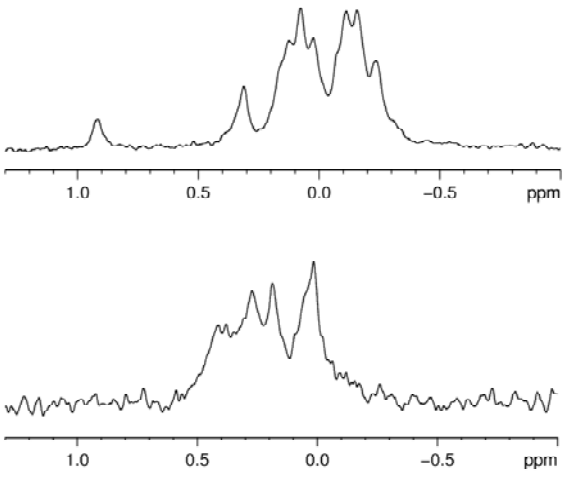


Figure 6-12: 1D ^{31}P NMR of α-AFB$_1$-FAPY in 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-3' (upper panel) and in 5'-$C^1T^2X^3A^4$-3' (lower panel). The phosphate resonance of C^{16} was deshielded to 0.9 ppm (upper panel).

Structural Refinement

From the NMR studies, 554 experimental restraints were derived that

could be used for refinement of the structure of α-AFB_1-FAPY modified dsDNA. In addition to the experimental restraints, 34 empirical distances based on hydrogen bonding geometries and 78 broad backbone torsion restraints were applied (Table 6-2). The proton-proton restraints consisted of 167 intra-residue and 155 inter-residue restraints. Vicinal (intra-residue) sugar ring proton distance restraints were discarded for refinement purposes; their corresponding NOE cross-peaks were included in CORMA calculations. There were 29 inter-residue restraints between α-AFB_1-FAPY and DNA. Table 6-2 summarizes the distribution of rMD restraints. A total of 82 restraints were used for refinement of α-AFB_1-FAPY modified ssDNA. In addition to the experimental restraints, 11 broad backbone torsion restraints encompassing standard A and B form values were applied to standard residues. Likewise, broad deoxyribose torsion angle restraints were applied that encompassed N and S ring pucker conformations for C^1, T^2, and A^4. The proton-proton restraints consisted of 52 intra-residue, and 4 inter-residue restraints defining the interaction of α-AFB_1-FAPY with DNA.

Table 6-2: Distribution of restraints applied to structural refinement.

	ds DNA	*ss DNA*
Assigned Resonances	233	39
Watson-Crick	34	0
Backbone Torsion	78	11
Ribose Torsion	120	15
Distance:	322	56
Inter-residue	155	4
Intra-residue	167	52
Total Restraints	554	82
Avg. Restraint per Residue	25	16

Stereo views of nine rMD refined solution structures for the α-AFB_1-FAPY modified duplex DNA are depicted in Figure 6-13. Individual structures were extracted from the final 100 ps of a 5 ns rMD trajectory. The root mean squared deviation (RMSD)

between the six core base pairs and AFB_1 of the resultant structures is 0.59 Å. The overall structure maintained Watson-Crick base pairing. The AFB_1 adduct was intercalated between X^5•C^{16} and T^4•A^{17}. The FAPY lesion is in a R_a configuration about the C5-N^5 bond; the formyl group has an *E* geometrical configuration. Stereo views of eight rMD refined solution structures for the α-AFB_1-FAPY modified single strand tetramer are depicted in Figure 6-14. Individual structures were extracted from the final 300 ps of a 5 ns rMD trajectory. The RMSD for the overall ensemble is 1.71 Å. Per residue RMSD are as follows: cytosine, 2.69 Å; thymine, 0.97 Å; FAPY, 0.24 Å; AFB_1, 0.43 Å; adenine, 0.43 Å. This is indicative of structural agreement at the lesion site while the terminal bases are structurally divergent. Similar to the dsDNA structure, the AFB_1 moiety was located between X^3 and T^2. The FAPY lesion is in a R_a configuration about the C5-N^5 bond; the formyl group is in an *E* geometrical configuration.

The 5 ns rMD trajectories of both ssDNA and dsDNA were analyzed for predicted hydrogen bonding motifs (Table 6-3). Occupancy was calculated from a cut-off distance of 3.5 Å and angle cut-off of 120 degrees. The amine proton A^6 H61 satisfied hydrogen bonding criteria with X^5 CHO 95% of the 5 ns trajectory in dsDNA. However, A^4 H61 and X^5 CHO were predicted to hydrogen bond 13% of the trajectory in ssDNA. Minimal hydrogen bonding with solvent was predicted for the various carbonyl oxygen atoms of AFB_1 in either ssDNA or dsDNA.

Table 6-3: Hydrogen bonding occupancy *

	ds DNA (%)	*ss DNA (%)*
X^n:O8 – A^{n+1}:H61	94.5	13.1
T^{15}:O4 – A^6:H62	99.9	n/a
AFB_1:O1-H_2O	8.6	2.5
AFB_1:O6A-H_2O	4.5	0.1
AFB_1:O7-H_2O	6.5	0.3
AFB_1:O11-H_2O	10.4	2.3

* Occupancy was calculated from 5 ns trajectories with a distance cutoff of 3.5 Å and angle cutoff of 120 degrees.

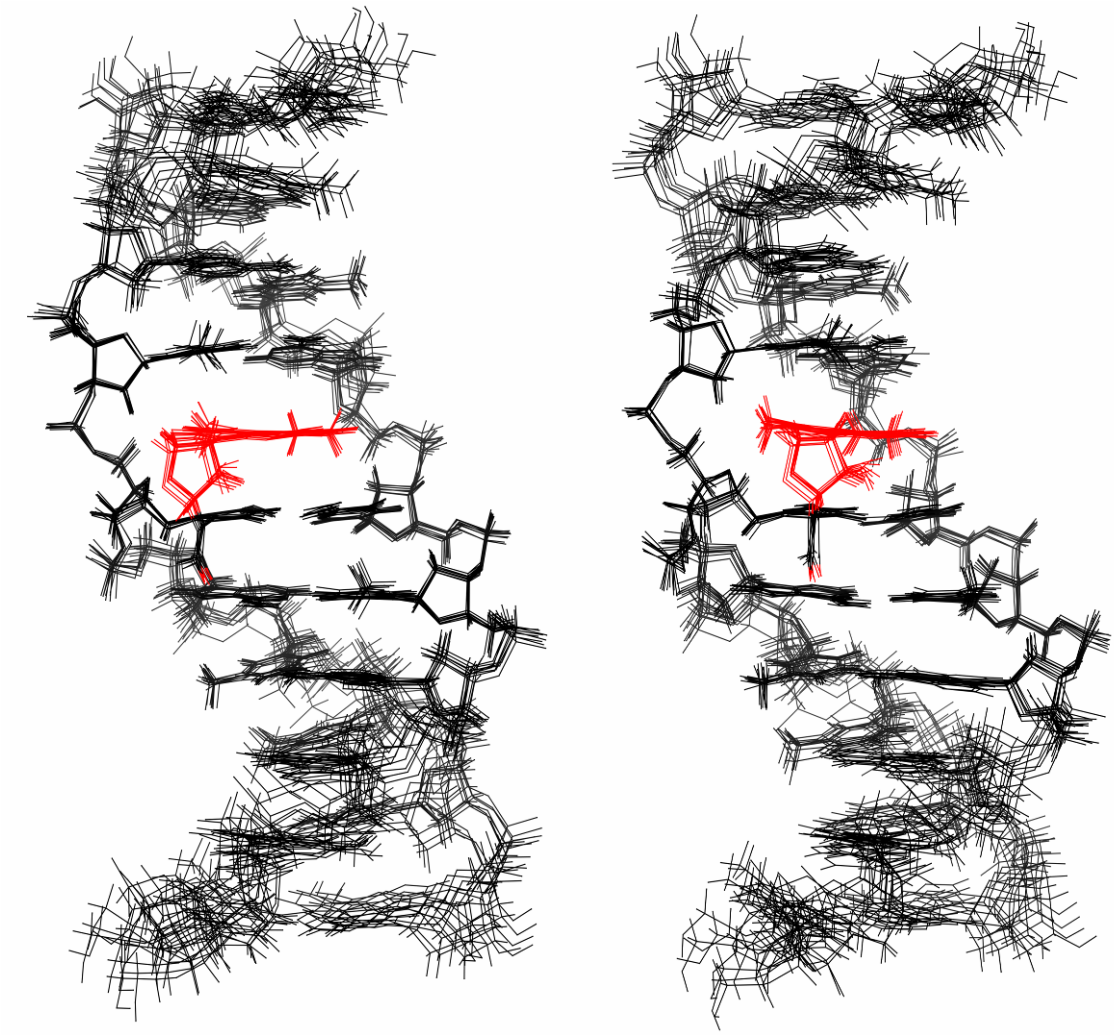


Figure 6-13: Stereo view of nine superimposed structures of 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' (X= α-AFB_1-FAPY) resulting from rMD calculations in explicit solvent. The AFB_1 moiety and the formyl oxygen are depicted in red. RMSD between the core six base pairs and AFB_1 of the resultant structures was 0.59 Å.

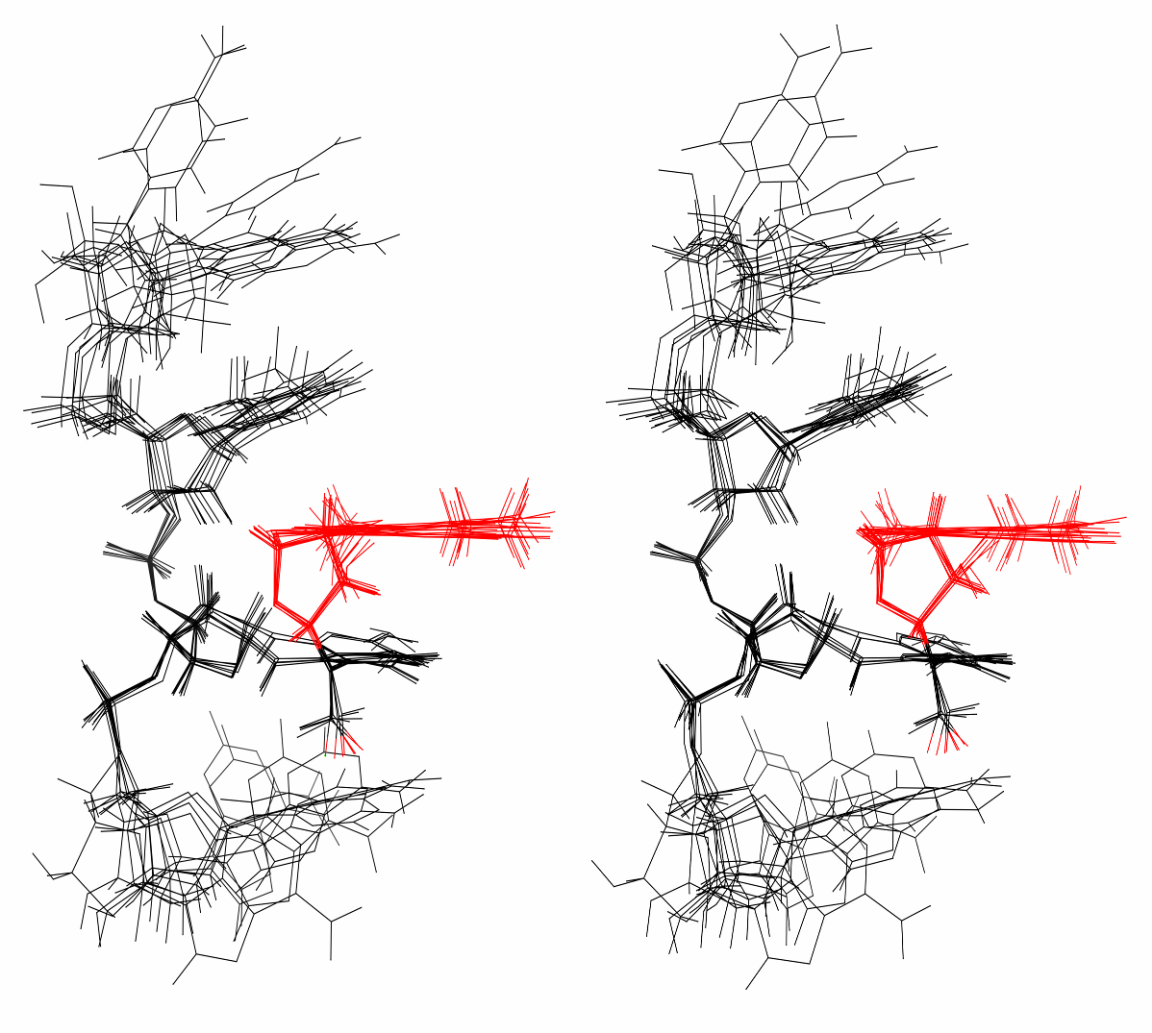


Figure 6-14: Stereo view of eight superimposed structures of 5'-CTXA-3' (X= α-AFB_1-FAPY) resulting from rMD calculations in explicit solvent. The AFB_1 moiety and the formyl oxygen are depicted in red. Pairwise RMSD was 1.71 Å. Per residue RMSD were as follows: cytosine, 2.69 Å; thymine, 0.97 Å; FAPY, 0.24 Å; AFB_1, 0.43 Å; adenine, 0.43 Å.

Accuracy of resultant structure ensembles was evaluated by calculation of the sixth root residuals between theoretical NOE intensities and experimental NMR data (Table 6-4). Both ssDNA and dsDNA refined ensembles were significantly improved as compared to their respective starting structures. Major improvement between the starting structures and the final refined structures was observed in both inter- and intra-residue NOEs. The final R_1^x value of the refined tetramer (ss-α-<rMD ensemble>) of 8.1 x 10^{-2} and 8.7 x 10^{-2} of the refined duplex structure, ds-α-<rMD ensemble>, suggested that the refined structures were in good agreement with the NOESY data. In addition, the sixth root residuals of the α-AFB_1-FAPY refined structures are in good agreement with the previously refined β-AFB_1-FAPY structure [115].

Table 6-4: Comparison of sixth root residual indexes, (R_1^X (x 10^2), for starting models and resulting rMD structures

	Intra R_1^X	*Inter* R_1^X	*Overall* R_1^X
ds-α-FAPY-i1	11.4	12.9	11.8
ds-α-FAPY-i2	11.2	10.9	11.1
ds-α-FAPY-i3	11.1	11.3	11.2
ds-α-FAPY-i4	11.2	10.9	11.1
ds-α-<rMD ensemble>	8.4	9.9	8.7
ss-α-FAPY-i1	8.5	18.8	9.2
ss-α-<rMD ensemble>	8.0	8.8	8.1
ds-β-FAPY-iA*	9.5	12.1	11.5
ds-β-FAPY-iB*	8.8	9.9	8.9
ds-b-FAPY-<rMDavg>*	8.5	9.5	8.7

* Results previously published Mao et al. [115].

Helicoidal Analysis

Helicoidal parameters were measured on an average of the ensemble structures of dsDNA and ssDNA in Figures 6-15 and 6-16 using CURVES [293]. Residues T^2, X^3, and A^4 in ssDNA were renumbered to T^4, X^5, and A^6 so that direct comparison may be made between ssDNA and dsDNA parameters. The helicoidal parameters of ssDNA C^1

were not compared with dsDNA but are reported in Appendix E. The comparison of base-axis parameters, x-displacement, y-displacement, tip, and inclination are graphically depicted in Figure 6-15 for α-AFB$_1$-FAPY modified ssDNA and dsDNA and the previously refined β-AFB$_1$-FAPY dsDNA structure [115]. The x-displacement of the X^5 was 1 Å relative to T^4 and A^6 in dsDNA α-AFB$_1$-FAPY; the x-displacement was mirrored in ssDNA α-AFB$_1$-FAPY with the exception that the ssDNA bases were all displaced -2 Å relative to dsDNA. Y-displacement of α-AFB$_1$-FAPY and β-AFB$_1$-FAPY were comparable in dsDNA; however, α-AFB$_1$-FAPY in ssDNA was displaced -4 Å relative to the duplexes. The comparison of inter-base parameters, shift, slide, rise, roll, twist, and tilt are graphically depicted in Figure 6-16. The inter-base rise was 1 Å less in ssDNA when compared to both duplex anomers. The duplexes were comparable in shift,

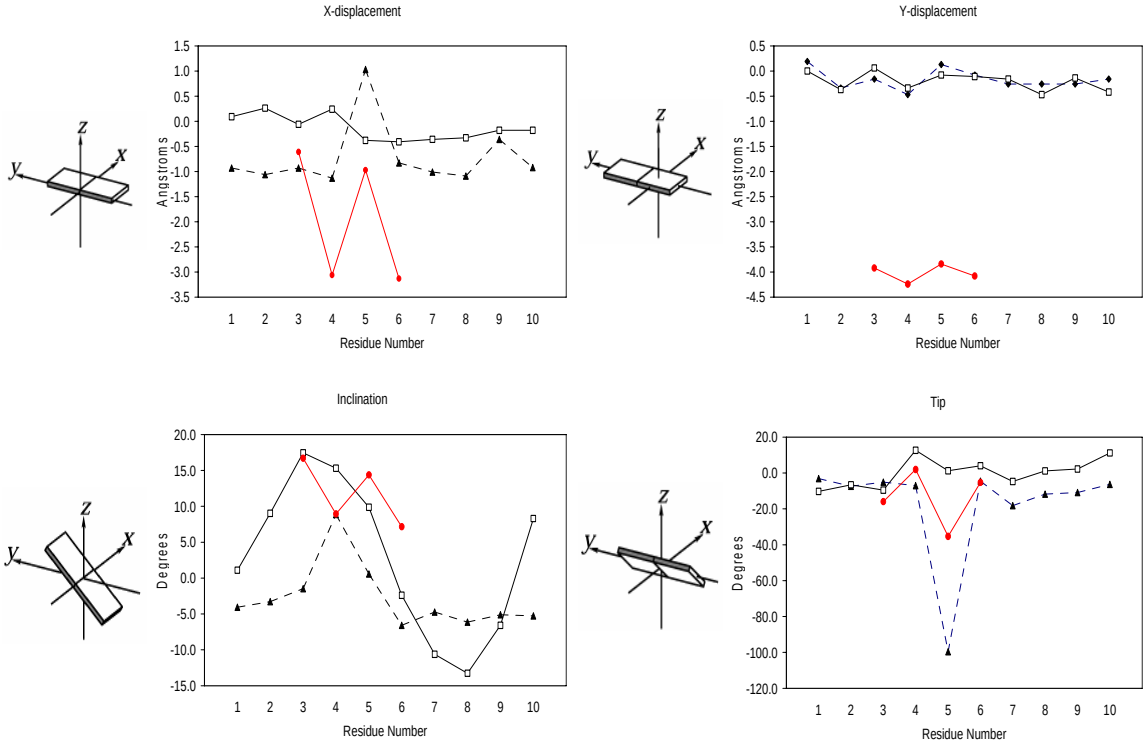


Figure 6-15: Global Base-axis helicoidal parameters. The rMD refined β-AFB$_1$-FAPY modified 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' structure (□) was compared with α-AFB$_1$-FAPY modified 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' (▲) and α-AFB$_1$-FAPY modified 5'-CTXA-3' (●).

slide, roll, and twist; however, the α-AFB$_1$-FAPY modified tetramer was divergent.

Helicoidal comparisons of α and β AFB$_1$-FAPY anomers in dsDNA reveal distinct alternations in backbone geometry (Appendix E). The α, γ, and ε torsion angles were comparable (± 30°); however, β and ζ differed >100°. Considering lesion deoxyribose torsions were loosely restrained, differences in ring pucker are not surprising and probably within experimental error. Lesion deoxyribose relative geometries were

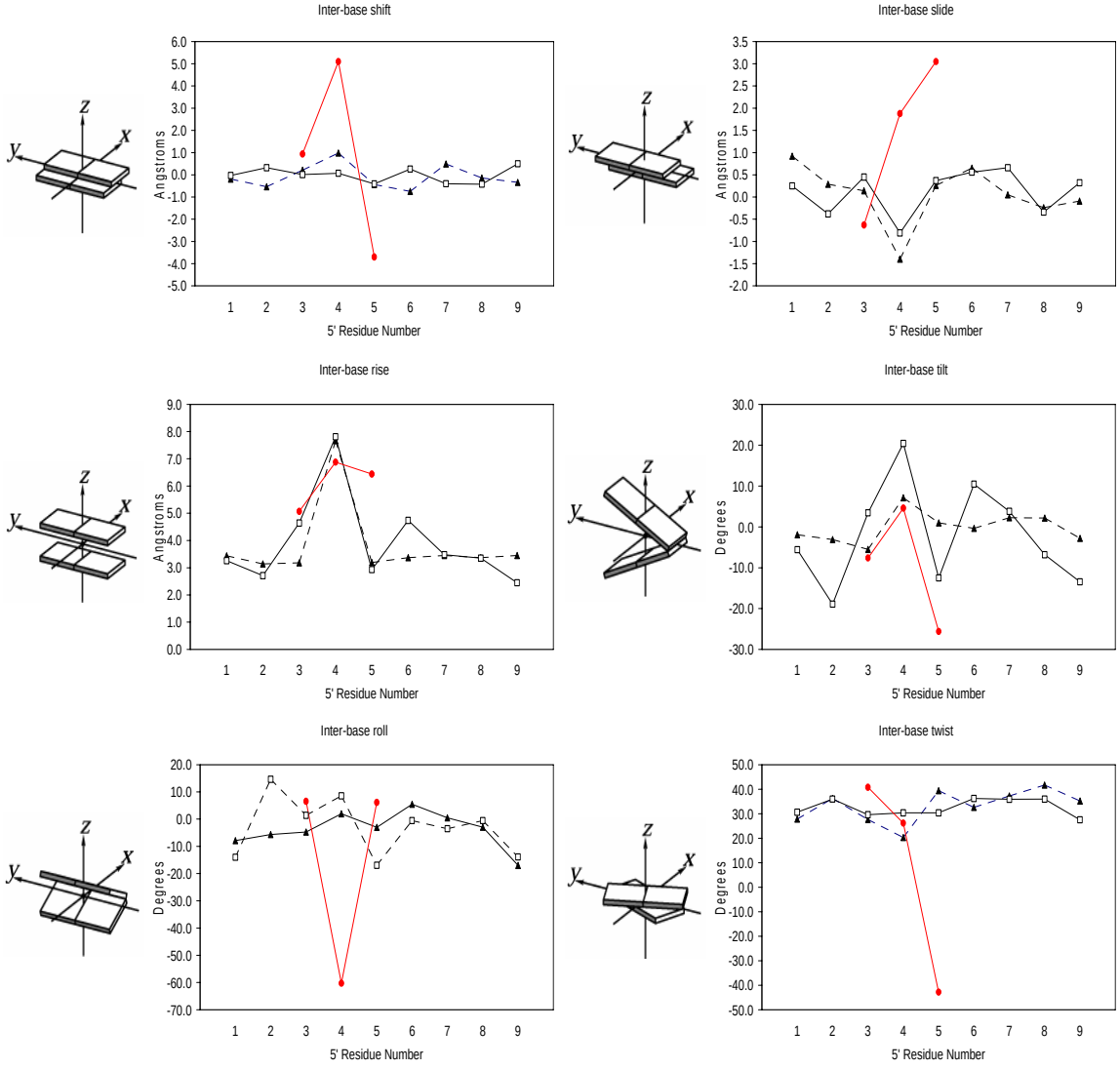


Figure 6-16: Global Inter-base helicoidal parameters. The rMD refined β-AFB$_1$-FAPY modified 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' structure (□) was compared with α-AFB$_1$-FAPY modified 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' (▲) and α-AFB$_1$-FAPY modified 5'-CTXA-3' (●).

significantly divergent (Figure 6-17). In the β structure, the sugar is approximately orthogonal to the FAPY base plane not unlike canonical DNA bases. In the dsDNA α structure, the deoxyribose is nearly parallel to the FAPY base and displaced toward the minor groove by ~ 2 Å.

Discussion

Duplex Stability

The thermal melting of AFB_1-FAPY modified duplexes offer key insights into altered stability of the helix. Figure 6-2, panel B indicates that a purified β-AFB_1-FAPY sample produces predominantly a single transition at 50 °C during the heating stage when the hybridization transitions from dsDNA to ssDNA. The ability of β-AFB_1-FAPY to stabilize dsDNA is consistent with previous reports [115, 116]. At 80 °C, a single strand environment shifts the equilibrium to favor α-AFB_1-FAPY. The complex nature of the sample at 80 °C is evidenced by the biphasic transition during the cooling phase. The transitions are observed at 50 °C and 23 °C representing β-AFB_1-FAPY and α-AFB_1-FAPY respectively. Likewise, the α-AFB_1-FAPY modified duplex produced biphasic transitions at 50 °C and 23 °C during the course of each

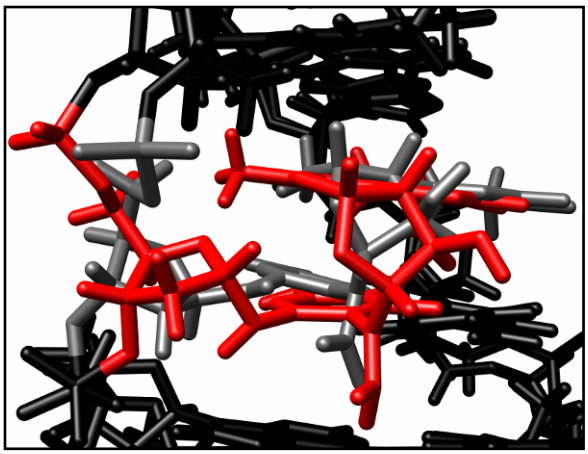


Figure 6-17: Superposition of α-AFB_1-FAPY lesion (red) with β-AFB_1-FAPY (grey) in dsDNA.

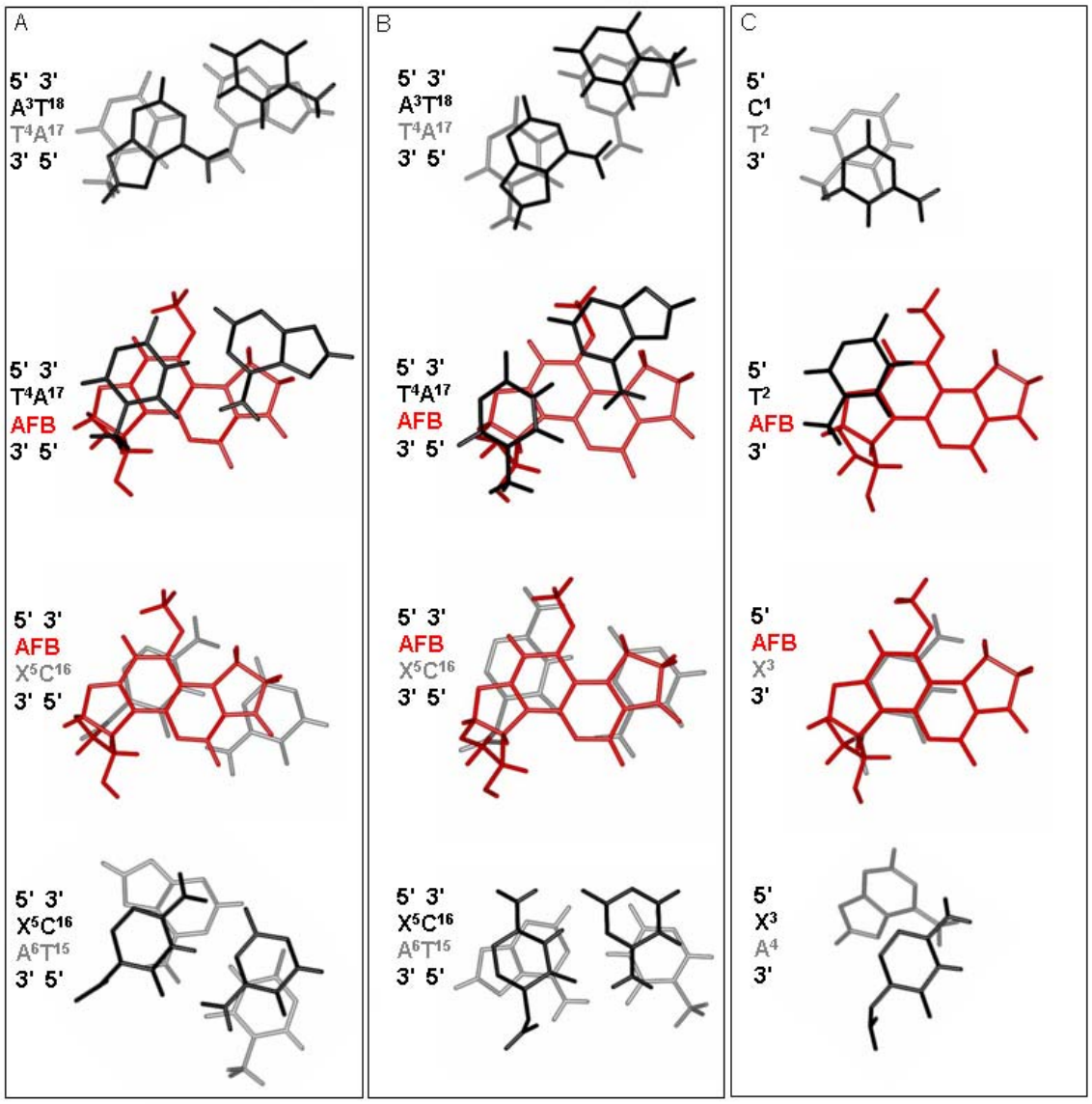


Figure 6-18: Stacking interactions of AFB_1-FAPY modified oligonucleotides. The neighbor stacking patterns of β-AFB_1-FAPY modified 5'-CTATXATTCA-3'•5'-TGAATCATAG-3' (panel A) haven been reported elsewhere (Mao et al. Biochemistry 37, 4374). Stacking patterns of α-AFB_1-FAPY in the same duplex sequence (panel B) were compared to α-AFB_1-FAPY placed in a tetramer (5'-CTXA-3') comprised of the same local sequence context (panel C).

heating and cooling stage (Figure 6-2, panel C). This is explained by transition of the less favored α-AFB_1-FAPY isomer at 23 °C to the more stable β-AFB_1-FAPY producing a second transition at 50 °C. The cooling stage of α-AFB_1-FAPY is comparable to that of β-AFB_1-FAPY. Comparison of T_m values with native DNA indicates that β-AFB_1-

FAPY stabilizes the duplex by 14 °C whereas α-AFB_1-FAPY destabilizes the duplex by 13 °C. Giri et al. have demonstrated the ability of β-AFB_1-FAPY to thermally stabilize, not just the nearest neighbor base-pairs, but also n+1, n+2, n-1, and n-2 base-pairs [116]. The α-AFB_1-FAPY affects the chemical environment of n+1, n+2, n-1, and n-2 base-pairs (Figure 6-6), suggesting that it may destabilize to the same extent that β-AFB_1-FAPY stabilizes. Attempts to repeat NMR melting studies with α-AFB_1-FAPY were complicated by rapid conversion to β-AFB_1-FAPY upon heating. Two major factors are involved in the stabilization of DNA: hydrogen bonding and stacking interactions [314-317]. Computational studies predicted that stacking forces are more significant than hydrogen bonding in duplex stabalizaiton [318-322]. This conclusion was recently substantiated experimentally [323, 324]. DNA intercalating agents often stabilize the helical structure by improved stacking interactions, but may destabilize a structure with poor stacking interactions as is the case with benzo[a]pyrene [325]. In the β anomer structure, optimal stacking interactions exist between the cyclopentone moiety of AFB_1 and C^{16} of the complementary strand (Figure 6-18). In the α anomer, minimal stacking of AFB_1 and C^{16} was observed. Additionally, stacking between X^5 and A^6 was lost in the α anomer relative to the β structure. This is expected to contribute to the improved thermal stability of the β anomer and reduced stability of the α anomer suggesting that AFB_1 acts similar to an inter-strand crosslink.

Lesion Structure

The forces driving anomeric equilibria in oligonucleotides are poorly understood. AFB_1-FAPY lesions prefer the α anomeric configuration by 2:1 in single stranded DNA while the β anomer is almost exclusively observed in double stranded DNA [31, 110,

115]. In AFB_1-FAPY nucleosides, the R_a atropisomer is favored over S_a by a factor of 7:1; the *Z* geometrical isomer was favored over *E* 3:1 [31]. In previous studies of β-AFB_1-FAPY [115] and now α-AFB_1-FAPY, only the R_a atropisomer is observed suggesting that S_a is only possible in the less conformationally defined nucleoside context. Equilibrium between R_a and S_a atropisomers in oligonucleotides would require the AFB_1 moiety to unstack from the 5' side of the FAPY base and re-intercalate on the 3' side. Unstacking and restacking of a large aromatic molecule like AFB_1 may represent a prohibitively large conformational energy barrier.

Although the *Z* geometrical isomer of AFB_1-FAPY is preferred in nucleoside 7:1 over *E*, *E* was exclusively observed in the dsDNA structures of the β anomer [115] and now the α anomer of AFB_1-FAPY. Stabilization of a single *E* geometrical isomer in both dsDNA structures is attributed to an intra-strand hydrogen bond between the FAPY formyl oxygen and the H61 proton of the 3'-adenine base [115]. This conclusion is supported by rMD trajectory analysis that predicts 95% hydrogen bond occupancy of X^5 H8 and A^6 H61 in dsDNA (Table 6-4). However, rMD trajectories predict a reduced occupancy of 13% for X^3 H8 and A^4 H61 in a single strand tetramer resulting from increased conformational mobility of the terminal A^4 residue. In AFB_1-FAPY nucleoside, formyl rotation caused doubling of ^{1}H NMR formyl resonances and 2D NOE chemical exchange cross peaks. A single formyl proton resonance was observed for α-AFB_1-FAPY in the tetramer δ 8.3 ppm at 5 °C (Figure 6-11). NMR experiments conducted at 30 °C reveal a sharp CHO resonance at δ 7.6 ppm that produces a weak NOE to AFB_1 H8. Chemical exchange cross peaks were not observed between formyl resonances in 2D NOE spectra of the tetramer. It is concluded that at low temperatures

the *E* formyl rotamer is favored in ssDNA. At the relatively high temperature of 30 °C, a temperature comparable to HPLC separation and rMD calculations, the *E* is preferred, but *Z* is also observed to a lesser extent. The Z isomer is postulated to represent the chromatographically observable species preceding the primary α anomer peak (Figure 6-1).

UV traces of the HPLC peaks agree with this conclusion. The UV absorbance of nucleic acids from 200 to 300 nm arise exclusively from electronic transitions of the planar purine and pyrimidine bases [326]. The electronic structure associated with the AFB_1 moiety (320-380 nm) is identical in the primary α anomer peak and the preceding shoulder (Figure 6-4, panel B); spectral differences are only found in the far UV region. Differences in the AFB_1 π–π* transitions are observed in the α and β anomers (Figure 6-4, panel A). Similarly, significant AFB_1 π–π* transitions have been observed in ECD spectra of R_a and S_a AFB_1 dibutyrates [31] [supplemental]. The transition observed in Figure 6-4, panel B cannot be directly correlated to a formyl flip; however, given the alternatives, this data supports structures that do not significantly perturb the AFB_1 moiety.

The inversion of geometrical isomer preference between nucleoside and oligonucleotides supports formyl hydrogen bonding. It is unknown if a 3'-cytosine would stabilize *E* or if 3'-guanine would stabilize *Z*. It is plausible that inter-residue formyl hydrogen bonding is a potential source of sequence related effects associated with FAPY lesions in DNA.

Helicoidal analyses of both anomers in duplex DNA as compared to the α anomer in ssDNA indicate structural differences. In dsDNA the anomers are similar in y-

displacement, inter-base shift, slide, rise, and twist (Figures 6-15 & 6-16), yet dsDNA structures differ in x-displacement and twist. The possibility was considered that the structure of α-AFB_1-FAPY in dsDNA may differ from the α-AFB_1-FAPY structure in a single strand environment, at least in some parameters. The α anomers have similar structural patterns yet maintain distinct differences. For example, both dsDNA and ssDNA α anomers produce an x-displacement of 2 Å relative to the 3' and 5' neighboring bases. However, the 3' and 5' neighboring base pairs of the single strand structure are displaced by 2 Å in the x direction relative to the corresponding bases in the duplex structures (Figure 6-15).

The key to the α anomer preference in ssDNA may lie in improved π–π orbital overlap of AFB_1 with neighboring aromatic bases. Stacking interactions of the AFB_1 moiety of the α anomers in ssDNA and dsDNA were compared to those of the β anomer in dsDNA (Figure 6-17). In the β anomer structure, optimal stacking interactions exist between the cyclopentone moiety of AFB_1 and C^{16} of the complementary strand and between the AFB_1 tetrahydrofuran and T^4. The stacking of the α-AFB_1-FAPY lesion is similar in ssDNA and dsDNA in that stacking is optimal between the benzene ring of AFB_1 and the intra-strand FAPY ring. Stacking of α-AFB_1-FAPY lesions differ in that the 5'-thymine is significantly displaced toward the minor groove in the tetramer. This is predicted to relax strain in the thymine-FAPY backbone linkage resulting from AFB_1 intercalation as is evidenced by a 1 Å reduction of base-base rise in the tetramer. The T^2 residue in ssDNA is allowed to explore this conformation because it is not stabilized by Watson-Crick hydrogen bonding with a complementary base.

Altered stacking of α and β AFB_1-FAPY anomers is supported by UV and ECD. Any electronic transition between 310-400 nm is expected to result form AFB_1 as DNA bases do not produce transitions in this range. A decrease in UV absorbance of the α-AFB_1-FAPY at 380 nm relative to the β anomer may be explained by altered AFB_1 stacking producing anomerically specific π–π* electronic transitions (Figure 6-4, panel A). Intercalators induce weak negative ECD transitions when polarized along the long axis of the intercalation site (perpendicular to the pseudo-dyad axis) and positive transitions when polarized along an axis perpendicular to long axis of the intercalation site (parallel to the pseudo-dyad axis) [327-329]. Intermediate values scale as the square of the cosine of the angle between the transition moment and the pseudo-dyad axis [327-329]. ECD spectra indicate a more negative AFB_1 induced molar ellipticity (310-400 nm) for the α anomer relative to the β anomer; this suggests that the AFB_1 moiety of the β anomer is less aligned with the long axis of the intercalation site than the α anomer. Of interest is the fact that relative magnitude of molar ellipticities is similar from 300 to 400 nm regardless of ssDNA or dsDNA. This would suggest that the AFB_1 orientation for α-AFB_1-FAPY is comparable in ssDNA and dsDNA. More pronounced ssDNA versus dsDNA ECD differences are attributed to winding angle (275 nm) and altered purine-pyrimidine stacking (210 – 250 nm). Differences in dsDNA ECD spectra are attributed to pre-melting transitions (260 and 230 nm) resulting from the reduced thermal stability of the α anomer in duplex [330].

Biochemical Implications

Structural analysis of α-AFB_1-FAPY in DNA is difficult due to instability of the anomeric configuration. NMR analysis of the α anomer in single strand DNA is problematic because the mobility of the chain causes only a few inter-residue NOEs to be observable. NMR analysis of the α anomer in ssDNA and dsDNA requires expeditious sample preparation and data acquisition. The resultant structure of α-AFB_1-FAPY in dsDNA represents a thermodynamically unfavorable isomeric configuration.

Toxicity of DNA containing α anomeric adenosine has been correlated to changes to the global duplex conformation, resulting from a kink of the helical axis [28]. Helicoidal comparisons of α and β AFB_1-FAPY anomers reveal distinct differences in backbone geometry, particularly in regard to β and ζ (Δ>100°). However, the most striking difference resides in the deoxyribose orientation of the lesion (Figure 6-18). In the β structure, the sugar is approximately orthogonal to the FAPY base plane, comparable to canonical DNA bases. In the α structure, the ribose is almost parallel to the FAPY base and displaced toward the minor groove. Disrupted backbone alignment of the α-AFB_1-FAPY adduct is expected to significantly contribute to the toxicity of this lesion.

The influence of varying DNA sequences on AFB_1 reactivity with DNA has been studied extensively [118-128]. Sequence-dependent repair of AFB_1-FAPY DNA lesions has received attention only recently; Oleykowski et al. have demonstrated the rate of incision with UvrABC endonuclease on AFB_1-FAPY can vary as much as 15-fold depending on sequence with 3'-adenine being the most resistant [189]. Similarly, repair of Me-dGuo-FAPY is sequence dependent [57, 68, 69]. Intra-strand formyl hydrogen

bonding is a sequence dependent structural effect that may influence repair of all FAPY lesions in DNA.

Measurement of AFB_1-FAPY anomer interconversion rates suggests that the event occurs on the order of minutes to hours at physiological pH. DNA replication occurs on the order of seconds per extension step. The highly mutagenic β-AFB_1-FAPY anomer is expected to be prevalent in genomic DNA prior to replication suggesting that replication enzymes would mostly encounter the β anomer. It is uncertain to what degree, if any, enzyme binding may catalyze interconversion. If base-base stacking interactions dictate preferred AFB_1-FAPY anomers, polymerase binding pockets may induce unexpected effects. Detailed structural analysis of AFB_1-FAPY anomers complexed with replication enzymes is necessary to address this possibility.

Although α-AFB_1-FAPY is not expected to exist at high levels in genomic DNA, it may be present in RNA. Site-specific mutagenesis experiments have addressed the effect of AFB_1 on replication; the role of AFB_1 anomers in translation should be thoroughly investigated. AFB_1 is known to disrupt protein synthesis [331-335]. Recently it was demonstrated that transcription of a *supF* gene fragment treated with aflatoxin B_1-8,9-epoxide was inhibited [336]. With regard to translation and RNA, both anomers of AFB_1-FAPY must be considered.

CHAPTER VII

PHYSICAL COMPARISON OF *CIS*-5*R*,6*S*-THYMINE GLYCOL LESIONS IN DUPLEX DNA OPPOSITE ADENOSINE AND GUANOSINE

Introduction

This chapter addresses the physical analysis of thymine glycol (*cis*-5*R*) base paired opposite adenosine and guanosine. Typically, oxidation of thymidine produces Tg base paired with adenine (Tg•A). However, Tg opposite guanine (Tg•G) may occur when a 5-methyl cytosine is oxidized to 5-methycytosine glycol that undergoes facile hydrolytic deamination to Tg [222, 223] as discussed in Chapter I. Physical characterization of the Tg•A and Tg•G interactions were investigated with NMR NOESY experiments as previously discussed in Chapter V. This chapter discusses the comparison of Tg•A and Tg•G measurements made by UV melting analysis, electronic circular dichroism (ECD), transverse relaxation (T_2), and longitudinal relaxation (T_1).

Results

Thermodynamics

The 5′-GTGCGTgGTTTGT-3′•5′-ACAAACACGCAC-3′ (Tg•A) and 5′-GTGCGTgGTTTGT-3′•5′-ACAAACGCGCAC-3′ (Tg•G) duplexes were analyzed by UV melting and compared to the unmodified 5′-GTGCGTGTTTGT-3′•5′-ACAAACACGCAC-3′ (T•A) duplex. All samples displayed a hyperchromic shift in UV

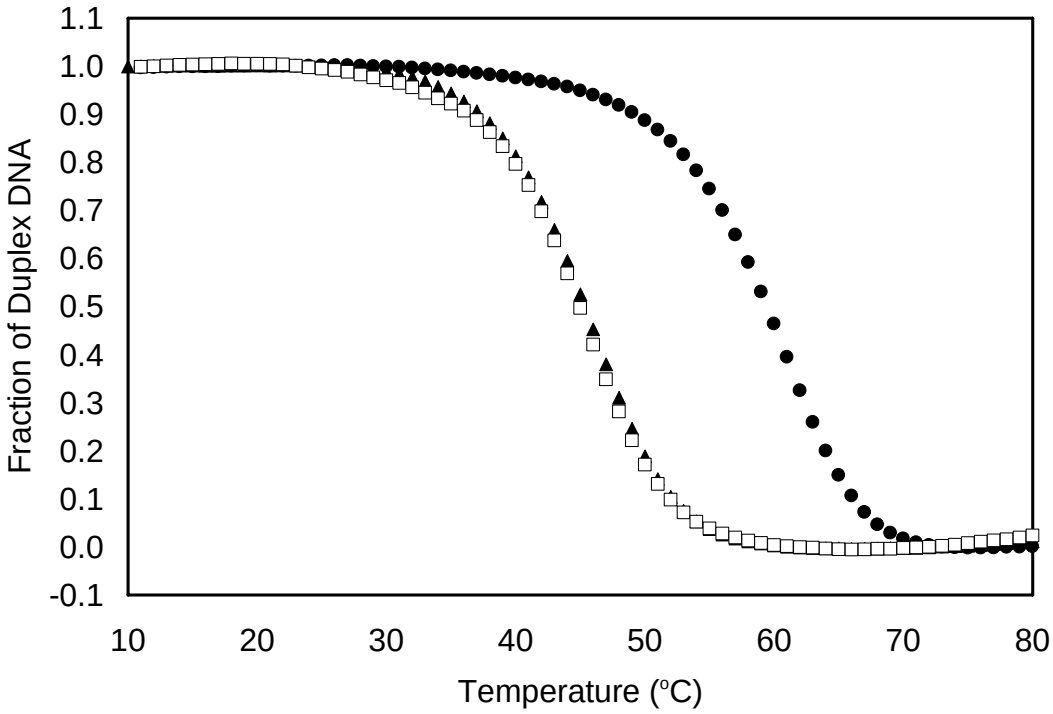


Figure 7-1: Hybridization plots (alpha curves) for the unmodified T•A sample (●), Tg•A (▲), and Tg•G (□).

absorption as temperature was increased from 5 to 80 °C. Average hybridization plots, α curves, are shown in Figure 7-1 [246]. The T_m values calculated for a concentration of 1.0 x 10^{-4} M are reported in Table 7-1. Both the Tg^6•A^{19} and Tg^6•G^{19} duplexes were destabilized 13 °C relative to the unmodified duplex, T^6•A^{19}.

Table 7-1: Thermodynamic parameters of DNA duplexes

Duplex	T_m* *(°C)*	*ΔH° (kcal/mol)*	*ΔS° (kcal/mol °K)*	$\Delta G°_{25°C}$ *(kcal/mol)*	$\Delta G°_{37°C}$ (kcal/mol)
T•A †	71.6	-87.3±2	-0.23±0.01	-18.1±0.2	-15.2±0.2
Tg•A †	58.3	-83.8±1	-0.23±0.01	-14.7±0.1	-11.9±0.1
Tg•G †	58.3	-87.5±5	-0.24±0.02	-15.0±0.3	-12.1±0.2
T•A ‡	73.4	-90.7±2	-0.24±0.02	-18.8±0.2	-15.9±0.2
Tg•A ‡	60.5	-81.1±2	-0.22±0.04	-14.9±0.4	-12.2±0.4
Tg•G ‡	60.0	-86.4±1	-0.24±0.02	-15.3±0.2	-12.5±0.2

Reported errors are standard deviations
*Calculated for 1.0 x 10^{-4} M DNA concentration
†Determined from individual melting curves
‡Determined from (T_m^{-1}) vs. ln(C_t/4) plots

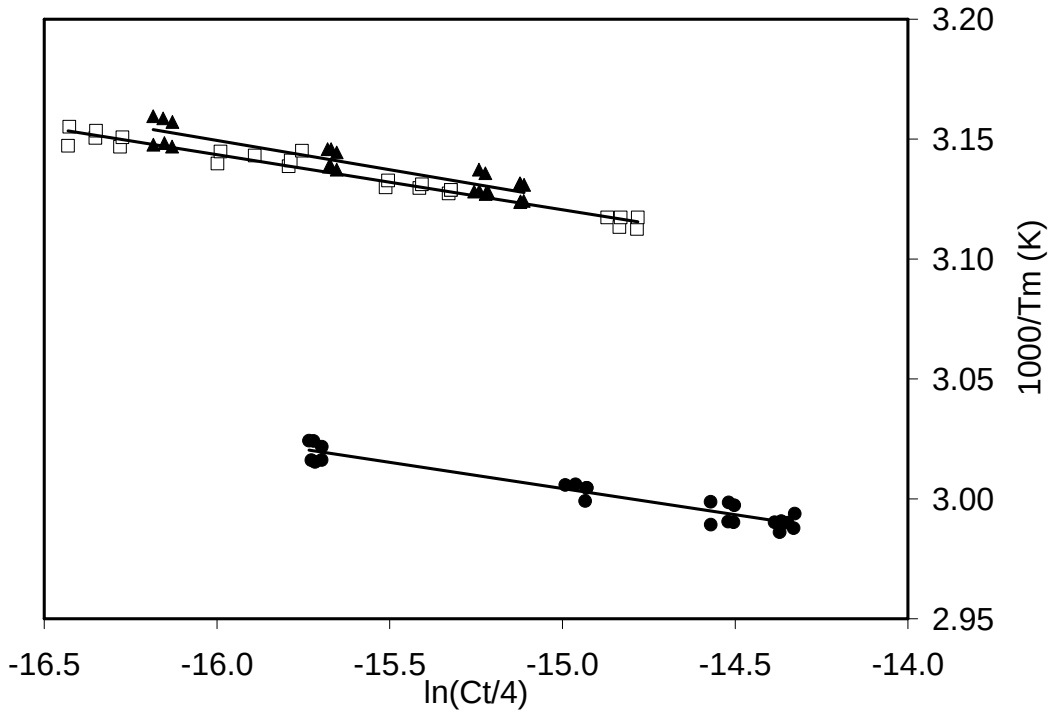


Figure 7-2: Plots of T_m^{-1} vs. $\ln(C_t/4)$ for the unmodified T•A sample (●), Tg•A (▲), and Tg•G (□). (C = concentration)

Thermodynamic values (ΔHº and ΔSº) were obtained by two methods, analysis of individual curves by the Meltwin application (v. 3.5, McDowell) and by linear regression analysis of van't Hoff plots (SigmaPlot v. 9.0). In van't Hoff plots, $(T_M)^{-1}$ vs. ln(C/4) produce a straight line where the slope equals R/ΔHº and the intercept equals ΔSº/ΔHº by Equation 2-1 (R = gas constant; 1.9872 cal K^{-1} mol^{-1}) (Figure 8-2). The Gibbs free energy with respect to duplex formation (ΔGº) was determined at 25 ºC and 37 ºC (Table 7-1). Error values are represented as standard deviations. Tg•A and Tg•G had ΔΔG values of 3 kcal/mol relative to T•A. Free energy values were negative, consistent with spontaneous formation of the duplexes at temperatures below the T_m of the samples. Thermodynamic values derived from van't Hoff and individual curve analysis were statistically indistinguishable (< 10% difference).

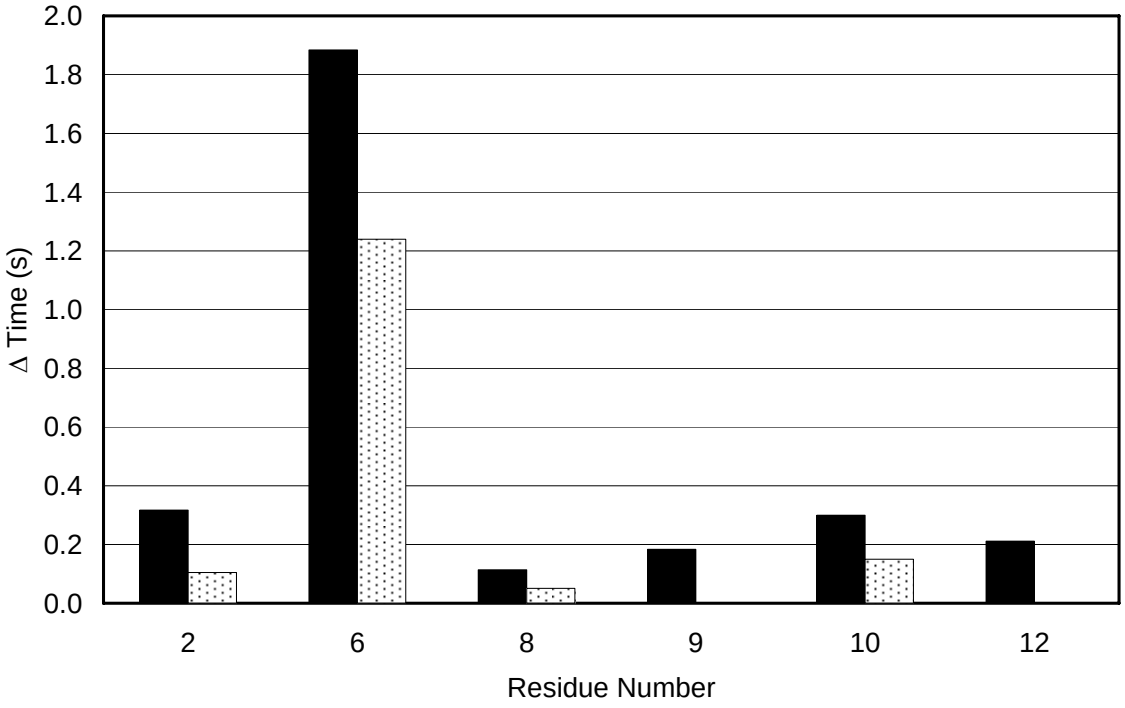


Figure 7-3: T^1 relaxation of thymine and thymine glycol CH_3 of Tg•A (■) and T•G (□) relative to unmodified T•A oligonucleotides.

T_1 Measurements

Unambiguous peak integration and intensity analysis of data collected in a pseudo 2D array was possible for sharp methyl signals. T_1 measurements on thymine and Tg CH_3 were measured (e.g. T^2, T^8, T^9, T^{10}, & T^{12}) in Tg•A and Tg•G duplexes. The difference in T_1 values relative to the duplex containing the unmodified T•A base pair were plotted (Figure 7-3). The T_1 relaxation of Tg CH_3 (Tg•A) was 1.9 s shorter than the corresponding T^6 CH_3 group of T•A. Tg CH_3 (Tg•A) relaxed an average of 1.4 s shorter than T^X CH_3 in Tg•A. T_1 relaxation of Tg CH_3 (Tg•G) was 1.25 s shorter relative to the control and relaxed an average of 1.0 s shorter than other T^X CH_3.

T_2 Measurements

T_2 measurements on thymine and Tg methyl groups (e.g. T^2, T^8, T^9, T^{10}, & T^{12}) in Tg•A and Tg•G duplexes were compared to those of T•A. The difference in T_2 values relative to the unmodified duplex containing the T•A base pair were plotted (Figure 7-4). The T_2 relaxation of Tg CH_3 (Tg•A) was 70 ms shorter than the corresponding T^6 CH_3 group of T•A. The T_2 of Tg CH_3 (Tg•A) relaxed approximately 60 ms shorter than T^X CH_3 in Tg•A. T_2 relaxation of Tg CH_3 (Tg•G) was 80 ms faster relative to the control and relaxed an average of 60 ms faster than T^X CH_3. Analysis of H1', H3', and H4' protons was not possible due to the inability to unambiguously resolve individual resonances during T_2 buildup periods.

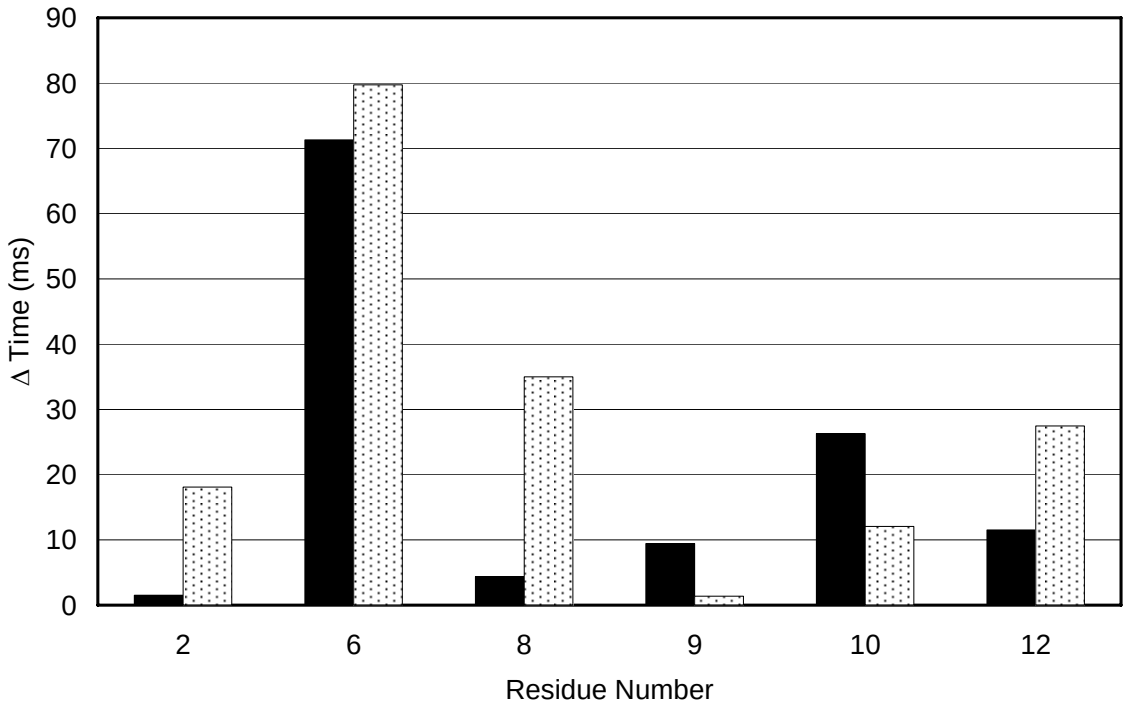


Figure 7-4: T^2 relaxation of thymine and thymine glycol CH_3 of Tg•A (■) and T•G (□) relative to unmodified T•A oligonucleotides.

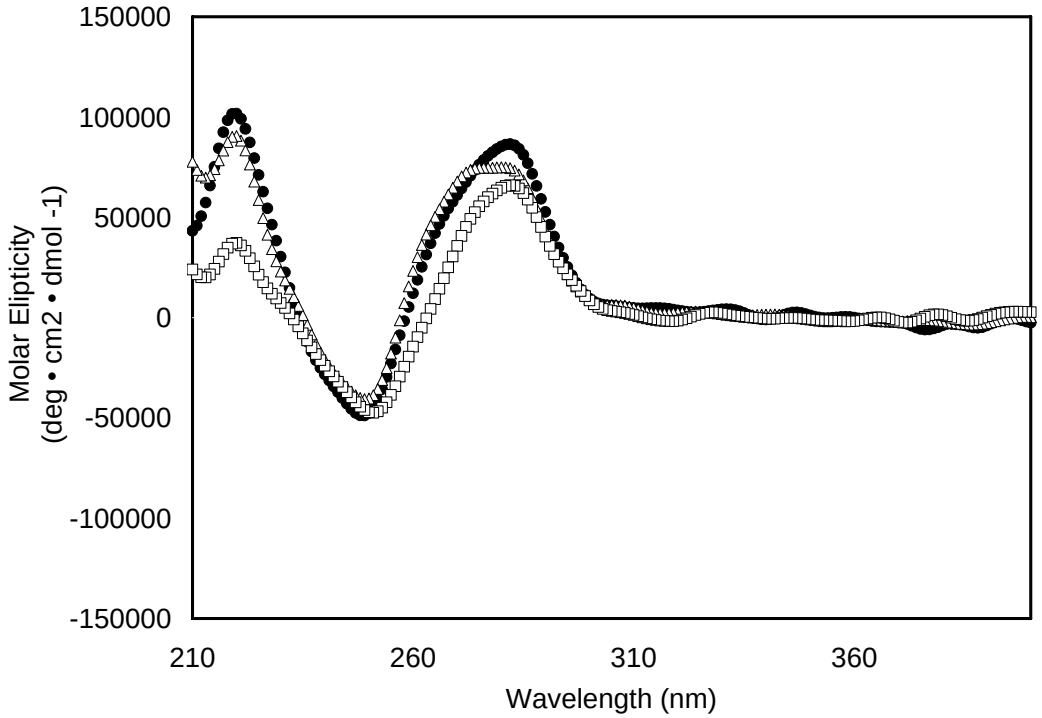


Figure 7-5: ECD spectra for the unmodified T•A sample (●), Tg•A (Δ), and Tg•G (□).

Electronic Circular Dichroism

Structural effects of Tg in duplex DNA were analyzed my ECD. There were small difference in ECD spectra of T•A, Tg•A, and Tg•G (Figure 7-5). Changes in molar elipticity were observed between 210 and 300 nm, there was no signal between 300 and 400 nm. Tg•A exhibited hypochromic shifts at 280 and 220 nm relative to T•A. The Tg•G duplex produced a hypochromic shift at 220 nm and there was an overall red shift between 250 and 280 nm.

Discussion

Thermodynamics

The DNA melting temperature (T_m) is defined as the temperature at which DNAs that melt by a one step process exist as 50% single strand and 50% duplex. The T_m is an excellent indication of the stability of a DNA duplex with respect to denaturation. Detailed theory and analysis of the thermodynamic properties of DNA have been discuss in detail elsewhere [246, 251]. In brief, thermodynamic values were comparable between van't Hoff and individual curve analysis (< 10%), indicating that both Tg-modified duplexes denature via one step transitions [246, 250]. In addition, there were no concentration dependent melting effects observed within the 0.1-1.0 μM range (Table 8-1). Free energy values (with respect to duplex formation) were calculated at 25 °C and 37 °C; the values at both temperatures were negative, consistent with spontaneous formation of the duplexes at temperatures below the sample T_m. The ΔΔG values of the modified duplexes result from reductions in duplex stability. Melting of Tg•G was enthalpically favored over Tg•A indicating a reduction of internal energy (U) of Tg•G relative to Tg•A. Melting of the Tg•G duplex is expected to be entropically favored over Tg•A, however, this conclusion can not be drawn from this data due to statistical error.

The thermodynamic measurements indicate that Tg•A and Tg•G have similar 13 ºC decreases in T_m. The presence of the 5*R*-Tg lesion perturbs the 5'-neighbor base pair G^5•C^{20}. The axial conformation of the Tg CH_3 group decreases stability of the G^5•C^{20} base pair (Chapter V). This is corroborated by the observation that the imino resonance attributed to base pair G^5•C^{20} broadens due to solvent exchange and disappears from the ^{1}H NMR spectrum ~ 35 °C lower in the 5*R*-Tg modified DNA as compared to the

corresponding unmodified oligodeoxynucleotide duplex (Figure 5-5). NMR analysis of both the exchangeable G^{19} H1 imino and A^{19} N^6 amine resonances suggest that these protons undergo increased exchange with solvent. The Tg^6 H3 was not identified, as a result of either exchange broadening or an upfield shift to a congested spectral region. Exchange broadening of nucleic acid imino and amino resonances is characteristic of reduced base pair hydrogen bonding. Taken together, this is expected to contribute to the observed 13 °C reduction in duplex T_m of Tg•A and Tg•G.

T_1 measurements

T_1 data indicates a significant decrease in longitudinal relaxation for the Tg CH_3 in both Tg•A and Tg•G (~1.5 s) (Figure 7-3). The Tg•A T_1 results are consistent with the structure resulting from rMD calculations (Chapter VIII). The Tg^6 CH_3 of Tg•A orients toward G^5 with numerous NOEs to G^5 and Tg^6 deoxyribose protons; all are potential sources of longitudinal relaxation. In contrast, a normal thymine has the CH_3 group facing into the major groove with fewer sources of longitudinal relaxation. It is possible that faster T_1 relaxation results from increased exposure to solvent. In the present work, this was minimized by dissolving the solute in D_2O; however, trace amounts of H_2O may significantly contribute to spin-lattice relaxation. Accelerated Tg^6 CH_3 (Tg•G) longitudinal relaxation may occur for similar reasons for those identified for the Tg•A duplex. Although the Tg•G structure is not yet refined, Tg CH_3 (Tg•G) has numerous NOEs to G^5 much like Tg•A (Chapter V, Appendix A). The structure of Tg•G is anticipated to be extrahelical producing an increased exposure to solvent relative to unmodified duplex. It is noteworthy that Tg^6 CH_3 (Tg•A) relaxes 0.6 s shorter than Tg^6 CH_3 (Tg•A). This attributed to a relative decrease of Tg•G exposure to solvent or

surrounding spin systems. Analysis of the Tg•G refined structure will be necessary to address specific spin-lattice relaxation sources.

T_2 measurements

T_2 data are indicative of a significant increase in transverse relaxation for Tg^6 CH_3 in Tg•A and Tg•G duplexes (Figure 7-4). Tg^6 CH_3 (Tg•A) and Tg^6 CH_3 (Tg•G) relaxed 70 and 80 ms shorter than T^6 CH_3 (T•A) respectively. Tg has a saturated six member ring, as opposed to an aromatic ring like thymidine. Saturated six member rings are anticipated to undergo puckering. The most direct interpretation of the faster T_2 for Tg CH_3 is as a consequence of the rapid puckering on the NMR time scale of the lesion's six member ring. However, there exists the possibility of increased backbone and sugar dynamics occurring in tandem with the re-puckering of the Tg ring. As Tg re-puckers the CH_3 group shifts from axial to equatorial; this change in conformation is likely accompanied by shifts in base stacking. Unfortunately, specific contributions to T_2 relaxation can not be ascertained from these experiments. Attempts to analyze relaxation of the Tg^6 deoxyribose protons were inconclusive due to the inability to assign relaxation times unambiguously. It is of interest that Tg^6 CH_3 (Tg•G) relaxes 10 ms shorter than Tg^6 CH_3 (Tg•A). This difference may result from increased Tg dynamics in Tg•G as a result of less favorable stacking in the mismatch sample.

Electronic Circular Dichroism

Electronic circular dichroism data was collect as a complement to NMR data. ECD spectra indicate small differences in the structure of T•A, Tg•A, and Tg•G duplexes (Figure 7-5). Characteristic of nucleic acids, UV electronic transitions were observed

between 210 and 300 nm [326], there was no absorbance between 300 and 400 nm as observed in AFB_1 modified DNA (Chapter VI). Tg•A exhibited hypochromic shifts at 280 and 220 nm relative to T•A. Differences in ECD spectra between Tg•A and T•A are attributed to the loss of aromaticity of the Tg base and the resultant alteration of stacking patterns of the Tg^6•A^{19} base pair. The hypochromic shift at 275 nm may be a result of a reduced winding angle of Tg•A relative to T•A [330, 337]. The Tg•G duplex produced a hypochromic shift at 220 nm and there was an overall red shift between 250 and 280 nm. ECD spectra represent a composite of multiple effects. Therefore, it is difficult to decipher Tg•G ECD data without the benefit of a refined structure. Differences between Tg•G and T•A not only result form the Tg modified base, but also the replacement of A^{17} with G^{17}. This represents multiple variables that may affect the electronic structure of the duplex. However, in light of thermodynamic data, it may be concluded that differences between Tg•A and Tg•G are not a consequence of reduced duplex T_m because these duplexes destabilized the duplex equally.

Summary

An explanation for the significant *cis-trans* equilibrium differences observed in the Tg•A and Tg•G duplexes must reside in altered contact between Tg and the complementary base. Tg•A has the correct geometry to form Watson-Crick hydrogen bonds, although NMR results indicate it is diminished (Chapter V), most likely as a consequence of steric clash between Tg^6 CH_3 and G^5. A possible explanation for the observed *cis-trans* equilibrium in the Tg•A sample may be reduced steric clash between G^5 and a *trans* Tg-*5R,6R* lesion where the CH_3 is in an equatorial conformation. Clark et al. have shown that an equatorial CH_3 produces less steric interaction with the 5'

neighbor. Therefore, in the Tg•A duplex, a thermodynamic drive to conserve Watson-Crick hydrogen bonding may facilitate epimerization to the sterically favorable *trans* epimer. Interestingly, DFT calculations predict that *trans* 5*R*-Tg with CH_3 in an equatorial position is the next most energetically favorable Tg configuration second only to *cis* 5*R*-Tg with CH_3 in the axial arrangement (Chapter V). In the Tg•G duplex there is no possibility for Watson-Crick hydrogen bonding, however a wobble base pair geometry is possible (Figure 7-6). A wobble base pair configuration would displace Tg further into the major groove than a Watson-Crick configuration. This may result in reduced steric clash between Tg CH_3 and G^5, thus, reducing the incentive to epimerize to *trans*.

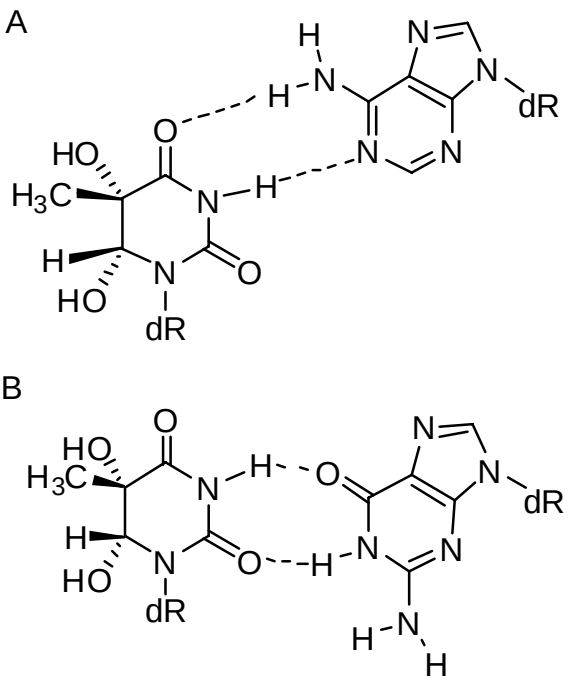


Figure 7-6: Thymine glycol base pair geometry. Watson-Crick base pair geometry (Panel A) was compared to that of a wobble base pair with Tg (Panel B).

Differences in Watson-Crick and wobble base geometry may explain deviations in physical measurements of Tg•A and Tg•G. ECD spectroscopy would be anticipated to be sensitive to altered electronic environments produced by differing base pair interactions. In the Tg•G wobble base pair scenario the Tg CH_3 would be displaced farther into the major groove. This may reduce longitudinal relaxation by limiting energy transfer to nearby spin systems explaining the reduced T_1 relaxation of Tg CH_3 of Tg•G relative to Tg•A. However, reduced stacking of Tg in a wobble base pair may allow for

faster repuckering of the Tg base explaining the increased transverse relaxation observed in Tg•G. Tg•A and Tg•G reduce duplex melting by the same amount. This indicates than neither base pair hydrogen bonds effectively as is evidenced by solvent exchange of G^{17} and A^{17} amino resonances. However, Tg^{6}•A^{17} and Tg^{6}•G^{17} base pairs are not disordered in NMR spectra, thus both base pairs are stabilized.

Thymine glycol is influenced by the complementary base. Repair of Tg is modulated by the presence of either a complementary dA or dG [224, 227]. The *cis-trans* equilibrium position of Tg is influenced by the opposing base (Chapter V). Therefore, the Tg lesion must "communicate" with the complementary base. Tg•A and Tg•G are ordered and destabilize dsDNA by the same magnitude, although Tg•G has comparatively less internal energy (U). NMR data indicates exchange broadening of amino protons expected to participate in hydrogen bonding. Taken together it is concluded that Tg•A and Tg•G must have differing steric and van der Waals interactions that explain their different properties.

CHAPTER VIII

NMR REFINED SOLUTION STRUCTURE OF *CIS*-5*R*,6*S*-THYMINE GLYCOL IN DNA

Introduction

This chapter address refinement and analysis of the solution structure of *cis*-5*R*,6*S*-thymine glycol (Tg) in the duplex 5'-GTGCGXGTTTGT-3'•5'-ACAAACACGCAC-3' (Tg•A). Thymine bases are easily oxidized by ionizing radiation or reactive oxygen species [196, 197]. Several products may result, the most stable include the Tg lesions [197, 198] (Scheme 1-5). Tg exists as a pair of *cis* enantiomers, each in equilibrium with its *trans* epimer. This equilibrium favors the *cis* isomers approximately 4:1 in nucleosides [37]. However, the *cis*-*trans* equilibrium is also contingent on the complementary base (Chapter V). In a Tg•A duplex, the 5*R*,6*S* isomer is preferred 7:3 over 5*R,*6*R*. Only one cross peak was observed between *trans*-5*R*,6*R*-Tg (6*R*-Tg) and DNA, therefore an accurate refined structure could not be determined.

Tg lesions are weakly mutagenic (MF< 1% causing T→C transitions) in *E. coli*, but represent a strong replication block [211, 213, 219]. However, Tg lesions are efficiently repaired *in vivo* [198, 224, 230, 312, 338]. Predictions from molecular modeling studies suggest that Tg could be extrahelical by nature of the loss of planarity of the base [198, 219]. The degree of extrahelicity is believed to be sequence dependent and the 3' base is believed to play a role in the degree of extrahelicity of Tg [219]. It has been proposed that lesion extrahelicity may be an important factor in DNA repair [339].

An NMR refined structure of Tg in a sequence flanked by two adenines (5'-AXA-3') has shown that the lesion is approximately half extrahelical [233]. Placing 5*R*-Tg in a 5'-GXC-3' context resulted in a disordered structure at the site of modification and the complementary base [232]. Structural studies of 5S-Tg have not been reported.

Results

Chemical Analysis

Mass spectrometric analysis of the Tg-adducted oligodeoxynucleotide 5'-GTGCGTgGTTTGT-3' using MALDI-TOF yielded a molecular ion peak (*m/z* 3732) in agreement with Tg modified 5'-GTGCGTgGTTTGT-3'. CGE analysis was indicative of a single Tg modified oligodeoxynucleotide that eluted 12.4 min after the internal standard; the complementary strand eluted 8.9 min after the internal standard. HPLC analysis also confirmed the homogeneity of the sample. CGE and HPLC analysis were repeated after NMR experiments to confirm there had been no degradation of sample purity. It was concluded that the Tg lesion did not undergo any detectable chemical change during the NMR studies. A 7:3 mixture of 6*S*-Tg and 6*R*-Tg *cis-trans* isomers was observed in solution by NMR (Chapter V). The major species was determined to be 5*R*,6*S* by distance filtering of observed NOE cross peaks.

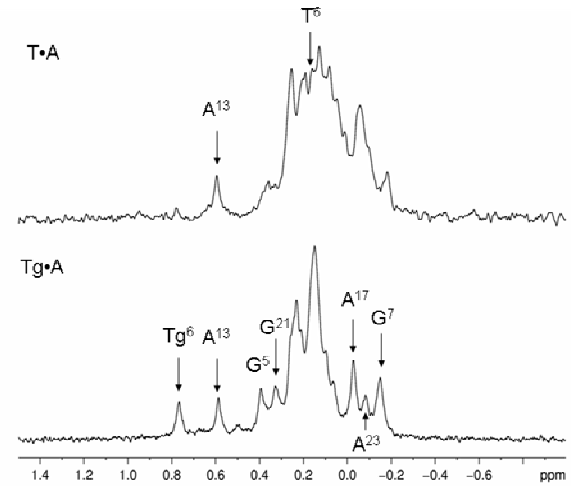


Figure 8-1: Phosphorus spectra of unmodified T•A (upper) and Tg•A (lower). Labels are indicative off 5' phosphate groups.

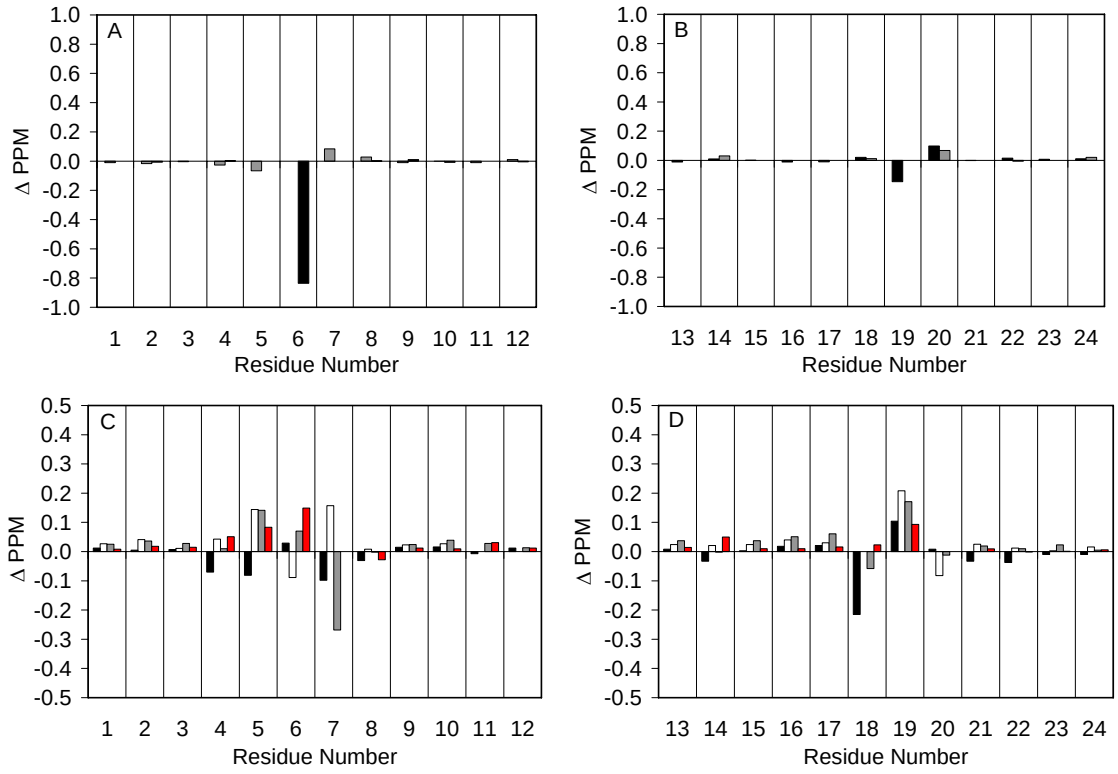


Figure 8-2: Chemical shift perturbation of non-exchangeable protons of Tg-*5R* modified 5'-$G^{1}T^{2}G^{3}C^{4}G^{5}\underline{Tg}^{6}G^{7}T^{8}T^{9}T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3' relative to unmodified duplex. Aromatic H8/H6 (■) and CH_3/H5 (■) resonances of the modified strand (panel A) are reported with the complementary strand (panel B). Deoxyribose H1' (■), H2' (□), H2" (■), and H3' (■) resonances of the modified strand (panel C) are reported with the complementary strand (panel D).

NMR Spectroscopy

NMR assignments of exchangeable, non-exchangeable, and Tg lesion proton resonances have been discussed in detail for the Tg•A duplex in Chapter V. Dispersion of ^{31}P resonances of the modified duplex (Tg•A) relative to the control (T•A) is indicative of backbone disturbance (Figure 8-1). The Tg^{6} and G^{5} 5' ^{31}P resonances were deshielded 0.6 and 0.2 ppm respectively. Inexplicably, the 3'-A^{13} resonance was deshielded an average of 0.4 ppm in both Tg•A and T•A. Although several phosphorus resonances were assigned in Tg•A, resonance overlap prohibited unambiguous resonance assignment in T•A.

Chemical Shift Perturbation

The ^{1}H chemical shifts of the modified duplex were compared to those of an unmodified duplex sample. Figure 8-2 indicates that chemical shift perturbation was localized to the lesion site and nearest neighbor base pairs. The Tg CH_3 was deshielded 0.8 ppm compared to T^6 CH_3 of the unmodified sample. Disturbance of G^5 and G^7 deoxyribose protons was on the order of ± 0.1 ppm. In addition, there were modest perturbations of deoxyribose protons of the complementary strand at residues A^{19} and C^{18} on the order of 0.2 ppm.

Torsion Angle Analysis

The preferred deoxyribose ring pucker was determined by two methods. First, ring conformations were approximated to be northern (N-type) or southern (S-type) by measurement of the mole fraction in the S configuration (X_S) (Table 8-1) [278]. The N-type mole fraction can be calculated by the standard sum rule for time-averaged physical properties [278]:

$$J_{obs} = X_s J_s + X_N J_N \qquad (8\text{-}1)$$

Second, ring pucker was determined by fitting ^{3}J ^{1}H coupling constants for deoxyribose protons by amplitude constrained multiplet evaluation (ACME) of COSY spectra [263]. Electronegitivity of substituent (EOS) Karplus curves were generated and converted to phase angle space assuming a maximum pucker amplitude (Φ) of 44 [277, 278]. $J_{1'2'}$, $J_{1'2''}$, and $J_{1'3'}$ were fit to the curve to determine phase angle ranges (ρ) for sugar rings. The coupling constants $^{3}J_{H1'\text{-}H2'}$ and $^{3}J_{H1'\text{-}H2''}$ could not be determined for G^{11} and T^{12}. Residues that had less than 50% X_S or indicated potential for C3' endo configuration from EOS Karplus curves were constrained accordingly during rMD calculations. The

deoxyribose of G^1, G^7, G^{11}, T^{12}, A^{13}, and C^{24} indicated a high probability for C3' endo conformation (ρ 0 – 210). The Tg^6 deoxyribose was mostly in a C2' endo conformation (70%) (ρ 125 – 210).

Table 8-1: Vicinal ^{1}H coupling constants were fit using amplitude constrained multiplet evaluation (ACME).

	$J_{H2'H2''}$	$J_{H1'H2'}$	$J_{H1'H2''}$	$J_{H2'H3'}$	X_S*		$J_{H2'H2''}$	$J_{H1'H2'}$	$J_{H1'H2''}$	$J_{H2'H3'}$	X_S*
G^1	-14.89	7.26	5.40	4.93	0.52	A^{13}	-17.16	7.71	5.11	4.86	0.54
T^2	-15.97	9.38	5.24	5.61	0.83	C^{14}	-14.82	9.23	4.54	6.48	0.69
G^3	-14.89	9.39	4.57	4.30	0.72	A^{15}	-14.86	9.24	4.53	4.31	0.69
C^4	-14.80	8.21	5.24	6.29	0.64	A^{16}	-17.35	9.54	4.58	4.82	0.75
G^5	-15.14	8.34	5.43	4.54	0.69	A^{17}	-15.88	8.87	5.49	5.23	0.79
Tg^6	-16.33	9.08	4.77	4.76	0.71	C^{18}	-15.26	8.55	4.86	5.53	0.64
G^7	-15.17	7.31	5.04	4.79	0.47	A^{19}	-14.79	8.64	5.53	5.55	0.76
T^8	-16.71	9.93	4.56	5.41	0.81	C^{20}	-15.53	8.99	5.11	5.66	0.75
T^9	-14.09	9.79	4.07	7.10	0.71	G^{21}	-14.68	9.40	4.36	3.90	0.69
T^{10}	-15.86	8.57	5.50	6.73	0.74	C^{22}	-14.91	8.36	5.25	6.45	0.67
G^{11}			9.12	5.05		A^{23}	-13.30	8.25	6.07	5.63	0.78
T^{12}			10.31	6.22		C^{24}	-14.31	6.99	5.82	7.07	0.54

* Mole fraction of S-type conformation ($X_S = ((J_{H1'H2'}+J_{H1'H2''})-9.4)/(15.7-9.4)$ [278]

The phosphodiester dihedral angle ε and glycosyl torsion angle χ were empirically assessed. The ε dihedral angle (C4'-C3'-O3'-P) was approximated by heteronuclear coupling constants ($^3J_{P-H3'}$) determined from constant time NOESY experiments [279, 280]; dihedral angles were determined by fitting *J* values to a Karplus relationship [281]. Karplus analysis yielded values for ε in the range of 193.0º ± 10º. No unusual $^3J_{P-H3'}$ couplings were observed; however, 1D ^{31}P chemical shifts were more dispersed than those of the unmodified oligonucleotide suggesting disruption of the backbone at α and ζ [340, 341]. Glycosyl torsion angles [χ] were estimated from NOESY data. The existence of cross peaks between pyrimidine H6 / purine H8 and anomeric H1' protons was consistent with the *anti* conformation of the glycosyl bond for non-terminal residues (χ = 160-340°) [282].

Structural Refinement

From the NMR studies, 675 experimental restraints were derived that could be used for structural refinement. In addition to the experimental restraints, 40 empirical distances based on hydrogen bonding geometries and 106 backbone torsion restraints were applied (Table 8-2). Proton-proton distance restraints were determined by MARDIGRAS calculation as discussed in Chapter II. The proton-proton restraints consisted of 213 intra-residue and 196 inter-residue restraints. Vicinal (intra-residue) sugar ring proton distance restraints were discarded for refinement purposes; their corresponding NOE cross-peaks were included in CORMA calculations. Analysis of exchangeable amino and imino protons indicated that Watson-Crick hydrogen bonding was disrupted for Tg^6•A^{19} and G^5•C^{20}. Therefore, empirical distance restraints were applied to the remaining ten base pairs but Tg^6•A^{19} and G^5•C^{20} were excluded. There were 16 proton-proton restraints between Tg H6 and Tg CH_3 to DNA (Table 8-3).

Table 8-2: Distribution of rMD restraints

Assigned Resonances	*222*
Watson-Crick	40
Backbone Torsion	106
Deoxyribose Torsion	120
Distance:	409
Inter-residue	196
Intra-residue	213
Total Restraints	675
Avg. Restraint per Residue	28

Table 8-3: Tg-(*5R,6S*) resonance assignments and inter/intra residue cross peaks.

Resonances	*δ (PPM)*	*Crosspeak with:*
H1’	5.74	Tg^6 H2’, Tg^6 H2’’, Tg^6 H3’, Tg^6 H4’, G^7 H8, G^7 H5’, Tg^6 CH_3
H2’	2.18	Tg^6 H1’, Tg^6 H2’’, Tg^6 H3’, Tg^6 H4’, Tg^6 H5’, Tg^6 CH_3, G^7 H8
H2’’	2.39	Tg^6 H1’, Tg^6 H2’, Tg^6 H3’, Tg^6 H4’, G^7 H8

H3'	4.68	Tg^6 H1', Tg^6 H2', Tg^6 H2'', Tg^6 H3', Tg^6 H4', Tg^6 H5', Tg^6 H5'', Tg^6 CH_3, Tg^6 H6, G^7 H8
H4'	4.15	Tg^6 H1', Tg^6 H2', Tg^6 H2'', Tg^6 H3', Tg^6 H5', Tg^6 CH_3
H5'	3.90	Tg^6 H2', Tg^6 H3', Tg^6 H4', Tg^6 H5'', Tg^6 CH_3
H5''	4.08	Tg^6 H5', Tg^6 H3', Tg^6 CH_3
H6	1.35	Tg^6 H3', Tg^6 CH_3
CH_3	0.59	G^5 H1', G^5 H2', G^5 H2'', G^5 H3', G^5 H4', G^5 H5', G^5 H8, Tg^6 H1', Tg^6 H2', Tg^6 H3', Tg^6 H4', Tg^6 H5', Tg^6 H5'', Tg^6 H6

Structure refinement by simulated annealing began with multiple starting trajectories. Thirty-two starting structures sampling typical A-form and B-form DNA parameters were used, sixteen of them had Tg^6 CH_3 in an axial position and sixteen had Tg^6 CH_3 in an equatorial position. Of the resultant structures, 29 converged with Tg^6 CH_3 in an axial position, 2 had Tg CH_3 in an equatorial conformation. Accuracy of resultant structures was evaluated by calculation of the sixth root residuals between theoretical NOE intensities and experimental NMR data. CORMA calculations indicated that the axial configuration better represented experimental data at the lesion site (R_1^X 9.82×10^{-2} equatorial vs. 7.37×10^{-2} axial). A representative structure taken from simulated annealing refinement was placed in a truncated octahedron water box for 5 ns of isothermic rMD refinement (300 K). An ensemble of ten structures was extracted from the final 100 ps of the rMD trajectory (Figure 8-3). The pair-wise RMSD between the heavy atoms of the ten structures was 0.59 Å. R_1^X values from CORMA calculations of individual residues indicate good agreement between the final structure and ^1H NOESY data (Figure 8-4). The ensemble had a total R_1^X value of 8.24×10^{-2}.

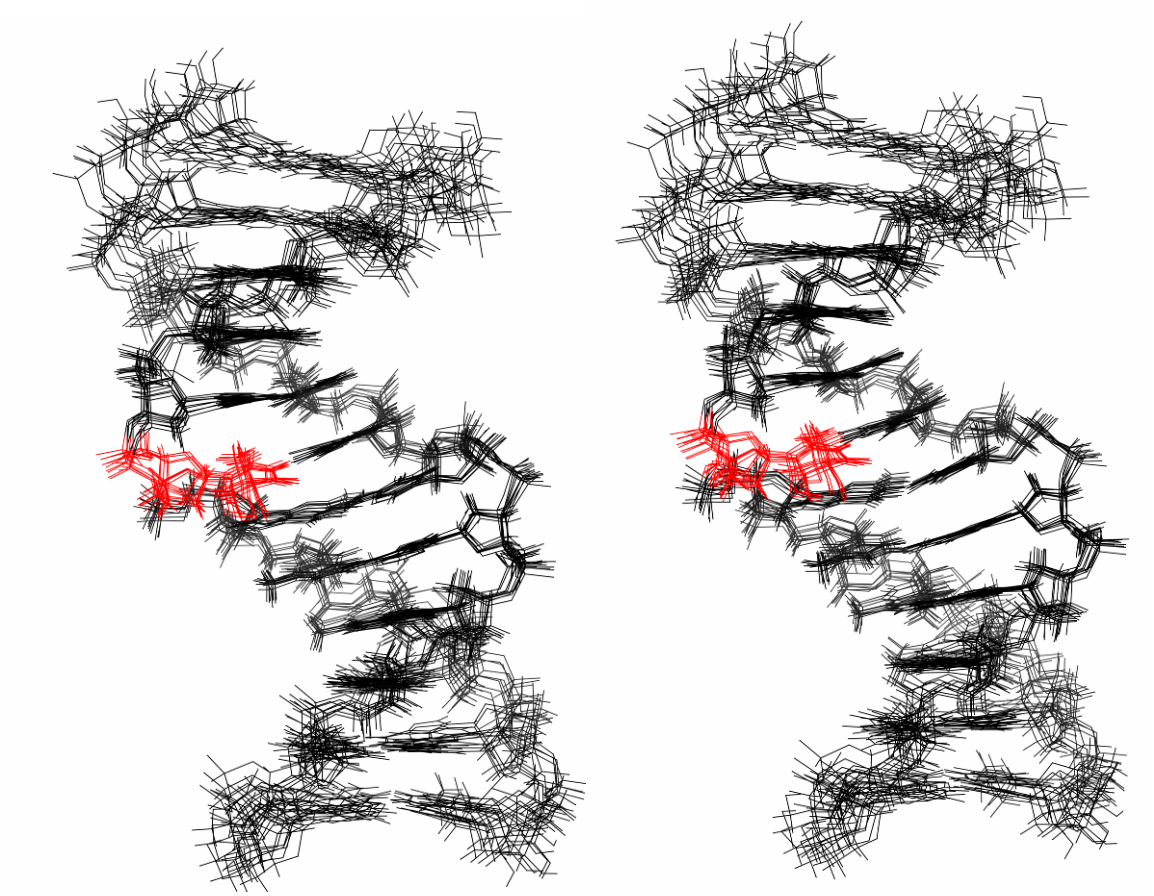


Figure 8-3: Stereo view of ten superimposed structures resulting from molecular dynamics calculations in explicit solvent. The RMSD between eight core base pairs of resultant structures was 0.58 Å.

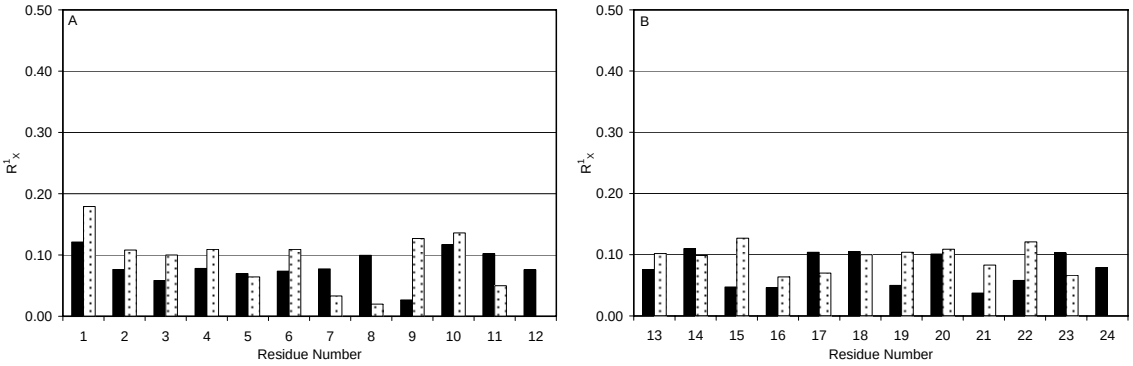


Figure 8-4: Complete relaxation matrix analysis R_x values for Tg-*5R* modified rMD ensemble of 5'-$G^1T^2G^3C^4G^5\underline{Tg}^6G^7T^8T^9T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3'. Per residue intra (■) and inter (□) R_x values for the modified strand are listed in panel A; values for the complementary strand are in panel B.

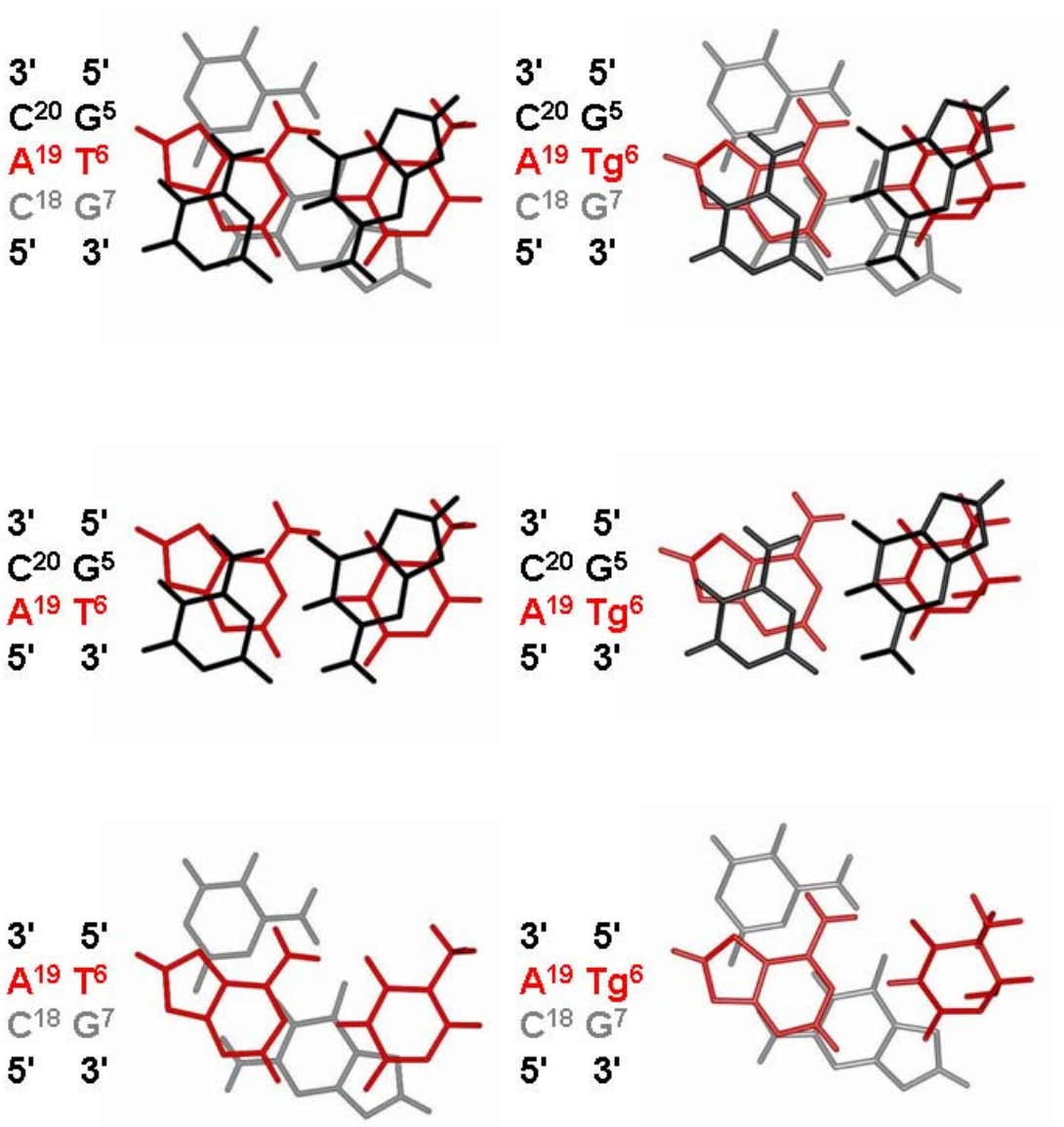


Figure 8-5: Nearest neighbor base pair stacking of unmodified duplex (left) compared to thymine glycol dulex (right).

The refined structure exhibited localized distortion at the lesion site. The Tg^6 lesion was partially extrahelical. The Tg^6 lesion tilted into the major groove and away from G^5. Although lacking the benefit of empirical Watson-Crick hydrogen bonding restraints, the Tg^6•A^{19} and G^5•C^{20} base pairs were ordered. Analysis of π–π stacking indicated that Tg^6 was not distorted laterally when compared to an unmodified duplex (Figure 8-5). The A^{19} residue remained ordered and stacked into the duplex.

Trajectory Analysis

In an attempt to better understand the origin of the increased Tg^6 molecular motion as evidenced by faster T^2 values (Figure 7-4), the isothermic rMD trajectory was analyzed. It was postulated that increased T^2 values for the lesion relative to unmodified thymine bases may have three possible sources: lateral base pair movement, glycosyl torsion angle fluctuation, and/or thymine glycol ring pucker. Base pair fluctuation was estimated by measurement of the distance between thymine/thymine glycol H3 and adenine N1 atoms during the 5 ns calculation (Table 8-4). The inter base pair distance between Tg^6 and A^{19} was 0.1 Å greater than unmodified T•A base pairs. However, the fluctuation of the distance was 15% greater than other T•A base pairs. The glycosidic torsion angle [χ] was tracked during the calculation. The fluctuation and absolute values of χ for the modified base were unremarkable. The average angle values for the C7-C6-C5-H6 torsion angle that define the thymine/thymine glycol bases were tracked. The absolute value for Tg^6 was 70° lower than standard thymine bases; this was expected considering the chemical modification of Tg^6. The amplitude of the ring motion was three fold greater in Tg^6.

The rMD trajectory was analyzed for predicted hydrogen bonds (Table 8-5). In spite of no restraints to preserved the Tg^6•A^{19} base pair, the simulation predicted conserved Watson-Crick hydrogen bonding. In addition, there was a high occupancy for the inter-strand hydrogen bond G^7 N7→Tg^6 HO6. The Tg^6 HO5 proton was predicted to hydrogen bond with solvent 14% of the trajectory.

Table 8-4: Average thymine and Tg glycosyl torsion angle, base torsion, and inter-strand base pair distance.*

Residue	*Inter-strand Base Pair Distance (H3→N1) (Å)*	*Glycosyl Torsion Angle (degrees)*	*Base Torsion Angle (C7-C5-C6-H6)(degrees)*
T^{2}	2.80±0.15	169±9	172±6
Tg^{6}	2.88±0.17	169±10	101±18
T^{8}	2.81±0.14	153±13	172±6
T^{9}	2.82±0.14	179±8	172±6
T^{10}	2.81±0.14	159±11	172±6
T^{12}	2.80±0.15	165±10	172±6

* The fluctuation of select torsion angles and atomic distances were measured during the course of 5 ns molecular dynamic simulations at 300 °K. Standard deviations represent amplitude of angle/distance flux.

Table 8-5: Hydrogen bonding occupancy *

	% Occupancy
A^{19}:N1→Tg^{6}:H3	99.9
Tg^{6}:O4→A^{19}:H61	98.0
G^{7}:N7→Tg^{6}:HO6	84.3
G^{5}:H1→C^{20}:NH1	99.8
G^{5}:NH1→C^{20}:O2	23.7
G^{5}:O6→C^{20}:H61	1.9
Tg^{6}:O2→A^{19}:H2	13.8
G^{7}:O6→Tg^{6}:HO6	0.9
G^{7}:O6→Tg^{6}:HO5	0.2
G^{7}:O6→A^{19}:H61	0.1
Tg^{6}:HO5→H_2O	14.0
G^{7}:O6→H_2O	6.9
A^{19}:H62→H_2O	7.1
Tg^{6}:O4→H_2O	3.9
Tg^{6}:O2→H_2O	5.2
G^{7}:N7→H_2O	1.9

* Occupancy was calculated from 5 ns trajectories with a distance cutoff of 3.5 Å and angle cutoff of 120 degrees.

Helicoidal Analysis

An analysis of the helicoidal properties was performed on an average structure obtained from the refined ensemble. Results are listed in detail in Appendix E. There was a slight bend in the DNA helix of 5-10° that originated at the lesion site. There was no significant disturbance in the X-displacement, Y-displacement, or Tip base pair parameters (Figure 8-6). Base-base parameters were normal with the exception of a 0.5

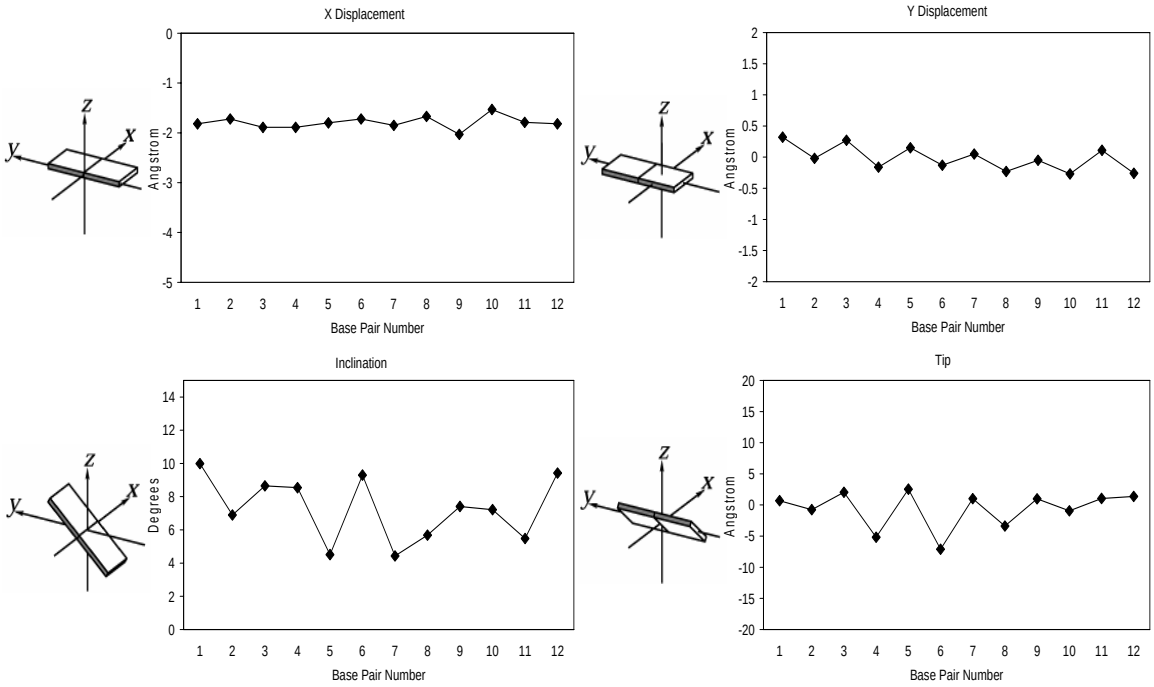


Figure 8-6: Global base pair helicoidal parameters. Parameters for the rMD refined *cis*-5*R*,6*S*-thymine glycol modified 5'-G^{1}T^{2}G^{3}C^{4}G^{5}<u>Tg</u>^{6}G^{7}T^{8}T^{9}T^{10}G^{11}T^{12}-3'•5'-A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}-3 duplex are listed.

Å displacement of Tg^{6} in regard to stagger and a 30° shift in propeller parameters (Figure 8-7). Global inter-base parameters indicated a 1 Å rise between Tg^{6} and G^{5} (Figure 8-8). In addition, there was an inter-base difference in tilt of 15° between Tg^{6} and G^{5}.

<u>Solvent Accessible Surface</u>

Solvent accessible surface (SAS) areas of the duplex were rendered with the program MSMS as a function of probe radius [342] (Figure 8-9). The SAS of the Tg base in the rMD refined structure relative to the SAS of a Tg base in a nucleoside was compared with residues A^{23}, T^{9}, and T^{12}. The terminal T^{12} residue was expected to have a high percent SAS whereas the interior T^{9} and A^{23} bases where expected to yield a reasonable approximation of canonical base SAS. The percent of surface area accessible to solvent was rendered as a function of probe radii. The average accessible surface area

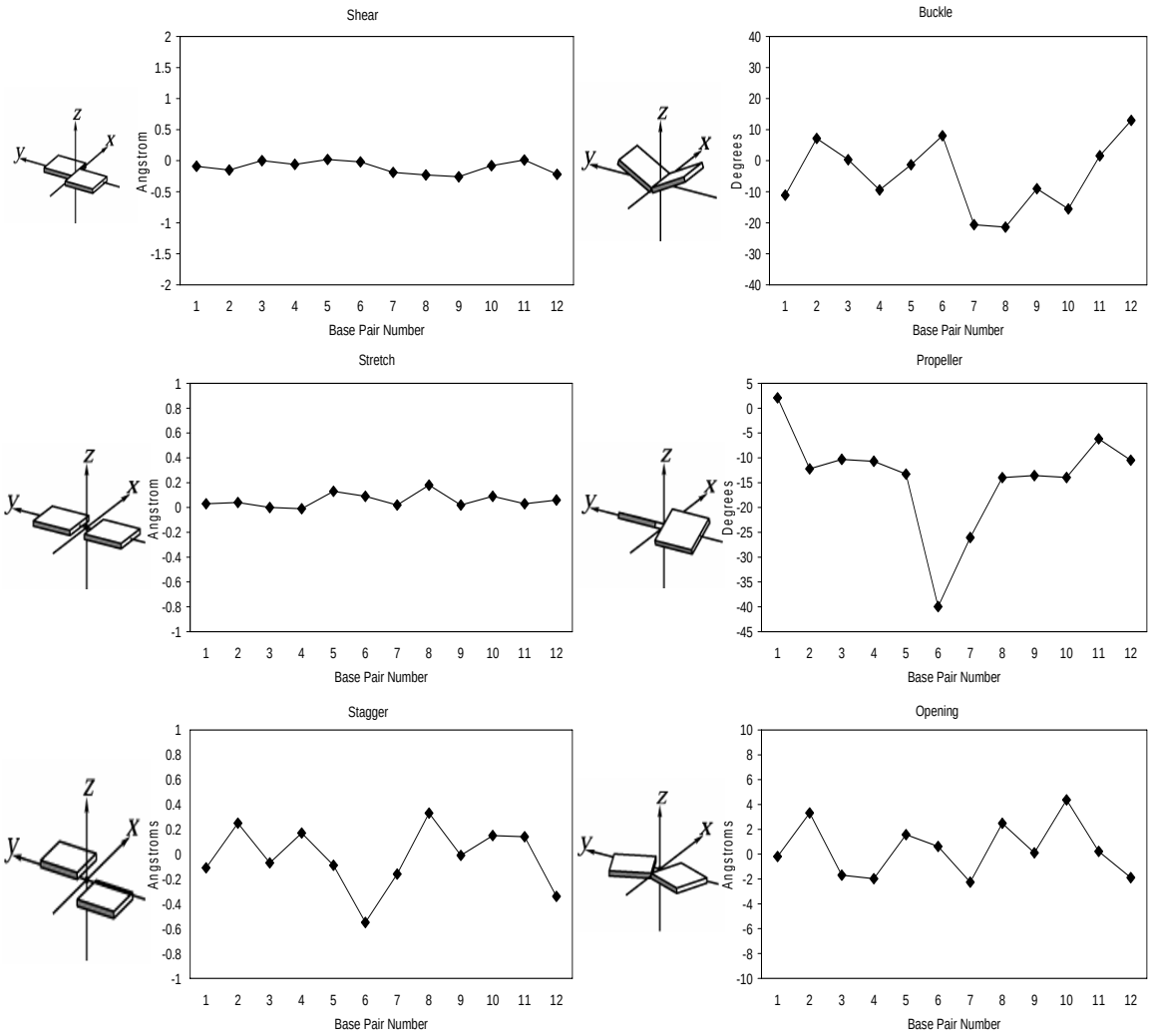


Figure 8-7: Global base-base helicoidal parameters. Parameters for the rMD refined *cis*-5*R*,6*S*-thymine glycol modified 5'-$G^{1}T^{2}G^{3}C^{4}G^{5}\underline{Tg}^{6}G^{7}T^{8}T^{9}T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3 duplex are listed.

of the Tg^{6} residue was approximately 20%. The terminal T^{12} residue had an elevated area exposed to solvent at an average of 21%. The SAS of Tg^{6} was comparable to that of the terminal T^{12} residue. The color coded Connolly surface of the residues of interest is depicted in Figure 8-9.

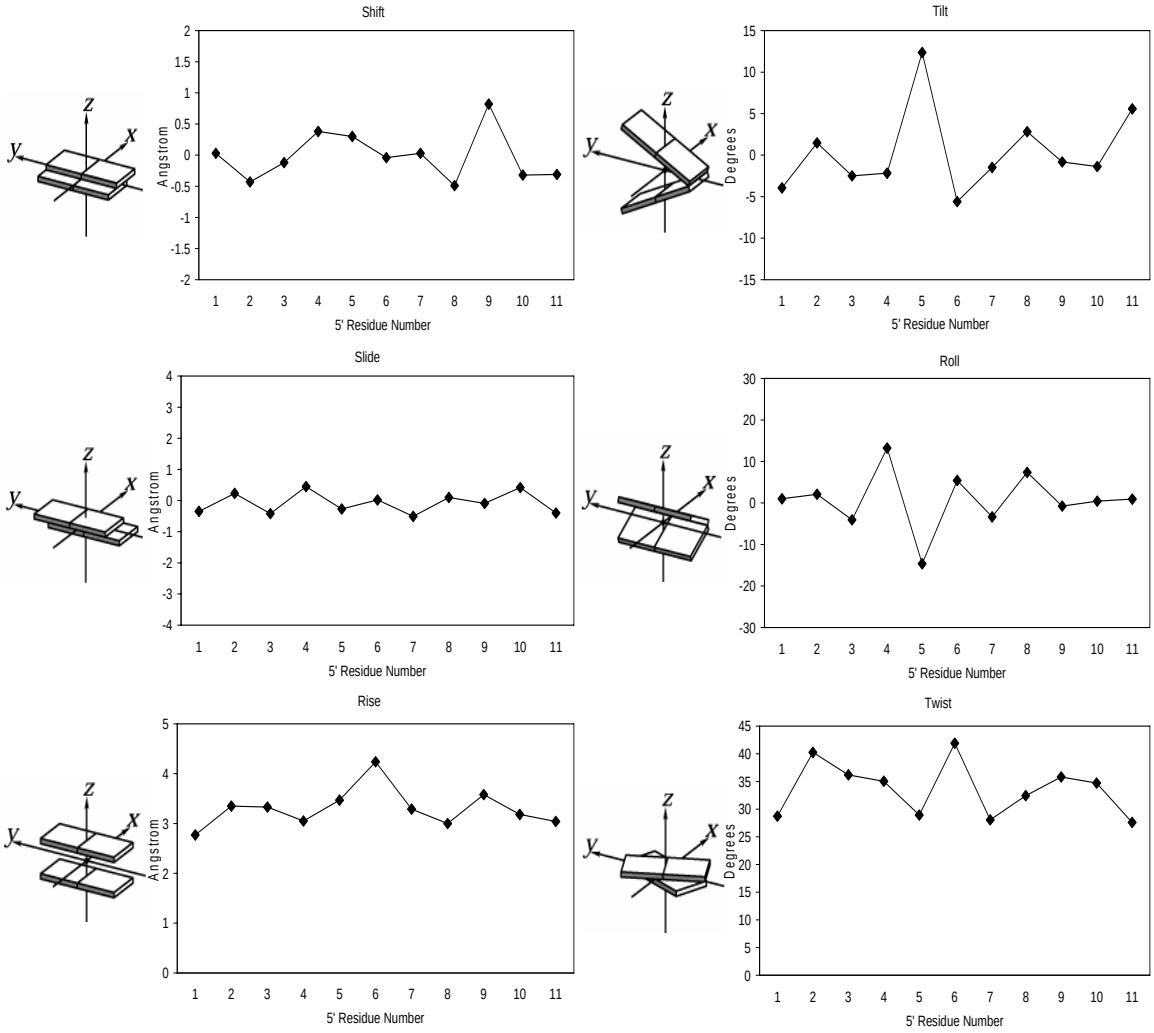


Figure 8-8: Global Inter-base helicoidal parameters. Parameters for the rMD refined *cis*-5*R*,6*S*-thymine glycol modified 5'-$G^{1}T^{2}G^{3}C^{4}G^{5}\underline{Tg}^{6}G^{7}T^{8}T^{9}T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3 duplex are listed.

Discussion

The refined structure confirmed that the Tg was partially extrahelical toward the major groove (Figure 8-3). NOE cross peaks between Tg^{6} CH_3 and the 5' neighbor, G^{5}, were consistent with a well ordered structure. Molecular modeling studies indicated an equatorial Tg^{6} CH_3 could be more easily incorporated into a DNA duplex [219]. An axial

Tg^6 CH_3 has greater steric interference with G^5 than an equatorial Tg^6 CH_3. CORMA calculations revealed that axial Tg^6 CH_3 better fit experimental NOESY data than an equatorial Tg^6 CH_3. This result was in agreement with crystallographic data [207, 234] and *ab initio* quantum calculations (Chapter V). The lack of Tg planarity causes the lesion to behave as a "steric wedge". Steric factors affect the ability of Tg^6 to efficiently Watson-Crick hydrogen bonding with A^{19} as is evidenced by thermodynamic data (Chapter VII).

Duplex stability is reduced in a Tg modified oligonucleotide (Chapter VII). Disruption of efficient Tg^6•A^{19} and G^5•C^{20} hydrogen bonding was given as an explanation for the reduced T_m. A first approximation would suggest that stacking of the Tg lesion is not significantly different from an unmodified duplex (Figure 8-5). However, helicoidal analysis indicates that π–π stacking is reduced in the modified structure as is evidenced by inter-base tilt (Figure 8-8) and base-base propeller twist (Figure 8-7). A reduction in π–π orbital overlap may also contribute to reduced duplex stability.

The present refined structure predicts a slight bending of the helical axis around the lesion site, as seen in Figure 8-3 and Appendix E. Helicoidal analysis indicated this bend to be 10° between terminal base pairs and 7° between penultimate base pairs. The loss of stability resulting from a Tg lesion is expected to control the degree of helical axis bending. This may be a consequence of steric interaction with Tg^6 CH_3 or an error resulting form under-refinement at the lesion site as a consequence of relatively few experimental restraints. Although DNA axis bending is probable, definitive conclusions

concerning bending are not possible without residual dipolar coupling restraints [343, 344].

Previously, the extrahelicity of Tg was determined in the 5'-AXA-3' sequence [233] by measuring the SAS of Tg and comparing this to other bases in the duplex. It was concluded that 5*R*-Tg was approximately half extrahelical (~50%). In modeling studies 5'-GXG-3' was predicted to have a greater extrahelicity than 5'-AXA-3' [219]. Although there is ambiguity in how 5'-AXA-3' SAS measurement was conducted, in 5'-GXG-3', the Tg lesion is 20 % exposed to solvent (Figure 8-9). The SAS of Tg^{6} in 5'-GXG-3' was comparable to the highly exposed, terminal T^{12} residue (22 %) and 15% greater than the well stacked T^{9} residue. The SAS of Tg in 5'-GXG-3' is less than in 5'-AXA-3' [233]; however, modeling studies predicted that Tg extrahelicity would be greater in 5'-GXG-3' [219]. The discrepancy between predicted

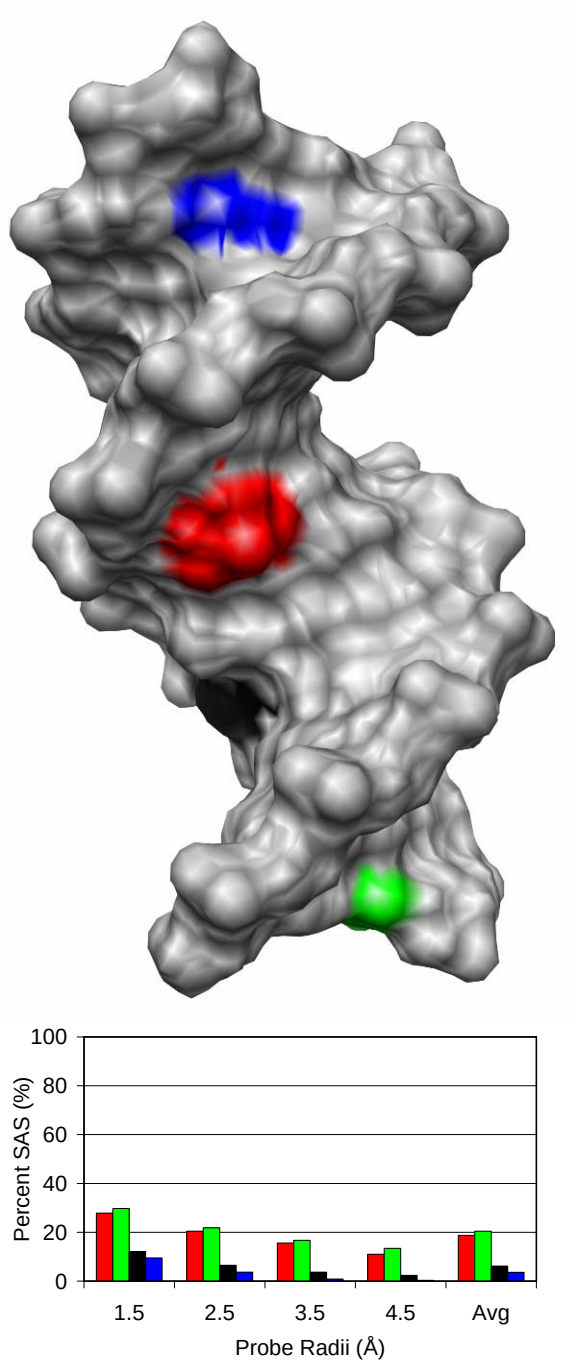


Figure 8-9: The solvent accessible surface (SAS) of select bases as a function of probe radius. The SAS of Tg^{6} (■) was compared to T^{9} (■), T^{12} (■), and A^{23} (■) (DNA SAS / Base SAS * 100). The color coded Connolly surface of the modified duplex is rendered in the upper panel.

extrahelicity and observed SAS of Tg in the 5'-GXG-3' sequence may be a result of experimental procedure. Modeling studies and structural refinement of Tg modified oligonucleotides were previously conducted in vacuo [219, 232, 233]. The present refinement of Tg in 5'-GXG-3' has been conducted using implicit and explicit solvent models [253, 283, 286, 287]. These models more accurately reflect behavior of biomolecules in solvated systems. However, it should be noted that the rMD refined 5'-GXG-3' structure would suggest conserved Watson-Crick hydrogen bonding in contradiction to thermodynamic and NMR experimental data. An explanation for this discrepancy may result from minimization of the AMBER energy function where hydrogen bonding of the G^5•C^{20} and Tg^6•A^{19} base pairs result in a lower system energy.

Restrained molecular dynamics trajectories were analyzed for predicted hydrogen bonds. Interestingly, 84 % occupancy was determined for Tg^6 OH6→G^7 N7. This hydrogen bond could not be confirmed experimentally. However, a similar bond would be expected in the 5'-AXA-3' context. An inter-residue hydrogen bond could stabilize the Tg lesion in multiple ways. First, epimerization at the chiral Tg C6 would favor the 6*S* configuration; a 6*R* would place the hydroxyl group on the 5' face of Tg where it would have an unfavorable steric clash with 5'-dG. A Tg^6 OH6→G^5 N7 hydrogen bond would not be possible because of improper geometry. Secondly, the Tg^6 OH6→G^7 N7 hydrogen bond may affect lesion extrahelicity. In 5'-GXC-3', a Tg^6 OH6→C7 hydrogen bond is not possible. As a likely consequence, structural studies of Tg modified 5'-GXC-3' found that the lesion and complementary base were disordered [232]. Perhaps repeating NMR analysis of a Tg modified oligonucleotide with a 3'-pyrimidine would offer new insights to lesion stability.

Pseudorotational analysis indicated that G^7 and the four terminal residues had an increased C3' endo (N-type) sugar population (Table 8-1). This observation for the terminal bases may be explained by increased dynamics and duplex fraying. It is not immediately obvious why G^7 has a 15 – 20 % increase in N-type conformation, but the fact that G^7 is adjacent to Tg^6 should be considered. Careful analysis of coupling constants of other Tg modified oligonucleotides is necessary to accurately interpret this observation.

Biological Significance

To understand why Tg lesions are efficiently repaired, it is necessary to understand what structural characteristics are recognized by the enzyme. There are multiple possibilities with Tg modified DNA: extrahelicity, reduced melting temperature, disrupted hydrogen bonding modes, and loss of base planarity. These properties are directly related to one another. For example, the loss of base planarity prohibits Tg^6 from effectively stacking. This results in the extrahelicity of the lesion and its inability to efficiently Watson-Crick hydrogen bond. The reduced stacking and hydrogen bonding causes the reduction of melting temperature and helix stability. The importance of steric contributions from Tg^6 CH_3 cannot be overlooked. While this work was in progress a crystal structure of a binary complex of the replicative RB69 DNA polymerase with a *cis*-5*R*,6*S*-Tg modified oligodeoxynucleotide was published [234]. The crystal structure indicates that Tg CH_3 hinders stacking of the 5' base causing replication to stall after

insertion of dA opposite Tg. In the present data Tg^{6} CH_3 interacts with the 5' G^{5} residue and effectively reduced the G^{5}•C^{20} base pair stability (Chapter V)

It has been reported that human endonuclease III (hNth) and human endonuclease like protein (hNeill) differentially repair Tg depending on whether the complementary base is guanine or adenine [224, 230]. Likewise, repair of Tg is modulated by lesion stereochemistry [224, 227, 228, 230]. Therefore, a variety of Tg lesions motifs are possible, i.e. 5*R*-Tg•A, 5*S*-Tg•A, 5*R*-Tg•G, 5*R,*6*R*-Tg•A, etc. In addition to the currently discussed Tg structure, only one other Tg solution structure exists [233]. To have a more complete understanding of how various Tg structural motifs are biologically interpreted, it would be useful to conduct structural studies of Tg modified duplexes that explore the various Tg lesion configurations and sequence related effects. Results maybe useful in explaining sequence related polymerase stalling [215]

CHAPTER IX

SUMMARY

The impetus for this dissertation was to develop, via NMR methodology, a better understanding of the effects of specific lesions on the three dimensional structure of DNA and, in turn, on biological processing. However, it became obvious in the early stages that structural results could not be accurately interpreted without knowledge of lesion equilibria. Three fluxional DNA lesions were studied: methyl formamidopyrimidine (Me-dGuo-FAPY), aflatoxin B_1 formamidopyrimidine (AFB_1-FAPY), and thymine glycol (Tg). These lesions differ in their chemical properties and demonstrate a range of structural scenarios that influence replication and repair.

Formamidopyrimidines

Comparison of methyl and AFB_1 substituted formamidopyrimidine adducts indicate lesion specific effects on FAPY equilibrium. The anomer equilibrium of AFB_1-FAPY favors the α anomer by 2:1 in ssDNA, whereas the α and β anomers of Me-dGuo-FAPY were present in approximately equal quantities in ssDNA. The difference in FAPY equilibrium of these two lesions is due to the large hydrophobic aromatic ring system of AFB_1. In ssDNA π–π orbital stacking is more effective with the α anomer than the β. In dsDNA, the AFB_1 prefers stacking with the complementary strand. With

Me-dGuo-FAPY the CH_3 group does not have π orbitals and can not stack; therefore, Me-dGuo-FAPY is free to equilibrate without stacking bias.

In contrast to the differential effects of stacking, AFB_1 and methyl-substituted FAPYs are postulated to have similar sequence-dependent properties. The central element of this comment is based on an observed inter-residue hydrogen bond between the FAPY formamide and a 3'-adenine N^6 amino group. This hydrogen bond was originally reported by Mao et al. [115] when studying β-AFB_1-FAPY in dsDNA and further corroborated in the present work (Chapter VI). Evidence supporting this conclusion includes deshielding of the A^6 NH1 chemical shift relative to unmodified DNA. Proton deshielding is consistent with electron withdrawing as a consequence of hydrogen bonding. In addition, AFB_1-FAPY nucleosides favor the formamide *Z* geometrical isomer over *E* by 2:1. In dsDNA, only the *E* isomer was observed. This is indicative of a driving force unique to DNA and absent in nucleosides where *E* is the preferred form. Preference of the *E* form is accounted for by a hydrogen bond between the formyl carbonyl group and the N^6 amino group of the 3' adenine (Chapter VI). Finally, hydrogen bond formation is supported by rMD calculations that predict 95% hydrogen bond occupancy. Circumstantial evidence for the existence of an inter-residue hydrogen bond was found when replacement of the 3'-dA with a 3'-dT resulted in a mixture of configurational isomers (Chapter IV). A 3'-dT will not hydrogen bond with the carbonyl group; therefore, the equilibrium of geometrical isomers may resemble the nucleoside equilibration more closely. This hydrogen bond was less stable in ssDNA samples containing α-AFB_1-FAPY (Chapter VI). Regardless, an inter-residue hydrogen bond represents a sequence specific property.

Only recently has the sequence-dependent repair of AFB_1-FAPY DNA lesions been addressed. Oleykowski et al. have demonstrated the rate of incision with UvrABC endonuclease on AFB_1-FAPY can vary as much as 15-fold depending on sequence with 3'-adenine being the most resistant [189]. Similarly, repair of Me-dGuo-FAPY is sequence dependent [34, 57, 68-70, 300] although these studies did not address a 3'-A. Intra-strand formyl hydrogen bonding is a structural effect that may influence repair of all FAPY lesions in DNA. Oleykowski's data [189] suggests that stabilization of formyl rotamers is correlated to AFB_1-FAPY repair efficiency. It is more difficult to offer conjecture on Me-dGuo-FAPY sequence specific repair because Me-dGuo-FAPY atropisomers may also be possible (Figure 1-3); thus far only the atropisomers of AFB_1-FAPY have been detected [31]. Although 3'-dA is expected to stabilize a R_a atropisomer and *E* geometrical isomer, it is unclear if a S_a atropisomer could be stabilized by an inter-residue hydrogen bond.

Thymine Glycol

It is estimated that human cells repair hundreds of Tg lesions per day [198]. The Tg lesion is a substrate for base excision repair, both in *E. coli* and in mammalian cells [307]. In *E. coli*, repair of Tg is initiated by endonuclease III (Nth) [308] and endonuclease VIII (Nei) [309]. The base excision repair of Tg lesions is dependant on both their stereochemical configurations (5R,6S vs. 5S,6R), and the identity of the complementary base [224]. Multiple studies have shown that thymine glycol inhibits DNA synthesis *in vivo* in most DNA sequences [211, 213-215]. However, polymerase read-through is possible in select sequences [215]. The presence of a pyrimidine 5' to Tg

enhances the probability of translesion synthesis more than a 5' purine [215-217]. The 3' base determines the extent of replication block [218].

Initially, the principal structural property associated with Tg lesions was their extrahelicity [233]. Based on solution structure and modeling studies [219, 231] the DNA sequence was postulated to alter the degree of extrahelicity and thereby explain observed sequence related biological effects [215-218]. More recently, a crystal structure of a *cis*-5*R*,6*S*-Tg modified oligonucleotide with the RB69 polymerase revealed that the Tg methyl group was in the axial conformation, hindering stacking of the adjacent 5´-template guanine [234]. Chapter VIII corroborates the fact that *cis*-5*R*,6*S*-Tg is extrahelical in a sequence other than that studied by Kung et al. [233]. In Chapter V it was demonstrated that the 6*R*-Tg to 6*S*-Tg equilibrium point was dependant on the complementary base.

Taken together it appears that Tg steric interactions may in fact be the principal determinant in sequence specific biological effects. For example, steric clash with the 5' and 3' neighbors may affect extrahelicity, while steric interactions with the complementary base affect 6*R*-Tg to 6*S*-Tg (*cis-trans*) epimerization. However, an additional sequence dependent factor may be intra-strand hydrogen bonding. Previous modeling studies [231] and the present work (Chapter VIII) predict a Tg^6 OH6→G^7 N7 hydrogen bond. Unfortunately, this hydrogen bond cannot be confirmed experimentally. However, it is interesting to note that initial structural studies of Tg with 3'-dC resulted in a disordered structure. A Tg^6 OH6→X^{n+1} N7 hydrogen bond may also contributed to stabilization of the Tg base so that an ordered structure was observed in 5'-ATgA-3' and 5'-GTgG-3' sequences but not 5'-GTgC-3' [232, 233] (Chapter VIII). The presence of 3'

pyrimidines produced increased replication blockage relative to 3' purines [218]. Sequence dependent toxicity may result, in part, to an inter-residue Tg-purine hydrogen bond where 3' pyrimidines produce unstable structures.

Despite the fact that Tg was a strong block to replication in *E. coli*, it was weakly mutagenic, causing < 0.5% T→C transitions [219]. It was concluded that the *cis*-5*R*,6*S* Tg adduct was displaced laterally toward the major groove as compared to an unmodified thymine, perhaps increasing the likelihood of G•Tg wobble pairing, potentially explaining the observed T→C transitions [219]. An interesting possibility is that the *trans*-5*R*,6*R* lesion is responsible for the < 0.5% mutagenicity reported in *E. coli*. Specifically, for the *trans*-5*R*,6*R* epimer, the equatorial conformation of the Tg CH_3 group is predicted to be more energetically favorable (Table 5-2). In the RB69-Tg crystal structure, stacking of the adjacent 5′-template guanine was hindered by an axial Tg CH_3 [234]. The equatorial orientation of the Tg CH_3 group in the *trans*-5*R*,6*R* epimer may allow stacking of the 5' neighbor template guanine, presumably facilitating incorporation of the next incoming nucleotide into the growing primer strand.

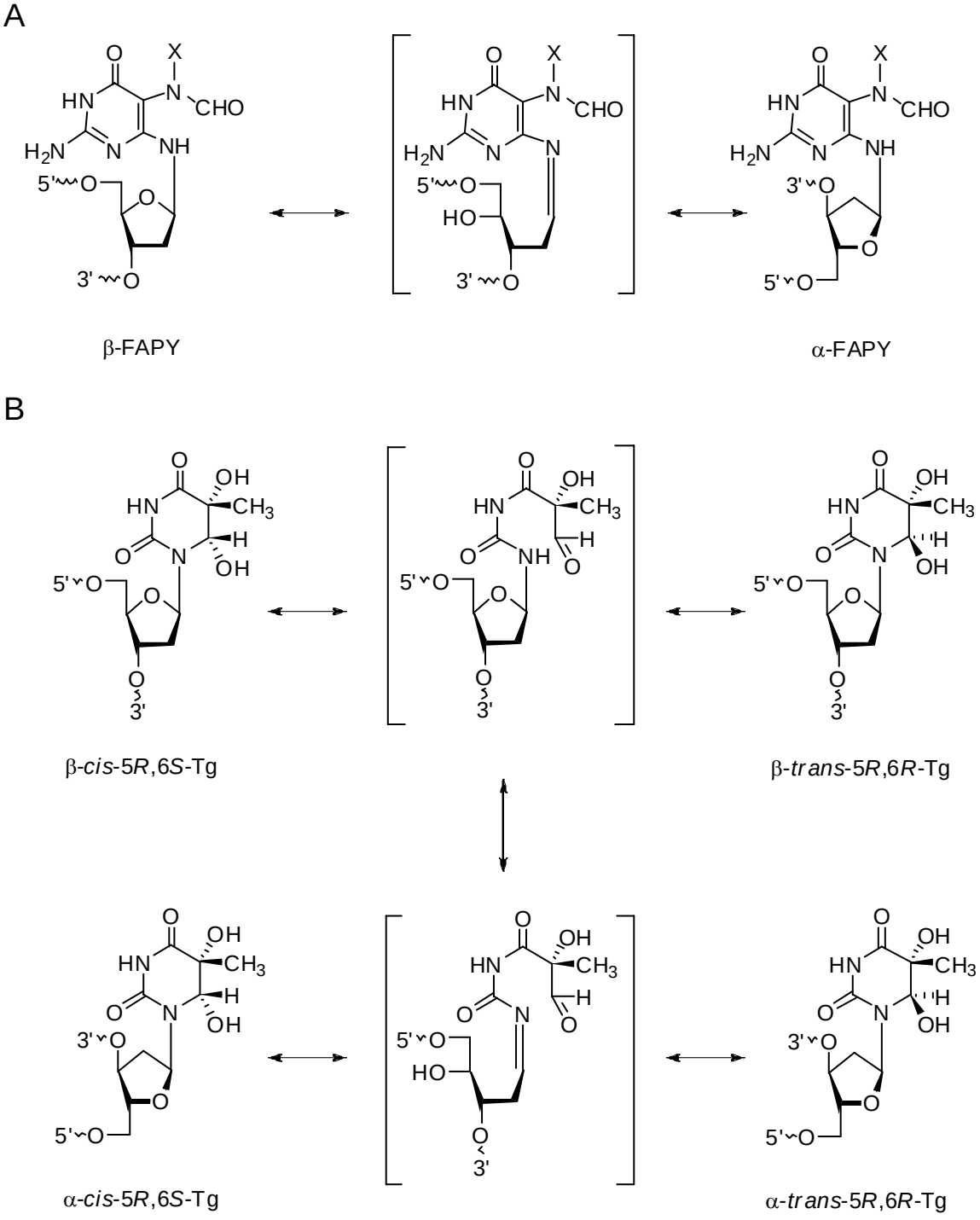
FAPY and Thymine Glycol Anomerization

It is of interest to compare the chemistry of the 5*R*-Tg adduct with that of the FAPY lesion. In oligodeoxynucleotides, two equilibrating AFB_1-FAPY species, separable by HPLC, were identified as anomers [31]. The configurational state of the AFB_1-FAPY adduct differs in ssDNA and dsDNA. This interconversion involves deoxyribose ring opening via formation of a transient iminium bond between C1' and the glycosidic nitrogen (Scheme 9-1) [47-50]. Similarly, the N1 position of the *cis-trans* Tg

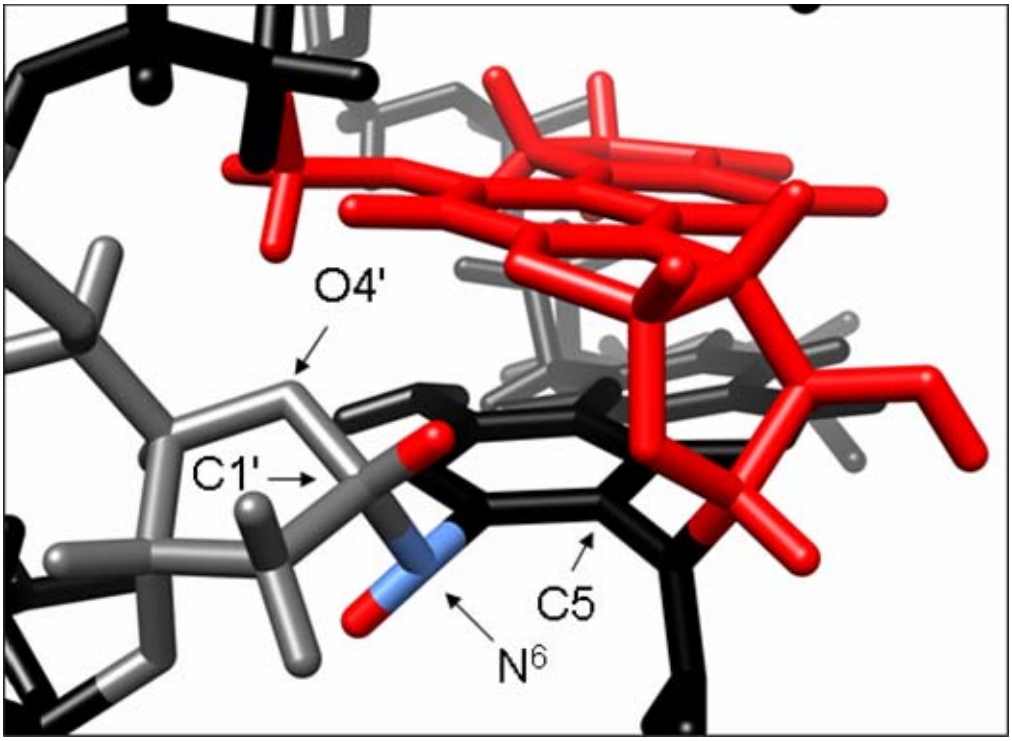
intermediate is a tertiary amine, allowing for the possibility of anomerization by the same pathway.

The rate of epimerization of *cis*-5*R*,6*S*-Tg was reported to be 5.84 x 10^{-3} min^{-1} at pH 7.4 and ambient temperature [37], which is fast compared to anomerization of FAPY lesions under the same conditions. HPLC and CGE monitoring of Tg modified ssDNA incubated under acidic conditions (pH 6.0) produced a single peak. The fact anomers were not observed suggests that there is a significant energy barrier to anomerization in Tg that is not present in FAPY samples.

Resonance delocalization may be a key aspect to anomerization of Tg and FAPY. In Tg the N1 lone pair is expected to delocalize with the adjacent carbonyl, much like in amides. If the nitrogen lone pair is delocalized, it will not be available to form the iminium bond necessary for deoxyribose ring opening (Scheme 9-1). This may represent a prohibitive barrier to anomerization. However, resonance stabilization also occurs in FAPYs. To estimate the extent of resonance delocalization, the length of the carbon nitrogen bond (C1'-N1 in Tg; C1'-N6 in FAPY) was determined by post DFT calculation of Tg and FAPY bases. The C-N bond of FAPY was calculated to be 1.36 Å as compared to 1.37 Å for the Tg intermediate (Appendix B). The 1% difference between the two measurements is inconclusive by itself, but would suggest the FAPY is more stable. This would indicate that FAPY will anomerize more slowly than the Tg intermediate. The pK_a of FAPY NH6 (21.6, -4.40) and Tg intermediate NH1 (9.26, -4.50) were estimated by partial charge distribution calculation [345-347] of the modified bases. These results



Scheme 9-1: Thymine glycol and FAPY lesion anomerization in DNA. Anomerization may occur in FAPY residues (X = AFB_1 or CH_3) by an iminium intermediate (Panel A). A similar iminium intermediate maybe possible during the *cis-trans* epimerization of thymine glycol lesions (Panel B).



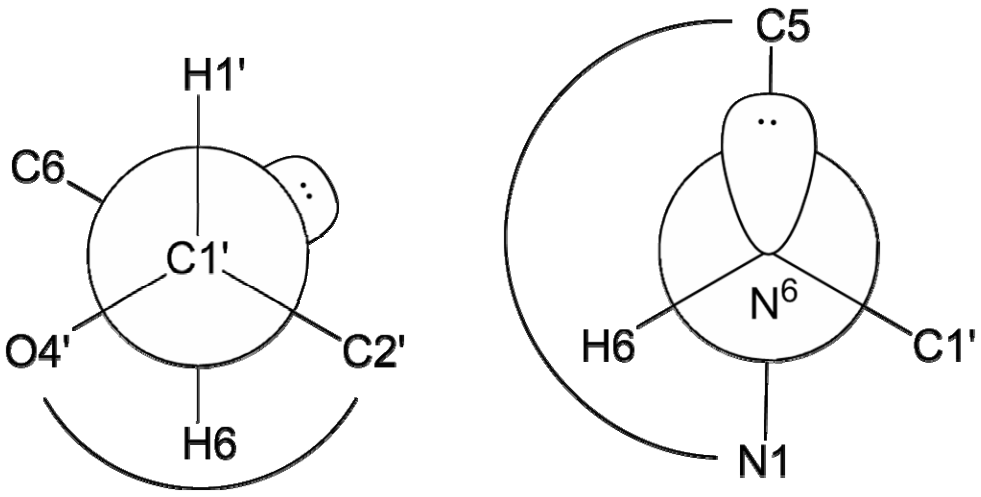


Figure 9-1: Newman projections of the α-AFB_1-FAPY C1'-N^6 and N^6-C6 bonds. Analysis of the rMD α-AFB_1-FAPY structure indicates that the N^6 lone pair is orthogonal to the plane of the FAPY base. This prohibits the lone pair from resonance delocalization with the ring, thus, it is free to participate in the transient iminium bond necessary for anomer epimerization.

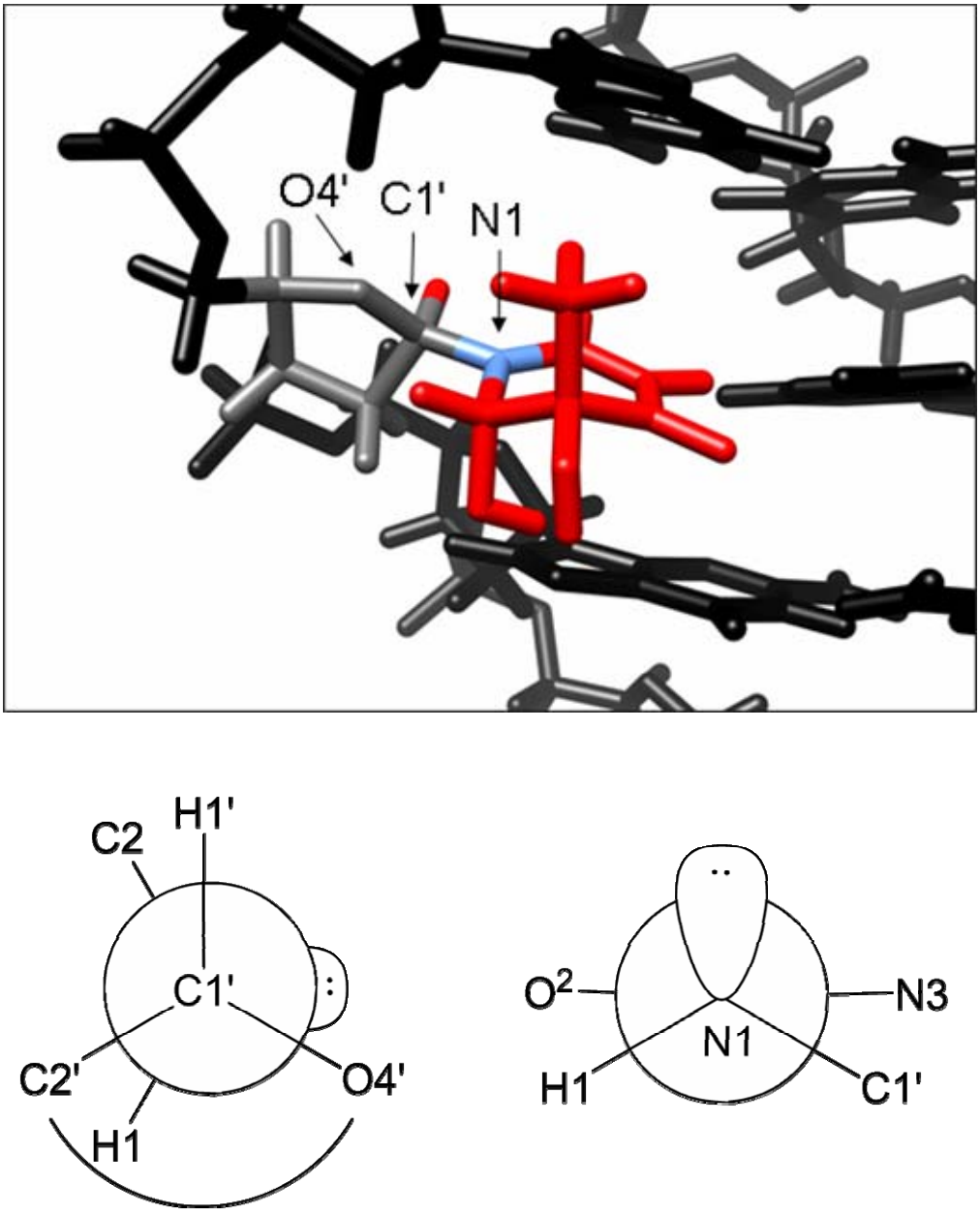


Figure 9-2: Newman projections of the C1'-N1 and N1-C2 bonds of the thymine glycol *cis-trans* intermediate. Analysis of the rMD *cis*-5*R*,6*S*-thymine glycol structure indicates that the N1 lone pair of the *cis-trans* Tg intermediate would be subject to resonance delocalization with the C2 carbonyl. Therefore, the N1 lone pair would be less available to participate in the C1'-N1 iminium bond necessary for anomer epimerization.

suggest that dissociation of FAPY NH6 may limit iminium bond formation prior to deoxyribose ring opening. Attempts to quantitate resonance stability relied on calculation of modified bases; it was considered that secondary structures maybe a factor in lesion specific anomerization. These data would suggest that FAPYs are less likely to anomerize than Tg contradictory to experimental results.

Careful analysis of the 3D structures of Tg and AFB_1-FAPY modified oligonucleotides suggests a stereoelectronic element to anomerization. Analysis of both α-AFB_1-FAPY and β-AFB_1-FAPY structures indicate the FAPY NH1 is *trans* to H1' in both anomer structures [115] (Chapter VI). Additionally, the N1 lone pair orbital is parallel to the plane of the FAPY base (Figure 9-1), indicating that N1 is sp3 hybridized. This configuration would prohibit lone pair conjugation with the ring system; however, it would be free to participate in iminium bond formation. Direct comparison with the Tg intermediate is not possible, but overall geometry in DNA is expected to be similar to a Tg lesion. In Tg the N1 is sp2 hybridized and postulated to remain sp2 hybridized in the Tg intermediate (Figure 9-2). Structural analysis suggests that the N1 lone pair orbital in the Tg intermediate is orthogonal to the Tg plane and has the proper geometry to conjugate with the adjacent carbonyl. This would reduce the availability of the lone pair to participate in an iminium bond and reduce the possibility of anomerization.

This stereo-electronic effect would be expected to be reduced in a highly mobile single strand or nucleoside environment. Although it has not been reported, Tg nucleoside may still anomerize. The fact that Tg anomers have not been observed suggests that the reaction is very slow. Considering that Tg lesions are efficiently

repaired and short lived *in vivo*, this introduces the question of whether the α anomer is biologically relevant in the case of slow anomerization.

APPENDIX A

NMR RESONANCE ASSIGNMENTS

Table A-1: Resonance assignments for Me-dGuo-FAPY in 5'-$A^1X^2C^3$-3' (furanose)

Spin System	*Residue*	*Atom*	*Nucleus*	*Chemical Shift(ppm)*
1	A^1	31P	^{31}P	-0.07
1	A^1	C1'	^{13}C	86.0
1	A^1	C2'	^{13}C	37.4
1	A^1	C3'	^{13}C	77.1
1	A^1	C4'	^{13}C	86.9
1	A^1	C5'	^{13}C	61.9
1	A^1	H1'	^{1}H	6.35
1	A^1	H2'	^{1}H	2.68
1	A^1	H2''	^{1}H	3.01
1	A^1	H3'	^{1}H	4.91
1	A^1	H4'	^{1}H	4.32
1	A^1	H5'	^{1}H	3.78
1	X^2	31P	^{31}P	-0.19
1	X^2	C1'	^{13}C	81.4
1	X^2	C2'	^{13}C	37.5
1	X^2	C3'	^{13}C	75.4
1	X^2	C4'	^{13}C	82.0
1	X^2	H1'	^{1}H	5.89
1	X^2	H2'	^{1}H	2.11
1	X^2	H2''	^{1}H	2.47
1	X^2	H3'	^{1}H	4.74
1	X^2	H4'	^{1}H	4.25
1	X^2	H5'	^{1}H	3.93
1	X^2	H5''	^{1}H	4.04
1	C^3	C1'	^{13}C	85.4
1	C^3	C2'	^{13}C	39.4
1	C^3	C3'	^{13}C	70.4
1	C^3	C4'	^{13}C	85.0
1	C^3	C5'	^{13}C	69.5
1	C^3	H1'	^{1}H	6.22
1	C^3	H2'	^{1}H	2.19
1	C^3	H2''	^{1}H	2.33
1	C^3	H3'	^{1}H	4.49
1	C^3	H4'	^{1}H	4.10

1	C^3	H5	^1H	5.96
1	C^3	H5'	^1H	3.89
1	C^3	H6	^1H	7.78
2	A^1	31P	^{31}P	-0.22
2	A^1	C1'	^{13}C	85.4
2	A^1	C2'	^{13}C	37.4
2	A^1	C3'	^{13}C	76.7
2	A^1	C5'	^{13}C	61.4
2	A^1	H1'	^1H	6.40
2	A^1	H2''	^1H	2.86
2	A^1	H3'	^1H	4.95
2	A^1	H4'	^1H	4.34
2	A^1	H5'	^1H	3.81
2	C^3	C1'	^{13}C	85.3
2	C^3	H1'	^1H	6.14
2	C^3	H2'	^1H	2.19
2	C^3	H2''	^1H	2.31
2	C^3	H3'	^1H	4.48
2	X^2	31P	^{31}P	-0.65
2	X^2	C1'	^{13}C	82.1
2	X^2	C2'	^{13}C	37.5
2	X^2	C3'	^{13}C	76.1
2	X^2	C4'	^{13}C	83.3
2	X^2	C5'	^{13}C	65.2
2	X^2	H1'	^1H	5.68
2	X^2	H2''	^1H	2.18
2	X^2	H3'	^1H	4.71
2	X^2	H4'	^1H	4.18
2	X^2	H5'	^1H	3.94
2	X^2	H5''	^1H	3.99
3	C^3	H1'	^1H	6.08
3	C^3	H2'	^1H	2.14
3	C^3	H2''	^1H	2.35
3	C^3	H3'	^1H	4.48
3	C^3	H4'	^1H	4.02
3	X^2	C4'	^{13}C	82.6
3	X^2	C5'	^{13}C	64.9
3	X^2	H1'	^1H	5.94
3	X^2	H3'	^1H	4.72
3	X^2	H4'	^1H	4.30
3	X^2	H5'	^1H	3.92
3	X^2	H5''	^1H	4.01
n	X^Y	CH_3	^1H	1.71
n	X^Y	CHO	^1H	8.30
n	X^Y	CHO	^1H	7.47

Table A-2: Resonance assignments for Me-dGuo-FAPY in 5'-$A^1X^2C^3$-3' (pyranose)

Spin System	*Residue*	*Atom*	*Nucleus*	*Chemical Shift(ppm)*
1	A^1	C1'	^{13}C	84.9
1	A^1	C2'	^{13}C	38.1
1	A^1	C3'	^{13}C	76.3
1	A^1	C4'	^{13}C	86.7
1	A^1	C5'	^{13}C	61.6
1	A^1	H1'	^{1}H	6.43
1	A^1	H2'	^{1}H	2.74
1	A^1	H2''	^{1}H	2.83
1	A^1	H3'	^{1}H	4.94
1	A^1	H4'	^{1}H	4.38
1	A^1	H5'	^{1}H	3.78
1	X^2	C1'	^{13}C	85.4
1	X^2	C2'	^{13}C	39.7
1	X^2	C3'	^{13}C	70.8
1	X^2	C4'	^{13}C	68.2
1	X^2	C5'	^{13}C	66.1
1	X^2	H1'	^{1}H	5.91
1	X^2	H2'	^{1}H	2.04
1	X^2	H2''	^{1}H	2.28
1	X^2	H3'	^{1}H	4.34
1	X^2	H4'	^{1}H	4.03
1	X^2	H5'	^{1}H	3.68
1	X^2	H5''	^{1}H	4.08
1	C^3	C1'	^{13}C	84.8
1	C^3	C2'	^{13}C	39.3
1	C^3	C3'	^{13}C	69.4
1	C^3	C5	^{13}C	95.7
1	C^3	H1'	^{1}H	6.09
1	C^3	H2'	^{1}H	2.12
1	C^3	H2''	^{1}H	2.25
1	C^3	H3'	^{1}H	4.43
1	C^3	H4'	^{1}H	4.04
1	C^3	H5	^{1}H	5.82
1	C^3	H5'	^{1}H	4.02
1	C^3	H6	^{1}H	7.69
1	C^3	31P	^{31}P	-0.45
2	X^2	C1'	^{13}C	74.6
2	X^2	C2'	^{13}C	34.7
2	X^2	C3'	^{13}C	71.0
2	X^2	C4'	^{13}C	69.7
2	X^2	C5'	^{13}C	69.4
2	X^2	H1'	^{1}H	5.54
2	X^2	H2'	^{1}H	2.03
2	X^2	H2''	^{1}H	2.21
2	X^2	H3'	^{1}H	4.68
2	X^2	H4'	^{1}H	4.25

2	X^2	H5'	^1H	3.89
2	X^2	P	^{31}P	-0.35
3	X^2	C1'	^{13}C	77.1
3	X^2	H1'	^1H	5.30
3	X^2	H2'	^1H	2.03
3	X^2	H2''	^1H	2.15
3	X^2	H3'	^1H	4.39
3	X^2	H5'	^1H	3.68
4	X^2	C1'	^{13}C	77.3
4	X^2	H1'	^1H	5.15
4	X^2	H2''	^1H	2.08
4	X^2	H3'	^1H	4.38
4	X^2	H4'	^1H	3.78
4	X^2	H5'	^1H	3.67
n	X^Y	CH_3	^1H	1.71
n	X^Y	CHO	^1H	8.30
n	X^Y	CHO	^1H	7.46

Table A-3: Resonance assignments for unmodified 5'-$C^1T^2A^3T^4G^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-3'

Primary Strand				**Complementary Strand**			
Residue	*Atom*	*Nucleus*	*Chemical Shift(ppm)*	*Residue*	*Atom*	*Nucleus*	*Chemical Shift(ppm)*
C^1	H1'	^1H	5.83	T^{11}	H1'	^1H	5.76
C^1	H2'	^1H	2.22	T^{11}	H2'	^1H	1.75
C^1	H2''	^1H	2.57	T^{11}	H2''	^1H	2.20
C^1	H3'	^1H	4.66	T^{11}	H3'	^1H	4.61
C^1	H4'	^1H	4.09	T^{11}	H4'	^1H	4.02
C^1	H5	^1H	5.91	T^{11}	H5'	^1H	3.64
C^1	H5''	^1H	3.80	T^{11}	H5''	^1H	3.62
C^1	H6	^1H	7.87	T^{11}	H6	^1H	7.36
C^1	C1'	^{13}C	83.7	T^{11}	CH_3	^1H	1.63
T^2	H1'	^1H	5.77	T^{11}	C1'	^{13}C	83.6
T^2	H2'	^1H	2.30	G^{12}	H1	^1H	12.68
T^2	H2''	^1H	2.61	G^{12}	H1'	^1H	5.34
T^2	H3'	^1H	4.92	G^{12}	H2'	^1H	2.73
T^2	H4'	^1H	4.23	G^{12}	H2''	^1H	2.77
T^2	H5'	^1H	4.23	G^{12}	H3'	^1H	4.97
T^2	H5''	^1H	4.07	G^{12}	H4'	^1H	4.30
T^2	H6	^1H	7.63	G^{12}	H5'	^1H	4.03
T^2	CH_3	^1H	1.71	G^{12}	H5''	^1H	3.94
T^2	C1'	^{13}C	83.2	G^{12}	H8	^1H	7.97
A^3	H1'	^1H	6.29	G^{12}	C1'	^{13}C	81.6

A^{3}	H2	^{1}H	7.36	A^{13}	H1'	^{1}H	6.00
A^{3}	H2'	^{1}H	2.71	A^{13}	H2	^{1}H	7.30
A^{3}	H2''	^{1}H	2.96	A^{13}	H2'	^{1}H	2.79
A^{3}	H3'	^{1}H	5.05	A^{13}	H2''	^{1}H	2.92
A^{3}	H4'	^{1}H	4.45	A^{13}	H3'	^{1}H	5.09
A^{3}	H5'	^{1}H	4.22	A^{13}	H4'	^{1}H	4.48
A^{3}	H5''	^{1}H	4.13	A^{13}	H5'	^{1}H	4.21
A^{3}	H61	^{1}H	7.69	A^{13}	H5''	^{1}H	4.15
A^{3}	H62	^{1}H	6.47	A^{13}	H62	^{1}H	6.00
A^{3}	H8	^{1}H	8.42	A^{13}	H8	^{1}H	8.23
A^{3}	C1'	^{13}C	82.2	A^{13}	C1'	^{13}C	81.9
T^{4}	H1'	^{1}H	5.66	A^{14}	H1'	^{1}H	6.21
T^{4}	H2'	^{1}H	1.96	A^{14}	H2	^{1}H	7.66
T^{4}	H2''	^{1}H	2.38	A^{14}	H2'	^{1}H	2.58
T^{4}	H3	^{1}H	13.33	A^{14}	H2''	^{1}H	2.94
T^{4}	H3'	^{1}H	4.85	A^{14}	H3'	^{1}H	4.84
T^{4}	H4'	^{1}H	4.16	A^{14}	H4'	^{1}H	4.49
T^{4}	H5'	^{1}H	4.15	A^{14}	H8	^{1}H	8.18
T^{4}	H6	^{1}H	7.11	A^{14}	H61	^{1}H	1.30
T^{4}	CH_3	^{1}H	1.42	A^{14}	H62	^{1}H	7.31
T^{4}	C1'	^{13}C	82.1	A^{14}	C1'	^{13}C	82.0
G^{5}	H1	^{1}H	12.2	T^{15}	H1'	^{1}H	5.86
G^{5}	H1'	^{1}H	5.62	T^{15}	H2'	^{1}H	2.01
G^{5}	H2'	^{1}H	2.68	T^{15}	H2''	^{1}H	2.57
G^{5}	H2''	^{1}H	2.77	T^{15}	H3	^{1}H	13.51
G^{5}	H3'	^{1}H	5.01	T^{15}	H3'	^{1}H	4.84
G^{5}	H4'	^{1}H	4.36	T^{15}	H4'	^{1}H	4.31
G^{5}	H5'	^{1}H	4.12	T^{15}	H5'	^{1}H	4.10
G^{5}	H5''	^{1}H	4.07	T^{15}	H6	^{1}H	7.10
G^{5}	H8	^{1}H	7.85	T^{15}	CH_3	^{1}H	1.30
G^{5}	C1'	^{13}C	81.1	T^{15}	C1'	^{13}C	82.2
A^{6}	H1'	^{1}H	6.22	C^{16}	H1'	^{1}H	5.65
A^{6}	H2	^{1}H	7.64	C^{16}	H2'	^{1}H	2.11
A^{6}	H2'	^{1}H	2.66	C^{16}	H2''	^{1}H	2.46
A^{6}	H2''	^{1}H	2.93	C^{16}	H3'	^{1}H	4.86
A^{6}	H3'	^{1}H	5.01	C^{16}	H4'	^{1}H	4.16
A^{6}	H4'	^{1}H	4.48	C^{16}	H41	^{1}H	6.73
A^{6}	H5'	^{1}H	4.22	C^{16}	H42	^{1}H	8.31
A^{6}	H5''	^{1}H	4.16	C^{16}	H5	^{1}H	5.58
A^{6}	H61	^{1}H	7.38	C^{16}	H5'	^{1}H	4.09
A^{6}	H62	^{1}H	5.93	C^{16}	H6	^{1}H	7.51
A^{6}	H8	^{1}H	8.19	C^{16}	C1'	^{13}C	83.4
A^{6}	C1'	^{13}C	82.4	A^{17}	H1'	^{1}H	6.18
T^{7}	H1'	^{1}H	5.96	A^{17}	H2	^{1}H	7.48
T^{7}	H2'	^{1}H	2.00	A^{17}	H2'	^{1}H	2.66
T^{7}	H2''	^{1}H	2.56	A^{17}	H2''	^{1}H	2.91
T^{7}	H3	^{1}H	13.76	A^{17}	H3'	^{1}H	5.01
T^{7}	H3'	^{1}H	4.84	A^{17}	H4'	^{1}H	4.41
T^{7}	H4'	^{1}H	4.18	A^{17}	H5'	^{1}H	4.17
T^{7}	H5'	^{1}H	4.13	A^{17}	H5''	^{1}H	4.09

T^{7}	H5''	^{1}H	4.09	A^{17}	H62	^{1}H	6.29
T^{7}	H6	^{1}H	7.20	A^{17}	H8	^{1}H	8.32
T^{7}	CH_3	^{1}H	1.29	A^{17}	C1'	^{13}C	82.3
T^{7}	C1'	^{13}C	82.4	T^{18}	H1'	^{1}H	5.52
T^{8}	H1'	^{1}H	6.08	T^{18}	H2'	^{1}H	1.97
T^{8}	H2'	^{1}H	2.04	T^{18}	H2''	^{1}H	2.30
T^{8}	H2''	^{1}H	2.45	T^{18}	H3	^{1}H	13.44
T^{8}	H3	^{1}H	13.88	T^{18}	H3'	^{1}H	4.84
T^{8}	H3'	^{1}H	4.87	T^{18}	H4'	^{1}H	4.40
T^{8}	H4'	^{1}H	4.21	T^{18}	H5'	^{1}H	4.24
T^{8}	H5'	^{1}H	4.14	T^{18}	H5''	^{1}H	4.07
T^{8}	H5''	^{1}H	4.08	T^{18}	H6	^{1}H	7.19
T^{8}	H6	^{1}H	7.39	T^{18}	CH_3	^{1}H	1.46
T^{8}	CH_3	^{1}H	1.59	T^{18}	C1'	^{13}C	82.0
T^{8}	C1'	^{13}C	82.5	A^{19}	H1'	^{1}H	6.05
C^{9}	H1'	^{1}H	5.70	A^{19}	H2'	^{1}H	2.67
C^{9}	H2'	^{1}H	2.02	A^{19}	H2''	^{1}H	2.87
C^{9}	H2''	^{1}H	2.31	A^{19}	H3'	^{1}H	5.01
C^{9}	H3'	^{1}H	4.83	A^{19}	H4'	^{1}H	4.39
C^{9}	H4'	^{1}H	4.08	A^{19}	H5'	^{1}H	4.13
C^{9}	H41	^{1}H	7.13	A^{19}	H5''	^{1}H	4.08
C^{9}	H42	^{1}H	8.51	A^{19}	H8	^{1}H	8.17
C^{9}	H5	^{1}H	5.72	A^{19}	C1'	^{13}C	82.2
C^{9}	H5'	^{1}H	4.03	G^{20}	H1'	^{1}H	5.99
C^{9}	H6	^{1}H	7.51	G^{20}	H2'	^{1}H	2.42
A^{10}	H1'	^{1}H	6.34	G^{20}	H2''	^{1}H	2.25
A^{10}	H2'	^{1}H	2.73	G^{20}	H3'	^{1}H	4.62
A^{10}	H2''	^{1}H	2.46	G^{20}	H4'	^{1}H	4.18
A^{10}	H3'	^{1}H	4.72	G^{20}	H5'	^{1}H	4.16
A^{10}	H4'	^{1}H	4.21	G^{20}	H5''	^{1}H	4.09
A^{10}	H5'	^{1}H	4.11	G^{20}	H8	^{1}H	7.69
A^{10}	H5''	^{1}H	4.08	G^{20}	C1'	^{13}C	81.7
A^{10}	H8	^{1}H	8.28				
A^{10}	C1'	^{13}C	82.7				

Table A-4: Resonance assignments for Me-dGuo-FAPY (furanose) modified 5'-$C^{1}T^{2}A^{3}T^{4}\underline{X}^{5}A^{6}T^{7}T^{8}C^{9}A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-3'

Primary Strand			**Complementary Strand**		
Residue	*Atom*	*Chemical Shift(ppm)*	*Residue*	*Atom*	*Chemical Shift(ppm)*
C^{1}	H1'	5.87	T^{11}	H1'	5.77
C^{1}	H2'1	2.21	T^{11}	H2'1	1.76
C^{1}	H2'2	2.57	T^{11}	H2'2	2.21

C^{1}	H3'	4.67	T^{11}	H3'	4.62
C^{1}	H5	5.94	T^{11}	H5'1	3.64
C^{1}	H5'1	3.79	T^{11}	H6	7.38
C^{1}	H5'2	3.84	T^{11}	CH_3	1.66
C^{1}	H6	7.88	G^{12}	H1'	5.35
T^{2}	H1'	5.74	G^{12}	H2'1	2.74
T^{2}	H2'1	2.30	G^{12}	H2'2	2.78
T^{2}	H2'2	2.59	G^{12}	H3'	4.98
T^{2}	H3'	4.91	G^{12}	H8	7.98
T^{2}	H6	7.61	A^{13}	H1'	6.01
T^{2}	CH_3	1.72	A^{13}	H2'1	2.78
A^{3}	H1'	6.36	A^{13}	H2'2	2.94
A^{3}	H2'1	2.85	A^{13}	H3'	5.08
A^{3}	H2'2	2.59	A^{13}	H8	8.23
A^{3}	H3'	5.06	A^{14}	H1'	6.20
A^{3}	H8	8.45	A^{14}	H2'1	2.57
T^{4}	H1'	6.03	A^{14}	H2'2	2.93
T^{4}	H2'1	1.83	A^{14}	H3'	5.01
T^{4}	H2'2	2.68	A^{14}	H8	8.17
T^{4}	H3'	4.85	T^{15}	H1'	5.97
T^{4}	H6	7.17	T^{15}	H2'1	2.00
T^{4}	CH_3	1.30	T^{15}	H2'2	2.48
X^{5}	CHO	8.29	T^{15}	H3'	4.84
X^{5}	H1'	5.96	T^{15}	H6	7.10
X^{5}	H2'1	2.51	T^{15}	CH_3	1.19
X^{5}	H3'	4.78	C^{16}	H1'	5.39
X^{5}	CH_3	3.53	C^{16}	H2'1	1.93
A^{6}	H1'	6.36	C^{16}	H2'2	2.31
A^{6}	H2'1	2.75	C^{16}	H3'	4.84
A^{6}	H2'2	2.99	C^{16}	H5	5.60
A^{6}	H3'	5.03	C^{16}	H6	7.50
A^{6}	H8	8.33	A^{17}	H1'	6.30
T^{7}	H1'	5.91	A^{17}	H2'1	2.96
T^{7}	H2'1	2.00	A^{17}	H2'2	2.83
T^{7}	H2'2	2.59	A^{17}	H3'	5.07
T^{7}	H3'	4.84	A^{17}	H8	8.27
T^{7}	H6	7.22	T^{18}	H1'	5.53
T^{7}	CH_3	1.44	T^{18}	H2'1	2.00
T^{8}	H1'	6.10	T^{18}	H2'2	2.28
T^{8}	H2'1	2.05	T^{18}	H3'	4.86
T^{8}	H2'2	2.47	T^{18}	H6	7.24
T^{8}	H3'	4.88	T^{18}	CH_3	1.54
T^{8}	H6	7.39	A^{19}	H1'	6.07
T^{8}	CH_3	1.61	A^{19}	H2'1	2.71
C^{9}	H1'	5.72	A^{19}	H2'2	2.89
C^{9}	H2'1	2.01	A^{19}	H3'	5.03
C^{9}	H2'2	2.31	A^{19}	H8	8.20
C^{9}	H3'	4.83	G^{20}	H1'	6.00
C^{9}	H5	5.73	G^{20}	H2'1	2.43
C^{9}	H6	7.51	G^{20}	H2'2	2.25

A^{10}	H1'	6.35	G^{20}	H3'	4.63
A^{10}	H2'1	2.74	G^{20}	H8	7.71
A^{10}	H2'2	2.46			
A^{10}	H3'	4.73			
A^{10}	H8	8.29			

Table A-5: Resonance assignments for α -AFB_1-FAPY modified 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-3'

Primary Strand				Complementary Strand			
Residue	*Atom*	*Nucleus*	*Chemical Shift(ppm)*	*Residue*	*Atom*	*Nucleus*	*Chemical Shift(ppm)*
C^1	H1'	^1H	5.81	T^{11}	H1'	^1H	5.73
C^1	H2'	^1H	2.19	T^{11}	H2'	^1H	1.70
C^1	H2''	^1H	2.52	T^{11}	H2''	^1H	2.14
C^1	H3'	^1H	4.62	T^{11}	H3'	^1H	4.55
C^1	H4'	^1H	4.05	T^{11}	H4'	^1H	3.96
C^1	H42	^1H	7.78	T^{11}	H5'	^1H	3.57
C^1	H5	^1H	5.86	T^{11}	H5''	^1H	3.58
C^1	H5'	^1H	3.74	T^{11}	H6	^1H	7.30
C^1	H5''	^1H	3.77	T^{11}	CH_3	^1H	1.57
C^1	H6	^1H	7.82	T^{11}	C1'	^{13}C	83.6
C^1	C1'	^{13}C	83.2	G^{12}	H1	^1H	12.67
T^2	H1'	^1H	5.79	G^{12}	H1'	^1H	5.33
T^2	H2'	^1H	2.24	G^{12}	H2'	^1H	2.67
T^2	H2''	^1H	2.54	G^{12}	H2''	^1H	2.72
T^2	H3	^1H	13.51	G^{12}	H3'	^1H	4.91
T^2	H3'	^1H	4.87	G^{12}	H4'	^1H	4.24
T^2	H4'	^1H	4.18	G^{12}	H5'	^1H	3.97
T^2	H5'	^1H	4.04	G^{12}	H5''	^1H	3.88
T^2	H5''	^1H	4.01	G^{12}	H8	^1H	7.91
T^2	H6	^1H	7.56	G^{12}	C1'	^{13}C	81.6
T^2	CH_3	^1H	1.65	A^{13}	H1'	^1H	5.93
T^2	C1'	^{13}C	83.2	A^{13}	H2	^1H	7.70
A^3	H1'	^1H	6.24	A^{13}	H2'	^1H	2.72
A^3	H2	^1H	7.44	A^{13}	H2''	^1H	2.82
A^3	H2'	^1H	2.57	A^{13}	H3'	^1H	5.01
A^3	H2''	^1H	2.89	A^{13}	H4'	^1H	4.41
A^3	H3'	^1H	4.98	A^{13}	H5'	^1H	4.21
A^3	H4'	^1H	4.39	A^{13}	H5''	^1H	4.09
A^3	H5'	^1H	4.23	A^{13}	H61	^1H	7.32
A^3	H5''	^1H	4.12	A^{13}	H62	^1H	6.04
A^3	H61	^1H	6.45	A^{13}	H8	^1H	8.14
A^3	H8	^1H	8.32	A^{13}	C1'	^{13}C	81.9

A^{3}	C1'	^{13}C	82.2	A^{14}	H1'	^{1}H	6.11
T^{4}	H1'	^{1}H	5.98	A^{14}	H2	^{1}H	7.69
T^{4}	H2'	^{1}H	2.23	A^{14}	H2'	^{1}H	2.50
T^{4}	H2''	^{1}H	2.36	A^{14}	H2''	^{1}H	2.80
T^{4}	H3	^{1}H	13.37	A^{14}	H3'	^{1}H	4.86
T^{4}	H3'	^{1}H	4.91	A^{14}	H4'	^{1}H	4.40
T^{4}	H4'	^{1}H	4.23	A^{14}	H5'	^{1}H	4.21
T^{4}	H5'	^{1}H	3.93	A^{14}	H5''	^{1}H	4.19
T^{4}	H5''	^{1}H	4.10	A^{14}	H61	^{1}H	7.34
T^{4}	H6	^{1}H	7.41	A^{14}	H62	^{1}H	6.03
T^{4}	CH_3	^{1}H	1.50	A^{14}	H8	^{1}H	8.03
T^{4}	C1'	^{13}C	82.1	A^{14} B	H8	^{1}H	8.05
X^{5}	H1'	^{1}H	6.33	A^{14}	C1'	^{13}C	82.1
X^{5}	H2'	^{1}H	2.93	T^{15}	H1'	^{1}H	5.73
X^{5}	H2''	^{1}H	2.79	T^{15}	H2'	^{1}H	2.07
X^{5}	H3	^{1}H	11.82	T^{15}	H2''	^{1}H	2.38
X^{5}	H3'	^{1}H	5.03	T^{15}	H3	^{1}H	13.69
X^{5}	H4'	^{1}H	4.62	T^{15}	H3'	^{1}H	4.69
X^{5}	H5'	^{1}H	4.20	T^{15}	H4'	^{1}H	4.19
X^{5}	H5''	^{1}H	4.45	T^{15}	H5'	^{1}H	4.08
X^{5}	H8	^{1}H	8.33	T^{15}	H5''	^{1}H	4.16
X^{5}	HN21	^{1}H	6.32	T^{15}	H6	^{1}H	6.95
X^{5}	HN22	^{1}H	6.78	T^{15}	CH_3	^{1}H	1.16
X^{5}	C1'	^{13}C	82.7	T^{15} B	H6	^{1}H	6.96
A^{6}	H1'	^{1}H	6.28	T^{15} B	CH_3	^{1}H	1.13
A^{6}	H2	^{1}H	7.73	T^{15}	C1'	^{13}C	82.2
A^{6}	H2'	^{1}H	2.74	C^{16}	H1'	^{1}H	5.81
A^{6}	H2''	^{1}H	2.88	C^{16}	H2'	^{1}H	1.85
A^{6}	H3'	^{1}H	5.01	C^{16}	H2''	^{1}H	2.42
A^{6}	H4'	^{1}H	4.44	C^{16}	H3'	^{1}H	4.71
A^{6}	H5'	^{1}H	4.00	C^{16}	H4'	^{1}H	4.20
A^{6}	H5''	^{1}H	4.12	C^{16}	H41	^{1}H	5.58
A^{6}	H61	^{1}H	7.33	C^{16}	H42	^{1}H	8.38
A^{6}	H62	^{1}H	6.11	C^{16}	H5	^{1}H	5.25
A^{6}	H8	^{1}H	8.30	C^{16}	H5'	^{1}H	4.04
A^{6}	C1'	^{13}C	82.6	C^{16}	H6	^{1}H	7.28
T^{7}	H1'	^{1}H	5.99	C^{16} B	H1'	^{1}H	5.83
T^{7}	H2'	^{1}H	2.03	C^{16} B	H2'	^{1}H	1.88
T^{7}	H2''	^{1}H	2.52	C^{16} B	H2''	^{1}H	2.43
T^{7}	H3	^{1}H	13.88	C^{16}	C1'	^{13}C	83.1
T^{7}	H3'	^{1}H	4.82	A^{17}	H1'	^{1}H	5.91
T^{7}	H4'	^{1}H	4.25	A^{17}	H2	^{1}H	7.72
T^{7}	H5'	^{1}H	4.18	A^{17}	H2'	^{1}H	2.69
T^{7}	H6	^{1}H	7.23	A^{17}	H2''	^{1}H	2.79
T^{7}	CH_3	^{1}H	1.26	A^{17}	H3'	^{1}H	4.84
T^{7}	C1'	^{13}C	82.5	A^{17}	H4'	^{1}H	4.19
T^{8}	H1'	^{1}H	6.03	A^{17}	H5'	^{1}H	3.86
T^{8}	H2'	^{1}H	2.01	A^{17}	H5''	^{1}H	4.11
T^{8}	H2''	^{1}H	2.39	A^{17}	H61	^{1}H	7.12
T^{8}	H3	^{1}H	13.92	A^{17}	H62	^{1}H	6.89

T^{8}	H3'	^{1}H	4.82	A^{17}	H8	^{1}H	8.33
T^{8}	H4'	^{1}H	4.17	A^{17}	C1'	^{13}C	82.1
T^{8}	H5'	^{1}H	4.04	T^{18}	H1'	^{1}H	5.58
T^{8}	H5''	^{1}H	4.08	T^{18}	H2'	^{1}H	1.94
T^{8}	H6	^{1}H	7.36	T^{18}	H2''	^{1}H	2.34
T^{8}	CH_3	^{1}H	1.56	T^{18}	H3	^{1}H	13.35
T^{8}	C1'	^{13}C	82.5	T^{18}	H3'	^{1}H	4.71
C^{9}	H1'	^{1}H	5.69	T^{18}	H4'	^{1}H	4.21
C^{9}	H2'	^{1}H	1.99	T^{18}	H5'	^{1}H	4.02
C^{9}	H2''	^{1}H	2.26	T^{18}	H5''	^{1}H	4.11
C^{9}	H3'	^{1}H	4.77	T^{18}	H6	^{1}H	7.26
C^{9}	H4'	^{1}H	4.03	T^{18}	CH_3	^{1}H	1.35
C^{9}	H41	^{1}H	7.15	T^{18}	C1'	^{13}C	83.3
C^{9}	H42	^{1}H	8.51	A^{19}	H1'	^{1}H	6.04
C^{9}	H5	^{1}H	5.70	A^{19}	H2'	^{1}H	2.63
C^{9}	H5'	^{1}H	4.00	A^{19}	H2''	^{1}H	2.83
C^{9}	H6	^{1}H	7.46	A^{19}	H3'	^{1}H	4.98
A^{10}	H1'	^{1}H	6.28	A^{19}	H4'	^{1}H	4.37
A^{10}	H2'	^{1}H	2.66	A^{19}	H5'	^{1}H	4.02
A^{10}	H2''	^{1}H	2.41	A^{19}	H5''	^{1}H	4.21
A^{10}	H3'	^{1}H	4.66	A^{19}	H8	^{1}H	8.14
A^{10}	H4'	^{1}H	4.15	A^{19}	C1'	^{13}C	82.2
A^{10}	H5'	^{1}H	4.05	G^{20}	H1	^{1}H	12.66
A^{10}	H5''	^{1}H	4.04	G^{20}	H1'	^{1}H	5.95
A^{10}	H8	^{1}H	8.22	G^{20}	H2'	^{1}H	2.38
A^{10}	C1'	^{13}C	82.8	G^{20}	H2''	^{1}H	2.20
				G^{20}	H3'	^{1}H	4.58
				G^{20}	H4'	^{1}H	4.13
AFB^{21}	H2A1	^{1}H	0.98	G^{20}	H5'	^{1}H	4.08
AFB^{21}	H2A2	^{1}H	1.65	G^{20}	H5''	^{1}H	4.21
AFB^{21}	H31	^{1}H	1.85	G^{20}	H8	^{1}H	7.66
AFB^{21}	H32	^{1}H	2.28	G^{20}	C1'	^{13}C	81.5
AFB^{21}	H5B	^{1}H	5.73				
AFB^{21}	H6a	^{1}H	6.14				
AFB^{21}	H8A	^{1}H	5.78				
AFB^{21}	H9	^{1}H	6.10				
AFB^{21}	H9a	^{1}H	3.47				
AFB^{21}	HM	^{1}H	3.51				
AFB^{21} B	H9a	^{1}H	3.42				

Table A-6: Resonance assignments for α -AFB_1-FAPY modified 5'-$C^{1}T^{2}X^{3}A^{4}$-5'

Residue	*Atom*	*Chemical Shift (ppm)*	*Residue*	*Atom*	*Chemical Shift (ppm)*
C^{1}	H1'	5.81	A^{4}	H1'	6.27

C^{1}	H2'	2.02	A^{4}	H2'	2.69
C^{1}	H2''	2.40	A^{4}	H2''	2.51
C^{1}	H3'	4.53	A^{4}	H3'	4.62
C^{1}	H4'	3.92	A^{4}	H4'	4.17
C^{1}	H5	5.54	A^{4}	H5'	4.00
C^{1}	H5'	3.58	A^{4}	H5''	3.96
C^{1}	H6	7.48	A^{4}	H8	8.29
T^{2}	H1'	6.19	AFB^{5}	H5	5.72
T^{2}	H2'	2.21	AFB^{5}	H5B	5.71
T^{2}	H2''	2.37	AFB^{5}	H6a	6.13
T^{2}	H3'	4.71	AFB^{5}	H8A	5.66
T^{2}	H4'	4.28	AFB^{5}	H9	6.16
T^{2}	H5'	3.92	AFB^{5}	H9a	3.66
T^{2}	H5''	3.89	AFB^{5}	CH^{3}	3.76
T^{2}	H6	7.54			
T^{2}	M7	1.70			
X^{3}	H1'	5.62			
X^{3}	H2'	2.30			
X^{3}	H2''	1.88			
X^{3}	H3'	4.54			
X^{3}	H4'	4.11			
X^{3}	H5'	3.82			
X^{3}	H5''	3.74			
X^{3}	H8	8.27			

Table A-7: Resonance assignments for unmodified 5'-$G^{1}T^{2}G^{3}C^{4}G^{5}T^{6}G^{7}T^{8}T^{9}T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3'

Primary Strand			**Complementary Strand**		
Residue	*Atom*	*Chemical Shift(ppm)*	*Residue*	*Atom*	*Chemical Shift(ppm)*
G^{1}	H1'	5.88	A^{13}	H1'	6.03
G^{1}	H2'	2.50	A^{13}	H2'	2.43
G^{1}	H2''	2.66	A^{13}	H2''	2.61
G^{1}	H3'	4.68	A^{13}	H3'	4.68
G^{1}	H5'	4.09	A^{13}	H4'	4.09
G^{1}	H8	7.82	A^{13}	H5'	3.94
T^{2}	CH_3	1.24	A^{13}	H8	8.04
T^{2}	H1'	5.78	C^{14}	H1'	6.03
T^{2}	H2'	2.08	C^{14}	H2'	1.86
T^{2}	H2''	2.42	C^{14}	H2''	2.08
T^{2}	H3	13.67	C^{14}	H3'	4.99
T^{2}	H3'	4.77	C^{14}	H42	8.20
T^{2}	H4'	4.12	C^{14}	H5	5.34
T^{2}	H5'	4.00	C^{14}	H6	7.26
T^{2}	H5''	3.95	A^{15}	H1'	5.62

T^{2}	H6	7.23	A^{15}	H2	7.10
G^{3}	H1	12.59	A^{15}	H2'	2.59
G^{3}	H1'	5.77	A^{15}	H2''	2.69
G^{3}	H2'	2.47	A^{15}	H3'	4.89
G^{3}	H2''	2.59	A^{15}	H4'	4.22
G^{3}	H22	6.38	A^{15}	H5'	4.06
G^{3}	H3'	4.85	A^{15}	H62	6.04
G^{3}	H4'	4.23	A^{15}	H8	8.06
G^{3}	H5'	4.07	A^{16}	H1'	5.73
G^{3}	H8	7.73	A^{16}	H2	6.99
C^{4}	H1'	5.55	A^{16}	H2'	2.49
C^{4}	H2'	1.89	A^{16}	H2''	2.69
C^{4}	H2''	2.25	A^{16}	H3'	4.91
C^{4}	H3'	4.69	A^{16}	H4'	4.29
C^{4}	H4'	4.05	A^{16}	H5'	4.09
C^{4}	H42	8.13	A^{16}	H5''	4.06
C^{4}	H5	5.15	A^{16}	H61	7.27
C^{4}	H5'	4.01	A^{16}	H62	6.00
C^{4}	H6	7.15	A^{16}	H8	7.96
C^{4}	H1'	5.55	A^{17}	H1'	5.89
G^{5}	H1	12.63	A^{17}	H2	7.39
G^{5}	H1'	5.81	A^{17}	H2'	2.38
G^{5}	H2'	2.45	A^{17}	H2''	2.68
G^{5}	H2''	2.63	A^{17}	H3'	4.84
G^{5}	H22	6.39	A^{17}	H4'	4.09
G^{5}	H3'	4.81	A^{17}	H5'	4.28
G^{5}	H4'	4.21	A^{17}	H61	7.19
G^{5}	H5'	3.91	A^{17}	H8	7.89
G^{5}	H5''	4.06	C^{18}	H1'	5.33
G^{5}	H8	7.72	C^{18}	H2'	1.70
T^{6}	CH_3	1.31	C^{18}	H2''	2.14
T^{6}	H1'	5.66	C^{18}	H3'	4.60
T^{6}	H2'	1.97	C^{18}	H4'	3.97
T^{6}	H2''	2.36	C^{18}	H42	7.82
T^{6}	H3	13.46	C^{18}	H5	5.00
T^{6}	H3'	4.72	C^{18}	H5'	3.85
T^{6}	H4'	4.05	C^{18}	H6	6.98
T^{6}	H5'	3.99	A^{19}	H1'	5.98
T^{6}	H6	7.00	A^{19}	H2	7.42
G^{7}	H1	12.36	A^{19}	H2'	2.47
G^{7}	H1'	5.83	A^{19}	H2''	2.68
G^{7}	H2'	2.65	A^{19}	H3'	4.83
G^{7}	H2''	2.45	A^{19}	H4'	3.97
G^{7}	H22	6.37	A^{19}	H5'	4.22
G^{7}	H3'	4.79	A^{19}	H61	5.80
G^{7}	H4'	4.12	A^{19}	H8	7.99
G^{7}	H5'	4.23	C^{20}	H1'	5.39
G^{7}	H5''	4.01	C^{20}	H2'	1.78
G^{7}	H8	7.68	C^{20}	H2''	2.15
T^{8}	CH_3	1.17	C^{20}	H3'	4.83

T^{8}	H1'	5.88	C^{20}	H4'	3.97
T^{8}	H2'	2.00	C^{20}	H42	7.99
T^{8}	H2''	1.46	C^{20}	H5	5.02
T^{8}	H3	13.86	C^{20}	H6	7.04
T^{8}	H3'	4.71	G^{21}	H1	12.71
T^{8}	H4'	4.02	G^{21}	H1'	5.69
T^{8}	H5'	4.12	G^{21}	H2'	2.45
T^{8}	H6	7.10	G^{21}	H2''	2.55
T^{9}	H1'	5.98	G^{21}	H22	6.30
T^{9}	H2'	2.45	G^{21}	H3'	4.81
T^{9}	H2''	1.57	G^{21}	H4'	4.19
T^{9}	H3	13.85	G^{21}	H5'	3.98
T^{9}	H3'	4.74	G^{21}	H8	7.67
T^{9}	H4'	4.07	C^{22}	H1'	5.47
T^{9}	H5'	4.02	C^{22}	H2'	1.86
T^{9}	H6	7.31	C^{22}	H2''	2.22
T^{10}	H1'	5.69	C^{22}	H3'	4.67
T^{10}	H2'	1.84	C^{22}	H4'	4.00
T^{10}	H2''	2.23	C^{22}	H42	8.25
T^{10}	H3	13.77	C^{22}	H5	5.25
T^{10}	H3'	4.75	C^{22}	H6	7.19
T^{10}	H4'	4.05	A^{23}	H1'	6.10
T^{10}	H5'	3.97	A^{23}	H2	7.68
T^{10}	H6	7.16	A^{23}	H2'	2.53
G^{11}	H1	12.54	A^{23}	H2''	2.72
G^{11}	H1'	5.93	A^{23}	H3'	4.85
G^{11}	H2'	2.57	A^{23}	H4'	4.24
G^{11}	H2''	1.50	A^{23}	H5'	4.11
G^{11}	H22	6.53	A^{23}	H5''	3.92
G^{11}	H3'	4.86	A^{23}	H8	8.11
G^{11}	H4'	4.26	C^{24}	H1'	5.95
G^{11}	H5'	3.95	C^{24}	H2'	1.95
G^{11}	H5''	4.10	C^{24}	H2''	1.98
G^{11}	H8	7.83	C^{24}	H3'	4.33
T^{12}	H1'	6.11	C^{24}	H4'	3.87
T^{12}	H2'	2.11	C^{24}	H5	5.29
T^{12}	H3'	4.40	C^{24}	H6	7.25
T^{12}	H4'	3.95			
T^{12}	H6	7.30			

Table A-8: Resonance assignments for *cis*-5*R*,6*S*-thymine glycol modified 5'-$G^{1}T^{2}G^{3}C^{4}G^{5}\underline{Tg}^{6}G^{7}T^{8}T^{9}T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}\underline{A}^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3'

Primary Strand	Complementary Strand

Residue	*Atom*	*Chemical Shift(ppm)*	*Residue*	*Atom*	*Chemical Shift(ppm)*
G^{1}	H1'	5.87	A^{13}	H1'	6.02
G^{1}	H2'	2.47	A^{13}	H2'	2.42
G^{1}	H2''	2.64	A^{13}	H2''	2.58
G^{1}	H3'	4.67	A^{13}	H3'	4.66
G^{1}	H4'	4.07	A^{13}	H4'	4.07
G^{1}	H5'	3.62	A^{13}	H5'	3.55
G^{1}	H5''	3.60	A^{13}	H5''	3.59
G^{1}	H8	7.79	A^{13}	H8	8.01
T^{2}	H1'	5.76	C^{14}	H1'	5.04
T^{2}	H2'	2.04	C^{14}	H2'	1.84
T^{2}	H2''	2.38	C^{14}	H2''	2.09
T^{2}	H3	13.67	C^{14}	H3'	4.62
T^{2}	H3'	4.75	C^{14}	H4'	3.93
T^{2}	H4'	4.09	C^{14}	H42	8.19
T^{2}	H5'	3.98	C^{14}	H5	5.37
T^{2}	H6	7.20	C^{14}	H5'	3.92
T^{2}	CH_3	1.24	C^{14}	H6	7.25
G^{3}	H1	12.62	A^{15}	H1'	5.61
G^{3}	H1'	5.75	A^{15}	H2	7.14
G^{3}	H2'	2.46	A^{15}	H2'	2.57
G^{3}	H2''	2.57	A^{15}	H2''	2.67
G^{3}	H22	6.36	A^{15}	H3'	4.88
G^{3}	H3'	4.83	A^{15}	H4'	4.20
G^{3}	H4'	4.22	A^{15}	H5'	3.84
G^{3}	H5'	3.93	A^{15}	H5''	3.95
G^{3}	H5''	4.00	A^{15}	H62	6.05
G^{3}	H8	7.71	A^{15}	H8	8.05
C^{4}	H1'	5.61	A^{16}	H1'	5.71
C^{4}	H2'	1.85	A^{16}	H2	7.06
C^{4}	H2''	2.25	A^{16}	H2'	2.46
C^{4}	H3'	4.64	A^{16}	H2''	2.65
C^{4}	H4'	4.04	A^{16}	H3'	4.89
C^{4}	H42	8.12	A^{16}	H4'	4.26
C^{4}	H5	5.14	A^{16}	H5'	4.03
C^{4}	H5'	4.00	A^{16}	H5''	4.05
C^{4}	H5''	4.08	A^{16}	H61	7.31
C^{4}	H6	7.11	A^{16}	H62	5.87
C^{4}	H1'	5.56	A^{16}	H8	7.94
G^{5}	H1	12.91	A^{17}	H1'	5.86
G^{5}	H1'	5.88	A^{17}	H2	7.44
G^{5}	H2'	2.31	A^{17}	H2'	2.36
G^{5}	H2''	2.48	A^{17}	H2''	2.63
G^{5}	H3'	4.73	A^{17}	H3'	4.81
G^{5}	H4'	4.15	A^{17}	H4'	4.25
G^{5}	H5'	3.95	A^{17}	H5'	4.07
G^{5}	H8	7.64	A^{17}	H62	5.70
Tg^{6}	H1'	5.63	A^{17}	H8	7.86
Tg^{6}	H2'	2.07	C^{18}	H1'	5.53

Tg^{6}	H2''	2.28	C^{18}	H2'	1.71
Tg^{6}	H3'	4.57	C^{18}	H2''	2.19
Tg^{6}	H4'	4.04	C^{18}	H3'	4.57
Tg^{6}	H5'	3.79	C^{18}	H4'	4.00
Tg^{6}	H5''	3.97	C^{18}	H41	7.03
Tg^{6}	H6 *cis*	4.58	C^{18}	H42	7.83
Tg^{6}	H6 *trans*	4.91	C^{18}	H5	5.02
Tg^{6}	CH_3 *cis*	0.49	C^{18}	H5'	3.95
Tg^{6}	CH_3 *trans*	1.24	C^{18}	H5''	4.07
G^{7}	H1	12.28	C^{18}	H6	6.99
G^{7}	H1'	5.91	A^{19}	H1'	5.87
G^{7}	H2'	2.50	A^{19}	H2	7.46
G^{7}	H2''	2.72	A^{19}	H2'	2.26
G^{7}	H22	6.52	A^{19}	H2''	2.51
G^{7}	H3'	4.78	A^{19}	H3'	4.74
G^{7}	H4'	4.27	A^{19}	H4'	4.11
G^{7}	H5'	3.94	A^{19}	H5'	3.89
G^{7}	H5''	4.02	A^{19}	H5''	3.94
G^{7}	H8	7.75	A^{19}	H62	6.14
T^{8}	H1'	5.91	A^{19}	H8	7.83
T^{8}	H2'	1.99	C^{20}	H1'	5.38
T^{8}	H2''	2.45	C^{20}	H2'	1.86
T^{8}	H3	13.79	C^{20}	H2''	2.16
T^{8}	H3'	4.73	C^{20}	H3'	4.64
T^{8}	H4'	4.09	C^{20}	H4'	3.97
T^{8}	H5'	4.02	C^{20}	H5	5.10
T^{8}	H6	7.12	C^{20}	H5'	3.92
T^{8}	CH_3	1.17	C^{20}	H6	7.12
T^{9}	H1'	5.95	G^{21}	H1	12.67
T^{9}	H2'	1.99	G^{21}	H1'	5.72
T^{9}	H2''	2.43	G^{21}	H2'	2.43
T^{9}	H3	13.85	G^{21}	H2''	2.54
T^{9}	H3'	4.73	G^{21}	H22	6.30
T^{9}	H4'	4.02	G^{21}	H3'	4.80
T^{9}	H6	7.29	G^{21}	H4'	4.18
T^{9}	CH_3	1.47	G^{21}	H5'	3.86
T^{10}	H1'	5.67	G^{21}	H5''	3.95
T^{10}	H2'	1.82	G^{21}	H8	7.65
T^{10}	H2''	2.20	C^{22}	H1'	5.50
T^{10}	H3	13.80	C^{22}	H2'	1.85
T^{10}	H3'	4.73	C^{22}	H2''	2.22
T^{10}	H4'	3.96	C^{22}	H3'	4.67
T^{10}	H5'	4.05	C^{22}	H4'	4.00
T^{10}	H6	7.15	C^{22}	H42	8.25
T^{10}	CH_3	1.57	C^{22}	H5	5.24
G^{11}	H1	12.53	C^{22}	H5'	4.03
G^{11}	H1'	5.93	C^{22}	H6	7.19
G^{11}	H2''	2.55	A^{23}	H1'	6.11
G^{11}	H22	6.53	A^{23}	H2	7.72
G^{11}	H3'	4.84	A^{23}	H2'	2.53

G^{11}	H4'	4.24	A^{23}	H2''	2.71
G^{11}	H5'	3.97	A^{23}	H3'	4.85
G^{11}	H8	7.81	A^{23}	H4'	4.23
T^{12}	H1'	6.09	A^{23}	H5'	3.95
T^{12}	H2''	2.10	A^{23}	H5''	3.99
T^{12}	H3'	4.39	A^{23}	H61	8.03
T^{12}	H4'	3.93	A^{23}	H8	8.10
T^{12}	H5'	3.96	C^{24}	H1'	5.95
T^{12}	H6	7.29	C^{24}	H2'	1.94
T^{12}	CH_3	1.50	C^{24}	H2''	1.99
			C^{24}	H3'	4.32
			C^{24}	H4'	3.87
			C^{24}	H5	5.31
			C^{24}	H5'	4.10
			C^{24}	H5''	3.91
			C^{24}	H6	7.25

Table A-9: Resonance assignments for *cis*-5*R*,6*S*-thymine glycol modified 5'-$G^1T^2G^3C^4G^5\underline{Tg}^6G^7T^8T^9T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}\underline{G}^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3'

Primary Strand			**Complementary Strand**		
Residue	*Atom*	*Chemical Shift(ppm)*	*Residue*	*Atom*	*Chemical Shift(ppm)*
G^1	H1'	5.85	A^{13}	H1'	6.01
G^1	H2'	2.48	A^{13}	H2	7.82
G^1	H2''	2.63	A^{13}	H2'	2.43
G^1	H3'	4.66	A^{13}	H2''	2.58
G^1	H4'	4.07	A^{13}	H3'	4.66
G^1	H5'	3.59	A^{13}	H4'	4.07
G^1	H5''	3.62	A^{13}	H5'	3.55
G^1	H8	7.79	A^{13}	H8	8.02
T^2	H1'	5.74	C^{14}	H1'	5.02
T^2	H2'	2.02	C^{14}	H2'	1.83
T^2	H2''	2.36	C^{14}	H2''	2.07
T^2	H3	13.62	C^{14}	H3'	4.62
T^2	H3'	4.74	C^{14}	H4'	3.93
T^2	H4'	4.08	C^{14}	H5	5.35
T^2	H5'	3.97	C^{14}	H5'	3.90
T^2	H6	7.20	C^{14}	H5''	3.88
T^2	CH_3	1.23	C^{14}	H6	7.25
G^3	H1	12.61	C^{14}	H62	8.16
G^3	H1'	5.73	A^{15}	H1'	5.59
G^3	H2'	2.45	A^{15}	H2	7.10
G^3	H2''	2.54	A^{15}	H2'	2.57

G^{3}	H22	6.37	A^{15}	H2''	2.66
G^{3}	H3'	4.82	A^{15}	H3'	4.88
G^{3}	H4'	4.21	A^{15}	H4'	4.20
G^{3}	H5'	3.84	A^{15}	H5'	3.84
G^{3}	H8	7.71	A^{15}	H5''	3.94
C^{4}	H1'	5.62	A^{15}	H62	6.00
C^{4}	H2'	1.71	A^{15}	H8	8.05
C^{4}	H2''	2.19	A^{16}	H1'	5.70
C^{4}	H3'	4.66	A^{16}	H2	7.03
C^{4}	H4'	4.02	A^{16}	H2'	2.46
C^{4}	H42	8.10	A^{16}	H2''	2.65
C^{4}	H5	5.15	A^{16}	H3'	4.89
C^{4}	H5'	3.97	A^{16}	H4'	4.26
C^{4}	H5''	4.05	A^{16}	H5'	4.04
C^{4}	H6	7.08	A^{16}	H5''	4.05
G^{5}	H1	12.82	A^{16}	H61	7.29
G^{5}	H1'	5.89	A^{16}	H62	5.82
G^{5}	H2'	2.40	A^{16}	H8	7.94
G^{5}	H2''	2.56	A^{17}	H1'	5.87
G^{5}	H3'	4.79	A^{17}	H2	7.47
G^{5}	H4'	4.18	A^{17}	H2'	2.35
G^{5}	H5'	3.91	A^{17}	H2''	2.62
G^{5}	H5''	3.87	A^{17}	H3'	4.80
G^{5}	H8	7.72	A^{17}	H4'	4.25
G^{5}	H5'	4.10	A^{17}	H5'	4.06
Tg^{6}	H1'	5.46	A^{17}	H62	5.67
Tg^{6}	H2'	2.06	A^{17}	H8	7.86
Tg^{6}	H2''	2.06	C^{18}	H1'	5.56
Tg^{6}	H3'	4.53	C^{18}	H2'	1.72
Tg^{6}	H4'	4.01	C^{18}	H2''	2.19
Tg^{6}	H6	4.70	C^{18}	H3'	4.53
Tg^{6}	CH_3	0.91	C^{18}	H4'	4.01
G^{7}	H1	12.67	C^{18}	H42	7.80
G^{7}	H1'	5.88	C^{18}	H5	4.99
G^{7}	H2'	2.49	C^{18}	H5'	3.92
G^{7}	H2''	2.70	C^{18}	H5''	4.05
G^{7}	H22	6.33	C^{18}	H6	6.95
G^{7}	H3'	4.77	G^{19}	H1'	5.60
G^{7}	H4'	4.22	G^{19}	H2'	2.15
G^{7}	H5'	3.84	G^{19}	H2''	2.36
G^{7}	H5''	3.84	G^{19}	H3'	4.68
G^{7}	H8	7.81	G^{19}	H4'	4.02
T^{8}	H1'	5.93	G^{19}	H5'	3.88
T^{8}	H2'	2.00	G^{19}	H5''	3.92
T^{8}	H2''	2.45	G^{19}	H8	7.42
T^{8}	H3	13.76	C^{20}	H1'	5.44
T^{8}	H3'	4.73	C^{20}	H2'	1.89
T^{8}	H4'	4.09	C^{20}	H2''	2.18
T^{8}	H5'	4.05	C^{20}	H3'	4.66
T^{8}	H5''	4.00	C^{20}	H4'	3.98

T^{8}	H6	7.13	C^{20}	H5	5.17
T^{8}	CH_3	1.16	C^{20}	H5'	3.88
T^{9}	H1'	5.96	C^{20}	H6	7.19
T^{9}	H2'	2.00	G^{21}	H1	12.78
T^{9}	H2''	2.44	G^{21}	H1'	5.73
T^{9}	H3	13.84	G^{21}	H2'	2.47
T^{9}	H3'	4.73	G^{21}	H2''	2.56
T^{9}	H4'	4.02	G^{21}	H22	6.30
T^{9}	H5'	3.97	G^{21}	H3'	4.81
T^{9}	H5''	4.02	G^{21}	H4'	4.20
T^{9}	H6	7.30	G^{21}	H5'	3.98
T^{9}	CH_3	1.47	G^{21}	H5''	3.84
T^{10}	H1'	5.67	G^{21}	H8	7.72
T^{10}	H2'	1.83	C^{22}	H1'	5.48
T^{10}	H2''	2.21	C^{22}	H2'	1.84
T^{10}	H3	13.76	C^{22}	H2''	2.21
T^{10}	H3'	4.73	C^{22}	H3'	4.66
T^{10}	H4'	4.05	C^{22}	H4'	4.02
T^{10}	H5'	3.96	C^{22}	H42	8.22
T^{10}	H5''	3.93	C^{22}	H5	5.28
T^{10}	H6	7.16	C^{22}	H6	7.19
T^{10}	CH_3	1.57	A^{23}	H1'	6.10
G^{11}	H1	12.50	A^{23}	H2	7.69
G^{11}	H1'	5.91	A^{23}	H2'	2.53
G^{11}	H2'	2.54	A^{23}	H2''	2.71
G^{11}	H22	6.49	A^{23}	H3'	4.85
G^{11}	H3'	4.84	A^{23}	H4'	4.24
G^{11}	H4'	4.24	A^{23}	H5'	3.99
G^{11}	H5'	3.97	A^{23}	H5''	3.94
G^{11}	H8	7.80	A^{23}	H8	8.10
T^{12}	H1'	6.08	C^{24}	H1'	5.92
T^{12}	H2'	2.10	C^{24}	H2'	1.94
T^{12}	H3'	4.38	C^{24}	H2''	1.99
T^{12}	H4'	3.93	C^{24}	H3'	4.31
T^{12}	H5'	4.08	C^{24}	H4'	3.86
T^{12}	H6	7.26	C^{24}	H5	5.24
T^{12}	CH_3	1.47	C^{24}	H5'	3.90
			C^{24}	H5''	3.87
			C^{24}	H6	7.21

APPENDIX B

MOLECULAR DYNAMICS TOPOLOGY OF NON-STANDARD BASES

cis-5R,6S-thymine glycol

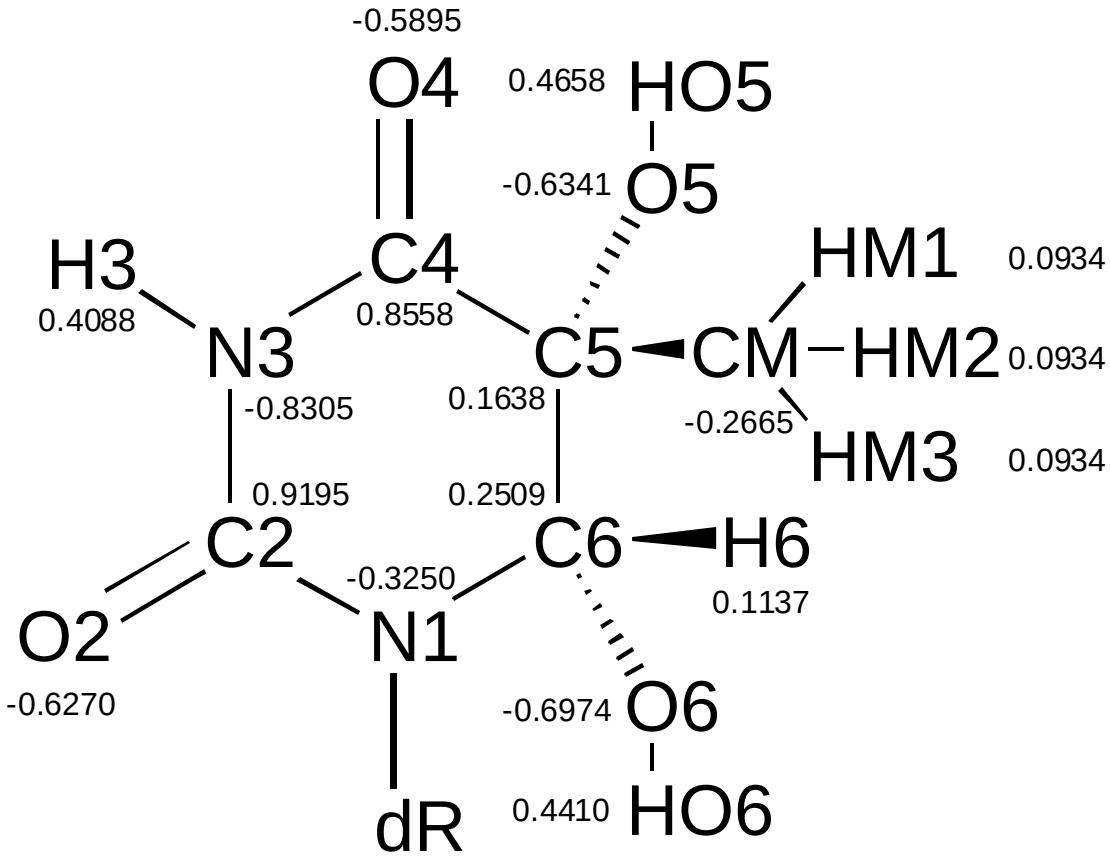


Figure B-1: Atomic charges on the cis-5R,6S-thymine glycol lesion calculated by Gaussian 03 and used in rMD simulations.

File B-1: Parameter and topology file of *cis*-5R,6S-thymine glycol lesion used in AMBER calculations

---

```
%VERSION  VERSION_STAMP = V0001.000  DATE = 01/30/08  16:25:27
%FLAG TITLE :  cis-5R,6S-thymine glycol (CH3 axial)
%FORMAT(20a4)
TG
%FLAG POINTERS
%FORMAT(10I8)
      36    10    14    23    32    35    57    48     0     0
     199     1    23    35    48    14    30     9    13     0
       0     0     0     0     0     0     0     0    36     0
       0
%FLAG ATOM_NAME
%FORMAT(20a4)
P   O2P O1P O5' C5' H5'1 H5'2 C4' H4' O4' C1' H1' C2' H2'1 H2'2 N1 C2 O2 N3 H3
C4 O4   C5   O5  HO5 CM  HM1 HM2 HM3 C6 H6 O6  HO6 C3' H3' O3'
%FLAG CHARGE
%FORMAT(5E16.8)
  2.12453796E+01 -1.41423270E+01 -1.41423270E+01 -9.02732742E+00
         -1.25733870E-01
  1.37396142E+00  1.37396142E+00  2.96841267E+00  2.14294248E+00
         -6.72585093E+00
  1.23911640E+00  3.28730292E+00 -1.55618442E+00  1.30836114E+00
          1.30836114E+00
 -6.21744876E+00  1.55800665E+01 -1.11156030E+01 -1.54215325E+01
          7.78766435E+00
  1.52976209E+01 -1.07620904E+01  1.03867110E-01 -1.26316984E+01
          8.97994944E+00
 -6.40513845E+00  2.09738673E+00  2.09738673E+00  2.09738673E+00
          8.63554797E+00
  1.15893828E+00 -1.21305851E+01  8.68474818E+00  1.29924999E+00
         1.79489655E+00
 -9.53390736E+00
%FLAG MASS
%FORMAT(5E16.8)
  3.09700000E+01  1.60000000E+01  1.60000000E+01  1.60000000E+01
1.20100000E+01
  1.00800000E+00  1.00800000E+00  1.20100000E+01  1.00800000E+00
1.60000000E+01
  1.20100000E+01  1.00800000E+00  1.20100000E+01  1.00800000E+00
1.00800000E+00
  1.40100000E+01  1.20100000E+01  1.60000000E+01  1.40100000E+01
1.00800000E+00
```

```
 1.20100000E+01  1.60000000E+01  1.20100000E+01  1.60000000E+01
1.00800000E+00
 1.20100000E+01  1.00800000E+00  1.00800000E+00  1.00800000E+00
1.20100000E+01
 1.00800000E+00  1.60000000E+01  1.00800000E+00  1.20100000E+01
1.00800000E+00
 1.60000000E+01
%FLAG ATOM_TYPE_INDEX
%FORMAT(10I8)
1     2     2     3     4     5     5     4     5     3
4     5     4     5     5     6     7     2     6     8
7     2     4     9    10     4     5     5     5     4
5     9    10     4     5     3
%FLAG NUMBER_EXCLUDED_ATOMS
%FORMAT(10I8)
 7     3     2     7    10     5     4    11     6    11
15     7     8     5     4    13     9     4     7     3
11     4    10     8     2     6     3     2     1     3
 2     1     1     2     1     1
%FLAG NONBONDED_PARM_INDEX
%FORMAT(10I8)
 1     2     4     7    11    16    22    29    37    46
 2     3     5     8    12    17    23    30    38    47
 4     5     6     9    13    18    24    31    39    48
 7     8     9    10    14    19    25    32    40    49
11    12    13    14    15    20    26    33    41    50
16    17    18    19    20    21    27    34    42    51
22    23    24    25    26    27    28    35    43    52
29    30    31    32    33    34    35    36    44    53
37    38    39    40    41    42    43    44    45    54
46    47    48    49    50    51    52    53    54    55
%FLAG RESIDUE_LABEL
%FORMAT(20a4)
TG
%FLAG RESIDUE_POINTER
%FORMAT(10I8)
      1
%FLAG BOND_FORCE_CONSTANT
%FORMAT(5E16.8)
 4.56400000E+02  3.11600000E+02  3.01500000E+02  3.37300000E+02
3.03100000E+02
 3.37300000E+02  3.30600000E+02  3.37300000E+02  4.78200000E+02
6.48000000E+02
 4.10200000E+02  3.28300000E+02  3.14100000E+02  3.69600000E+02
%FLAG BOND_EQUIL_VALUE
%FORMAT(5E16.8)
```

 1.50300000E+00 1.63600000E+00 1.43900000E+00 1.09200000E+00
1.53500000E+00
 1.09200000E+00 1.46000000E+00 1.09200000E+00 1.34500000E+00
1.21400000E+00
 1.00900000E+00 1.50800000E+00 1.42600000E+00 9.74000000E-01
%FLAG ANGLE_FORCE_CONSTANT
%FORMAT(5E16.8)
 7.76000000E+01 4.51000000E+01 4.31000000E+01 5.09000000E+01
6.78000000E+01
 4.64000000E+01 6.32000000E+01 3.94000000E+01 6.21000000E+01
5.09000000E+01
 7.01650000E+01 4.64000000E+01 6.39000000E+01 6.40000000E+01
4.64000000E+01
 4.98000000E+01 6.59000000E+01 3.94000000E+01 7.58000000E+01
7.54000000E+01
 7.02490000E+01 4.92000000E+01 6.74000000E+01 6.79000000E+01
6.84000000E+01
 6.38000000E+01 6.80000000E+01 4.71000000E+01 6.77000000E+01
5.11000000E+01
%FLAG ANGLE_EQUIL_VALUE
%FORMAT(5E16.8)
 2.05041368E+00 2.04587583E+00 2.03627651E+00 1.89717371E+00
1.89228679E+00
 1.92073567E+00 1.93085858E+00 1.89106506E+00 1.97937875E+00
1.89717371E+00
 1.95520593E+00 1.92073567E+00 2.11795796E+00 1.96559065E+00
1.92073567E+00
 1.91113635E+00 1.95703853E+00 1.89106506E+00 2.12982620E+00
1.94953361E+00
 1.96253632E+00 2.06751792E+00 2.08793828E+00 2.00974750E+00
1.91043822E+00
 1.92911325E+00 2.14867576E+00 1.88774893E+00 1.90991462E+00
1.91113635E+00
%FLAG DIHEDRAL_FORCE_CONSTANT
%FORMAT(5E16.8)
 3.83000000E-01 1.05000000E+00 1.56000000E-01 0.00000000E+00
2.50000000E+00
 1.67000000E-01 0.00000000E+00 1.00000000E+00 1.05000000E+01
%FLAG DIHEDRAL_PERIODICITY
%FORMAT(5E16.8)
 3.00000000E+00 2.00000000E+00 3.00000000E+00 2.00000000E+00
2.00000000E+00
 3.00000000E+00 2.00000000E+00 2.00000000E+00 2.00000000E+00
%FLAG DIHEDRAL_PHASE
%FORMAT(5E16.8)

 0.00000000E+00 3.14159400E+00 0.00000000E+00 0.00000000E+00
3.14159400E+00
 0.00000000E+00 3.14159400E+00 3.14159400E+00 3.14159400E+00
%FLAG SOLTY
%FORMAT(5E16.8)
 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00
0.00000000E+00
 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00
0.00000000E+00
 0.00000000E+00 0.00000000E+00 0.00000000E+00
%FLAG LENNARD_JONES_ACOEF
%FORMAT(5E16.8)
 6.02589390E+06 1.64263766E+06 3.79876399E+05 1.58759528E+06
3.70622491E+05
 3.61397723E+05 2.54188684E+06 6.47841731E+05 6.28541240E+05
1.04308023E+06
 2.54238516E+05 5.44261042E+04 5.33379252E+04 9.71708117E+04
7.51607703E+03
 2.45746558E+06 6.06829342E+05 5.89818288E+05 9.95480466E+05
8.96776989E+04
 9.44293233E+05 2.25370349E+06 5.74393458E+05 5.57281136E+05
9.24822270E+05
 8.61541883E+04 8.82619071E+05 8.19971662E+05 8.41065839E+03
1.02595236E+03
 1.03954408E+03 2.56678134E+03 1.07193646E+02 2.12601181E+03
2.27577561E+03
 1.39982777E-01 1.98683736E+06 4.71003287E+05 4.58874091E+05
7.91544157E+05
 6.82786631E+04 7.44975864E+05 7.01803794E+05 1.40467023E+03
5.81803229E+05
 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00
0.00000000E+00
 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00
0.00000000E+00
%FLAG LENNARD_JONES_BCOEF
%FORMAT(5E16.8)
 2.19561270E+03 1.16041466E+03 5.64885984E+02 1.08210555E+03
5.29252520E+02
 4.95732238E+02 1.22636552E+03 6.26720080E+02 5.85549272E+02
6.75612247E+02
 2.38716853E+02 1.11805549E+02 1.04986921E+02 1.26919150E+02
2.17257828E+01
 1.34630496E+03 6.77220874E+02 6.33305958E+02 7.36907417E+02
1.36131731E+02
 8.01323529E+02 1.08732781E+03 5.55666448E+02 5.19163331E+02
5.99015525E+02

```
 1.12529845E+02  6.53361429E+02  5.31102864E+02  4.34187588E+01
1.53505284E+01
 1.46567808E+01  2.06278363E+01  2.59456373E+00  2.09604198E+01
1.82891803E+01
 9.37598976E-02  1.27682121E+03  6.29300710E+02  5.89183300E+02
6.93079947E+02
 1.25287818E+02  7.50714425E+02  6.14502845E+02  1.79702257E+01
6.99746810E+02
 0.00000000E+00  0.00000000E+00  0.00000000E+00  0.00000000E+00
0.00000000E+00
 0.00000000E+00  0.00000000E+00  0.00000000E+00  0.00000000E+00
0.00000000E+00
%FLAG BONDS_INC_HYDROGEN
%FORMAT(10I8)
  12    15     4    12    18     4    21    24     4    30
  33     6    36    39     8    36    42     8    54    57
  11    69    72    14    75    78     8    75    81     8
  75    84     8    87    90     6    93    96    14    99
102     4
%FLAG BONDS_WITHOUT_HYDROGEN
%FORMAT(10I8)
   0     3     1     0     6     1     0     9     2     9
  12     3    12    21     5    21    27     3    21    99
   5    27    30     3    30    36     5    30    45     7
  36    99     5    45    48     9    45    87     7    48
  51    10    48    54     9    54    60     9    60    63
  10    60    66    12    66    69    13    66    75     5
  66    87     5    87    93    13    99   105     3
%FLAG ANGLES_INC_HYDROGEN
%FORMAT(10I8)
   9    12    15     4     9    12    18     4    12    21
  24     6    15    12    18     8    15    12    21     6
  18    12    21     6    21    99   102     6    24    21
  27     4    24    21    99     6    27    30    33    10
  30    36    39    12    30    36    42    12    33    30
  36    15    33    30    45    16    36    99   102     6
  39    36    42    18    39    36    99    12    42    36
  99    12    45    87    90    16    48    54    57    22
  57    54    60    22    66    69    72    28    66    75
  78    12    66    75    81    12    66    75    84    12
  66    87    90    15    78    75    81    18    78    75
  84    18    81    75    84    18    87    93    96    28
  90    87    93    30   102    99   105     4
%FLAG ANGLES_WITHOUT_HYDROGEN
%FORMAT(10I8)
   0     9    12     1     3     0     6     2     3     0
```

```
  9    3    6    0    9    3    9   12   21    5
 12   21   27    5   12   21   99    7   21   27
 30    9   21   99   36    7   21   99  105    5
 27   21   99    5   27   30   36    5   27   30
 45   11   30   36   99    7   30   45   48   13
 30   45   87   14   36   30   45   17   36   99
105    5   45   48   51   19   45   48   54   20
 45   87   66   17   45   87   93   21   48   45
 87   13   48   54   60   23   51   48   54   19
 54   60   63   19   54   60   66   24   60   66
 69   25   60   66   75   26   60   66   87   26
 63   60   66   27   66   87   93   29   69   66
 75   29   69   66   87   29   75   66   87    7
%FLAG DIHEDRALS_INC_HYDROGEN
%FORMAT(10I8)
  0    9   12   15    1    0    9   12   18    1
  9   12   21   24    3   12   21   99  102    3
 15   12   21   24    3   15   12   21   27    3
 15   12   21   99    3   18   12   21   24    3
 18   12   21   27    3   18   12   21   99    3
 21   27   30   33    1   21   99   36   39    3
 21   99   36   42    3   24   21   27   30    1
 24   21   99   36    3   24   21   99  102    3
 24   21   99  105    3   27   21   99  102    3
 27   30   36   39    3   27   30   36   42    3
 30   36   99  102    3   30   45   87   90    4
 33   30   36   39    3   33   30   36   42    3
 33   30   36   99    3   33   30   45   48    4
 33   30   45   87    4   39   36   30   45    3
 39   36   99  102    3   39   36   99  105    3
 42   36   30   45    3   42   36   99  102    3
 42   36   99  105    3   45   48   54   57    5
 45   87   93   96    6   48   45   87   90    4
 51   48   54   57    5   57   54   60   63    5
 57   54   60   66    5   60   66   69   72    6
 60   66   75   78    3   60   66   75   81    3
 60   66   75   84    3   60   66   87   90    3
 66   87   93   96    6   69   66   75   78    3
 69   66   75   81    3   69   66   75   84    3
 69   66   87   90    3   72   69   66   75    6
 72   69   66   87    6   75   66   87   90    3
 78   75   66   87    3   81   75   66   87    3
 84   75   66   87    3   90   87   93   96    6
 48   60  -54  -57    8
%FLAG DIHEDRALS_WITHOUT_HYDROGEN
%FORMAT(10I8)
```

```
  0     9    12    21     1     3     0     9    12     2
  6     0     9    12     2     9    12    21    27     3
  9    12    21    99     3    12    21    27    30     1
 12    21    99    36     3    12    21    99   105     3
 21    27   -30    36     1    21    27    30    45     1
 21    99   -36    30     3    27    21   -99    36     3
 27    21    99   105     3    27    30   -36    99     3
 27    30    45    48     4    27    30    45    87     4
 30    27   -21    99     1    30    36    99   105     3
 30    45    48    51     5    30    45    48    54     5
 30    45    87    66     4    30    45    87    93     4
 36    30    45    48     4    36    30    45    87     4
 45    30    36    99     3    45    48    54    60     5
 45    87   -66    60     3    45    87    66    69     3
 45    87    66    75     3    48    45    87    66     4
 48    45    87    93     4    48    54    60    63     5
 48    54   -60    66     5    51    48    45    87     5
 51    48    54    60     5    54    48    45    87     5
 54    60    66    69     7    54    60    66    75     7
 54    60   -66    87     7    60    66    87    93     3
 63    60    66    69     7    63    60    66    75     7
 63    60    66    87     7    69    66    87    93     3
 75    66    87    93     3    48    30   -45   -87     8
 45    54   -48   -51     9    54    66   -60   -63     9
%FLAG EXCLUDED_ATOMS_LIST
%FORMAT(10I8)
   2     3     4     5     6     7     8     3     4     5
   4     5     5     6     7     8     9    10    34     6
   7     8     9    10    11    13    34    35    36     7
   8     9    10    34     8     9    10    34     9    10
  11    12    13    14    15    16    34    35    36    10
  11    13    34    35    36    11    12    13    14    15
  16    17    30    34    35    36    12    13    14    15
  16    17    18    19    23    30    31    32    34    35
  36    13    14    15    16    17    30    34    14    15
  16    17    30    34    35    36    15    16    34    35
  36    16    34    35    36    17    18    19    20    21
  23    24    26    30    31    32    33    34    18    19
  20    21    22    23    30    31    32    19    20    21
  30    20    21    22    23    24    26    30    21    22
  23    22    23    24    25    26    27    28    29    30
  31    32    23    24    26    30    24    25    26    27
  28    29    30    31    32    33    25    26    27    28
  29    30    31    32    26    30    27    28    29    30
  31    32    28    29    30    29    30    30    31    32
  33    32    33    33     0    35    36    36     0
```

%FLAG HBOND_ACOEF
%FORMAT(5E16.8)

%FLAG HBOND_BCOEF
%FORMAT(5E16.8)

%FLAG HBCUT
%FORMAT(5E16.8)

%FLAG AMBER_ATOM_TYPE
%FORMAT(20a4)
P O2 O2 OS CT H1 H1 CT H1 OS CT H2 CT HC HC N C O N H
C O CT OH HO CT HC HC HC CT H2 OH HO CT H1 OS
%FLAG TREE_CHAIN_CLASSIFICATION
%FORMAT(20a4)
M E E M M E E M E S 3 E B E E B B E B E
B E B S E 3 E E E B E S E M E M
%FLAG JOIN_ARRAY
%FORMAT(10I8)
    0   0   0   0   0   0   0   0   0   0
    0   0   0   0   0   0   0   0   0   0
    0   0   0   0   0   0   0   0   0   0
    0   0   0   0   0   0
%FLAG IROTAT
%FORMAT(10I8)
    0   0   0   0   0   0   0   0   0   0
    0   0   0   0   0   0   0   0   0   0
    0   0   0   0   0   0   0   0   0   0
    0   0   0   0   0   0
%FLAG RADIUS_SET
%FORMAT(1a80)
modified Bondi radii (mbondi)
%FLAG RADII
%FORMAT(5E16.8)
 1.85000000E+00 1.50000000E+00 1.50000000E+00 1.50000000E+00
1.70000000E+00
 1.30000000E+00 1.30000000E+00 1.70000000E+00 1.30000000E+00
1.50000000E+00
 1.70000000E+00 1.30000000E+00 1.70000000E+00 1.30000000E+00
1.30000000E+00
 1.55000000E+00 1.70000000E+00 1.50000000E+00 1.55000000E+00
1.30000000E+00
 1.70000000E+00 1.50000000E+00 1.70000000E+00 1.50000000E+00
8.00000000E-01
 1.70000000E+00 1.30000000E+00 1.30000000E+00 1.30000000E+00
1.70000000E+00

```
 1.30000000E+00  1.50000000E+00  8.00000000E-01  1.70000000E+00
1.30000000E+00
 1.50000000E+00
%FLAG SCREEN
%FORMAT(5E16.8)
 8.60000000E-01  8.50000000E-01  8.50000000E-01  8.50000000E-01  7.20000000E-01
 8.50000000E-01  8.50000000E-01  7.20000000E-01  8.50000000E-01  8.50000000E-01
 7.20000000E-01  8.50000000E-01  7.20000000E-01  8.50000000E-01  8.50000000E-01
 7.90000000E-01  7.20000000E-01  8.50000000E-01  7.90000000E-01  8.50000000E-01
 7.20000000E-01  8.50000000E-01  7.20000000E-01  8.50000000E-01  8.50000000E-01
 7.20000000E-01  8.50000000E-01  8.50000000E-01  8.50000000E-01  7.20000000E-01
 8.50000000E-01  8.50000000E-01  8.50000000E-01  7.20000000E-01  8.50000000E-01
 8.50000000E-01
```

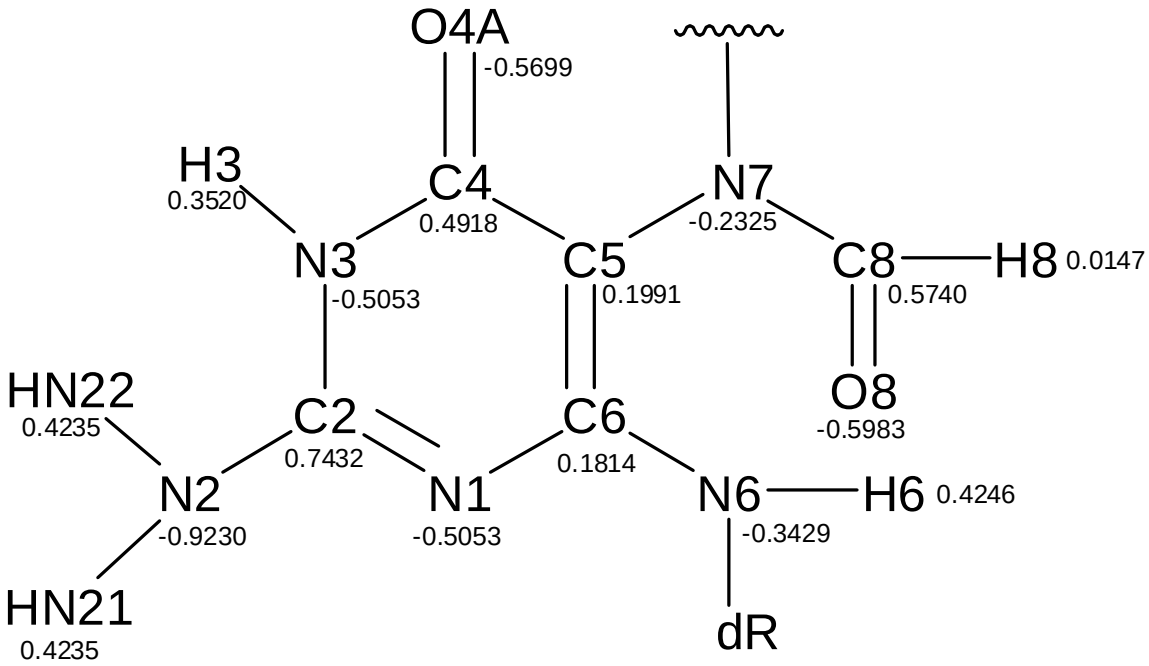


Figure B-2: Atomic charges for the formamidopyrimidine base calculated by Gaussian 03 and used in rMD simulations.

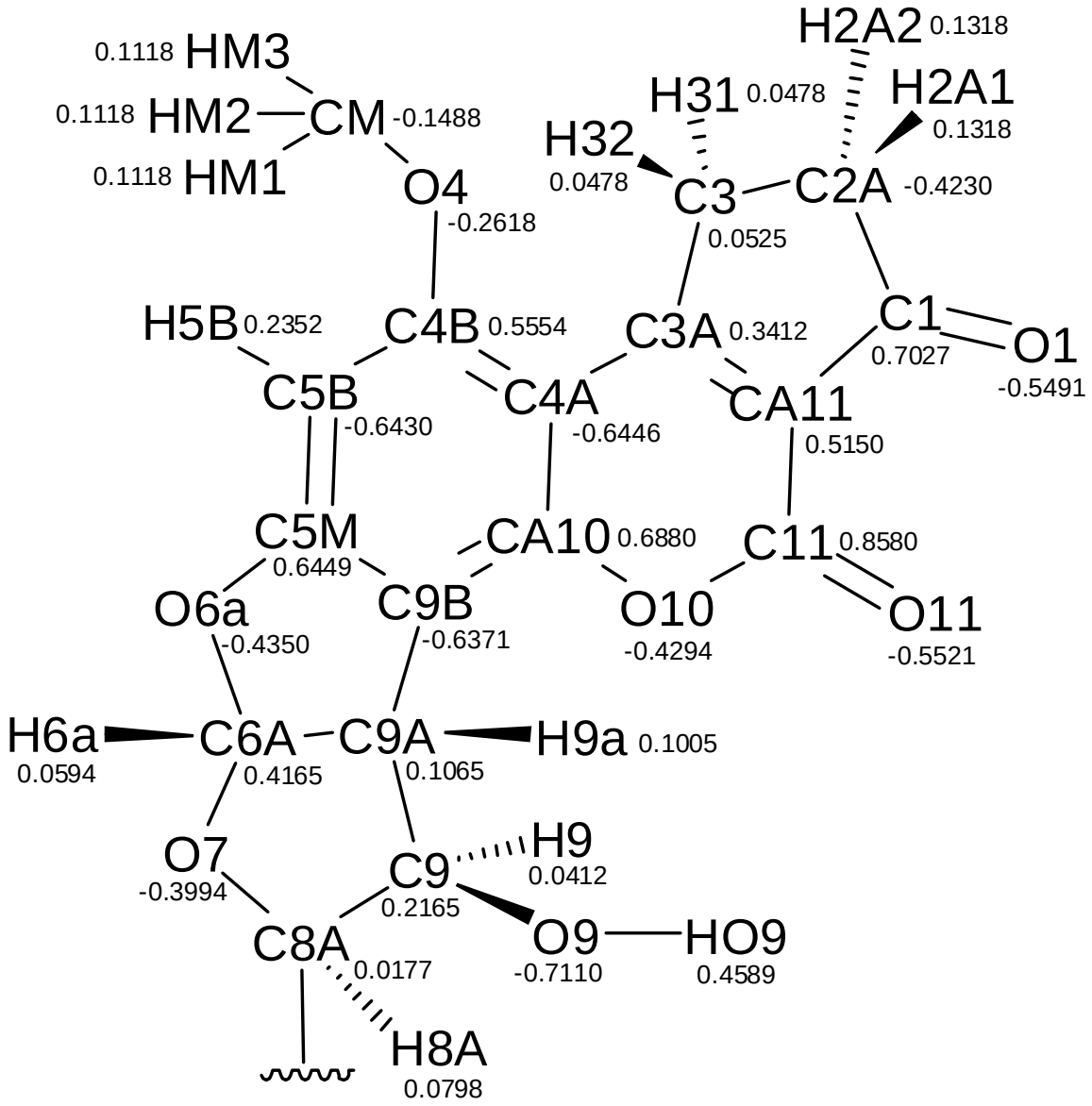


Figure B-3: Atomic charges for the aflatoxin B_1 adduct calculated by Gaussian 03 and used in rMD simulations.

File B-2: Parameter and topology file of formamidopyrimidine base of the AFB_1-FAPY lesion used in AMBER calculations

---

```
%VERSION  VERSION_STAMP = V0001.000  DATE = 01/31/08  15:09:23
%FLAG TITLE      FAPY Base (no AFB)
%FORMAT(20a4)
FB
%FLAG POINTERS
%FORMAT(10I8)
  35      8     12     24     28     33     51     44      0      0
 177      1     24     33     44     21     40     18     16      0
   0      0      0      0      0      0      0      0     35      0
   0
%FLAG ATOM_NAME
%FORMAT(20a4)
P   O1P O2P O5' C5' H5'1H5'2C4' H4' O4' C1' N6  N7  C5  C4  O4A N3  H3  C2  N2
HN21HN22N1  C6  C3' H3' C2' H2'1H2'2O3' H1' C8  O8  H8  H6
%FLAG CHARGE
%FORMAT(5E16.8)
  2.12453796E+01 -1.41423270E+01 -1.41423270E+01 -9.02732742E+00
        -1.25733870E-01
  1.37396142E+00  1.37396142E+00  2.96841267E+00  2.14294248E+00
        -6.72585093E+00
  6.52358340E-01 -6.24842667E+00 -4.23668475E+00  3.62805993E+00
         8.96172714E+00
 -1.03848888E+01 -9.20772819E+00  6.41424960E+00  1.35428134E+01
        -1.68191829E+01
  7.71714405E+00  7.71714405E+00 -9.20772819E+00  3.30552522E+00
         1.29924999E+00
  1.79489655E+00 -1.55618442E+00  1.30836114E+00  1.30836114E+00
        -9.53390736E+00
  2.71512270E-01  1.04596002E+01 -1.09024021E+01  2.67867810E-01
         7.73718858E+00
%FLAG MASS
%FORMAT(5E16.8)
  3.09700000E+01  1.60000000E+01  1.60000000E+01  1.60000000E+01
1.20100000E+01
  1.00800000E+00  1.00800000E+00  1.20100000E+01  1.00800000E+00
1.60000000E+01
  1.20100000E+01  1.40100000E+01  1.40100000E+01  1.20100000E+01
1.20100000E+01
  1.60000000E+01  1.40100000E+01  1.00800000E+00  1.20100000E+01
1.40100000E+01
```

```
 1.00800000E+00  1.00800000E+00  1.40100000E+01  1.20100000E+01
1.20100000E+01
 1.00800000E+00  1.20100000E+01  1.00800000E+00  1.00800000E+00
1.60000000E+01
 1.00800000E+00  1.20100000E+01  1.60000000E+01  1.00800000E+00
1.00800000E+00
%FLAG ATOM_TYPE_INDEX
%FORMAT(10I8)
     1     2     2     3     4     5     5     4     5     3
     4     6     6     7     7     2     6     8     7     6
     8     8     6     7     4     5     4     5     5     3
     5     7     2     5     8
%FLAG NUMBER_EXCLUDED_ATOMS
%FORMAT(10I8)
     7     3     2     7    10     5     4    11     6    11
    12    12     9    11     8     4     7     3     5     4
     2     1     2     4     6     4     5     3     2     1
     1     2     1     1     1
%FLAG NONBONDED_PARM_INDEX
%FORMAT(10I8)
     1     2     4     7    11    16    22    29     2     3
     5     8    12    17    23    30     4     5     6     9
    13    18    24    31     7     8     9    10    14    19
    25    32    11    12    13    14    15    20    26    33
    16    17    18    19    20    21    27    34    22    23
    24    25    26    27    28    35    29    30    31    32
    33    34    35    36
%FLAG RESIDUE_LABEL
%FORMAT(20a4)
FB
%FLAG RESIDUE_POINTER
%FORMAT(10I8)
     1
%FLAG BOND_FORCE_CONSTANT
%FORMAT(5E16.8)
 4.56400000E+02  3.11600000E+02  3.01500000E+02  3.37300000E+02
3.03100000E+02
 3.20600000E+02  3.37300000E+02  4.49000000E+02  3.94100000E+02
4.50000000E+02
 4.49900000E+02  4.18300000E+02  6.48000000E+02  4.11100000E+02
4.06600000E+02
 4.11100000E+02  4.49000000E+02  4.92900000E+02  4.31600000E+02
3.37300000E+02
 4.50000000E+02
%FLAG BOND_EQUIL_VALUE
%FORMAT(5E16.8)
```

1.50300000E+00 1.63600000E+00 1.43900000E+00 1.09200000E+00
1.53500000E+00
1.47000000E+00 1.09200000E+00 1.36400000E+00 1.01800000E+00
1.36000000E+00
1.40600000E+00 1.42900000E+00 1.21400000E+00 1.40000000E+00
1.01100000E+00
1.39100000E+00 1.36400000E+00 1.33600000E+00 1.37600000E+00
1.09200000E+00
1.09000000E+00
%FLAG ANGLE_FORCE_CONSTANT
%FORMAT(5E16.8)
7.76000000E+01 4.51000000E+01 4.31000000E+01 5.09000000E+01
6.78000000E+01
4.64000000E+01 6.32000000E+01 3.94000000E+01 6.21000000E+01
7.03760000E+01
5.09000000E+01 6.28200000E+01 4.71000000E+01 4.64000000E+01
4.94000000E+01
6.62000000E+01 6.86000000E+01 7.39000000E+01 6.92990000E+01
7.41000000E+01
7.41000000E+01 7.41000000E+01 7.28000000E+01 7.00000000E+01
7.00000000E+01
6.79000000E+01 4.88000000E+01 6.43000000E+01 7.50000000E+01
7.31250000E+01
7.17000000E+01 4.76000000E+01 4.91000000E+01 6.94260000E+01
7.28000000E+01
4.13000000E+01 4.91000000E+01 4.64000000E+01 3.94000000E+01
5.54000000E+01
%FLAG ANGLE_EQUIL_VALUE
%FORMAT(5E16.8)
2.05041368E+00 2.04587583E+00 2.03627651E+00 1.89717371E+00
1.89228679E+00
1.92073567E+00 1.93085858E+00 1.89106506E+00 1.97937875E+00
1.92614619E+00
1.89717371E+00 2.05582421E+00 1.91846674E+00 1.92073567E+00
1.91637234E+00
1.92649525E+00 2.07659363E+00 1.95372240E+00 2.07127038E+00
2.13384046E+00
1.97100117E+00 2.14518510E+00 2.07903710E+00 1.98810540E+00
1.97955329E+00
2.10661331E+00 2.05948940E+00 2.18323330E+00 2.14413791E+00
1.98793087E+00
2.15757695E+00 2.08182962E+00 2.08357495E+00 1.92108473E+00
2.07449924E+00
1.86977203E+00 2.08357495E+00 1.92073567E+00 1.89106506E+00
2.09439600E+00
%FLAG DIHEDRAL_FORCE_CONSTANT

%FORMAT(5E16.8)
 3.83000000E-01 1.05000000E+00 1.56000000E-01 3.00000000E-01 3.00000000E-
01
 4.00000000E+00 6.65000000E+00 4.75000000E+00 2.17500000E+00 3.00000000E-
01
 1.45000000E+00 3.50000000E-01 6.25000000E-01 4.15000000E+00
4.80000000E+00
 1.00000000E+00 1.05000000E+01 1.10000000E+00
%FLAG DIHEDRAL_PERIODICITY
%FORMAT(5E16.8)
 3.00000000E+00 2.00000000E+00 3.00000000E+00 3.00000000E+00
2.00000000E+00
 2.00000000E+00 2.00000000E+00 2.00000000E+00 2.00000000E+00
2.00000000E+00
 2.00000000E+00 4.00000000E+00 2.00000000E+00 2.00000000E+00
2.00000000E+00
 2.00000000E+00 2.00000000E+00 2.00000000E+00
%FLAG DIHEDRAL_PHASE
%FORMAT(5E16.8)
 0.00000000E+00 3.14159400E+00 0.00000000E+00 0.00000000E+00
3.14159400E+00
 3.14159400E+00 3.14159400E+00 3.14159400E+00 3.14159400E+00
0.00000000E+00
 3.14159400E+00 3.14159400E+00 3.14159400E+00 3.14159400E+00
3.14159400E+00
 3.14159400E+00 3.14159400E+00 3.14159400E+00
%FLAG SOLTY
%FORMAT(5E16.8)
 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00
0.00000000E+00
 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00
0.00000000E+00
 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00
0.00000000E+00
 0.00000000E+00
%FLAG LENNARD_JONES_ACOEF
%FORMAT(5E16.8)
 6.02589390E+06 1.64263766E+06 3.79876399E+05 1.58759528E+06
3.70622491E+05
 3.61397723E+05 2.54188684E+06 6.47841731E+05 6.28541240E+05
1.04308023E+06
 2.54238516E+05 5.44261042E+04 5.33379252E+04 9.71708117E+04
7.51607703E+03
 2.45746558E+06 6.06829342E+05 5.89818288E+05 9.95480466E+05
8.96776989E+04

```
 9.44293233E+05  2.25370349E+06  5.74393458E+05  5.57281136E+05
9.24822270E+05
 8.61541883E+04  8.82619071E+05  8.19971662E+05  8.41065839E+03
1.02595236E+03
 1.03954408E+03  2.56678134E+03  1.07193646E+02  2.12601181E+03
2.27577561E+03
 1.39982777E-01
%FLAG LENNARD_JONES_BCOEF
%FORMAT(5E16.8)
 2.19561270E+03  1.16041466E+03  5.64885984E+02  1.08210555E+03
5.29252520E+02
 4.95732238E+02  1.22636552E+03  6.26720080E+02  5.85549272E+02
6.75612247E+02
 2.38716853E+02  1.11805549E+02  1.04986921E+02  1.26919150E+02
2.17257828E+01
 1.34630496E+03  6.77220874E+02  6.33305958E+02  7.36907417E+02
1.36131731E+02
 8.01323529E+02  1.08732781E+03  5.55666448E+02  5.19163331E+02
5.99015525E+02
 1.12529845E+02  6.53361429E+02  5.31102864E+02  4.34187588E+01
1.53505284E+01
 1.46567808E+01  2.06278363E+01  2.59456373E+00  2.09604198E+01
1.82891803E+01
 9.37598976E-02
%FLAG BONDS_INC_HYDROGEN
%FORMAT(10I8)
    12    15     4    12    18     4    21    24     4    30
    90     7    33   102     9    48    51    15    57    60
     9    57    63     9    72    75     4    78    81    20
    78    84    20    93    99    21
%FLAG BONDS_WITHOUT_HYDROGEN
%FORMAT(10I8)
     0     3     1     0     6     1     0     9     2     9
    12     3    12    21     5    21    27     3    21    72
     5    27    30     3    30    33     6    30    78     5
    33    69     8    36    39     8    36    93    10    39
    42    11    39    69    12    42    45    13    42    48
    14    48    54    16    54    57    17    54    66    18
    66    69    19    72    78     5    72    87     3    93
    96    13
%FLAG ANGLES_INC_HYDROGEN
%FORMAT(10I8)
     9    12    15     4     9    12    18     4    12    21
    24     6    15    12    18     8    15    12    21     6
    18    12    21     6    21    72    75     6    24    21
    27     4    24    21    72     6    27    30    90    11
```

```
  30   33  102   13   30   78   81   14   30   78
  84   14   33   30   90   15   36   93   99   21
  42   48   51   27   51   48   54   32   54   57
  60   33   54   57   63   33   60   57   63   36
  69   33  102   37   72   78   81   14   72   78
  84   14   75   72   78    6   75   72   87    4
  78   30   90   38   81   78   84   39   96   93
  99   40
%FLAG ANGLES_WITHOUT_HYDROGEN
%FORMAT(10I8)
   0    9   12    1    3    0    6    2    3    0
   9    3    6    0    9    3    9   12   21    5
  12   21   27    5   12   21   72    7   21   27
  30    9   21   72   78    7   21   72   87    5
  27   21   72    5   27   30   33   10   27   30
  78    5   30   33   69   12   30   78   72    7
  33   30   78   16   33   69   39   17   33   69
  66   18   36   39   42   19   36   39   69   17
  36   93   96   20   39   36   93   22   39   42
  45   23   39   42   48   24   39   69   66   25
  42   39   69   26   42   48   54   28   45   42
  48   29   48   54   57   30   48   54   66   31
  54   66   69   34   57   54   66   35   78   72
  87    5
%FLAG DIHEDRALS_INC_HYDROGEN
%FORMAT(10I8)
   0    9   12   15    1    0    9   12   18    1
   9   12   21   24    3   12   21   72   75    3
  15   12   21   24    3   15   12   21   27    3
  15   12   21   72    3   18   12   21   24    3
  18   12   21   27    3   18   12   21   72    3
  21   27   30   90    1   21   72   78   81    3
  21   72   78   84    3   24   21   27   30    1
  24   21   72   75    3   24   21   72   78    3
  24   21   72   87    3   27   21   72   75    3
  27   30   33  102    4   27   30   78   81    3
  27   30   78   84    3   30   78   72   75    3
  33   30   78   81    3   33   30   78   84    3
  39   36   93   99   10   39   42   48   51   11
  39   42  -48   51   12   39   69   33  102    5
  45   42   48   51   11   45   42  -48   51   12
  48   54   57   60    5   48   54   57   63    5
  51   48   54   57   13   51   48   54   66   13
  60   57   54   66    5   63   57   54   66    5
  66   69   33  102    5   69   33   30   90    4
  72   78   30   90    3   75   72   78   81    3
```

```
   75    72    78    84     3    78    30    33   102     4
   81    78    30    90     3    81    78    72    87     3
   84    78    30    90     3    84    78    72    87     3
   90    30    33   102     4    30    69   -33  -102    16
   42    54   -48   -51    16    54    60   -57   -63    16
   36    99   -93   -96    17
%FLAG DIHEDRALS_WITHOUT_HYDROGEN
%FORMAT(10I8)
    0     9    12    21     1     3     0     9    12     2
    6     0     9    12     2     9    12    21    27     3
    9    12    21    72     3    12    21    27    30     1
   12    21    72    78     3    12    21    72    87     3
   21    27    30    33     1    21    27   -30    78     1
   21    72   -78    30     3    27    21   -72    78     3
   27    21    72    87     3    27    30    33    69     4
   27    30   -78    72     3    30    27   -21    72     1
   30    33    69    39     5    30    33    69    66     5
   30    78    72    87     3    33    30    78    72     3
   33    69    39    36     6    33    69    39    42     7
   33    69    66    54     8    36    39    42    45     9
   36    39    42    48     9    36    39    69    66     6
   39    36    93    96     5    39    42    48    54    11
   39    42   -48    54    12    39    69   -66    54     8
   42    39    36    93     5    42    39    69    66     7
   42    48    54    57    13    42    48   -54    66    13
   45    42    39    69     9    45    42    48    54    11
   45    42   -48    54    12    48    42    39    69     9
   48    54   -66    69    14    57    54    66    69    15
   69    33    30    78     4    69    39    36    93     5
   39    48   -42   -45    17    57    48   -54   -66    18
%FLAG EXCLUDED_ATOMS_LIST
%FORMAT(10I8)
    2     3     4     5     6     7     8     3     4     5
    4     5     5     6     7     8     9    10    25     6
    7     8     9    10    11    25    26    27    30     7
    8     9    10    25     8     9    10    25     9    10
   11    12    25    26    27    28    29    30    31    10
   11    25    26    27    30    11    12    24    25    26
   27    28    29    30    31    35    12    14    23    24
   25    26    27    28    29    30    31    35    13    14
   15    19    23    24    25    27    28    29    31    35
   14    15    16    17    23    24    32    33    34    15
   16    17    18    19    23    24    32    33    34    35
   16    17    18    19    20    23    24    32    17    18
   19    24    18    19    20    21    22    23    24    19
   20    23    20    21    22    23    24    21    22    23
```

```
    24    22    23    23    24    35    27    31    32    35
    26    27    28    29    30    31    27    28    29    30
    28    29    30    31    35    29    30    31    30    31
     0    35    33    34    34     0     0
%FLAG HBOND_ACOEF
%FORMAT(5E16.8)

%FLAG HBOND_BCOEF
%FORMAT(5E16.8)

%FLAG HBCUT
%FORMAT(5E16.8)

%FLAG AMBER_ATOM_TYPE
%FORMAT(20a4)
P   O2  O2  OS  CT  H1  H1  CT  H1  OS  CT  N2  N2  CM  C   O   NA  H   CA  N2
H   H   NC  CM  CT  H1  CT  HC  HC  OS  H2  C   O   HA  H
%FLAG TREE_CHAIN_CLASSIFICATION
%FORMAT(20a4)
M   E   E   M   M   E   E   M   E   S   3   B   S   B   B   E   B   E   S   B
E   E   E   B   M   E   B   E   E   M   E   B   E   E   E
%FLAG JOIN_ARRAY
%FORMAT(10I8)
    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0
%FLAG IROTAT
%FORMAT(10I8)
    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0    0    0    0    0    0
    0    0    0    0    0
%FLAG RADIUS_SET
%FORMAT(1a80)
modified Bondi radii (mbondi)
%FLAG RADII
%FORMAT(5E16.8)
 1.85000000E+00 1.50000000E+00 1.50000000E+00 1.50000000E+00
1.70000000E+00
 1.30000000E+00 1.30000000E+00 1.70000000E+00 1.30000000E+00
1.50000000E+00
 1.70000000E+00 1.55000000E+00 1.55000000E+00 1.70000000E+00
1.70000000E+00
 1.50000000E+00 1.55000000E+00 1.30000000E+00 1.70000000E+00
1.55000000E+00
```

```
 1.30000000E+00  1.30000000E+00  1.55000000E+00  1.70000000E+00
1.70000000E+00
 1.30000000E+00  1.70000000E+00  1.30000000E+00  1.30000000E+00
1.50000000E+00
 1.30000000E+00  1.70000000E+00  1.50000000E+00  1.30000000E+00
1.30000000E+00
%FLAG SCREEN
%FORMAT(5E16.8)
 8.60000000E-01  8.50000000E-01  8.50000000E-01  8.50000000E-01  7.20000000E-01
 8.50000000E-01  8.50000000E-01  7.20000000E-01  8.50000000E-01  8.50000000E-01
 7.20000000E-01  7.90000000E-01  7.90000000E-01  7.20000000E-01  7.20000000E-01
 8.50000000E-01  7.90000000E-01  8.50000000E-01  7.20000000E-01  7.90000000E-01
 8.50000000E-01  8.50000000E-01  7.90000000E-01  7.20000000E-01  7.20000000E-01
 8.50000000E-01  7.20000000E-01  8.50000000E-01  8.50000000E-01  8.50000000E-01
 8.50000000E-01  7.20000000E-01  8.50000000E-01  8.50000000E-01  8.50000000E-01
```

File B-3: Parameter and topology file of aflatoxin B_1 adduct of the FAPY-AFB_1 lesion used in AMBER calculations

---

```
%VERSION  VERSION_STAMP = V0001.000  DATE = 01/31/08  15:09:09
%FLAG TITLE      AFB adduct only
%FORMAT(20a4)
FA
%FLAG POINTERS
%FORMAT(10I8)
      37       7      13      28      30      44      46      70       0       0
     209       1      28      44      70      18      39      13      12       0
       0       0       0       0       0       0       0       0      37       0
       0
%FLAG ATOM_NAME
%FORMAT(20a4)
O11 C11 O10 CA10C9B C9A C9  O9  HO9 C8A O7  H8A H9  H9a C6A H6a O6A
C5M C5B H5B
C4B O4  CM  HM1 HM2 HM3 C4A C3A CA11C1  O1  C2A H2A1H2A2C3  H31 H32
%FLAG CHARGE
%FORMAT(5E16.8)
 -1.00596207E+01  1.56350978E+01 -7.82520229E+00  1.25365780E+01
         -1.16103384E+01
```

```
 1.94103940E+00  3.94458128E+00 -1.29566020E+01  8.36148458E+00
        3.23081379E-01
-7.27743995E+00  1.45322843E+00  7.51123206E-01  1.83079448E+00
        7.58904128E+00
 1.08313351E+00 -7.92578939E+00  1.17511968E+01 -1.17176678E+01
        4.28515607E+00
 1.01213943E+01 -4.77005147E+00 -2.71220713E+00  2.03761759E+00
        2.03761759E+00
 2.03761759E+00 -1.17464590E+01  6.21763098E+00 -9.38484895E+00
        1.28040813E+01
-1.00056827E+01 -7.70748623E+00  2.40224581E+00  2.40224581E+00
        9.56670750E-01
 8.71937055E-01  8.71937055E-01
%FLAG MASS
%FORMAT(5E16.8)
 1.60000000E+01  1.20100000E+01  1.60000000E+01  1.20100000E+01
1.20100000E+01
 1.20100000E+01  1.20100000E+01  1.60000000E+01  1.00800000E+00
1.20100000E+01
 1.60000000E+01  1.00800000E+00  1.00800000E+00  1.00800000E+00
1.20100000E+01
 1.00800000E+00  1.60000000E+01  1.20100000E+01  1.20100000E+01
1.00800000E+00
 1.20100000E+01  1.60000000E+01  1.20100000E+01  1.00800000E+00
1.00800000E+00
 1.00800000E+00  1.20100000E+01  1.20100000E+01  1.20100000E+01
1.20100000E+01
 1.60000000E+01  1.20100000E+01  1.00800000E+00  1.00800000E+00
1.20100000E+01
 1.00800000E+00  1.00800000E+00
%FLAG ATOM_TYPE_INDEX
%FORMAT(10I8)
    1    2    3    2    2    4    4    5    6    4
    3    7    7    7    4    7    3    2    2    7
    2    3    4    7    7    7    2    2    2    2
    1    4    7    7    4    7    7
%FLAG NUMBER_EXCLUDED_ATOMS
%FORMAT(10I8)
    6   10    9   14   16   14   11    7    2    7
    7    2    2    4    4    2    4    5    6    3
    9    6    4    2    1    1    7    9    8    7
    4    5    4    3    2    1    1
%FLAG NONBONDED_PARM_INDEX
%FORMAT(10I8)
    1    2    4    7   11   16   22    2    3    5
    8   12   17   23    4    5    6    9   13   18
```

```
    24      7      8      9     10     14     19     25     11     12
    13     14     15     20     26     16     17     18     19     20
    21     27     22     23     24     25     26     27     28
%FLAG RESIDUE_LABEL
%FORMAT(20a4)
FA
%FLAG RESIDUE_POINTER
%FORMAT(10I8)
      1
%FLAG BOND_FORCE_CONSTANT
%FORMAT(5E16.8)
  6.48000000E+02  4.11300000E+02  4.49900000E+02  3.92600000E+02
4.11700000E+02
  4.78400000E+02  3.28300000E+02  4.18300000E+02  3.03100000E+02
3.37300000E+02
  3.14100000E+02  3.37300000E+02  3.69600000E+02  3.01500000E+02
3.37300000E+02
  3.92600000E+02  3.44300000E+02  3.28300000E+02
%FLAG BOND_EQUIL_VALUE
%FORMAT(5E16.8)
  1.21400000E+00  1.34300000E+00  1.40600000E+00  1.35700000E+00
1.43400000E+00
  1.38700000E+00  1.50800000E+00  1.42900000E+00  1.53500000E+00
1.09200000E+00
  1.42600000E+00  1.09200000E+00  9.74000000E-01  1.43900000E+00
1.09200000E+00
  1.35700000E+00  1.08700000E+00  1.50800000E+00
%FLAG ANGLE_FORCE_CONSTANT
%FORMAT(5E16.8)
  7.62000000E+01  7.28000000E+01  6.82000000E+01  6.79000000E+01
6.66000000E+01
  7.16800000E+01  7.12000000E+01  7.12000000E+01  6.43000000E+01
6.77000000E+01
  6.72000000E+01  6.60000000E+01  6.37000000E+01  4.70000000E+01
7.12000000E+01
  6.43000000E+01  6.77000000E+01  6.32000000E+01  4.64000000E+01
6.78000000E+01
  4.64000000E+01  4.64000000E+01  4.71000000E+01  5.11000000E+01
6.21000000E+01
  5.09000000E+01  5.09000000E+01  7.17000000E+01  6.42000000E+01
7.12000000E+01
  5.03000000E+01  5.03000000E+01  6.42000000E+01  3.94000000E+01
6.35830000E+01
  4.72000000E+01  6.38000000E+01  6.80000000E+01  3.94000000E+01
%FLAG ANGLE_EQUIL_VALUE
%FORMAT(5E16.8)
```

 2.13680752E+00 2.07903710E+00 1.86575777E+00 2.10661331E+00
2.07484830E+00
 1.95712580E+00 2.11935422E+00 2.11935422E+00 2.15408629E+00
1.93801443E+00
 2.09387240E+00 2.09614133E+00 1.93661817E+00 1.92841512E+00
2.11935422E+00
 2.15408629E+00 1.90991462E+00 1.93085858E+00 1.92073567E+00
1.89228679E+00
 1.92073567E+00 1.92073567E+00 1.88774893E+00 1.91113635E+00
1.97937875E+00
 1.89717371E+00 1.89717371E+00 1.92405179E+00 1.95634040E+00
2.11935422E+00
 2.08916001E+00 2.08916001E+00 1.95634040E+00 1.89106506E+00
2.03182592E+00
 1.91427794E+00 1.92911325E+00 2.14867576E+00 1.89106506E+00
%FLAG DIHEDRAL_FORCE_CONSTANT
%FORMAT(5E16.8)
 2.70000000E+00 2.17500000E+00 1.05000000E+00 6.65000000E+00
       0.00000000E+00
 4.00000000E+00 3.62500000E+00 1.56000000E-01 1.67000000E-01
       3.83000000E-01
 0.00000000E+00 1.05000000E+01 1.10000000E+00
%FLAG DIHEDRAL_PERIODICITY
%FORMAT(5E16.8)
 2.00000000E+00 2.00000000E+00 2.00000000E+00 2.00000000E+00
2.00000000E+00
 2.00000000E+00 2.00000000E+00 3.00000000E+00 3.00000000E+00
3.00000000E+00
 2.00000000E+00 2.00000000E+00 2.00000000E+00
%FLAG DIHEDRAL_PHASE
%FORMAT(5E16.8)
 3.14159400E+00 3.14159400E+00 3.14159400E+00 3.14159400E+00
0.00000000E+00
 3.14159400E+00 3.14159400E+00 0.00000000E+00 0.00000000E+00
0.00000000E+00
 3.14159400E+00 3.14159400E+00 3.14159400E+00
%FLAG SOLTY
%FORMAT(5E16.8)
 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00
0.00000000E+00
 0.00000000E+00 0.00000000E+00 0.00000000E+00 0.00000000E+00
0.00000000E+00
 0.00000000E+00 0.00000000E+00
%FLAG LENNARD_JONES_ACOEF
%FORMAT(5E16.8)

```
 3.79876399E+05  5.74393458E+05  8.19971662E+05  3.70622491E+05
5.57281136E+05
 3.61397723E+05  6.47841731E+05  9.24822270E+05  6.28541240E+05
1.04308023E+06
 4.71003287E+05  7.01803794E+05  4.58874091E+05  7.91544157E+05
5.81803229E+05
 0.00000000E+00  0.00000000E+00  0.00000000E+00  0.00000000E+00
0.00000000E+00
 0.00000000E+00  5.44261042E+04  8.61541883E+04  5.33379252E+04
9.71708117E+04
 6.82786631E+04  0.00000000E+00  7.51607703E+03
%FLAG LENNARD_JONES_BCOEF
%FORMAT(5E16.8)
 5.64885984E+02  5.55666448E+02  5.31102864E+02  5.29252520E+02
5.19163331E+02
 4.95732238E+02  6.26720080E+02  5.99015525E+02  5.85549272E+02
6.75612247E+02
 6.29300710E+02  6.14502845E+02  5.89183300E+02  6.93079947E+02
6.99746810E+02
 0.00000000E+00  0.00000000E+00  0.00000000E+00  0.00000000E+00
0.00000000E+00
 0.00000000E+00  1.11805549E+02  1.12529845E+02  1.04986921E+02
1.26919150E+02
 1.25287818E+02  0.00000000E+00  2.17257828E+01
%FLAG BONDS_INC_HYDROGEN
%FORMAT(10I8)
    15    39    10    18    36    12    21    24    13    27
    33    12    42    45    15    54    57    17    66    69
    12    66    72    12    66    75    12    93    96    10
    93    99    10   102   105    10   102   108    10
%FLAG BONDS_WITHOUT_HYDROGEN
%FORMAT(10I8)
     0     3     1     3     6     2     3    84     3     6
     9     4     9    12     5     9    78     6    12    15
     7    12    51     8    15    18     9    15    42     9
    18    21    11    18    27     9    27    30    14    30
    42    14    42    48    14    48    51    16    51    54
     5    54    60     6    60    63     4    60    78     6
    63    66    14    78    81     5    81    84     8    81
   102     7    84    87     3    87    90     1    87    93
    18    93   102     9
%FLAG ANGLES_INC_HYDROGEN
%FORMAT(10I8)
    12    15    39    14    15    18    36    19    15    42
    45    21    18    15    39    22    18    21    24    23
    18    27    33    19    21    18    36    24    27    18
```

```
    36    19    30    27    33    26    30    42    45    27
    39    15    42    22    45    42    48    27    51    54
    57    31    57    54    60    32    63    66    69    26
    63    66    72    26    63    66    75    26    69    66
    72    34    69    66    75    34    72    66    75    34
    81   102   105    14    81   102   108    14    87    93
    96    36    87    93    99    36    93   102   105    22
    93   102   108    22    96    93    99    39    96    93
   102    22    99    93   102    22   105   102   108    39
%FLAG ANGLES_WITHOUT_HYDROGEN
%FORMAT(10I8)
     0     3     6     1     0     3    84     2     3     6
     9     3     3    84    81     4     3    84    87     5
     6     3    84     6     6     9    12     7     6     9
    78     8     9    12    15     9     9    12    51    10
     9    78    60    11     9    78    81    12    12     9
    78    12    12    15    18    13    12    15    42    13
    12    51    48    15    12    51    54    10    15    12
    51    16    15    18    21    17    15    18    27    18
    15    42    30    20    15    42    48    20    18    15
    42    18    18    27    30    20    21    18    27    17
    27    30    42    25    30    42    48    28    42    48
    51    29    48    51    54    30    51    54    60    12
    54    60    63     8    54    60    78    11    60    63
    66    33    60    78    81    12    63    60    78     8
    78    81    84    10    78    81   102     9    81    84
    87     4    81   102    93    13    84    81   102    16
    84    87    90     2    84    87    93    35    87    93
   102    37    90    87    93    38
%FLAG DIHEDRALS_INC_HYDROGEN
%FORMAT(10I8)
     9    12    15    39     5    12    15    18    36     8
    12    15    42    45     8    12    51    54    57     4
    15    18    21    24     9    15    18    27    33     8
    18    15    42    45     8    21    18    15    39     8
    21    18    27    33     8    24    21    18    27     9
    24    21    18    36     9    27    18    15    39     8
    27    30    42    45    10    30    27    18    36     8
    30    42    15    39     8    33    27    18    36     8
    33    27    30    42    10    36    18    15    39     8
    36    18    15    42     8    39    15    12    51     5
    39    15    42    45     8    39    15    42    48     8
    45    42    48    51    10    48    51    54    57     4
    57    54    60    63     4    57    54    60    78     4
    60    63    66    69    10    60    63    66    72    10
    60    63    66    75    10    78    81   102   105     5
```

```
    78    81   102   108     5    81   102    93    96     8
    81   102    93    99     8    84    81   102   105     5
    84    81   102   108     5    84    87    93    96    11
    84    87    93    99    11    87    93   102   105     8
    87    93   102   108     8    90    87    93    96    11
    90    87    93    99    11    96    93   102   105     8
    96    93   102   108     8    99    93   102   105     8
    99    93   102   108     8    51    60   -54   -57    13
%FLAG DIHEDRALS_WITHOUT_HYDROGEN
%FORMAT(10I8)
     0     3     6     9     1     0     3    84    81     2
     0     3    84    87     2     3     6     9    12     3
     3     6     9    78     3     3    84   -81    78     4
     3    84    81   102     4     3    84    87    90     2
     3    84    87    93     2     6     3    84    81     2
     6     3    84    87     2     6     9    12    15     4
     6     9    12    51     4     6     9    78    60     4
     6     9   -78    81     4     9     6     3    84     1
     9    12    15    18     5     9    12    15    42     5
     9    12    51    48     4     9    12    51    54     6
     9    78   -60    54     7     9    78    60    63     4
     9    78   -81    84     6     9    78    81   102     4
    12     9    78    60     7    12     9    78    81     7
    12    15    18    21     8    12    15    18    27     8
    12    15    42    30     8    12    15   -42    48     8
    12    51   -48    42     3    12    51   -54    60     6
    15    12     9    78     4    15    12   -51    48     4
    15    12    51    54     4    15    18   -27    30     8
    15    42   -30    27    10    15    42   -48    51    10
    18    15    12    51     5    18    15   -42    30     8
    18    15    42    48     8    18    27   -30    42    10
    21    18    15    42     8    21    18    27    30     8
    27    18   -15    42     8    27    30    42    48    10
    30    42    48    51    10    42    15   -12    51     5
    42    48    51    54     3    48    51    54    60     4
    51    12     9    78     6    51    54    60    63     4
    51    54   -60    78     7    54    60    63    66     3
    54    60    78    81     7    60    78    81    84     6
    60    78    81   102     4    63    60    78    81     4
    66    63    60    78     3    78    81    84    87     4
    78    81   102    93     5    81    84    87    90     2
    81    84   -87    93     2    81   102   -93    87     8
    84    81  -102    93     5    84    87   -93   102    11
    87    84   -81   102     4    90    87    93   102    11
     0     3   -84    -6    12    84    93   -87   -90    12
%FLAG EXCLUDED_ATOMS_LIST
```

```
%FORMAT(10I8)
       2       3       4      28      29      30       3       4       5      27
      28      29      30      31      32      35       4       5       6      18
      21      27      28      29      30       5       6       7      14      15
      17      18      19      21      22      27      28      29      35       6
       7       8      10      11      13      14      15      16      17      18
      19      20      21      27      28       7       8       9      10      11
      12      13      14      15      16      17      18      19      27       8
       9      10      11      12      13      14      15      16      17      18
       9      10      11      12      13      14      15      10      13      11
      12      13      14      15      16      17      12      13      14      15
      16      17      18      13      15      14      15      15      16      17
      18      16      17      18      19      17      18      18      19      20
      21      19      20      21      22      27      20      21      22      23
      27      28      21      22      27      22      23      24      25      26
      27      28      29      35      23      24      25      26      27      28
      24      25      26      27      25      26      26       0      28      29
      30      32      35      36      37      29      30      31      32      33
      34      35      36      37      30      31      32      33      34      35
      36      37      31      32      33      34      35      36      37      32
      33      34      35      33      34      35      36      37      34      35
      36      37      35      36      37      36      37      37       0
%FLAG HBOND_ACOEF
%FORMAT(5E16.8)

%FLAG HBOND_BCOEF
%FORMAT(5E16.8)

%FLAG HBCUT
%FORMAT(5E16.8)

%FLAG AMBER_ATOM_TYPE
%FORMAT(20a4)
O   C   OS  CA  CB  CT  CT  OH  HO  CT  OS  H1  H1  HC  CT  H2  OS  CB  CA  HA
CA  OS  CT  H1  H1  H1  CA  CB  CB  C   O   CT  HC  HC  CT  HC  HC
%FLAG TREE_CHAIN_CLASSIFICATION
%FORMAT(20a4)
E   B   S   M   M   M   M   S   E   M   S   E   E   E   B   E   S   S   B   E
S   S   3   E   E   E   M   M   S   B   E   B   E   E   M   E   E
%FLAG JOIN_ARRAY
%FORMAT(10I8)
       0       0       0       0       0       0       0       0       0       0
       0       0       0       0       0       0       0       0       0       0
       0       0       0       0       0       0       0       0       0       0
       0       0       0       0       0       0       0
%FLAG IROTAT
```

%FORMAT(10I8)
0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0
%FLAG RADIUS_SET
%FORMAT(1a80)
modified Bondi radii (mbondi)
%FLAG RADII
%FORMAT(5E16.8)
1.50000000E+00 1.70000000E+00 1.50000000E+00 1.70000000E+00
1.70000000E+00
1.70000000E+00 1.70000000E+00 1.50000000E+00 8.00000000E-01
1.70000000E+00
1.50000000E+00 1.30000000E+00 1.30000000E+00 1.30000000E+00
1.70000000E+00
1.30000000E+00 1.50000000E+00 1.70000000E+00 1.70000000E+00
1.30000000E+00
1.70000000E+00 1.50000000E+00 1.70000000E+00 1.30000000E+00
1.30000000E+00
1.30000000E+00 1.70000000E+00 1.70000000E+00 1.70000000E+00
1.70000000E+00
1.50000000E+00 1.70000000E+00 1.30000000E+00 1.30000000E+00
1.70000000E+00
1.30000000E+00 1.30000000E+00
%FLAG SCREEN
%FORMAT(5E16.8)
8.50000000E-01 7.20000000E-01 8.50000000E-01 7.20000000E-01 7.20000000E-01
7.20000000E-01 7.20000000E-01 8.50000000E-01 8.50000000E-01 7.20000000E-01
8.50000000E-01 8.50000000E-01 8.50000000E-01 8.50000000E-01 7.20000000E-01
8.50000000E-01 8.50000000E-01 7.20000000E-01 7.20000000E-01 8.50000000E-01
7.20000000E-01 8.50000000E-01 7.20000000E-01 8.50000000E-01 8.50000000E-01
8.50000000E-01 7.20000000E-01 7.20000000E-01 7.20000000E-01 7.20000000E-01
8.50000000E-01 7.20000000E-01 8.50000000E-01 8.50000000E-01 7.20000000E-01
8.50000000E-01 8.50000000E-01

APPENDIX C

EXPERIMENTAL DISTANCE RESTRAINTS

Table C-1: α -AFB_1-FAPY modified 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-3'

	Residue Number	Residue Name	Atom Name	Residue Number	Residue Name	Atom Name	Lower Bound	Upper Bound
Class 1	12	GUA	H1	9	CYT	H42	3.05	3.91
	12	GUA	H1	9	CYT	H5	4.14	6.12
	1	CYT	H1'	1	CYT	H3'	3.43	4.72
	1	CYT	H1'	1	CYT	H4'	2.63	3.63
	1	CYT	H1'	1	CYT	H5'1	2.82	5.61
	1	CYT	H1'	2	THY	H6	1.80	3.80
	10	ADE	H1'	10	ADE	H8	2.87	4.12
	11	THY	H1'	11	THY	H3'	3.25	4.82
	11	THY	H1'	12	GUA	H8	3.74	5.21
	12	GUA	H1'	12	GUA	H4'	2.83	3.54
	12	GUA	H1'	12	GUA	H5'1	2.81	5.52
	12	GUA	H1'	12	GUA	H5'2	2.52	5.40
	12	GUA	H1'	12	GUA	H8	3.29	6.85
	12	GUA	H1'	13	ADE	H8	3.15	5.60
	13	ADE	H1'	13	ADE	H4'	2.93	3.45
	13	ADE	H1'	13	ADE	H8	3.20	6.75
	13	ADE	H1'	14	ADE	H8	3.14	4.55
	14	ADE	H1'	14	ADE	H4'	2.94	3.47
	14	ADE	H1'	15	THY	H6	3.19	3.78
	14	ADE	H1'	14	ADE	H8	3.38	4.03
	15	THY	H1'	15	THY	H6	3.58	6.48
	16	CYT	H1'	17	ADE	H2	2.95	3.48
	16	CYT	H1'	16	CYT	H2'1	2.85	5.35
	16	CYT	H1'	17	ADE	H5'1	3.13	4.61
	17	ADE	H1'	18	THY	H6	3.03	3.40
	18	THY	H1'	18	THY	H6	3.48	5.23
	18	THY	H1'	19	ADE	H8	1.80	4.06
	19	ADE	H1'	19	ADE	H3'	2.77	5.65
	19	ADE	H1'	19	ADE	H4'	2.75	4.12
	19	ADE	H1'	19	ADE	H8	3.17	6.09
	19	ADE	H1'	20	GUA	H8	1.80	4.00
	2	THY	H1'	2	THY	H2'1	2.53	5.25
	2	THY	H1'	2	THY	H3'	3.06	3.59
	2	THY	H1'	2	THY	H6	2.85	5.72

20	GUA	H1'	20	GUA	H3'	3.29	5.43
20	GUA	H1'	20	GUA	H8	3.22	5.38
3	ADE	H1'	3	ADE	H3'	3.33	5.04
3	ADE	H1'	3	ADE	H4'	2.88	3.27
3	ADE	H1'	3	ADE	H5'2	1.80	5.84
3	ADE	H1'	3	ADE	H8	3.24	4.22
3	ADE	H1'	4	THY	M7	4.15	5.29
4	THY	H1'	17	ADE	H2	3.49	4.74
4	THY	H1'	4	THY	H6	3.39	5.73
4	THY	H1'	21	FA	M	3.58	4.15
5	FB	H1'	5	FB	H2'1	3.05	4.34
5	FB	H1'	5	FB	H3'	3.78	4.76
5	FB	H1'	5	FB	H8	3.96	5.31
6	ADE	H1'	6	ADE	H3'	3.29	4.21
6	ADE	H1'	6	ADE	H4'	2.97	3.46
6	ADE	H1'	7	THY	H4'	2.96	3.82
6	ADE	H1'	7	THY	H6	3.14	4.09
6	ADE	H1'	7	THY	M7	3.95	4.56
7	THY	H1'	7	THY	H3'	3.27	5.52
7	THY	H1'	7	THY	H6	3.30	4.49
7	THY	H1'	8	THY	H6	3.02	3.88
8	THY	H1'	8	THY	H3'	3.41	5.81
8	THY	H1'	8	THY	H6	2.58	3.70
8	THY	H1'	9	CYT	H6	3.38	6.05
9	CYT	H1'	10	ADE	H8	3.56	4.51
17	ADE	H2	4	THY	H5'1	3.46	5.24
6	ADE	H2	5	FB	H3	4.57	6.68
1	CYT	H2'1	1	CYT	H6	2.63	3.63
11	THY	H2'1	12	GUA	H8	1.80	3.99
18	THY	H2'1	19	ADE	H8	3.14	4.19
19	ADE	H2'1	20	GUA	H8	2.70	5.53
2	THY	H2'1	3	ADE	H8	2.93	5.93
3	ADE	H2'1	3	ADE	H4'	3.10	5.42
3	ADE	H2'1	4	THY	H6	1.80	5.11
3	ADE	H2'1	4	THY	M7	3.46	3.99
5	FB	H2'1	5	FB	H5'2	4.21	5.55
6	ADE	H2'1	6	ADE	H4'	3.57	5.91
6	ADE	H2'1	7	THY	H6	2.66	5.94
6	ADE	H2'1	7	THY	M7	3.25	4.41
8	THY	H2'1	8	THY	H5'2	2.68	5.78
9	CYT	H2'1	10	ADE	H8	3.31	5.02
1	CYT	H2'2	1	CYT	H3'	2.75	4.86
1	CYT	H2'2	1	CYT	H4'	2.20	3.46
1	CYT	H2'2	2	THY	H5'2	2.31	5.26
1	CYT	H2'2	1	CYT	H6	2.70	5.68
11	THY	H2'2	12	GUA	H4'	2.30	5.85
11	THY	H2'2	11	THY	H5'1	2.18	5.80
19	ADE	H2'2	19	ADE	H4'	2.59	5.73
19	ADE	H2'2	20	GUA	H8	2.30	3.66
20	GUA	H2'2	20	GUA	H4'	3.00	5.03

20	GUA	H2'2	20	GUA	H8	2.61	6.32
3	ADE	H2'2	3	ADE	H4'	3.20	4.20
6	ADE	H2'2	7	THY	H4'	2.51	2.97
6	ADE	H2'2	7	THY	M7	3.44	4.27
7	THY	H2'2	8	THY	H5'1	2.91	3.17
7	THY	H2'2	7	THY	H6	2.78	5.71
7	THY	H2'2	8	THY	M7	3.24	6.51
8	THY	H2'2	8	THY	H3'	2.15	5.12
8	THY	H2'2	8	THY	M7	3.37	6.73
4	THY	H3	17	ADE	H61	1.80	3.92
4	THY	H3	17	ADE	H62	3.44	6.10
4	THY	H3	21	FA	M	5.78	9.44
5	FB	H3	6	ADE	H61	4.18	5.13
5	FB	H3	21	FA	M	5.25	7.44
8	THY	H3	9	CYT	H42	4.34	4.79
1	CYT	H3'	1	CYT	H5	3.79	6.61
1	CYT	H3'	2	THY	H6	3.22	4.93
10	ADE	H3'	10	ADE	H8	2.95	3.77
11	THY	H3'	11	THY	H6	3.09	5.02
11	THY	H3'	12	GUA	H8	3.27	3.86
12	GUA	H3'	13	ADE	H8	2.29	5.97
14	ADE	H3'	14	ADE	H4'	2.80	3.90
14	ADE	H3'	15	THY	H6	3.48	4.80
14	ADE	H3'	14	ADE	H8	3.74	6.34
15	THY	H3'	15	THY	H6	3.06	4.38
18	THY	H3'	19	ADE	H8	3.62	4.90
2	THY	H3'	3	ADE	H8	3.34	6.73
20	GUA	H3'	20	GUA	H8	2.95	4.73
3	ADE	H3'	4	THY	H6	3.84	6.13
3	ADE	H3'	4	THY	M7	4.14	4.82
4	THY	H3'	4	THY	H6	4.16	6.47
6	ADE	H3'	7	THY	H6	3.71	5.92
7	THY	H3'	7	THY	H6	3.26	6.23
8	THY	H3'	9	CYT	H5	3.43	6.77
8	THY	H3'	8	THY	H6	2.91	6.40
8	THY	H3'	9	CYT	H6	2.25	6.63
9	CYT	H3'	10	ADE	H8	3.98	5.66
1	CYT	H4'	1	CYT	H6	3.42	6.42
1	CYT	H4'	2	THY	M7	3.53	6.86
10	ADE	H4'	10	ADE	H8	3.54	5.49
12	GUA	H4'	12	GUA	H8	2.83	5.25
18	THY	H4'	18	THY	H6	3.51	4.60
2	THY	H4'	2	THY	H6	3.01	6.66
3	ADE	H4'	3	ADE	H8	3.91	6.15
4	THY	H4'	4	THY	H6	3.52	4.60
6	ADE	H4'	6	ADE	H8	2.81	5.99
7	THY	H4'	7	THY	H6	3.61	6.77
1	CYT	H5	2	THY	M7	3.85	4.62
16	CYT	H5	15	THY	H6	3.37	4.53
16	CYT	H5	15	THY	M7	4.43	5.92

	9	CYT	H5	8	THY	M7	3.71	4.42
	1	CYT	H5'1	1	CYT	H6	2.91	4.80
	15	THY	H5'1	15	THY	H6	3.43	6.14
	2	THY	H5'1	2	THY	H6	2.66	5.10
	4	THY	H5'1	4	THY	H6	2.89	6.76
	9	CYT	H5'1	9	CYT	H6	2.79	5.33
	1	CYT	H5'2	1	CYT	H6	2.46	3.50
	4	THY	H5'2	4	THY	H6	3.02	6.37
	1	CYT	H6	2	THY	M7	3.44	4.18
	7	THY	H6	7	THY	M7	2.95	4.38
	7	THY	H6	8	THY	M7	3.33	4.80
	14	ADE	H8	15	THY	M7	3.80	4.29
	6	ADE	H8	7	THY	M7	3.71	4.30
	21	FA	H9a	4	THY	M7	1.80	3.91
Class 2	12	GUA	H1	8	THY	H3	3.91	5.30
	12	GUA	H1	9	CYT	H41	3.23	5.14
	10	ADE	H1'	10	ADE	H3'	2.65	5.71
	11	THY	H1'	11	THY	H4'	2.51	3.02
	11	THY	H1'	11	THY	H5'1	3.30	4.99
	11	THY	H1'	11	THY	M7	4.00	7.09
	12	GUA	H1'	12	GUA	H3'	3.69	5.68
	13	ADE	H1'	13	ADE	H3'	3.06	5.70
	14	ADE	H1'	14	ADE	H2	4.04	5.50
	14	ADE	H1'	14	ADE	H2'1	2.75	5.25
	14	ADE	H1'	14	ADE	H3'	3.75	5.93
	14	ADE	H1'	15	THY	M7	4.43	7.18
	15	THY	H1'	15	THY	H3'	3.48	5.06
	16	CYT	H1'	21	FA	H2A1	3.37	6.29
	16	CYT	H1'	16	CYT	H3'	3.45	5.45
	17	ADE	H1'	17	ADE	H3'	3.31	5.79
	17	ADE	H1'	17	ADE	H5'1	2.97	5.02
	17	ADE	H1'	17	ADE	H8	3.12	6.27
	18	THY	H1'	18	THY	H2'1	2.70	5.48
	18	THY	H1'	18	THY	H3'	3.35	4.45
	18	THY	H1'	18	THY	H4'	2.96	3.47
	2	THY	H1'	2	THY	H2'2	1.70	2.50
	20	GUA	H1'	20	GUA	H4'	2.45	3.02
	3	ADE	H1'	4	THY	H6	3.34	4.07
	4	THY	H1'	4	THY	H2'1	2.82	4.58
	4	THY	H1'	4	THY	H3'	3.62	5.74
	5	FB	H1'	5	FB	H2'2	3.33	4.23
	6	ADE	H1'	6	ADE	H2'1	2.70	5.15
	6	ADE	H1'	6	ADE	H2'2	2.33	3.23
	7	THY	H1'	7	THY	H2'1	2.38	5.21
	7	THY	H1'	8	THY	M7	3.82	5.06
	8	THY	H1'	8	THY	H2'1	2.35	5.56
	9	CYT	H1'	9	CYT	H3'	3.11	5.83
	13	ADE	H2	14	ADE	H62	3.24	5.01
	17	ADE	H2	4	THY	H3	1.80	3.09
	10	ADE	H2'1	10	ADE	H4'	2.75	5.87

11	THY	H2'1	11	THY	H4'	3.39	5.54
11	THY	H2'1	11	THY	H5'1	2.47	4.58
12	GUA	H2'1	12	GUA	H5'1	2.14	4.68
12	GUA	H2'1	12	GUA	H5'2	2.80	5.37
13	ADE	H2'1	14	ADE	H8	2.58	4.61
14	ADE	H2'1	14	ADE	H4'	3.01	5.38
14	ADE	H2'1	15	THY	H6	2.78	4.90
15	THY	H2'1	16	CYT	H5	3.28	4.73
16	CYT	H2'1	17	ADE	H5'1	2.90	3.99
17	ADE	H2'1	18	THY	H6	3.52	6.30
18	THY	H2'1	18	THY	M7	3.43	4.76
2	THY	H2'1	2	THY	H4'	3.34	4.99
2	THY	H2'1	21	FA	M	4.16	4.63
4	THY	H2'1	4	THY	H5'1	1.63	5.55
5	FB	H2'1	5	FB	H3'	2.98	4.70
7	THY	H2'1	7	THY	M7	2.38	4.38
9	CYT	H2'1	9	CYT	H4'	2.29	5.36
10	ADE	H2'2	10	ADE	H8	2.70	5.58
11	THY	H2'2	11	THY	H4'	2.64	4.83
11	THY	H2'2	12	GUA	H8	2.86	3.45
12	GUA	H2'2	12	GUA	H4'	2.28	5.73
14	ADE	H2'2	15	THY	M7	3.37	5.04
15	THY	H2'2	15	THY	H6	2.74	4.89
16	CYT	H2'2	17	ADE	H5'1	2.06	3.32
17	ADE	H2'2	18	THY	H6	1.80	3.30
18	THY	H2'2	18	THY	H6	2.08	5.35
18	THY	H2'2	19	ADE	H8	1.80	3.50
20	GUA	H2'2	20	GUA	H3'	1.97	5.51
3	ADE	H2'2	4	THY	M7	1.80	4.11
4	THY	H2'2	21	FA	M	3.69	7.28
5	FB	H2'2	3	ADE	H4'	2.73	3.43
6	ADE	H2'2	6	ADE	H4'	3.30	4.85
6	ADE	H2'2	7	THY	H6	2.43	3.53
6	ADE	H2'2	6	ADE	H8	2.69	4.54
8	THY	H2'2	9	CYT	H5	2.57	5.28
8	THY	H2'2	9	CYT	H6	2.45	3.25
9	CYT	H2'2	9	CYT	H4'	2.25	4.00
9	CYT	H2'2	9	CYT	H6	2.20	4.88
9	CYT	H2'2	10	ADE	H8	1.80	3.78
21	FA	H2A1	16	CYT	H3'	3.08	5.50
21	FA	H2A1	21	FA	H31	2.08	4.16
21	FA	H2A1	16	CYT	H5	3.11	5.50
21	FA	H2A1	17	ADE	H8	3.15	4.12
21	FA	H2A2	21	FA	M	3.73	4.40
15	THY	H3	5	FB	H3	1.80	4.43
15	THY	H3	16	CYT	H42	1.80	5.33
15	THY	H3	14	ADE	H61	1.80	3.61
5	FB	H3	16	CYT	H41	1.80	4.19
5	FB	H3	16	CYT	H5	1.80	6.71
5	FB	H3	5	FB	HN22	1.80	6.34

	8	THY	H3	13	ADE	H62	1.80	4.98
	1	CYT	H3'	1	CYT	H5'1	2.60	4.76
	1	CYT	H3'	1	CYT	H5'2	2.41	3.81
	1	CYT	H3'	1	CYT	H6	3.06	4.00
	12	GUA	H3'	12	GUA	H5'2	2.20	5.74
	12	GUA	H3'	12	GUA	H8	3.50	6.51
	13	ADE	H3'	14	ADE	H8	3.62	6.65
	14	ADE	H3'	15	THY	M7	3.80	4.80
	17	ADE	H3'	17	ADE	H5'1	2.22	4.83
	17	ADE	H3'	18	THY	H6	3.08	6.31
	19	ADE	H3'	20	GUA	H8	3.64	6.67
	2	THY	H3'	2	THY	H6	3.24	6.84
	6	ADE	H3'	7	THY	M7	4.22	5.76
	8	THY	H3'	8	THY	M7	3.54	6.60
	9	CYT	H3'	9	CYT	H6	2.80	5.89
	11	THY	H4'	11	THY	H6	3.52	6.09
	13	ADE	H4'	13	ADE	H8	2.51	6.44
	14	ADE	H4'	14	ADE	H8	3.49	5.05
	15	THY	H4'	15	THY	H6	3.39	4.68
	17	ADE	H4'	17	ADE	H8	2.75	5.00
	19	ADE	H4'	19	ADE	H8	3.16	6.67
	8	THY	H4'	8	THY	H6	3.47	6.67
	8	THY	H4'	8	THY	M7	3.71	7.64
	10	ADE	H5'1	10	ADE	H8	3.24	4.47
	12	GUA	H5'1	12	GUA	H8	3.28	6.14
	13	ADE	H5'1	13	ADE	H8	1.98	6.39
	14	ADE	H5'1	14	ADE	H8	2.73	5.32
	19	ADE	H5'1	19	ADE	H8	3.27	6.36
	20	GUA	H5'1	20	GUA	H8	3.10	6.49
	12	GUA	H5'2	12	GUA	H8	2.88	6.62
	13	ADE	H5'2	13	ADE	H8	2.23	4.60
	21	FA	H5B	4	THY	H6	3.53	4.36
	4	THY	H6	21	FA	H6a	1.80	3.97
	4	THY	H6	21	FA	H9a	3.48	5.00
	8	THY	H6	8	THY	M7	2.83	3.33
	21	FA	H6a	5	FB	H8	3.44	5.50
Class 3	12	GUA	H1	13	ADE	H61	3.00	4.45
	1	CYT	H1'	1	CYT	H2'2	2.04	2.86
	1	CYT	H1'	1	CYT	H6	2.81	3.70
	10	ADE	H1'	10	ADE	H2'1	2.42	4.94
	10	ADE	H1'	10	ADE	H4'	2.61	3.39
	12	GUA	H1'	12	GUA	H2'1	2.48	4.96
	15	THY	H1'	16	CYT	H5	1.80	6.51
	16	CYT	H1'	16	CYT	H6	3.47	6.60
	17	ADE	H1'	18	THY	M7	1.80	5.24
	18	THY	H1'	18	THY	H2'2	2.23	3.15
	19	ADE	H1'	19	ADE	H5'1	2.06	4.86
	2	THY	H1'	2	THY	H4'	2.50	3.99
	20	GUA	H1'	20	GUA	H2'2	1.89	3.34
	3	ADE	H1'	3	ADE	H2'1	2.73	4.59

3	ADE	H1'	3	ADE	H2'2	2.12	2.96
4	THY	H1'	21	FA	H5B	2.87	3.64
5	FB	H1'	21	FA	H9a	3.75	5.60
6	ADE	H1'	6	ADE	H8	3.35	4.08
7	THY	H1'	7	THY	H4'	2.71	3.13
8	THY	H1'	8	THY	H4'	1.97	3.30
3	ADE	H2	18	THY	H3	1.80	3.49
6	ADE	H2	15	THY	H3	2.85	3.31
1	CYT	H2'1	1	CYT	H3'	2.17	2.89
1	CYT	H2'1	1	CYT	H4'	2.24	3.45
1	CYT	H2'1	2	THY	H6	2.00	2.70
10	ADE	H2'1	10	ADE	H3'	2.15	3.34
12	GUA	H2'1	12	GUA	H4'	2.22	4.08
13	ADE	H2'1	13	ADE	H4'	2.93	5.02
13	ADE	H2'1	13	ADE	H5'1	1.93	4.97
14	ADE	H2'1	14	ADE	H8	2.46	3.06
14	ADE	H2'1	15	THY	M7	3.29	4.52
15	THY	H2'1	15	THY	M7	3.54	5.54
16	CYT	H2'1	16	CYT	H3'	2.24	3.06
16	CYT	H2'1	16	CYT	H5	3.16	6.04
16	CYT	H2'1	16	CYT	H6	1.80	4.46
19	ADE	H2'1	19	ADE	H4'	2.29	5.80
20	GUA	H2'1	20	GUA	H8	2.02	2.84
4	THY	H2'1	4	THY	H3'	2.25	3.69
11	THY	H2'2	11	THY	H3'	2.45	3.89
11	THY	H2'2	11	THY	H6	2.65	4.85
12	GUA	H2'2	12	GUA	H8	2.09	3.80
16	CYT	H2'2	21	FA	H2A1	3.40	5.74
19	ADE	H2'2	19	ADE	H3'	2.00	5.03
2	THY	H2'2	2	THY	H3'	2.04	2.90
3	ADE	H2'2	3	ADE	H3'	2.07	4.21
3	ADE	H2'2	4	THY	H6	1.80	3.24
4	THY	H2'2	4	THY	H3'	2.36	4.28
4	THY	H2'2	4	THY	H6	2.33	5.25
7	THY	H2'2	8	THY	H6	1.80	2.93
8	THY	H2'2	8	THY	H6	2.34	5.62
5	FB	H3	16	CYT	H42	1.80	5.60
5	FB	H3	5	FB	HN21	2.51	3.35
1	CYT	H3'	2	THY	M7	2.95	5.00
15	THY	H3'	16	CYT	H5	1.80	5.40
16	CYT	H3'	16	CYT	H6	2.17	3.12
5	FB	H3'	6	ADE	H4'	2.50	2.96
1	CYT	H4'	1	CYT	H5'1	1.95	4.09
20	GUA	H4'	20	GUA	H8	3.32	4.90
9	CYT	H4'	9	CYT	H6	2.46	5.90
6	ADE	H5'1	6	ADE	H8	3.82	6.66
15	THY	H5'2	15	THY	H6	3.15	6.49
20	GUA	H5'2	20	GUA	H8	2.20	6.34
18	THY	H6	18	THY	M7	2.89	3.47
4	THY	H6	4	THY	M7	3.16	3.74

	21	FA	H6a	21	FA	H8A	2.94	3.34
	21	FA	H6a	21	FA	H9a	2.66	3.16
	21	FA	H6a	4	THY	M7	1.80	3.81
	17	ADE	H8	18	THY	M7	1.80	4.33
	3	ADE	H8	4	THY	M7	3.33	3.97
Class 4	10	ADE	H1'	10	ADE	H2'2	1.91	2.75
	11	THY	H1'	11	THY	H2'1	2.64	5.10
	13	ADE	H1'	13	ADE	H2'2	2.08	2.85
	14	ADE	H1'	15	THY	H5'1	1.80	5.55
	14	ADE	H1'	15	THY	H5'2	2.74	4.77
	15	THY	H1'	15	THY	H4'	3.27	3.99
	19	ADE	H1'	19	ADE	H5'2	2.73	5.10
	4	THY	H1'	4	THY	H2'2	2.16	2.88
	8	THY	H1'	8	THY	H2'2	2.03	2.83
	9	CYT	H1'	9	CYT	H2'1	2.45	4.71
	9	CYT	H1'	9	CYT	H2'2	2.03	2.88
	9	CYT	H1'	9	CYT	H4'	2.39	3.24
	13	ADE	H2	8	THY	H3	1.80	3.39
	14	ADE	H2	7	THY	H3	1.80	3.91
	10	ADE	H2'1	10	ADE	H8	2.10	2.78
	11	THY	H2'1	11	THY	H3'	2.18	2.89
	12	GUA	H2'1	12	GUA	H3'	2.05	3.92
	12	GUA	H2'1	12	GUA	H8	1.80	2.96
	13	ADE	H2'1	13	ADE	H8	1.64	2.44
	15	THY	H2'1	16	CYT	H6	2.34	3.20
	17	ADE	H2'1	18	THY	M7	1.80	7.19
	2	THY	H2'1	2	THY	H3'	2.28	3.77
	3	ADE	H2'1	3	ADE	H3'	2.17	3.09
	4	THY	H2'1	4	THY	H6	2.12	3.10
	7	THY	H2'1	7	THY	H6	2.08	2.86
	8	THY	H2'1	8	THY	H6	1.84	2.51
	9	CYT	H2'1	9	CYT	H6	1.71	2.54
	1	CYT	H2'2	2	THY	M7	3.22	7.21
	10	ADE	H2'2	10	ADE	H4'	2.29	3.41
	12	GUA	H2'2	12	GUA	H3'	1.82	2.80
	13	ADE	H2'2	13	ADE	H4'	3.07	5.56
	14	ADE	H2'2	15	THY	H6	2.33	3.33
	17	ADE	H2'2	18	THY	M7	1.80	4.13
	18	THY	H2'2	18	THY	H3'	2.31	3.77
	2	THY	H2'2	2	THY	H4'	2.05	2.88
	9	CYT	H2'2	9	CYT	H3'	2.42	4.29
	21	FA	H2A1	21	FA	H2A2	2.54	3.03
	7	THY	H3	14	ADE	H61	1.80	4.86
	7	THY	H3	14	ADE	H62	1.80	5.53
	8	THY	H3	13	ADE	H61	1.80	2.75
	11	THY	H3'	11	THY	H5'1	2.14	3.04
	12	GUA	H3'	12	GUA	H5'1	2.10	4.09
	13	ADE	H3'	13	ADE	H8	2.69	6.58
	17	ADE	H3'	18	THY	M7	3.97	6.19
	19	ADE	H3'	19	ADE	H4'	1.72	2.73

	19	ADE	H3'	19	ADE	H8	2.59	5.39
	3	ADE	H3'	2	THY	H4'	2.36	2.87
	3	ADE	H3'	3	ADE	H8	2.90	4.60
	6	ADE	H3'	6	ADE	H5'1	2.54	5.42
	6	ADE	H3'	6	ADE	H8	3.04	4.62
	9	CYT	H5	8	THY	H6	3.18	4.46
	8	THY	H5'1	8	THY	H6	3.23	5.69
	10	ADE	H5'2	10	ADE	H8	3.83	6.67
	8	THY	H5'2	8	THY	H6	3.26	5.10
	11	THY	H6	11	THY	M7	2.75	3.35
	17	ADE	H61	4	THY	M7	1.80	4.46
Class 5	1	CYT	H1'	1	CYT	H2'1	2.35	4.24
	11	THY	H1'	11	THY	H2'2	2.07	2.77
	11	THY	H1'	11	THY	H6	2.78	3.83
	12	GUA	H1'	12	GUA	H2'2	2.13	2.82
	13	ADE	H1'	13	ADE	H2'1	2.27	5.10
	14	ADE	H1'	14	ADE	H2'2	2.11	3.09
	15	THY	H1'	15	THY	H2'1	2.89	4.48
	15	THY	H1'	15	THY	H2'2	2.25	2.99
	16	CYT	H1'	16	CYT	H2'2	2.11	2.90
	17	ADE	H1'	17	ADE	H2'1	2.28	5.02
	17	ADE	H1'	17	ADE	H2'2	2.01	3.44
	17	ADE	H1'	17	ADE	H4'	2.42	3.01
	17	ADE	H1'	18	THY	H5'1	2.72	4.19
	19	ADE	H1'	19	ADE	H2'1	2.10	4.33
	19	ADE	H1'	19	ADE	H2'2	1.97	3.02
	2	THY	H1'	3	ADE	H8	2.33	6.22
	20	GUA	H1'	20	GUA	H2'1	2.22	4.92
	4	THY	H1'	4	THY	H4'	2.20	2.92
	5	FB	H1'	5	FB	H4'	3.50	4.12
	5	FB	H1'	5	FB	H5'1	2.81	4.77
	7	THY	H1'	7	THY	H2'2	2.23	3.06
	8	THY	H1'	9	CYT	H1'	3.14	5.39
	8	THY	H1'	9	CYT	H5'1	2.51	3.32
	9	CYT	H1'	10	ADE	H4'	3.19	4.15
	9	CYT	H1'	9	CYT	H5'1	2.60	5.10
	9	CYT	H1'	9	CYT	H6	2.45	5.95
	13	ADE	H2	14	ADE	H5'2	3.19	4.38
	1	CYT	H2'1	1	CYT	H2'2	1.52	2.22
	1	CYT	H2'1	2	THY	M7	2.39	4.07
	10	ADE	H2'1	10	ADE	H2'2	1.46	2.25
	11	THY	H2'1	11	THY	H2'2	1.60	2.43
	11	THY	H2'1	11	THY	H6	1.91	2.71
	13	ADE	H2'1	13	ADE	H3'	1.83	3.16
	14	ADE	H2'1	14	ADE	H2'2	1.48	2.51
	14	ADE	H2'1	14	ADE	H3'	2.04	3.05
	15	THY	H2'1	15	THY	H2'2	1.62	2.69
	15	THY	H2'1	15	THY	H3'	2.37	3.06
	15	THY	H2'1	14	ADE	H5'2	2.19	3.06
	15	THY	H2'1	15	THY	H6	2.12	3.01

16	CYT	H2'1	16	CYT	H2'2	1.73	2.58
17	ADE	H2'1	17	ADE	H3'	1.92	3.18
17	ADE	H2'1	17	ADE	H8	2.05	3.72
18	THY	H2'1	18	THY	H2'2	1.58	2.46
18	THY	H2'1	18	THY	H3'	2.20	3.62
18	THY	H2'1	18	THY	H6	2.18	3.32
19	ADE	H2'1	19	ADE	H2'2	1.30	2.49
19	ADE	H2'1	19	ADE	H3'	1.75	2.68
19	ADE	H2'1	19	ADE	H8	1.92	2.64
2	THY	H2'1	2	THY	H2'2	1.36	2.51
2	THY	H2'1	2	THY	H6	1.93	3.03
20	GUA	H2'1	20	GUA	H2'2	1.30	2.04
20	GUA	H2'1	20	GUA	H3'	1.88	4.28
3	ADE	H2'1	3	ADE	H2'2	1.46	2.63
3	ADE	H2'1	3	ADE	H8	1.98	2.67
5	FB	H2'1	5	FB	H2'2	1.42	2.99
6	ADE	H2'1	6	ADE	H2'2	1.36	2.42
6	ADE	H2'1	6	ADE	H8	2.15	2.99
7	THY	H2'1	7	THY	H2'2	1.68	2.43
7	THY	H2'1	8	THY	M7	2.47	3.38
8	THY	H2'1	8	THY	H2'2	1.44	2.19
8	THY	H2'1	8	THY	H3'	1.76	2.53
8	THY	H2'1	8	THY	H4'	2.95	4.92
9	CYT	H2'1	9	CYT	H2'2	1.49	2.28
9	CYT	H2'1	9	CYT	H3'	1.68	2.97
9	CYT	H2'1	10	ADE	H5'1	2.02	4.25
10	ADE	H2'2	10	ADE	H3'	2.30	3.55
13	ADE	H2'2	13	ADE	H3'	1.74	4.31
13	ADE	H2'2	13	ADE	H8	1.76	4.89
13	ADE	H2'2	14	ADE	H8	2.23	3.03
14	ADE	H2'2	14	ADE	H3'	2.42	4.77
15	THY	H2'2	15	THY	H3'	2.53	3.30
15	THY	H2'2	16	CYT	H4'	2.19	2.76
16	CYT	H2'2	16	CYT	H3'	2.26	3.18
16	CYT	H2'2	16	CYT	H6	1.93	5.00
17	ADE	H2'2	17	ADE	H3'	2.07	4.18
17	ADE	H2'2	16	CYT	H4'	2.02	2.88
17	ADE	H2'2	17	ADE	H8	2.19	3.77
2	THY	H2'2	2	THY	H6	2.03	4.71
4	THY	H2'2	4	THY	H5'2	1.68	4.90
5	FB	H2'2	5	FB	H3'	1.60	3.54
7	THY	H2'2	8	THY	H5'2	2.43	5.48
21	FA	H2A1	21	FA	H32	1.85	4.81
21	FA	H2A1	16	CYT	H6	3.50	6.29
21	FA	H2A2	21	FA	H31	2.13	3.34
21	FA	H2A2	21	FA	H32	1.68	2.84
15	THY	H3	7	THY	H3	3.17	3.95
10	ADE	H3'	10	ADE	H4'	2.41	3.46
10	ADE	H3'	10	ADE	H5'2	2.14	2.79
11	THY	H3'	11	THY	H4'	2.28	2.97

12	GUA	H3'	12	GUA	H4'	1.83	2.80
13	ADE	H3'	13	ADE	H4'	2.15	2.93
14	ADE	H3'	14	ADE	H5'2	1.90	3.08
16	CYT	H3'	16	CYT	H4'	1.72	2.59
17	ADE	H3'	18	THY	H5'1	1.44	2.41
17	ADE	H3'	17	ADE	H8	2.73	6.87
18	THY	H3'	18	THY	H6	2.46	5.00
19	ADE	H3'	19	ADE	H5'1	2.07	4.51
20	GUA	H3'	20	GUA	H4'	2.17	3.27
20	GUA	H3'	20	GUA	H5'2	2.46	4.99
4	THY	H3'	4	THY	H5'2	2.34	3.28
5	FB	H3'	5	FB	H4'	3.07	4.18
6	ADE	H3'	6	ADE	H4'	2.35	4.24
6	ADE	H3'	6	ADE	H5'2	1.75	3.42
7	THY	H3'	8	THY	H4'	1.61	2.44
9	CYT	H3'	9	CYT	H4'	1.93	2.95
11	THY	H4'	11	THY	H5'1	1.77	3.25
12	GUA	H4'	12	GUA	H5'1	1.63	3.53
12	GUA	H4'	12	GUA	H5'2	1.80	4.62
16	CYT	H4'	16	CYT	H6	2.50	5.00
19	ADE	H4'	19	ADE	H5'1	1.68	2.71
6	ADE	H4'	6	ADE	H5'1	2.06	4.96
6	ADE	H4'	6	ADE	H5'2	1.42	3.43
1	CYT	H5	1	CYT	H6	2.24	2.79
16	CYT	H5	16	CYT	H6	2.49	3.14
9	CYT	H5	9	CYT	H6	2.15	2.73
11	THY	H5'1	11	THY	H6	1.66	6.03
11	THY	H5'2	11	THY	H6	2.86	4.02
14	ADE	H5'2	14	ADE	H8	3.11	5.74
21	FA	H5B	21	FA	M	2.49	3.25
1	CYT	H6	2	THY	H6	3.33	6.17
11	THY	H6	12	GUA	H8	3.61	6.61
15	THY	H6	16	CYT	H6	3.03	5.02
15	THY	H6	14	ADE	H8	3.72	5.57
15	THY	H6	15	THY	M7	3.04	4.10
18	THY	H6	19	ADE	H8	3.49	4.83
2	THY	H6	3	ADE	H8	3.03	6.18
2	THY	H6	2	THY	M7	2.84	3.65
4	THY	H6	3	ADE	H8	3.14	6.35
7	THY	H6	6	ADE	H8	3.65	5.93
9	CYT	H6	10	ADE	H8	3.27	6.21
12	GUA	H8	13	ADE	H8	3.48	6.39
19	ADE	H8	20	GUA	H8	3.27	5.35
5	FB	H8	21	FA	H8A	2.21	3.32
21	FA	H9	21	FA	H9a	3.16	5.12

Table C-2: α-AFB_1-FAPY modified 5'-$C^1T^2X^3A^4$-3'

	Residue Number	*Residue Name*	*Atom Name*	*Residue Number*	*Residue Name*	*Atom Name*	*Lower Bound*	*Upper Bound*
Class 1	2	THY	H1'	5	FA	M	4.12	6.16
	1	CYT	H2''	1	CYT	H6	3.60	6.66
	3	FB	H2''	3	FB	H4'	2.49	4.49
	1	CYT	H3'	1	CYT	H6	2.89	4.00
	1	CYT	H5	1	CYT	H6	2.63	2.92
	1	CYT	H1'	1	CYT	H2'	2.73	4.07
	1	CYT	H1'	1	CYT	H2''	1.80	2.91
	1	CYT	H1'	1	CYT	H6	3.28	4.09
	2	THY	H1'	2	THY	H2'	2.84	3.50
	2	THY	H1'	2	THY	H2''	2.35	3.04
	2	THY	H1'	2	THY	H6	3.38	5.19
	3	FB	H1'	3	FB	H2''	2.42	3.64
	3	FB	H1'	3	FB	H3'	3.79	5.38
	4	ADE	H1'	4	ADE	H4'	2.66	3.20
	1	CYT	H2'	1	CYT	H3'	1.80	4.25
	1	CYT	H2'	1	CYT	H6	2.89	3.87
	3	FB	H2'	5	FA	H5B	2.80	3.80
	4	ADE	H2'	4	ADE	H4'	2.52	5.10
	1	CYT	H2''	1	CYT	H3'	1.93	2.65
	1	CYT	H2''	1	CYT	H4'	2.57	4.10
	3	FB	H2''	3	FB	H3'	2.28	2.86
	3	FB	H2''	5	FA	H5B	2.93	4.84
	4	ADE	H2''	4	ADE	H4'	2.55	5.31
	1	CYT	H3'	1	CYT	H5'	2.37	3.16
	1	CYT	H5'	1	CYT	H6	2.95	3.48
	5	FA	H9	5	FA	H9a	3.03	4.72
	2	THY	H1'	2	THY	H4'	3.25	4.13
	3	FB	H1'	3	FB	H2'	2.03	2.68
	3	FB	H1'	3	FB	H5'	1.84	2.53
	4	ADE	H1'	4	ADE	H2'	2.66	4.74
	2	THY	H2'	2	THY	H3'	2.18	2.86
	3	FB	H2'	3	FB	H3'	2.26	2.98
	4	ADE	H2'	4	ADE	H3'	2.27	3.06
	2	THY	H2''	2	THY	H3'	2.26	2.91
	4	ADE	H2''	4	ADE	H3'	2.70	4.27
	2	THY	H3'	2	THY	H4'	2.73	4.96
	2	THY	H3'	2	THY	H6	2.86	4.52
	4	ADE	H3'	4	ADE	H4'	2.54	4.00
	2	THY	H6	2	THY	M7	2.90	3.36
	5	FA	H6a	5	FA	H9a	2.59	2.96
	4	ADE	H1'	4	ADE	H2''	2.04	2.72
	2	THY	H2'	2	THY	H6	2.15	2.88
	2	THY	H2''	2	THY	H6	2.45	3.22
	4	ADE	H2''	4	ADE	H8	2.46	4.96

	1	CYT	H3'	1	CYT	H4'	2.19	2.76
	2	THY	H3'	2	THY	H5'	1.81	3.80
	2	THY	H3'	2	THY	H5''	2.26	4.35
	3	FB	H3'	3	FB	H5'	1.90	2.93
	4	ADE	H3'	4	ADE	H5'	2.20	2.98
	5	FA	H5B	5	FA	M	2.84	3.41
	5	FA	H6a	5	FA	H8A	3.34	4.55
	1	CYT	H2'	1	CYT	H2''	1.45	2.16
	2	THY	H2'	2	THY	H2''	1.55	2.21
	3	FB	H2'	3	FB	H2''	1.55	2.39
	4	ADE	H2'	4	ADE	H2''	1.50	2.14
	4	ADE	H2'	4	ADE	H8	1.88	3.10
	3	FB	H3'	3	FB	H4'	2.28	3.76
	1	CYT	H4'	1	CYT	H5'	2.02	3.25
	2	THY	H4'	2	THY	H5'	1.76	2.86
	2	THY	H4'	2	THY	H5''	2.10	4.50
	3	FB	H4'	3	FB	H5''	1.75	3.28
	3	FB	H4'	3	FB	H5'	1.60	3.48
	4	ADE	H4'	4	ADE	H5'	1.74	3.64
	4	ADE	H4'	4	ADE	H5''	1.59	3.12
	2	THY	H5'	2	THY	H6	2.50	4.02
	3	FB	H8	5	FA	H8A	2.14	2.76

Table C-3: *cis*-5*R*,6*S*-thymine glycol modified 5'-$G^{1}T^{2}G^{3}C^{4}G^{5}\underline{Tg}^{6}G^{7}T^{8}T^{9}T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}\underline{A}^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3'

	Residue Number	*Residue Name*	*Atom Name*	*Residue Number*	*Residue Name*	*Atom Name*	*Lower Bound*	*Upper Bound*
Class 1	11	GUA	H1	11	GUA	H22	1.80	4.14
	11	GUA	H1	10	THY	H3	1.80	6.98
	11	GUA	H1	14	CYT	H42	1.80	5.20
	21	GUA	H1	4	CYT	H42	1.80	5.04
	3	GUA	H1	22	CYT	H42	1.80	5.07
	3	GUA	H1	4	CYT	H42	1.80	6.84
	7	GUA	H1	17	ADE	H2	1.80	6.81
	7	GUA	H1	7	GUA	H22	1.80	4.00
	7	GUA	H1	8	THY	H3	1.80	6.86
	7	GUA	H1	18	CYT	H42	1.80	4.56
	1	GUA	H1'	2	THY	H6	2.86	5.82
	1	GUA	H1'	1	GUA	H8	2.15	6.09
	1	GUA	H1'	2	THY	M7	4.44	6.49
	10	THY	H1'	10	THY	H3'	2.08	5.48
	10	THY	H1'	10	THY	H6	2.51	6.00
	10	THY	H1'	11	GUA	H8	3.47	5.40

11	GUA	H1'	11	GUA	H3'	2.06	4.26
11	GUA	H1'	11	GUA	H8	1.87	6.02
12	THY	H1'	12	THY	H3'	2.42	5.67
13	ADE	H1'	13	ADE	H3'	2.29	4.94
14	CYT	H1'	14	CYT	H3'	2.16	3.74
14	CYT	H1'	14	CYT	H6	2.89	5.64
14	CYT	H1'	15	ADE	H8	2.17	5.77
15	ADE	H1'	15	ADE	H3'	3.50	4.23
15	ADE	H1'	15	ADE	H8	2.43	6.07
15	ADE	H1'	16	ADE	H8	2.95	5.46
16	ADE	H1'	16	ADE	H3'	2.05	5.47
16	ADE	H1'	16	ADE	H8	1.99	6.12
16	ADE	H1'	17	ADE	H8	2.13	6.16
17	ADE	H1'	17	ADE	H3'	2.17	5.73
17	ADE	H1'	18	CYT	H6	3.02	6.28
17	ADE	H1'	17	ADE	H8	2.36	4.88
18	CYT	H1'	18	CYT	H3'	2.01	5.58
18	CYT	H1'	18	CYT	H6	2.64	4.38
19	ADE	H1'	20	CYT	H5	2.81	6.63
19	ADE	H1'	19	ADE	H8	2.50	5.61
2	THY	H1'	2	THY	H3'	2.01	4.07
2	THY	H1'	2	THY	H6	2.46	6.54
20	CYT	H1'	20	CYT	H6	2.59	5.06
20	CYT	H1'	21	GUA	H8	2.47	5.26
21	GUA	H1'	21	GUA	H3'	1.88	5.30
21	GUA	H1'	22	CYT	H6	2.08	4.20
21	GUA	H1'	21	GUA	H8	1.95	5.72
22	CYT	H1'	22	CYT	H3'	2.17	5.32
22	CYT	H1'	22	CYT	H6	2.79	6.59
23	ADE	H1'	23	ADE	H3'	2.36	5.62
23	ADE	H1'	24	CYT	H5'1	1.95	5.38
23	ADE	H1'	24	CYT	H6	2.65	5.55
23	ADE	H1'	23	ADE	H8	2.33	5.94
24	CYT	H1'	24	CYT	H3'	2.13	5.48
24	CYT	H1'	24	CYT	H6	2.52	4.78
3	GUA	H1'	3	GUA	H3'	2.26	5.35
3	GUA	H1'	4	CYT	H6	2.61	6.79
4	CYT	H1'	4	CYT	H3'	2.10	3.66
4	CYT	H1'	4	CYT	H6	2.70	4.39
4	CYT	H1'	5	GUA	H8	2.43	4.69
5	GUA	H1'	6	TG	H6	2.61	5.50
5	GUA	H1'	5	GUA	H8	2.43	5.75
5	GUA	H1'	6	TG	M	2.79	6.64
7	GUA	H1'	8	THY	M7	2.84	6.29
8	THY	H1'	9	THY	M7	4.03	6.22
9	THY	H1'	9	THY	H3'	2.05	5.58
9	THY	H1'	10	THY	M7	4.66	6.73
15	ADE	H2	10	THY	H3	3.24	4.76
16	ADE	H2	9	THY	H3	3.29	4.94
17	ADE	H2	8	THY	H3	3.18	4.61

	23	ADE	H2	2	THY	H3	3.12	5.54
	17	ADE	H2'1	18	CYT	H5	3.24	4.85
	21	GUA	H2'1	22	CYT	H5	2.64	6.32
	5	GUA	H2'1	5	GUA	H4'	3.13	4.75
	13	ADE	H2'2	14	CYT	H5	1.83	4.78
	23	ADE	H2'2	24	CYT	H5	2.61	6.18
	3	GUA	H2'2	4	CYT	H5	2.19	4.00
	1	GUA	H3'	1	GUA	H8	2.15	6.44
	1	GUA	H3'	2	THY	M7	2.78	5.47
	10	THY	H3'	11	GUA	H8	2.61	6.34
	10	THY	H3'	10	THY	M7	2.59	7.53
	11	GUA	H3'	11	GUA	H8	1.83	5.00
	11	GUA	H3'	12	THY	M7	3.87	7.51
	12	THY	H3'	12	THY	H6	2.10	4.69
	13	ADE	H3'	13	ADE	H8	2.00	6.48
	14	CYT	H3'	14	CYT	H6	1.90	6.10
	15	ADE	H3'	15	ADE	H8	2.63	6.62
	16	ADE	H3'	17	ADE	H8	2.76	6.35
	17	ADE	H3'	18	CYT	H6	2.16	5.89
	17	ADE	H3'	17	ADE	H8	1.97	6.64
	18	CYT	H3'	18	CYT	H6	2.58	6.29
	18	CYT	H3'	19	ADE	H8	2.44	5.43
	19	ADE	H3'	19	ADE	H8	2.08	6.06
	2	THY	H3'	2	THY	H6	2.14	5.83
	2	THY	H3'	2	THY	M7	5.01	6.50
	21	GUA	H3'	22	CYT	H5	3.53	6.15
	21	GUA	H3'	22	CYT	H6	2.46	5.59
	21	GUA	H3'	21	GUA	H8	1.95	5.82
	22	CYT	H3'	23	ADE	H8	2.81	6.57
	23	ADE	H3'	23	ADE	H8	2.69	5.66
	24	CYT	H3'	24	CYT	H6	1.89	3.77
	3	GUA	H3'	4	CYT	H5	2.96	6.15
	3	GUA	H3'	3	GUA	H8	1.93	5.81
	5	GUA	H3'	5	GUA	H8	2.19	6.16
	5	GUA	H3'	6	TG	M	2.68	7.54
	7	GUA	H3'	7	GUA	H8	1.82	6.50
	7	GUA	H3'	8	THY	M7	3.02	7.17
	2	THY	H4'	2	THY	H6	2.17	6.35
	5	GUA	H4'	6	TG	M	3.28	6.37
	7	GUA	H4'	8	THY	M7	2.81	6.13
	18	CYT	H5	17	ADE	H8	2.37	5.86
	20	CYT	H5	19	ADE	H8	2.72	4.50
	22	CYT	H5	21	GUA	H8	2.89	5.45
	13	ADE	H5'1	13	ADE	H8	1.96	6.20
	2	THY	H5'1	2	THY	M7	3.56	7.65
	8	THY	H5'1	8	THY	M7	4.90	7.21
	2	THY	H6	3	GUA	H8	2.80	5.90
	20	CYT	H6	19	ADE	H8	2.76	5.00
	1	GUA	H8	2	THY	M7	2.90	7.51
Class 2	3	GUA	H1	23	ADE	H2	1.80	6.92

1	GUA	H1'	1	GUA	H4'	1.98	3.94
11	GUA	H1'	11	GUA	H4'	1.88	4.36
11	GUA	H1'	12	THY	M7	2.98	7.70
13	ADE	H1'	13	ADE	H4'	2.01	4.02
13	ADE	H1'	13	ADE	H5'1	2.47	4.28
13	ADE	H1'	14	CYT	H6	2.28	5.02
14	CYT	H1'	15	ADE	H5'1	2.98	5.27
15	ADE	H1'	15	ADE	H4'	2.92	4.72
16	ADE	H1'	16	ADE	H4'	1.92	4.14
17	ADE	H1'	17	ADE	H4'	2.00	4.30
18	CYT	H1'	19	ADE	H2	3.52	6.00
18	CYT	H1'	19	ADE	H8	2.09	5.97
19	ADE	H1'	20	CYT	H1'	2.60	4.04
19	ADE	H1'	20	CYT	H6	2.74	6.07
2	THY	H1'	2	THY	H2'1	1.86	3.77
2	THY	H1'	3	GUA	H5'1	2.50	5.13
20	CYT	H1'	21	GUA	H5'1	2.57	4.61
21	GUA	H1'	21	GUA	H4'	2.58	4.86
21	GUA	H1'	22	CYT	H5	2.45	5.63
22	CYT	H1'	23	ADE	H8	2.01	5.18
23	ADE	H1'	23	ADE	H4'	2.29	5.78
3	GUA	H1'	3	GUA	H4'	2.85	5.59
5	GUA	H1'	5	GUA	H4'	2.87	4.57
6	TG	H1'	6	TG	M	3.38	7.60
7	GUA	H1'	7	GUA	H3'	1.87	5.74
7	GUA	H1'	7	GUA	H4'	2.91	5.56
8	THY	H1'	8	THY	H3'	1.89	5.31
17	ADE	H2	9	THY	H3	1.80	5.84
14	CYT	H2'1	15	ADE	H8	2.73	4.95
18	CYT	H2'1	18	CYT	H4'	2.50	5.48
19	ADE	H2'1	20	CYT	H5	2.82	4.96
2	THY	H2'1	3	GUA	H8	2.89	5.54
2	THY	H2'1	2	THY	M7	2.53	6.54
20	CYT	H2'1	20	CYT	H5	1.94	5.50
22	CYT	H2'1	22	CYT	H5	2.29	4.86
23	ADE	H2'1	24	CYT	H5	3.21	4.33
24	CYT	H2'1	24	CYT	H4'	2.42	5.92
3	GUA	H2'1	4	CYT	H5	2.97	6.45
4	CYT	H2'1	4	CYT	H5	2.18	4.60
6	TG	H2'1	6	TG	H4'	2.50	5.48
6	TG	H2'1	6	TG	H5'2	2.49	5.87
15	ADE	H2'2	15	ADE	H4'	3.04	5.59
2	THY	H2'2	2	THY	H4'	2.01	4.14
20	CYT	H2'2	21	GUA	H5'1	2.35	5.25
22	CYT	H2'2	22	CYT	H5	2.20	6.56
23	ADE	H2'2	23	ADE	H4'	3.10	5.68
5	GUA	H2'2	5	GUA	H4'	2.71	5.77
5	GUA	H2'2	6	TG	H5'2	3.68	5.56
5	GUA	H2'2	6	TG	H6	3.06	6.53
6	TG	H2'2	6	TG	H5'2	3.22	4.23

10	THY	H3	15	ADE	H61	1.80	6.49
10	THY	H3	15	ADE	H62	1.80	6.58
2	THY	H3	23	ADE	H61	1.80	6.23
2	THY	H3	23	ADE	H62	1.80	6.86
8	THY	H3	16	ADE	H61	1.80	4.29
9	THY	H3	16	ADE	H61	1.80	5.97
9	THY	H3	16	ADE	H62	1.80	6.43
11	GUA	H3'	12	THY	H6	2.40	5.60
19	ADE	H3'	20	CYT	H5	2.30	5.73
2	THY	H3'	3	GUA	H8	2.99	6.13
23	ADE	H3'	24	CYT	H6	3.29	6.09
3	GUA	H3'	4	CYT	H6	2.81	6.46
7	GUA	H3'	8	THY	H6	1.90	6.17
8	THY	H3'	9	THY	M7	2.86	7.23
16	ADE	H4'	16	ADE	H8	4.93	6.84
17	ADE	H4'	17	ADE	H8	2.20	5.97
18	CYT	H4'	18	CYT	H6	2.70	4.85
19	ADE	H4'	19	ADE	H8	1.96	6.83
21	GUA	H4'	21	GUA	H8	2.96	6.48
23	ADE	H4'	23	ADE	H8	3.87	6.84
5	GUA	H4'	5	GUA	H8	2.29	6.55
6	TG	H4'	6	TG	M	5.36	7.22
7	GUA	H4'	7	GUA	H8	2.73	6.83
14	CYT	H5	13	ADE	H8	2.48	6.57
24	CYT	H5	23	ADE	H8	2.39	6.41
4	CYT	H5	3	GUA	H8	3.33	6.59
1	GUA	H5'1	1	GUA	H8	1.98	5.62
12	THY	H5'1	12	THY	H6	2.08	5.49
15	ADE	H5'1	15	ADE	H8	2.07	6.80
19	ADE	H5'1	19	ADE	H8	2.15	6.29
20	CYT	H5'1	20	CYT	H6	2.28	5.52
21	GUA	H5'1	21	GUA	H8	1.87	5.83
23	ADE	H5'1	23	ADE	H8	3.34	6.42
24	CYT	H5'1	24	CYT	H6	2.34	5.55
5	GUA	H5'1	6	TG	M	3.44	6.89
7	GUA	H5'1	7	GUA	H8	3.02	5.99
1	GUA	H5'2	1	GUA	H8	2.10	6.04
15	ADE	H5'2	15	ADE	H8	2.15	6.52
18	CYT	H5'2	18	CYT	H6	2.39	6.08
23	ADE	H5'2	23	ADE	H8	2.95	4.98
6	TG	H5'1	6	TG	M	3.84	7.78
10	THY	H6	11	GUA	H8	3.91	6.18
12	THY	H6	11	GUA	H8	3.56	6.36
18	CYT	H6	17	ADE	H8	3.75	5.33
18	CYT	H6	19	ADE	H8	3.00	5.15
20	CYT	H6	21	GUA	H8	3.47	6.33
22	CYT	H6	21	GUA	H8	2.86	5.81
22	CYT	H6	23	ADE	H8	2.78	6.80
24	CYT	H6	23	ADE	H8	3.01	5.75
4	CYT	H6	3	GUA	H8	2.73	5.74

	4	CYT	H6	5	GUA	H8	2.86	5.93
	8	THY	H6	7	GUA	H8	2.10	6.62
Class 3	10	THY	H1'	10	THY	H4'	1.67	3.80
	12	THY	H1'	12	THY	H4'	2.14	3.51
	12	THY	H1'	12	THY	H6	2.21	6.46
	13	ADE	H1'	14	CYT	H5'1	1.81	4.69
	18	CYT	H1'	18	CYT	H4'	1.84	4.96
	18	CYT	H1'	18	CYT	H5'1	2.94	5.70
	18	CYT	H1'	19	ADE	H5'1	2.13	4.14
	2	THY	H1'	2	THY	H2'2	1.79	2.65
	20	CYT	H1'	20	CYT	H3'	2.03	5.18
	20	CYT	H1'	20	CYT	H4'	1.70	4.02
	22	CYT	H1'	22	CYT	H4'	2.05	3.56
	22	CYT	H1'	23	ADE	H5'1	3.14	4.32
	23	ADE	H1'	24	CYT	H5	2.81	5.87
	24	CYT	H1'	24	CYT	H4'	1.88	4.52
	3	GUA	H1'	4	CYT	H5	3.26	4.68
	3	GUA	H1'	4	CYT	H5'1	2.32	3.23
	4	CYT	H1'	5	GUA	H5'1	2.66	3.53
	5	GUA	H1'	5	GUA	H3'	2.09	5.64
	6	TG	H1'	6	TG	H4'	2.05	3.27
	6	TG	H1'	6	TG	H6	1.78	5.00
	6	TG	H1'	7	GUA	H8	2.09	3.65
	7	GUA	H1'	7	GUA	H8	1.88	5.97
	9	THY	H1'	9	THY	H4'	2.08	4.38
	9	THY	H1'	10	THY	H6	2.13	4.40
	9	THY	H1'	9	THY	H6	1.77	4.31
	16	ADE	H2	10	THY	H3	4.67	6.41
	1	GUA	H2'1	1	GUA	H3'	2.25	4.46
	1	GUA	H2'1	2	THY	H6	2.38	4.26
	1	GUA	H2'1	2	THY	M7	3.53	7.92
	13	ADE	H2'1	14	CYT	H5	2.06	4.53
	13	ADE	H2'1	14	CYT	H6	2.64	5.16
	15	ADE	H2'1	15	ADE	H4'	2.90	4.97
	15	ADE	H2'1	16	ADE	H8	2.27	5.84
	16	ADE	H2'1	16	ADE	H4'	2.05	5.50
	17	ADE	H2'1	17	ADE	H5'1	2.26	4.23
	18	CYT	H2'1	18	CYT	H5	1.99	4.76
	19	ADE	H2'1	19	ADE	H4'	1.98	5.58
	2	THY	H2'1	2	THY	H4'	2.17	4.27
	21	GUA	H2'1	21	GUA	H4'	2.54	5.63
	22	CYT	H2'1	22	CYT	H4'	3.20	4.62
	23	ADE	H2'1	23	ADE	H4'	2.55	5.66
	24	CYT	H2'1	24	CYT	H5	1.95	4.71
	3	GUA	H2'1	3	GUA	H4'	2.50	4.56
	3	GUA	H2'1	4	CYT	H6	2.54	4.32
	4	CYT	H2'1	4	CYT	H4'	3.23	4.57
	6	TG	H2'1	6	TG	M	2.13	7.21
	7	GUA	H2'1	7	GUA	H4'	2.95	5.17
	8	THY	H2'1	8	THY	H4'	2.29	5.54

	8	THY	H2'1	8	THY	M7	2.59	5.88
	1	GUA	H2'2	2	THY	H5'1	2.03	5.92
	1	GUA	H2'2	2	THY	H6	2.55	5.17
	1	GUA	H2'2	1	GUA	H8	1.98	4.21
	13	ADE	H2'2	13	ADE	H5'1	2.07	4.98
	14	CYT	H2'2	15	ADE	H8	2.74	5.96
	16	ADE	H2'2	16	ADE	H4'	1.97	3.01
	17	ADE	H2'2	18	CYT	H5	2.41	4.22
	18	CYT	H2'2	19	ADE	H5'1	2.30	4.42
	19	ADE	H2'2	20	CYT	H5	2.41	5.24
	21	GUA	H2'2	21	GUA	H4'	2.31	4.20
	24	CYT	H2'2	24	CYT	H4'	1.95	4.41
	3	GUA	H2'2	3	GUA	H4'	2.58	3.55
	3	GUA	H2'2	4	CYT	H6	2.15	3.17
	7	GUA	H2'2	7	GUA	H4'	2.39	4.71
	8	THY	H2'2	8	THY	H4'	1.94	4.80
	1	GUA	H3'	1	GUA	H5'1	2.00	4.14
	13	ADE	H3'	14	CYT	H6	2.31	5.58
	14	CYT	H3'	15	ADE	H8	3.14	6.49
	15	ADE	H3'	15	ADE	H4'	1.95	3.52
	20	CYT	H3'	20	CYT	H6	2.05	5.17
	20	CYT	H3'	21	GUA	H8	3.19	5.62
	22	CYT	H3'	22	CYT	H6	2.50	5.48
	24	CYT	H3'	24	CYT	H4'	2.06	4.29
	4	CYT	H3'	4	CYT	H6	2.25	6.20
	4	CYT	H3'	5	GUA	H8	2.42	6.32
	6	TG	H3'	6	TG	H4'	2.11	5.45
	1	GUA	H4'	1	GUA	H8	2.26	6.49
	10	THY	H4'	10	THY	H6	2.94	6.11
	11	GUA	H4'	11	GUA	H8	2.41	6.86
	15	ADE	H4'	15	ADE	H8	2.33	6.81
	3	GUA	H4'	3	GUA	H8	2.76	6.79
	4	CYT	H4'	4	CYT	H6	1.67	5.00
	8	THY	H4'	8	THY	H6	2.27	4.49
	10	THY	H5'1	10	THY	H6	1.91	5.05
	11	GUA	H5'1	11	GUA	H8	2.39	5.37
	17	ADE	H5'1	17	ADE	H8	2.04	6.59
	18	CYT	H5'1	18	CYT	H6	2.11	4.62
	3	GUA	H5'1	3	GUA	H8	2.05	5.42
	6	TG	H5'1	6	TG	M	2.66	7.83
	3	GUA	H5'2	3	GUA	H8	2.41	6.60
	7	GUA	H5'2	7	GUA	H8	2.12	5.00
	14	CYT	H6	13	ADE	H8	2.99	5.68
	14	CYT	H6	15	ADE	H8	2.81	6.38
	11	GUA	H8	12	THY	M7	3.37	6.98
Class 4	11	GUA	H1'	12	THY	H6	2.42	5.56
	14	CYT	H1'	14	CYT	H2'1	1.92	4.32
	14	CYT	H1'	14	CYT	H4'	1.77	5.01
	16	ADE	H1'	17	ADE	H5'1	2.61	4.03
	19	ADE	H1'	19	ADE	H2	3.57	6.09

2	THY	H1'	2	THY	H4'	1.85	3.45
2	THY	H1'	2	THY	H5'1	2.25	3.75
2	THY	H1'	3	GUA	H8	2.11	4.20
20	CYT	H1'	20	CYT	H5'1	3.30	3.61
22	CYT	H1'	22	CYT	H2'1	2.03	5.51
3	GUA	H1'	3	GUA	H2'1	2.03	3.98
3	GUA	H1'	3	GUA	H8	2.07	4.22
6	TG	H1'	7	GUA	H5'1	2.22	3.60
7	GUA	H1'	8	THY	H6	1.80	6.59
8	THY	H1'	8	THY	H6	2.88	6.91
8	THY	H1'	9	THY	H6	2.46	5.08
9	THY	H1'	10	THY	H5'1	2.17	4.97
10	THY	H2'1	10	THY	H4'	1.71	2.98
10	THY	H2'1	11	GUA	H8	2.56	4.75
13	ADE	H2'1	13	ADE	H5'1	2.45	3.48
16	ADE	H2'1	17	ADE	H8	2.49	6.56
17	ADE	H2'1	17	ADE	H4'	2.08	4.69
17	ADE	H2'1	18	CYT	H6	2.17	4.15
18	CYT	H2'1	19	ADE	H8	2.25	3.75
19	ADE	H2'1	20	CYT	H6	2.27	4.43
2	THY	H2'1	2	THY	H3'	1.86	3.39
20	CYT	H2'1	20	CYT	H4'	1.95	3.85
21	GUA	H2'1	22	CYT	H6	2.20	5.65
22	CYT	H2'1	23	ADE	H8	2.49	4.14
23	ADE	H2'1	24	CYT	H6	2.46	5.11
4	CYT	H2'1	5	GUA	H8	2.42	3.68
5	GUA	H2'1	6	TG	M	2.94	6.35
6	TG	H2'1	7	GUA	H8	2.28	4.82
7	GUA	H2'1	8	THY	H6	1.61	3.65
7	GUA	H2'1	8	THY	M7	2.89	6.44
8	THY	H2'1	8	THY	H5'1	2.19	5.34
1	GUA	H2'2	1	GUA	H3'	1.94	2.97
1	GUA	H2'2	2	THY	M7	3.01	4.69
13	ADE	H2'2	14	CYT	H6	2.27	4.20
13	ADE	H2'2	13	ADE	H8	2.27	6.46
15	ADE	H2'2	15	ADE	H8	2.05	6.60
17	ADE	H2'2	17	ADE	H4'	2.29	5.26
17	ADE	H2'2	18	CYT	H6	2.12	6.14
18	CYT	H2'2	18	CYT	H6	1.82	4.29
18	CYT	H2'2	19	ADE	H8	2.04	4.04
19	ADE	H2'2	19	ADE	H4'	2.00	5.89
2	THY	H2'2	2	THY	H3'	1.72	3.17
2	THY	H2'2	2	THY	H6	1.75	3.59
2	THY	H2'2	3	GUA	H8	2.12	5.93
20	CYT	H2'2	20	CYT	H4'	2.16	3.68
21	GUA	H2'2	22	CYT	H5	2.01	4.83
21	GUA	H2'2	22	CYT	H6	2.14	4.15
22	CYT	H2'2	22	CYT	H6	1.89	5.84
22	CYT	H2'2	23	ADE	H8	1.80	3.91
23	ADE	H2'2	24	CYT	H6	2.01	4.70

	23	ADE	H2'2	23	ADE	H8	1.91	6.07
	24	CYT	H2'2	24	CYT	H3'	2.09	5.24
	4	CYT	H2'2	4	CYT	H4'	2.17	4.20
	4	CYT	H2'2	4	CYT	H6	1.84	5.32
	5	GUA	H2'2	5	GUA	H8	2.00	4.59
	5	GUA	H2'2	6	TG	M	3.27	6.22
	6	TG	H2'2	6	TG	H4'	2.35	4.68
	6	TG	H2'2	7	GUA	H8	2.24	5.12
	7	GUA	H2'2	7	GUA	H8	2.79	6.12
	8	THY	H2'2	8	THY	H6	1.92	4.12
	8	THY	H2'2	9	THY	H6	1.71	3.63
	1	GUA	H3'	2	THY	H6	2.51	5.50
	12	THY	H3'	12	THY	H4'	2.05	4.91
	12	THY	H3'	13	ADE	H4'	3.27	4.55
	15	ADE	H3'	15	ADE	H5'1	2.24	4.61
	15	ADE	H3'	16	ADE	H8	2.51	5.58
	16	ADE	H3'	16	ADE	H4'	1.91	3.65
	16	ADE	H3'	16	ADE	H5'1	2.21	5.05
	16	ADE	H3'	16	ADE	H8	1.91	5.22
	18	CYT	H3'	18	CYT	H5'2	1.92	3.49
	2	THY	H3'	2	THY	H4'	1.65	3.61
	24	CYT	H3'	24	CYT	H5'1	2.43	4.82
	24	CYT	H3'	24	CYT	H5'2	2.04	4.85
	4	CYT	H3'	4	CYT	H5'2	1.81	4.06
	5	GUA	H3'	5	GUA	H4'	2.35	4.56
	7	GUA	H3'	7	GUA	H4'	1.96	3.33
	12	THY	H4'	12	THY	H6	2.19	5.34
	9	THY	H4'	9	THY	H6	2.08	4.54
	14	CYT	H5'1	14	CYT	H6	2.12	5.27
	2	THY	H5'1	2	THY	H6	2.19	5.78
	4	CYT	H5'1	4	CYT	H6	2.74	4.37
	5	GUA	H5'1	5	GUA	H8	2.28	5.53
	4	CYT	H5'2	4	CYT	H6	2.14	6.01
	9	THY	H6	10	THY	M7	2.28	5.24
	5	GUA	H8	6	TG	M	2.31	5.09
	7	GUA	H8	8	THY	M7	2.76	4.99
Class 5	1	GUA	H1'	2	THY	H5'1	1.87	3.46
	10	THY	H1'	10	THY	H2'1	1.79	3.53
	10	THY	H1'	10	THY	H2'2	1.87	2.64
	11	GUA	H1'	11	GUA	H2'2	1.87	3.44
	12	THY	H1'	12	THY	H2'2	1.77	3.21
	13	ADE	H1'	13	ADE	H2'1	1.90	4.21
	13	ADE	H1'	13	ADE	H2'2	1.70	2.56
	13	ADE	H1'	13	ADE	H8	1.98	5.25
	14	CYT	H1'	14	CYT	H2'2	1.93	2.63
	15	ADE	H1'	15	ADE	H2'1	2.60	5.58
	15	ADE	H1'	15	ADE	H2'2	2.27	2.64
	16	ADE	H1'	16	ADE	H2'1	1.82	5.63
	16	ADE	H1'	16	ADE	H2'2	1.71	2.55
	17	ADE	H1'	17	ADE	H2'1	2.46	4.92

17	ADE	H1'	17	ADE	H2'2	1.85	2.54
18	CYT	H1'	18	CYT	H2'1	1.83	5.49
18	CYT	H1'	18	CYT	H2'2	1.97	2.74
19	ADE	H1'	19	ADE	H2'1	1.64	4.61
19	ADE	H1'	19	ADE	H2'2	1.94	2.58
19	ADE	H1'	19	ADE	H3'	1.85	5.59
19	ADE	H1'	19	ADE	H4'	1.79	4.02
20	CYT	H1'	20	CYT	H2'1	1.82	4.50
20	CYT	H1'	20	CYT	H2'2	1.93	2.54
21	GUA	H1'	21	GUA	H2'1	1.77	5.73
21	GUA	H1'	21	GUA	H2'2	1.69	2.57
21	GUA	H1'	22	CYT	H5'1	1.85	3.49
22	CYT	H1'	22	CYT	H2'2	1.97	2.51
23	ADE	H1'	23	ADE	H2'1	1.98	5.16
23	ADE	H1'	23	ADE	H2'2	2.17	2.78
24	CYT	H1'	24	CYT	H2'1	2.28	5.74
24	CYT	H1'	24	CYT	H2'2	1.80	2.69
3	GUA	H1'	3	GUA	H2'2	1.66	2.52
4	CYT	H1'	4	CYT	H2'1	2.10	4.90
4	CYT	H1'	4	CYT	H2'2	2.02	2.58
4	CYT	H1'	4	CYT	H4'	1.97	3.20
5	GUA	H1'	5	GUA	H2'1	1.80	5.36
5	GUA	H1'	5	GUA	H2'2	1.69	2.29
6	TG	H1'	6	TG	H2'1	1.93	5.55
6	TG	H1'	6	TG	H2'2	2.11	3.14
7	GUA	H1'	7	GUA	H2'1	1.75	2.81
7	GUA	H1'	7	GUA	H2'2	1.88	3.13
8	THY	H1'	8	THY	H2'1	1.98	3.58
8	THY	H1'	8	THY	H2'2	2.00	5.77
8	THY	H1'	8	THY	H4'	1.65	2.70
9	THY	H1'	9	THY	H2'2	1.67	3.07
1	GUA	H2'1	1	GUA	H2'2	1.54	2.88
1	GUA	H2'1	1	GUA	H8	2.14	4.50
10	THY	H2'1	10	THY	H2'2	1.52	3.03
10	THY	H2'1	10	THY	H6	1.80	5.12
13	ADE	H2'1	13	ADE	H3'	1.71	2.88
13	ADE	H2'1	13	ADE	H4'	2.08	3.41
13	ADE	H2'1	13	ADE	H8	1.84	3.62
14	CYT	H2'1	14	CYT	H2'2	1.63	3.04
14	CYT	H2'1	14	CYT	H3'	1.96	3.12
14	CYT	H2'1	14	CYT	H4'	2.04	3.14
14	CYT	H2'1	14	CYT	H6	1.75	3.77
15	ADE	H2'1	15	ADE	H2'2	1.44	3.04
15	ADE	H2'1	15	ADE	H3'	2.15	3.48
15	ADE	H2'1	15	ADE	H8	2.17	5.60
16	ADE	H2'1	16	ADE	H3'	1.88	3.55
16	ADE	H2'1	16	ADE	H8	1.82	2.91
17	ADE	H2'1	17	ADE	H2'2	1.59	3.01
17	ADE	H2'1	17	ADE	H3'	1.89	3.04
17	ADE	H2'1	17	ADE	H8	1.87	4.96

18	CYT	H2'1	18	CYT	H2'2	1.68	2.85
18	CYT	H2'1	18	CYT	H3'	1.94	3.81
18	CYT	H2'1	18	CYT	H6	2.00	4.70
19	ADE	H2'1	19	ADE	H2'2	1.54	2.83
19	ADE	H2'1	19	ADE	H3'	1.90	3.73
19	ADE	H2'1	19	ADE	H8	2.15	5.12
2	THY	H2'1	2	THY	H2'2	1.71	3.11
2	THY	H2'1	2	THY	H6	1.86	5.40
20	CYT	H2'1	20	CYT	H2'2	1.57	3.04
20	CYT	H2'1	20	CYT	H3'	1.66	3.28
20	CYT	H2'1	20	CYT	H6	1.77	4.49
20	CYT	H2'1	21	GUA	H8	2.31	4.28
21	GUA	H2'1	21	GUA	H3'	1.74	3.56
21	GUA	H2'1	21	GUA	H8	1.74	2.97
22	CYT	H2'1	22	CYT	H2'2	1.68	2.88
22	CYT	H2'1	22	CYT	H3'	2.19	3.84
22	CYT	H2'1	22	CYT	H6	1.92	4.95
23	ADE	H2'1	23	ADE	H2'2	1.43	3.01
23	ADE	H2'1	23	ADE	H3'	1.87	3.57
23	ADE	H2'1	23	ADE	H8	2.09	6.44
24	CYT	H2'1	24	CYT	H3'	1.92	3.31
24	CYT	H2'1	24	CYT	H6	2.09	3.13
3	GUA	H2'1	3	GUA	H3'	1.79	3.71
3	GUA	H2'1	3	GUA	H8	1.68	3.23
4	CYT	H2'1	4	CYT	H2'2	1.61	2.85
4	CYT	H2'1	4	CYT	H6	2.03	5.56
5	GUA	H2'1	5	GUA	H2'2	1.50	3.09
5	GUA	H2'1	5	GUA	H3'	1.74	3.51
5	GUA	H2'1	5	GUA	H8	1.98	6.10
6	TG	H2'1	6	TG	H2'2	1.56	2.92
6	TG	H2'1	6	TG	H3'	1.87	3.46
7	GUA	H2'1	7	GUA	H3'	1.83	3.99
7	GUA	H2'1	7	GUA	H8	1.97	3.02
8	THY	H2'1	8	THY	H2'2	1.49	3.35
8	THY	H2'1	8	THY	H3'	1.77	3.01
8	THY	H2'1	8	THY	H6	2.01	6.55
8	THY	H2'1	9	THY	H6	2.03	3.46
8	THY	H2'1	9	THY	M7	2.93	5.06
9	THY	H2'1	10	THY	H6	2.05	4.28
9	THY	H2'1	10	THY	M7	1.63	4.32
10	THY	H2'2	10	THY	H3'	1.63	3.47
10	THY	H2'2	10	THY	H4'	1.90	5.94
10	THY	H2'2	10	THY	H6	1.80	5.67
10	THY	H2'2	11	GUA	H8	2.38	3.84
11	GUA	H2'2	11	GUA	H3'	1.57	3.36
11	GUA	H2'2	11	GUA	H4'	2.10	4.78
11	GUA	H2'2	12	THY	H6	2.58	4.42
11	GUA	H2'2	11	GUA	H8	1.87	4.39
11	GUA	H2'2	12	THY	M7	3.47	5.77
12	THY	H2'2	12	THY	H3'	1.88	3.42

12	THY	H2'2	12	THY	H4'	2.25	3.93
13	ADE	H2'2	13	ADE	H3'	2.05	4.36
14	CYT	H2'2	14	CYT	H3'	1.71	3.15
14	CYT	H2'2	14	CYT	H6	1.68	4.78
15	ADE	H2'2	15	ADE	H3'	2.10	3.42
15	ADE	H2'2	16	ADE	H8	2.42	5.25
16	ADE	H2'2	16	ADE	H3'	1.84	3.96
16	ADE	H2'2	16	ADE	H8	2.02	5.04
16	ADE	H2'2	17	ADE	H8	1.54	3.23
17	ADE	H2'2	17	ADE	H3'	1.64	2.87
17	ADE	H2'2	17	ADE	H5'1	1.74	4.27
17	ADE	H2'2	17	ADE	H8	1.70	4.72
18	CYT	H2'2	18	CYT	H3'	1.90	3.38
18	CYT	H2'2	18	CYT	H4'	1.93	5.88
19	ADE	H2'2	19	ADE	H3'	1.82	2.73
19	ADE	H2'2	19	ADE	H8	1.89	4.31
20	CYT	H2'2	20	CYT	H3'	1.71	3.66
20	CYT	H2'2	20	CYT	H6	1.69	4.42
20	CYT	H2'2	21	GUA	H8	2.17	4.31
21	GUA	H2'2	21	GUA	H3'	1.72	3.02
21	GUA	H2'2	21	GUA	H8	2.32	5.44
22	CYT	H2'2	22	CYT	H3'	2.19	3.43
22	CYT	H2'2	22	CYT	H4'	2.60	5.71
23	ADE	H2'2	23	ADE	H3'	2.06	3.92
24	CYT	H2'2	24	CYT	H6	2.28	5.46
3	GUA	H2'2	3	GUA	H8	2.09	6.38
4	CYT	H2'2	4	CYT	H3'	2.54	4.31
4	CYT	H2'2	5	GUA	H8	2.41	3.53
5	GUA	H2'2	5	GUA	H3'	1.88	3.02
6	TG	H2'2	6	TG	H3'	1.62	3.86
7	GUA	H2'2	7	GUA	H3'	2.20	5.22
7	GUA	H2'2	8	THY	H6	1.90	4.22
7	GUA	H2'2	8	THY	M7	2.32	4.49
8	THY	H2'2	8	THY	H3'	1.69	2.79
8	THY	H2'2	9	THY	M7	2.87	7.19
9	THY	H2'2	10	THY	H6	1.91	5.48
9	THY	H2'2	9	THY	H6	1.83	5.93
9	THY	H2'2	10	THY	M7	3.41	5.76
1	GUA	H3'	1	GUA	H5'2	1.66	4.01
10	THY	H3'	10	THY	H4'	1.48	2.88
10	THY	H3'	10	THY	H5'1	1.66	4.09
10	THY	H3'	10	THY	H6	2.23	6.65
13	ADE	H3'	13	ADE	H4'	1.92	3.57
13	ADE	H3'	13	ADE	H5'1	1.81	2.94
14	CYT	H3'	14	CYT	H4'	1.70	3.35
15	ADE	H3'	15	ADE	H5'2	1.85	3.67
16	ADE	H3'	16	ADE	H5'2	1.76	3.47
17	ADE	H3'	17	ADE	H4'	1.81	4.24
17	ADE	H3'	17	ADE	H5'1	1.92	3.42
18	CYT	H3'	18	CYT	H4'	1.78	3.32

18	CYT	H3'	18	CYT	H5'1	1.72	3.28
19	ADE	H3'	19	ADE	H4'	1.70	2.85
19	ADE	H3'	19	ADE	H5'1	1.77	3.47
19	ADE	H3'	20	CYT	H6	2.25	5.35
20	CYT	H3'	20	CYT	H4'	1.55	2.87
20	CYT	H3'	20	CYT	H5'1	1.88	3.28
21	GUA	H3'	21	GUA	H4'	2.03	3.60
21	GUA	H3'	21	GUA	H5'1	2.52	5.20
21	GUA	H3'	21	GUA	H5'2	1.74	3.32
22	CYT	H3'	22	CYT	H4'	2.07	3.39
22	CYT	H3'	22	CYT	H5'1	1.54	2.03
23	ADE	H3'	23	ADE	H4'	2.10	2.93
23	ADE	H3'	23	ADE	H5'2	1.37	3.58
3	GUA	H3'	3	GUA	H4'	2.15	3.38
3	GUA	H3'	3	GUA	H5'1	2.24	3.76
3	GUA	H3'	3	GUA	H5'2	1.76	3.39
4	CYT	H3'	4	CYT	H4'	1.94	4.59
7	GUA	H3'	7	GUA	H5'1	1.82	2.92
7	GUA	H3'	7	GUA	H5'2	2.09	4.23
8	THY	H3'	9	THY	H6	1.71	6.59
9	THY	H3'	9	THY	H4'	1.57	2.73
9	THY	H3'	10	THY	H6	1.90	4.58
9	THY	H3'	9	THY	H6	2.08	5.44
1	GUA	H4'	1	GUA	H5'1	1.63	2.56
13	ADE	H4'	13	ADE	H5'1	1.75	2.66
15	ADE	H4'	15	ADE	H5'1	1.98	4.44
15	ADE	H4'	15	ADE	H5'2	1.81	2.93
17	ADE	H4'	17	ADE	H5'1	1.58	2.64
21	GUA	H4'	21	GUA	H5'1	1.66	4.16
21	GUA	H4'	21	GUA	H5'2	1.59	3.27
3	GUA	H4'	3	GUA	H5'1	1.48	3.42
3	GUA	H4'	3	GUA	H5'2	1.92	4.50
5	GUA	H4'	5	GUA	H5'1	1.70	2.71
7	GUA	H4'	7	GUA	H5'1	2.09	5.72
7	GUA	H4'	7	GUA	H5'2	1.61	3.02
14	CYT	H5	14	CYT	H6	1.79	2.82
18	CYT	H5	18	CYT	H6	1.86	2.71
20	CYT	H5	20	CYT	H6	1.87	2.69
22	CYT	H5	22	CYT	H6	1.82	2.73
24	CYT	H5	24	CYT	H6	1.94	2.82
4	CYT	H5	4	CYT	H6	1.80	3.03
6	TG	H5'1	6	TG	H5'2	1.57	2.05
10	THY	H6	10	THY	M7	2.28	3.82
12	THY	H6	12	THY	M7	2.22	3.95
2	THY	H6	2	THY	M7	2.28	3.74
8	THY	H6	8	THY	M7	2.23	3.68
8	THY	H6	9	THY	M7	2.86	6.81
9	THY	H6	9	THY	M7	2.17	3.52
7	GUA	H8	6	TG	M	5.29	7.38

APPENDIX D

AMBER MOLECULAR DYNAMICS CONTROL FILES

File D-1: Generic simulated annealing PBS script for use with the ACCRE cluster (Advanced Computing Center for Research and Education, Vanderbilt University). This file is used with a "foreach" loop script for submission of

---

```
#!/bin/csh
#
#PBS -M kyle.l.brown@vanderbilt.edu
#PBS -l nodes=1:ppn=1:x86
#PBS -l walltime=00:50:00
#PBS -l cput=0:50:00
#PBS -l mem=500mb
#PBS -o accre.output
#PBS -j oe
#

cd ~/"project_name"

date >> nodes.list
echo "$STRUCT ran on the following node:" >> nodes.list
cat $PBS_NODEFILE >> nodes.list

cat << eof > $STRUCT.in

 simulated annealing protocol, 22 ps

 &cntrl
    nstlim=22000, pencut=-0.001, nmropt=1,
    ntpr=200, ntt=1, ntwx=200,
    cut=12.0, ntb=0, vlimit=10, rgbmax=12.0,
    ntb=0, ntc=2, ntf=2,
    igb=1, saltcon=0.2, offset=0.13,
 /
 &ewald
 /
#
# simulated annealing algorithm:
# H bonds held w/ SHAKE
# generalized born solvent model
#
#from steps 0 to 5000: heat the system to 600K
#from steps 5001-18000: re-cool to low temperatures with long tautp
#from steps 18001-22000: final cooling with short tautp
#
 &wt type='TEMP0', istep1=0,istep2=5000,value1=600.,
```

```
            value2=600.,    /
 &wt type='TEMP0', istep1=5001, istep2=18000, value1=600.0,
            value2=100.0,     /
 &wt type='TEMP0', istep1=18001, istep2=22000, value1=0.0,
            value2=0.0,     /

 &wt type='TAUTP', istep1=0,istep2=5000,value1=0.4,
            value2=0.4,     /
 &wt type='TAUTP', istep1=5001,istep2=18000,value1=1.0,
            value2=0.4,     /
 &wt type='TAUTP', istep1=18001,istep2=19000,value1=1.0,
            value2=0.8,     /
 &wt type='TAUTP', istep1=19001,istep2=22000,value1=0.1,
            value2=0.01,    /

 &wt type='REST', istep1=0,istep2=3000,value1=0.1,
            value2=1.0,  /
 &wt type='REST', istep1=3001,istep2=22000,value1=1.0,
            value2=1.0,  /

 &wt type='END'  /
LISTOUT=POUT
DISANG=RST
eof

/usr/local/structbio/amber9/exe/sander.nopar -O -i $STRUCT.in -p
file.prmtop -c $STRUCT.x -o $STRUCT.out -r $STRUCT.rst -x $STRUCT.trj

rm -f $STRUCT.in

/usr/local/structbio/amber9/exe/ambpdb -p file.prmtop <$STRUCT.rst>
$STRUCT.pdb
```

File D-2: Control script for equilibration of structures in explicit solvent. A detailed explanation of explicit solvent equilibration prior to production calculations maybe found in Chapter II.

```
#!/bin/csh

#This is used to equilibrate a system for refinement in explicit
solvent

set pmemd=/sb/apps/amber9/x86_64/exe/pmemd

#Hold solute fixed and allow water to equilibrate around molecule
cat<<eof>eq1.in
initial minimisation solvent + ions, dna fixed
 &cntrl
  imin   = 1,
  maxcyc = 1000,
  ncyc   = 500,
  ntb    = 1,
  ntr    = 1,
  cut    = 10
 /
Hold the DNA fixed
500.0
RES 1 24
END
END
eof

$pmemd -O -i eq1.in -o eq1.out -p fam1.prmtop -c fam1.inpcrd -r
eq1out.rst -ref fam1.inpcrd

\rm eq1.in

#allow solute to adjust to water
cat<<eof>eq2.in
tgwater: initial minimisation whole system water and dna
 &cntrl
  imin   = 1,
  maxcyc = 2500,
  ncyc   = 1000,
  ntb    = 1,
  ntr    = 0,
  cut    = 10
 /
eof

$pmemd -O -i eq2.in -o eq2.out -p fam1.prmtop -c eq1out.rst -r
eq2out.rst

\rm eq2.in

cat<<eof>eq3.in
#tgwater heat to 300K: 20ps MD NO res on DNA
 &cntrl
  imin   = 0,
```

```
  irest  = 0,
  ntx    = 1,
  ntb    = 1,
  cut    = 10,
  ntr    = 1,
  ntc    = 2,
  ntf    = 2,
  tempi  = 0.0,
  temp0  = 300.0,
  ntt    = 3,
  gamma_ln = 1.0,
  nstlim = 10000, dt = 0.002,
  ntpr = 100, ntwx = 100, ntwr = 1000
 /
keep DNA fixed with weak restraints
10.0
RES 1 24
END
END
eof

$pmemd -O -i eq3.in -o eq3.out -p fam1.prmtop -c eq2out.rst -r
eq3out.rst -x eq3.traj -ref eq2out.rst

/rm eq3.in

cat<<eof>eq4.in

#300K, 100ps MD, NMR RST turned on slowly
 &cntrl
  imin = 0, irest = 1, ntx = 7,
  ntb = 2, pres0 = 1.0, ntp = 1,
  taup = 2.0,
  cut = 15,
  nmropt = 1,
  pencut =-0.001,
  ntpr   = 200,
  ntc = 2, ntf = 2,
  tempi = 300.0, temp0 = 300.0,
  ntt = 3, gamma_ln = 1.0,
  nstlim = 50000, dt = 0.002,
  ntpr = 100, ntwx = 100, ntwr = 1000
 /
 &wt type='REST', istep1=0,istep2=30000,value1=0.1,
            value2=1.0,  /
 &wt type='REST', istep1=30001,istep2=50000,value1=1.0,
            value2=1.0,  /
 &wt type='END' /
DISANG=RST
LISTOUT=POUT
END
END
eof

$pmemd -O -i eq4.in -o eq4.out -p fam1.prmtop -c eq3out.rst -r
eq4out.rst -x eq4.traj
/rm eq4.in
```

File D-3: Control script for explicit solvent production calculations

```
#!/bin/csh

cat << eof > prd.in
#4mer production run at 300K: 100ps MD, NMR RST turned on
 &cntrl
  imin = 0, irest = 1, ntx = 7,
  ntb = 2, pres0 = 1.0, ntp = 1,
  taup = 2.0,
  cut = 7,
  nmropt = 1,
  pencut =-0.001,
  ntpr   = 200,
  ntc = 2, ntf = 2,
  tempi = 300.0, temp0 = 300.0,
  ntt = 3, gamma_ln = 1.0,
  nstlim = 50000, dt = 0.002,
  ntpr = 100, ntwx = 100, ntwr = 1000
 /
 &wt type='END' /
DISANG=RST
LISTOUT=POUT
END
END

eof
set MDSTARTJOB=1
set MDENDJOB=10
set MDCURRENTJOB=$MDSTARTJOB
set MDINPUT=1

while ( $MDCURRENTJOB <= $MDENDJOB )
   echo -n "Job $MDCURRENTJOB started at: "
   date
   @ MDINPUT = $MDCURRENTJOB - 1

   /sb/apps/amber9/x86_64/exe/pmemd -O -i prd.in \
                            -o fam1_$MDCURRENTJOB.out \
                            -p fam1.prmtop \
                            -c fam1_$MDINPUT.rst \
                            -r fam1_$MDCURRENTJOB.rst \
                            -x fam1_$MDCURRENTJOB.traj
   gzip -9 -v fam1_$MDCURRENTJOB.traj
   echo -n "Job $MDCURRENTJOB finished at: "
   date
   @ MDCURRENTJOB = $MDCURRENTJOB + 1
end

rm prd.in
```

APPENDIX E

HELICOIDAL ANALYSIS RESULTS

Figure E-1: Helicoidal parameter definitions.

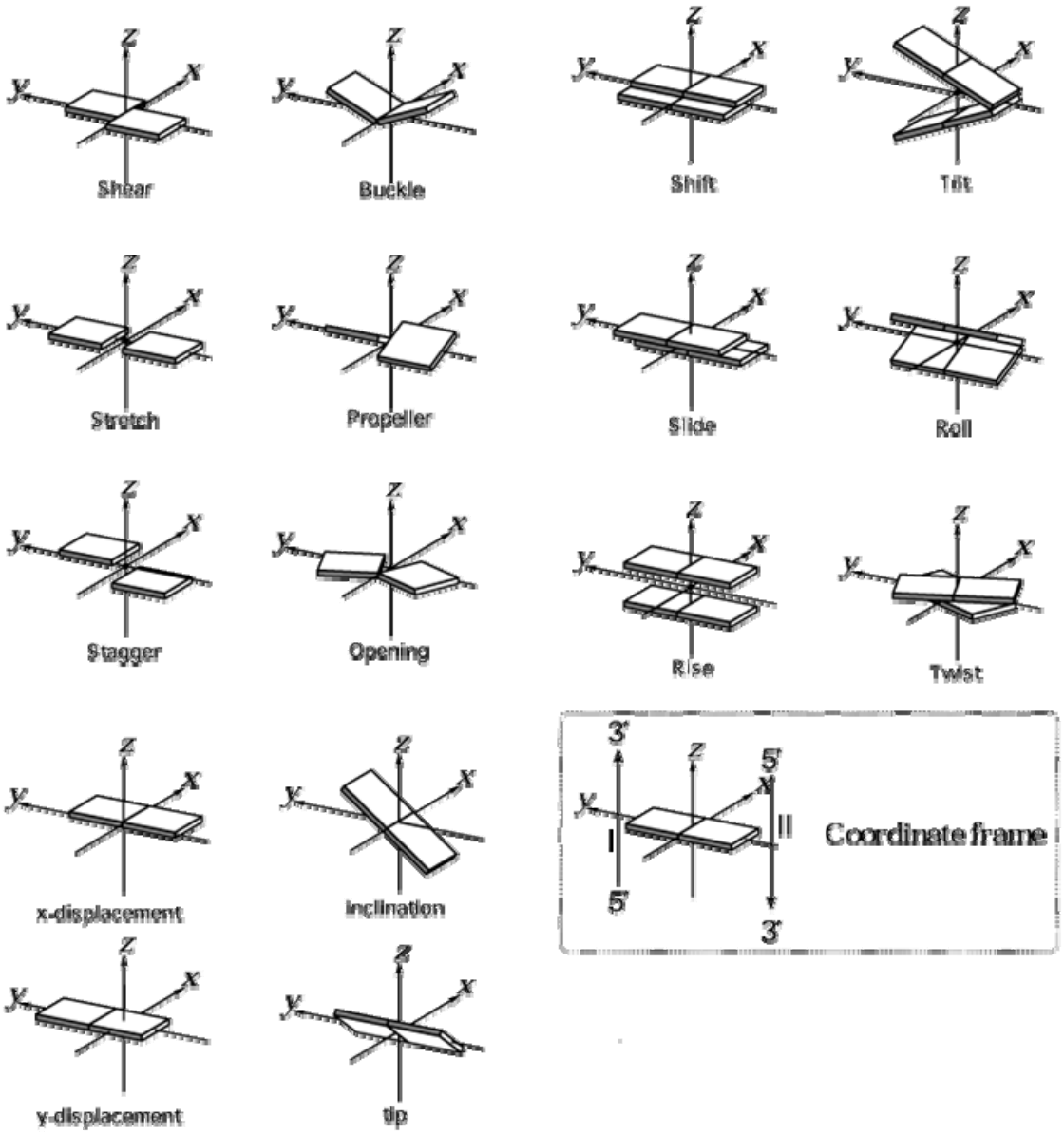


Table E-1: Global base-axis parameters for 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-5' (X = α-$AFB_1$-FAPY)

1st Strand	*X-disp (dx)*	*Y-disp (dy)*	*Inclination (eta)*	*Tip (theta)*	*Bc*	*Tc*
C^1	-0.93	0.19	-4.07	-3.11	3	0
T^2	-1.06	-0.34	-3.27	-7.44	4	-27
A^3	-0.93	-0.16	-1.48	-5.12	2	32
T^4	-1.13	-0.47	8.86	-7.04	4	-30
X^5	1.03	0.13	0.58	-99.6	1	12
A^6	-0.83	-0.08	-6.57	-4.54	2	22
T^7	-1.01	-0.26	-4.73	-18.25	4	-24
T^8	-1.09	-0.26	-6.13	-11.76	4	-18
C^9	-0.36	-0.26	-5.1	-10.92	3	-12
A^{10}	-0.92	-0.16	-5.27	-6.37	2	0
2nd Strand	*X-disp (dx)*	*Y-disp (dy)*	*Inclination (eta)*	*Tip (theta)*	*Bc*	*Tc*
G^{20}	-0.93	-0.23	-4.18	6.54	1	0
A^{19}	-0.96	0.33	-3.85	0.02	2	27
T^{18}	-1.05	0.19	-4.27	4.24	4	-32
A^{17}	-0.99	0.64	-7.5	-2.05	2	30
C^{16}	-0.74	0.01	-7.62	-12.74	3	-12
T^{15}	-0.9	0.14	-2.97	-8.93	4	-22
A^{14}	-0.97	0.27	-3.11	0.49	2	24
A^{13}	-0.68	0.27	1.54	-4.14	2	18
G^{13}	-0.94	0.15	5.81	0.77	1	12
T^{11}	-1.12	0.13	4.17	-11.04	4	0

Table E-2: Global base pair-axis parameters for 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-5' (X = α-$AFB_1$-FAPY)

Base Pair	*X-disp (dx)*	*Y-disp (dy)*	*Inclination (eta)*	*Tip (theta)*	*Bc*	*Tc*
C^1-G^{20}	-0.93	0.21	-4.12	-4.83	3	0
T^2-A^{19}	-1.01	-0.33	-3.56	-3.73	4	-27
A^3-T^{18}	-0.99	-0.18	-2.88	-4.68	2	32
T^4-A^{17}	-1.06	-0.56	0.68	-2.5	4	-30
X^5-C^{16}	0.14	0.06	-3.52	-43.43	1	12
A^6-T^{15}	-0.87	-0.11	-4.77	2.2	2	22
T^7-A^{14}	-0.99	-0.27	-3.92	-9.37	4	-24
T^8-A^{13}	-0.89	-0.26	-2.29	-3.81	4	-18
C^9-G^{12}	-0.65	-0.2	0.35	-5.85	3	-12
A^{10}-T^{11}	-1.02	-0.14	-0.55	2.34	2	0
Average:	-0.83	-0.18	-2.46	-7.37		

Table E-3: Global base-base parameters for 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-5' (X = α-$AFB_1$-FAPY)

Base Pair	*Shear (Sx)*	*Stretch (Sy)*	*Stagger (Sz)*	*Buckle (kappa)*	*Propel (omega)*	*Opening (sigma)*	*Bc*	*Tc*

C^{1}-G^{20}	0	-0.05	0.07	0.11	3.43	-3.12	3	0
T^{2}-A^{19}	-0.1	-0.01	-0.07	0.58	-7.42	0.72	4	-27
A^{3}-T^{18}	0.12	0.03	-0.03	2.78	-0.88	1.55	2	32
T^{4}-A^{17}	-0.13	0.17	-0.29	16.36	-9.09	7.61	4	-30
X^{5}-C^{16}	1.77	0.14	-0.12	8.2	-112.34	32.45	1	12
A^{6}-T^{15}	0.06	0.06	0.14	-3.6	-13.48	2.03	2	22
T^{7}-A^{14}	-0.04	0.01	-0.08	-1.62	-17.76	-0.99	4	-24
T^{8}-A^{13}	-0.41	0.01	-0.08	-7.67	-15.9	0.65	4	-18
C^{9}-G^{12}	0.58	-0.11	0.09	-10.92	-10.15	-1.61	3	-12
A^{10}-T^{11}	0.19	-0.03	-0.13	-9.43	-17.42	0.69	2	0
Average:	0.2	0.02	-0.05	-0.52	-20.1	4		

Table E-4: Global inter-base parameters for 5'-$C^{1}T^{2}A^{3}T^{4}\underline{X}^{5}A^{6}T^{7}T^{8}C^{9}A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-5' (X = α-$AFB_1$-FAPY)

1st Strand	*Shift (Dx)*	*Slide (Dy)*	*Rise (Dz)*	*Tilt (tau)*	*Roll (rho)*	*Twist (Omega)*	*Dc*
C^{1}-T^{2}	-0.29	-0.88	3.3	-1.46	-2.94	31.67	-3
T^{2}-A^{3}	-0.32	-0.25	3.19	-0.9	12.19	37.03	10
A^{3}-T^{4}	-0.05	0	2.92	8.08	-3.48	33.68	7
T^{4}-X^{5}	2.89	1.37	7.85	-1.04	-105.24	45.13	-9
X^{5}-A^{6}	-2.15	-0.35	3.45	-10.79	101.87	8.94	2
A^{6}-T^{7}	-0.85	-0.7	3.15	1.62	-9.74	29.56	7
T^{7}-T^{8}	0.12	-0.05	3.45	-3.81	1.36	38.85	-4
T^{8}-C^{9}	0.85	0.12	3.54	-1.09	8.7	39.47	-2
C^{9}-A^{10}	-0.73	0.18	3.23	-1.3	9.72	37.47	9
2nd Strand	*Shift (Dx)*	*Slide (Dy)*	*Rise (Dz)*	*Tilt (tau)*	*Roll (rho)*	*Twist (Omega)*	*Dc*
G^{20}-A^{19}	-0.19	0.92	3.44	-1.94	-7.9	27.83	3
A^{19}-T^{18}	-0.54	0.29	3.14	-3.1	-5.66	36.2	10
T^{18}-A^{17}	0.21	0.14	3.18	-5.5	-4.73	27.62	7
A^{17}-C^{16}	0.98	-1.4	7.67	7.12	1.98	20.28	9
C^{16}-T^{15}	-0.44	0.26	3.19	1	-3	39.37	-2
T^{15}-A^{14}	-0.75	0.65	3.37	-0.36	5.46	32.57	7
A^{14}-A^{13}	0.49	0.05	3.45	2.24	0.5	37.21	4
A^{13}-G^{12}	-0.14	-0.24	3.37	2.16	-2.95	41.73	2
G^{12}-T^{11}	-0.34	-0.09	3.45	-2.79	-16.98	35.17	-9

Table E-5: Global inter-base pair parameters for 5'-$C^{1}T^{2}A^{3}T^{4}\underline{X}^{5}A^{6}T^{7}T^{8}C^{9}A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-5' (X = α-$AFB_1$-FAPY)

Residue	*Shift (Dx)*	*Slide (Dy)*	*Rise (Dz)*	*Tilt (tau)*	*Roll (rho)*	*Twist (Omega)*	*Dc*
C^{1}-T^{2}	-0.24	-0.9	3.37	-1.7	2.48	29.75	-3
T^{2}-A^{3}	-0.43	-0.27	3.16	-2	8.92	36.61	10
A^{3}-T^{4}	0.08	-0.07	3.05	1.29	0.63	30.65	7
T^{4}-X^{5}	1.93	1.39	7.76	3.04	-53.61	32.71	-9
X^{5}-A^{6}	-1.3	-0.31	3.32	-4.89	52.43	24.15	2
A^{6}-T^{7}	-0.8	-0.68	3.26	0.63	-7.6	31.07	7

T^7-T^8	0.31	-0.05	3.45	-0.79	0.43	38.03	-4
T^8-C^9	0.36	0.18	3.45	0.53	5.82	40.6	-2
C^9-A^{10}	-0.54	0.13	3.34	-2.04	13.35	36.32	9
Average:	-0.07	-0.06	3.8	-0.66	2.54	33.32	

Table E-6: Local inter-base parameters for 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-5' (X = α-$AFB_1$-FAPY)

1st Strand	*Shift (Dx)*	*Slide (Dy)*	*Rise (Dz)*	*Tilt (tau)*	*Roll (rho)*	*Twist (Omega)*	*Dc*
C^1-T^2	0.04	-1.58	3.26	1.38	-4.81	31.72	-3
T^2-A^3	0.15	-1.04	3.32	3.16	10.57	37.32	10
A^3-T^4	0.38	-0.39	2.93	11.06	-1.15	32.82	7
T^4-X^5	6.79	4.61	2.85	69.38	-81.48	47.67	-9
X^5-A^6	1.35	-0.99	5.82	3.39	100.92	21.07	2
A^6-T^7	-0.16	-1.37	3.24	7.45	-12.15	29.11	7
T^7-T^8	1.11	-1.11	3.1	6.1	-2.58	37.84	-4
T^8-C^9	1.62	-0.71	3.35	6.57	4.81	39.52	-2
C^9-A^{10}	-0.1	-0.56	3.3	4.3	6.12	37.61	9
2nd Strand	*Shift (Dx)*	*Slide (Dy)*	*Rise (Dz)*	*Tilt (tau)*	*Roll (rho)*	*Twist (Omega)*	*Dc*
G^{20}-A^{19}	0.02	-1.63	3.35	-0.35	5.8	27.96	3
A^{19}-T^{18}	-0.25	-1.11	3.27	-1.81	3.3	36.81	10
T^{18}-A^{17}	0.47	-0.95	3.07	-4.82	1.66	27.42	7
A^{17}-C^{16}	-0.03	0.09	7.55	6.7	-3.28	19.52	9
C^{16}-T^{15}	-0.96	-1.09	3.08	-7.56	-0.97	39.01	-2
T^{15}-A^{14}	-0.87	-1.27	3.35	-2.76	-6.78	32.67	7
A^{14}-A^{13}	0.5	-0.67	3.4	0.83	-1.23	37.14	4
A^{13}-G^{12}	-0.12	-0.09	3.49	0.73	5.87	41.09	2
G^{12}-T^{11}	-0.54	-0.35	3.59	-5.89	19.41	34.69	-9

Table E-7: Local inter-base pair parameters for 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-5' (X = α-$AFB_1$-FAPY)

Residue	*Shift (Dx)*	*Slide (Dy)*	*Rise (Dz)*	*Tilt (tau)*	*Roll (rho)*	*Twist (Omega)*	*Dc*
C^1-T^2	0.03	-1.6	3.3	0.45	0.49	29.8	-3
T^2-A^3	-0.05	-1.07	3.3	0.63	6.96	37.12	10
A^3-T^4	0.42	-0.67	3.01	3.17	0.55	29.85	7
T^4-X^5	3.59	3.52	6.91	40.02	-38.46	21.63	-9
X^5-A^6	0.1	-0.77	3.97	-20.63	47.48	42.15	2
A^6-T^7	-0.51	-1.32	3.31	2.36	-9.46	31.21	7
T^7-T^8	0.81	-0.9	3.26	3.5	-1.93	37.89	-4
T^8-C^9	0.75	-0.4	3.51	3.68	5.34	40.47	-2
C^9-A^{10}	-0.31	-0.46	3.44	-0.74	12.82	36.81	9
Average:	0.54	-0.41	3.78	3.61	2.64	34.1	

Table E-8: Global axis curvature parameters for 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-5' (X = α-$AFB_1$-FAPY)

Residue	*Ax*	*Ay*	*Ainc*	*Atip*	*Adis*	*Angle*	*Path*	*Dc*
C^1-T^2	-0.16	-0.36	-2.26	1.38	0.39	2.65	3.39	-3
T^2-A^3	-0.45	-0.43	-2.69	9.87	0.62	10.23	3.21	10
A^3-T^4	0.15	0.31	-2.27	-1.56	0.35	2.75	3.07	7
T^4-X^5	0.73	0.77	7.24	-12.67	1.06	14.59	7.77	-9
X^5-A^6	-0.29	-0.14	-3.64	6.81	0.32	7.72	3.33	2
A^6-T^7	-0.68	-0.52	-0.22	3.97	0.85	3.98	3.37	7
T^7-T^8	0.2	-0.06	-2.42	-5.12	0.21	5.66	3.46	-4
T^8-C^9	0.12	0.12	-2.11	7.86	0.17	8.14	3.45	-2
C^9-A^{10}	-0.17	0.07	-1.14	5.16	0.18	5.29	3.34	9
Overall axis bend => UU=12.98 PP=9.27								

Table E-9: Backbone parameters for 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-5' (X = α-$AFB_1$-FAPY)

Residue	*C1′-C2′*	*C2′-C3′*	*Phase*	*Amplitude*	*Pucker*	*C1′*	*C2′*	*C3′*
C^1	22.99	-22.59	159.8	24.15	C2'-endo	104.4	105.5	105
T^2	36.36	-25.28	132.67	37.63	C1'-exo	102.3	104.4	103.7
A^3	30.29	-26.58	151.69	30.62	C2'-endo	104	105	103.1
T^4	34.89	-27.77	143.12	35.3	C1'-exo	104.3	102.5	104.9
X^5	34.12	-24.64	134.94	35.12	C1'-exo	102.3	104.8	104.1
A^6	26.11	-28.75	171.45	29.33	C2'-endo	104.9	104.8	103.7
T^7	31.76	-19.45	124.85	34.3	C1'-exo	103	105.1	104.1
T^8	30.97	-18.27	122.82	33.93	C1'-exo	102.8	105.5	104.2
C^9	34.5	-22.36	128.37	36.36	C1'-exo	102.4	105	103.7
A^{10}	14.17	-6.5	112.35	17.2	C1'-exo	105.2	107.1	105.4
	Chi	*Gamma*	*Delta*	*Epsilon*	*Zeta*	*Alpha*	*Beta*	
	C1′-N	*C5′-C4′*	*C4′-C3′*	*C3′-O3′*	*O3′-P*	*P-O5′*	*O5′-C5′*	
C^1	-139.6	173.28	134.5	-169.01	-84.77	-69.47	174.5	
T^2	-125.78	48.48	126.55	-170.6	-100.83	-71.42	-173.63	
A^3	-109.29	44.66	134.16	-174.7	-104.97	-66.37	168.88	
T^4	-106.68	57.98	136.47	-148.71	-69.69	-81	-178.73	
X^5	80.7	48.82	131.6	-107.79	122.51	-86.3	168.4	
A^6	-117.63	56.42	142.36	-180	-95.72	-64.82	178.24	
T^7	-120.67	54.99	118.97	177.3	-90.77	-58.69	173.74	
T^8	-125.23	51.62	119.09	-174.34	-85.61	-62.59	170.79	
C^9	-116.82	52.29	122.5	-172.11	-111.23	-70.83	-178.56	
A^{10}	-119.35	52.12	116.52	.....	.....	.....	.....	
	C1′-C2′	*C2′-C3′*	*Phase*	*Amplitude*	*Pucker*	*C1′*	*C2′*	*C3′*
G^{20}	27.05	-8.51	103.55	36.97	O1'-endo	104.1	103.4	106.1
A^{19}	32.12	-27.84	149.81	32.55	C2'-endo	103.8	104.3	103.8
T^{18}	34.56	-25.3	136.74	35.18	C1'-exo	103.1	104.3	103.5
A^{17}	32.99	-28.11	148.92	33.22	C2'-endo	103.6	104.6	103
C^{16}	-3.24	21.25	49.96	33.39	C4'-exo	104.9	106.1	103.5
T^{15}	24.58	-5.36	98.47	36.38	O1'-endo	103.1	106.5	104.1
A^{14}	25.88	-26.56	165.39	27.78	C2'-endo	105.1	104.8	104.1
A^{13}	31.01	-33.84	171.29	34.7	C2'-endo	104.8	103.4	102.8

G^{13}	38.77	-33.48	150.52	39.09	C2'-endo	103.2	102.7	102.5
T^{11}	23.91	-9.29	107.57	31.05	O1'-endo	103.8	105.3	106.7

	Chi C1'-N	*Gamma C5'-C4'*	*Delta C4'-C3'*	*Epsilon C3'-O3'*	*Zeta O3'-P*	*Alpha P-O5'*	*Beta O5'-C5'*
G^{20}	-130.92	49.75	105.53	.....	.....	.....	.....
A^{19}	-115.74	48.77	134.04	-173.24	-93.71	-71.96	173.2
T^{18}	-120.34	58.87	126.55	179.45	-98.95	-65.94	-175.64
A^{17}	-95.09	65.11	133.57	178.42	-114.35	-63.43	169.69
C^{16}	-138.35	50.1	87.13	-159.48	-87.49	-54.69	-168.02
T^{15}	-134.65	54.92	102.57	-175.61	-85.95	-66.73	169.57
A^{14}	-116.97	52.8	137.4	-176.81	-88.07	-64.22	167.07
A^{13}	-113.09	58.48	148.1	-172.91	-116.03	-69.05	-179.21
G^{13}	-99.67	49.11	139.63	-168.3	-131.44	-68.2	169.25
T^{11}	-143.54	176.66	116.76	-150.74	-78.91	-81.97	-179.89

Table E-10: Global base-axis parameters for 5'-$C^1T^2\underline{X}^3A^4$-3 (X = α-AFB_1-FAPY)

Residue	*X-disp (dx)*	*Y-disp (dy)*	*Inclination (eta)*	*Tip (theta)*	*Bc*	*Tc*
C^1	-0.61	-3.92	16.7	-15.96	3	0
T^2	-3.06	-4.24	8.94	1.9	4	-25
X^3	-0.97	-3.84	14.38	-35.31	1	12
A^4	-3.13	-4.08	7.16	-5.26	2	0

Table E-11: Global inter-base parameters for 5'-$C^1T^2\underline{X}^3A^4$-3 (X = α-AFB_1-FAPY)

Residue	*Shift (Dx)*	*Slide (Dy)*	*Rise (Dz)*	*Tilt (tau)*	*Roll (rho)*	*Twist (Omega)*	*Dc*
C^1-T^2	-2.44	-0.32	5.2	-13.03	-0.65	36.45	-3
T^2-X^3	2.13	0.49	8.06	-3.45	-65.22	17.86	-9
X^3-A^4	-2.13	-0.13	5.8	-10.43	18.32	-52.07	2

Table E-12: Local inter-base parameters for 5'-$C^1T^2\underline{X}^3A^4$-3 (X = α-AFB_1-FAPY)

Residue	*Shift (Dx)*	*Slide (Dy)*	*Rise (Dz)*	*Tilt (tau)*	*Roll (rho)*	*Twist (Omega)*	*Dc*
C^1-T^2	0.94	-0.63	5.07	-7.6	6.53	40.78	-3
T^2-X^3	5.1	1.88	6.88	4.6	-60.25	26.18	-9
X^3-A^4	-3.7	3.05	6.44	-25.59	6.14	-42.79	2

Table E-13: Global axis curvature parameters for 5'-$C^1T^2\underline{X}^3A^4$-3 (X = α-AFB_1-FAPY)

Residue	*Ax*	*Ay*	*Ainc*	*Atip*	*Adis*	*Angle*	*Path*	*Dc*
C^1-T^2	0.01	0	-5.27	-18.5	0.01	19.23	5.13	-3
T^2-X^3	0.04	0.09	-8.89	-28.02	0.1	29.37	7.79	-9
X^3-A^4	0.03	0.1	-3.22	-11.72	0.11	12.15	5.77	2
Overall axis bend => UU=59.46 PP=101.41								

Table E-14: Global base-axis parameters for 5'-$G^1T^2G^3C^4G^5\underline{X}^6G^7T^8T^9T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3' (X = 5*R*-thymine glycol)

Residue	*X-disp (dx)*	*Y-disp (dy)*	*Inclination (eta)*	*Tip (theta)*	*Bc*	*Tc*
G^1	-1.86	0.33	4.43	1.73	1	0
T^2	-1.80	0	10.48	-6.86	4	-29
G^3	-1.89	0.27	8.81	-3.13	1	15
C^4	-1.92	-0.16	3.81	-10.53	3	-13
G^5	-1.79	0.22	3.85	-4.1	1	14
X^6	-1.73	-0.09	13.32	-27.08	4	-29
G^7	-1.94	0.06	-5.88	-12	1	16
T^8	-1.79	-0.14	-5.02	-10.37	4	-23
T^9	-2.16	-0.04	2.89	-5.8	4	-20
T^{10}	-1.57	-0.22	-0.56	-7.92	4	-26
G^{11}	-1.79	0.13	6.27	-2.05	1	16
T^{12}	-1.92	-0.23	15.9	-3.86	4	0
	X-disp (dx)	*Y-disp (dy)*	*Inclination (eta)*	*Tip (theta)*	*Bc*	*Tc*
C^{24}	-1.77	-0.3	15.55	0.35	3	0
A^{23}	-1.65	0.05	3.31	-5.35	2	29
C^{22}	-1.89	-0.27	8.49	-7.19	3	-15
G^{21}	-1.86	0.15	13.26	-0.16	1	13
C^{20}	-1.81	-0.09	5.17	-9.19	3	-14
A^{19}	-1.71	0.17	5.28	-12.87	2	29
C^{18}	-1.75	-0.04	14.74	-14.04	3	-16
A^{17}	-1.56	0.32	16.39	-3.62	2	23
A^{16}	-1.9	0.05	11.93	-7.78	2	20
A^{15}	-1.49	0.31	15	-6.04	2	26
C^{14}	-1.8	-0.1	4.69	-4.14	3	-16
A^{13}	-1.71	0.29	2.93	-6.61	2	0

Table E-15: Global base pair-axis parameters for 5'-$G^1T^2G^3C^4G^5\underline{X}^6G^7T^8T^9T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3' (X = 5*R*-thymine glycol)

Base pair	*X-disp (dx)*	*Y-disp (dy)*	*Inclination (eta)*	*Tip (theta)*	*Bc*	*Tc*
G^1-C^{24}	-1.82	0.32	9.99	0.69	1	0
T^2-A^{23}	-1.72	-0.02	6.9	-0.75	4	-29
G^3-C^{22}	-1.89	0.27	8.65	2.03	1	15
C^4-G^{21}	-1.89	-0.16	8.54	-5.18	3	-13
G^5-C^{20}	-1.8	0.15	4.51	2.54	1	14
X^6-A^{19}	-1.72	-0.13	9.3	-7.11	4	-29
G^7-C^{18}	-1.85	0.05	4.43	1.02	1	16
T^8-A^{17}	-1.67	-0.23	5.68	-3.38	4	-23
T^9-A^{16}	-2.03	-0.05	7.41	0.99	4	-20
T^{10}-A^{15}	-1.53	-0.27	7.22	-0.94	4	-26

G^{11}-C^{14}	-1.79	0.11	5.48	1.05	1	16
T^{12}-A^{13}	-1.82	-0.26	9.42	1.38	4	0
Average:	-1.79	-0.02	7.29	-0.64		

Table E-16: Global base-base parameters for 5'-$G^1T^2G^3C^4G^5\underline{X}^6G^7T^8T^9T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3' (X = 5*R*-thymine glycol)

Base pair	*Shear (Sx)*	*Stretch (Sy)*	*Stagger (Sz)*	*Buckle (kappa)*	*Propel (omega)*	*Opening (sigma)*	*Bc*	*Tc*
G^1-C^{24}	-0.09	0.03	-0.11	-11.12	2.08	-0.19	1	0
T^2-A^{23}	-0.15	0.04	0.25	7.17	-12.2	3.32	4	-29
G^3-C^{22}	0	0	-0.07	0.31	-10.31	-1.69	1	15
C^4-G^{21}	-0.06	-0.01	0.17	-9.45	-10.69	-1.97	3	-13
G^5-C^{20}	0.02	0.13	-0.09	-1.32	-13.29	1.57	1	14
X^6-A^{19}	-0.02	0.09	-0.55	8.04	-39.95	0.62	4	-29
G^7-C^{18}	-0.19	0.02	-0.16	-20.63	-26.05	-2.26	1	16
T^8-A^{17}	-0.23	0.18	0.33	-21.41	-13.99	2.48	4	-23
T^9-A^{16}	-0.26	0.02	-0.01	-9.04	-13.58	0.11	4	-20
T^{10}-A^{15}	-0.08	0.09	0.15	-15.56	-13.96	4.37	4	-26
G^{11}-C^{14}	0.01	0.03	0.14	1.58	-6.19	0.22	1	16
T^{12}-A^{13}	-0.22	0.06	-0.34	12.97	-10.47	-1.89	4	0
Average:	-0.11	0.06	-0.03	-4.87	-14.05	0.39		

Table E-17: Global inter-base parameters for 5'-$G^1T^2G^3C^4G^5\underline{X}^6G^7T^8T^9T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3' (X = 5*R*-thymine glycol)

1st Strand	*Shift (Dx)*	*Slide (Dy)*	*Rise (Dz)*	*Tilt (tau)*	*Roll (rho)*	*Twist (Omega)*	*Dc*
G^1/T^2	0	-0.35	2.96	5.2	-6.14	30.48	6
T^2/G^3	-0.36	0.2	3.19	-1.96	3.04	37.75	-9
G^3/C^4	-0.14	-0.43	3.45	-7.38	-4.26	36.05	5
C^4/G^5	0.42	0.52	2.92	1.88	11.91	36.8	8
G^5/X^6	0.27	-0.3	3.24	17.04	-27.97	28.46	6
X^6/G^7	-0.13	-0.01	4.44	-19.93	12.38	40.48	-9
G^7/T^8	0.01	-0.43	3.53	-1.88	2.67	30.43	6
T^8/T^9	-0.5	0.02	2.83	9	7.58	31.25	-4
T^9/T^{10}	0.91	-0.05	3.66	-4.08	-0.94	37.93	-4
T^{10}/G^{11}	-0.28	0.39	3.17	7.19	4.32	32.62	-9
G^{11}/T^{12}	-0.42	-0.39	2.8	11.26	-1.23	26.55	6
2nd Strand	*Shift (Dx)*	*Slide (Dy)*	*Rise (Dz)*	*Tilt (tau)*	*Roll (rho)*	*Twist (Omega)*	*Dc*
C^{24}/A^{23}	0.06	0.36	2.59	-13.1	-8.14	26.97	-6
A^{23}/C^{22}	-0.51	-0.25	3.52	4.9	-1.15	42.75	9
C^{22}/G^{21}	-0.09	0.42	3.2	2.39	3.89	36.33	5
G^{21}/C^{20}	0.34	-0.38	3.18	-6.25	-14.52	33.27	8
C^{20}/A^{19}	0.32	0.25	3.7	7.68	1.32	29.41	-6
A^{19}/C^{18}	0.04	-0.05	4.04	8.73	1.53	43.35	9

C^{18}/A^{17}	0.05	0.58	3.04	-1.1	9.39	25.69	-6
A^{17}/A^{16}	-0.47	-0.18	3.17	-3.37	-7.16	33.62	4
A^{16}/A^{15}	0.73	0.12	3.5	2.45	0.56	33.67	4
A^{15}/C^{14}	-0.36	-0.45	3.18	-9.96	3.45	36.78	9
C^{14}/A^{13}	-0.2	0.42	3.28	-0.12	-3.06	28.66	-6

Table E-18: Global inter-base pair parameters for 5'-$G^1T^2G^3C^4G^5\underline{X}^6G^7T^8T^9T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3' (X = 5*R*-thymine glycol)

Residue	*Shift (Dx)*	*Slide (Dy)*	*Rise (Dz)*	*Tilt (tau)*	*Roll (rho)*	*Twist (Omega)*	*Dc*
G^1/T^2	0.03	-0.35	2.77	-3.95	1	28.72	6
T^2/G^3	-0.43	0.23	3.35	1.47	2.1	40.25	-9
G^3/C^4	-0.12	-0.42	3.33	-2.5	-4.07	36.19	5
C^4/G^5	0.38	0.45	3.05	-2.18	13.22	35.04	8
G^5/X^6	0.3	-0.27	3.47	12.36	-14.65	28.93	6
X^6/G^7	-0.04	0.02	4.24	-5.6	5.42	41.91	-9
G^7/T^8	0.03	-0.51	3.29	-1.49	-3.36	28.06	6
T^8/T^9	-0.49	0.1	3	2.82	7.37	32.44	-4
T^9/T^{10}	0.82	-0.09	3.58	-0.82	-0.75	35.8	-4
T^{10}/G^{11}	-0.32	0.42	3.18	-1.39	0.44	34.7	-9
G^{11}/T^{12}	-0.31	-0.4	3.04	5.57	0.92	27.6	6
Average:	-0.01	-0.08	3.3	0.39	0.69	33.61	

Table E-19: Local inter-base parameters for 5'-$G^1T^2G^3C^4G^5\underline{X}^6G^7T^8T^9T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3' (X = 5*R*-thymine glycol)

1st Strand	*Shift (Dx)*	*Slide (Dy)*	*Rise (Dz)*	*Tilt (tau)*	*Roll (rho)*	*Twist (Omega)*	*Dc*
G^1/T^2	0.01	-0.84	3.19	6.33	-1.94	29.66	6
T^2/G^3	-0.16	-0.52	3.32	1.33	9.07	37.47	-9
G^3/C^4	0.29	-1.13	3.7	-2.98	-0.05	36.1	5
C^4/G^5	0.77	-0.53	3.07	6.49	13.98	36.06	8
G^5/X^6	1.06	-0.5	3.22	24.81	-22.3	24.19	6
X^6/G^7	1.55	-1.23	4.12	-4.33	13.78	45.17	-9
G^7/T^8	0.66	-1.73	3.35	3.96	-0.25	30.19	6
T^8/T^9	-0.1	-1.13	2.95	13.07	6.77	29.95	-4
T^9/T^{10}	1.42	-1.17	3.58	0.6	-0.05	38.1	-4
T^{10}/G^{11}	-0.04	-0.46	3.17	9.68	5.65	32.07	-9
G^{11}/T^{12}	-0.3	-0.68	3.03	12.33	3.87	25.48	6
2nd Strand	Shift (Dx)	Slide (Dy)	Rise (Dz)	Tilt (tau)	Roll (rho)	Twist (Omega)	Dc
C^{24}/A^{23}	-0.05	-0.74	2.84	-14.08	12.2	25.93	-6
A^{23}/C^{22}	-0.99	-0.75	3.49	-0.02	5.33	42.91	9
C^{22}/G^{21}	-0.38	-0.88	3.52	-0.01	3.17	35.09	5
G^{21}/C^{20}	0.16	-0.23	3.48	-8.69	19.29	31.81	8
C^{20}/A^{19}	-0.35	-0.86	3.62	1.95	1.28	30.62	-6

A^{19}/C^{18}	-0.93	-0.6	4.01	-1.99	5.79	43.98	9
C^{18}/A^{17}	-0.4	-0.39	3.3	-4.86	-2.35	24	-6
A^{17}/A^{16}	-0.64	-0.08	3.35	-6.42	15.02	31.63	4
A^{16}/A^{15}	0.36	-0.25	3.77	-1.58	7.2	32.54	4
A^{15}/C^{14}	-0.46	-0.07	3.19	-12.57	2.85	35.7	9
C^{14}/A^{13}	-0.45	-1.08	3.34	-2.72	4.85	28.45	-6

Table E-20: Local inter-base pair parameters for 5'-$G^1T^2G^3C^4G^5\underline{X}^6G^7T^8T^9T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3' (X = 5*R*-thymine glycol)

Residue	*Shift (Dx)*	*Slide (Dy)*	*Rise (Dz)*	*Tilt (tau)*	*Roll (rho)*	*Twist (Omega)*	*Dc*
G^1/T^2	-0.03	-0.79	3	-3.73	5.3	27.78	6
T^2/G^3	-0.56	-0.63	3.43	0.64	7.21	40.42	-9
G^3/C^4	-0.06	-1.01	3.63	-1.52	1.58	35.79	5
C^4/G^5	0.48	-0.39	3.29	-1.05	16.7	34.15	8
G^5/X^6	0.4	-0.67	3.47	13.43	-10.57	27.51	6
X^6/G^7	0.26	-0.92	4.26	-3.34	9.88	46.21	-9
G^7/T^8	0.09	-1.04	3.43	-0.46	-1.29	26.81	6
T^8/T^9	-0.34	-0.62	3.2	3.47	11.12	31.17	-4
T^9/T^{10}	0.89	-0.71	3.74	-0.42	3.62	35.76	-4
T^{10}/G^{11}	-0.24	-0.26	3.18	-1.48	4.53	33.7	-9
G^{11}/T^{12}	-0.36	-0.86	3.18	4.81	4.36	26.89	6
Average:	0.05	-0.72	3.44	0.94	4.77	33.29	

Table E-21: Global axis curvature parameters for 5'-$G^1T^2G^3C^4G^5\underline{X}^6G^7T^8T^9T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3' (X = 5*R*-thymine glycol)

Residue	*Ax*	*Ay*	*Ainc*	*Atip*	*Adis*	*Angle*	*Path*	*Dc*
G^1/T^2	-0.07	-0.01	-0.85	2.45	0.07	2.59	2.77	6
T^2/G^3	-0.27	-0.07	-0.28	-0.68	0.28	0.74	3.36	-9
G^3/C^4	-0.11	0	-2.38	3.14	0.11	3.94	3.33	5
C^4/G^5	0.29	0.14	1.84	5.49	0.32	5.79	3.06	8
G^5/X^6	0.21	0.01	7.57	-4.99	0.21	9.07	3.47	6
X^6/G^7	0.09	-0.16	-0.73	-2.71	0.18	2.8	4.24	-9
G^7/T^8	-0.14	-0.23	-2.74	1.04	0.27	2.93	3.3	6
T^8/T^9	-0.13	-0.08	1.09	3.01	0.16	3.2	3.01	-4
T^9/T^{10}	0.32	0.13	-0.63	1.18	0.35	1.34	3.6	-4
T^{10}/G^{11}	-0.06	0.04	0.35	-1.55	0.07	1.59	3.18	-9
G^{11}/T^{12}	-0.29	-0.03	1.63	0.59	0.29	1.74	3.05	6
Overall axis bend => UU=10.46 PP=7.76								

Table E-22: Backbone parameters for 5'-$G^1T^2G^3C^4G^5\underline{X}^6G^7T^8T^9T^{10}G^{11}T^{12}$-3'•5'-$A^{13}C^{14}A^{15}A^{16}A^{17}C^{18}A^{19}C^{20}G^{21}C^{22}A^{23}C^{24}$-3' (X = 5*R*-thymine glycol)

Residue	*C1′-C2′*	*C2′-C3′*	*Phase*	*Amplitude*	*Pucker*	*C1′*	*C2′*	*C3′*
G^{1}	28.29	-31.88	172.32	32.2	C2'-endo	104.1	104.3	104
T^{2}	35.9	-27.55	139.94	36.36	C1'-exo	102.6	104.3	103.3
G^{3}	31.54	-33.64	168.32	34.66	C2'-endo	104.1	103.4	103
C^{4}	32.14	-26.22	144.47	32.49	C2'-endo	103.4	104.7	104
G^{5}	36.45	-30.36	146.55	36.77	C2'-endo	102.6	104	102.9
X^{6}	35.73	-18.33	116.98	41.12	C1'-exo	102.8	102.8	105.4
G^{7}	30.1	-17.79	122.49	33.25	C1'-exo	102.6	105.8	104.2
T^{8}	-6.5	20.95	42.22	28.31	C4'-exo	104.2	108.3	103.8
T^{9}	31.61	-28.25	152.04	32.22	C2'-endo	103.6	104.4	104
T^{10}	12.85	9.12	75.76	37.35	O1'-endo	104.1	106.5	104.1
G^{11}	-3.34	14.98	47.16	22.08	C4'-exo	103.6	110.9	103.3
T^{12}	-8.73	22.71	37.63	28.74	C4'-exo	104.4	107.5	104.3

	Chi	*Gamma*	*Delta*	*Epsilon*	*Zeta*	*Alpha*	*Beta*
	C1′-N	*C5′-C4′*	*C4′-C3′*	*C3′-O3′*	*O3′-P*	*P-O5′*	*O5′-C5′*
G^{1}	-106.41	-174.89	147.49	-174.53	-102.02	-74.47	179.88
T^{2}	-110.89	60.89	130.48	-174.32	-113.19	-76.97	-175.1
G^{3}	-110.18	59.93	145.37	-177.95	-101.74	-74.08	-174.81
C^{4}	-120.81	58.41	132.82	-165.59	-92.05	-80.21	179.42
G^{5}	-104.24	53.03	135.69	-158.47	-91.9	-107.14	115.74
X^{6}	-136.04	153.7	120.14	-161.15	-89.06	-81.25	177.02
G^{7}	-129.97	61.77	118.73	-169.61	-102.58	127.03	-166.1
T^{8}	-156.12	-169.58	93.86	-162.89	-78.19	-71.24	177.96
T^{9}	-122.41	67.64	136.07	-163.9	-88.82	-78	168.49
T^{10}	-138.54	58.67	91.47	-166.71	-89.95	-80.82	171.84
G^{11}	-134.32	64.95	100.63	-168.87	-82.63	-79.1	173.99
T^{12}	-135.01	63.23	92.04	.....	.....	.....	.....

	C1′-C2′	*C2′-C3′*	*Phase*	*Amplitude*	*Pucker*	*C1′*	*C2′*	*C3′*
C^{24}	-5.25	20.36	45.61	29.62	C4'-exo	105	107.6	101.9
A^{23}	-5.79	21.09	44.38	29.55	C4'-exo	104	109.8	102.5
C^{22}	31.68	-25.05	141.68	32.06	C1'-exo	102.8	105.9	103.6
G^{21}	29.56	-33.37	173.55	33.89	C2'-endo	104.6	103.7	103.2
C^{20}	3.65	16.54	61.42	35.08	C4'-exo	104.3	107.1	103.7
A^{19}	29.82	-24.1	143.38	30.21	C1'-exo	103.4	105.4	104
C^{18}	30.98	-31.23	162.85	32.89	C2'-endo	103.8	104.1	103.7
A^{17}	36.08	-29.65	145.58	36.4	C2'-endo	103.1	103.5	103.5
A^{16}	29.52	-30.45	165.66	31.79	C2'-endo	104.6	103.9	103.6
A^{15}	26.28	-24.15	154.74	26.76	C2'-endo	103.5	107.2	104
C^{14}	18.47	1.87	86.96	35.56	O1'-endo	102.9	108.3	103.6
A^{13}	30.53	-27.26	152.03	31.09	C2'-endo	103.3	105.1	103.1

	Chi	*Gamma*	*Delta*	*Epsilon*	*Zeta*	*Alpha*	*Beta*
	C1′-N	*C5′-C4′*	*C4′-C3′*	*C3′-O3′*	*O3′-P*	*P-O5′*	*O5′-C5′*
C^{24}	-132.33	62.34	93.6	.....	.....	.....	.....
A^{23}	-138	59.53	93.46	-165.25	-77.42	-78.22	174.81
C^{22}	-116.28	58.77	131.29	-168.84	-97.35	-89.2	174.55
G^{21}	-107.42	63.21	147.66	-175.27	-101.31	-76.7	-177.86
C^{20}	-143.47	64.98	89.04	-166.44	-86.25	-72.4	-174.03
A^{19}	-103.53	46.19	129.11	-171.7	-100.86	-76.71	162.9
C^{18}	-148.32	170.07	144.72	-150	-78.03	-87.21	-173.11
A^{17}	-94.25	51.77	135.93	-151.26	-81.36	-101.69	63

A^{16}	-104.16	57.69	142.62	-168.03	-94.43	-79.95	-177.3
A^{15}	-109.97	65.71	134.45	-173.37	-106.07	-73.6	-172.05
C^{14}	-132.26	63.79	96.97	-173.96	-92.6	-68.94	176.66
A^{13}	-104.02	66.56	132.65	-171.96	-96.33	-77.5	168.74

APPENDIX F

PDB COORDINATE FILES

File F-1: Average structure of rMD refined α-AFB_1-FAPY modified duplex 5'-$C^1T^2A^3T^4\underline{X}^5A^6T^7T^8C^9A^{10}$-3'•5'-$T^{11}G^{12}A^{13}A^{14}T^{15}C^{16}A^{17}T^{18}A^{19}G^{20}$-3'

```
ATOM      1  H5T DC5     1      42.058  34.351  35.928  1.00  0.00
ATOM      2  O5* DC5     1      41.860  34.034  36.827  1.00  0.00
ATOM      3  C5* DC5     1      40.486  33.649  36.869  1.00  0.00
ATOM      4 1H5* DC5     1      40.220  33.297  37.836  1.00  0.00
ATOM      5 2H5* DC5     1      40.266  32.850  36.219  1.00  0.00
ATOM      6  C4* DC5     1      39.592  34.800  36.503  1.00  0.00
ATOM      7  H4* DC5     1      39.712  35.542  37.228  1.00  0.00
ATOM      8  O4* DC5     1      39.996  35.316  35.239  1.00  0.00
ATOM      9  C1* DC5     1      38.934  35.352  34.282  1.00  0.00
ATOM     10  H1* DC5     1      38.544  36.352  34.237  1.00  0.00
ATOM     11  N1  DC5     1      39.403  34.903  32.951  1.00  0.00
ATOM     12  C6  DC5     1      40.260  33.861  32.848  1.00  0.00
ATOM     13  H6  DC5     1      40.617  33.387  33.730  1.00  0.00
ATOM     14  C5  DC5     1      40.685  33.431  31.563  1.00  0.00
ATOM     15  H5  DC5     1      41.365  32.621  31.464  1.00  0.00
ATOM     16  C4  DC5     1      40.181  34.134  30.468  1.00  0.00
ATOM     17  N4  DC5     1      40.557  33.757  29.212  1.00  0.00
ATOM     18 1H4  DC5     1      41.129  32.938  29.018  1.00  0.00
ATOM     19 2H4  DC5     1      40.171  34.165  28.362  1.00  0.00
ATOM     20  N3  DC5     1      39.375  35.140  30.579  1.00  0.00
ATOM     21  C2  DC5     1      38.981  35.547  31.795  1.00  0.00
ATOM     22  O2  DC5     1      38.244  36.490  31.829  1.00  0.00
ATOM     23  C3* DC5     1      38.135  34.440  36.367  1.00  0.00
ATOM     24  H3* DC5     1      37.926  33.483  36.771  1.00  0.00
ATOM     25  C2* DC5     1      37.862  34.465  34.875  1.00  0.00
ATOM     26 1H2* DC5     1      37.908  33.468  34.472  1.00  0.00
ATOM     27 2H2* DC5     1      36.891  34.864  34.657  1.00  0.00
ATOM     28  O3* DC5     1      37.376  35.427  37.047  1.00  0.00
ATOM     29  P   DT      2      35.846  35.279  37.348  1.00  0.00
ATOM     30  O1P DT      2      35.548  36.180  38.424  1.00  0.00
ATOM     31  O2P DT      2      35.547  33.887  37.513  1.00  0.00
ATOM     32  O5* DT      2      35.097  35.773  36.054  1.00  0.00
ATOM     33  C5* DT      2      35.096  37.144  35.671  1.00  0.00
ATOM     34 1H5* DT      2      36.100  37.503  35.632  1.00  0.00
ATOM     35 2H5* DT      2      34.581  37.704  36.411  1.00  0.00
ATOM     36  C4* DT      2      34.422  37.358  34.318  1.00  0.00
ATOM     37  H4* DT      2      34.323  38.411  34.148  1.00  0.00
ATOM     38  O4* DT      2      35.244  36.820  33.275  1.00  0.00
ATOM     39  C1* DT      2      34.434  36.231  32.263  1.00  0.00
ATOM     40  H1* DT      2      34.059  37.021  31.648  1.00  0.00
ATOM     41  N1  DT      2      35.170  35.227  31.462  1.00  0.00
```

```
ATOM     42  C6  DT      2      35.757  34.159  32.067  1.00  0.00
ATOM     43  H6  DT      2      35.730  34.072  33.129  1.00  0.00
ATOM     44  C5  DT      2      36.414  33.156  31.298  1.00  0.00
ATOM     45  C7  DT      2      37.066  31.981  31.986  1.00  0.00
ATOM     46 1H7  DT      2      36.616  31.049  31.637  1.00  0.00
ATOM     47 2H7  DT      2      38.132  31.959  31.749  1.00  0.00
ATOM     48 3H7  DT      2      36.939  32.059  33.069  1.00  0.00
ATOM     49  C4  DT      2      36.491  33.286  29.899  1.00  0.00
ATOM     50  O4  DT      2      37.056  32.487  29.207  1.00  0.00
ATOM     51  N3  DT      2      35.881  34.392  29.375  1.00  0.00
ATOM     52  H3  DT      2      35.904  34.491  28.383  1.00  0.00
ATOM     53  C2  DT      2      35.198  35.355  30.079  1.00  0.00
ATOM     54  O2  DT      2      34.671  36.253  29.490  1.00  0.00
ATOM     55  C3* DT      2      33.057  36.674  34.165  1.00  0.00
ATOM     56  H3* DT      2      32.745  36.209  35.078  1.00  0.00
ATOM     57  C2* DT      2      33.283  35.663  33.062  1.00  0.00
ATOM     58 1H2* DT      2      33.555  34.735  33.501  1.00  0.00
ATOM     59 2H2* DT      2      32.422  35.510  32.458  1.00  0.00
ATOM     60  O3* DT      2      32.086  37.641  33.754  1.00  0.00
ATOM     61  P   DA      3      30.535  37.352  33.703  1.00  0.00
ATOM     62  O1P DA      3      29.872  38.544  34.087  1.00  0.00
ATOM     63  O2P DA      3      30.263  36.166  34.429  1.00  0.00
ATOM     64  O5* DA      3      30.193  37.077  32.219  1.00  0.00
ATOM     65  C5* DA      3      30.196  38.107  31.261  1.00  0.00
ATOM     66 1H5* DA      3      31.182  38.502  31.158  1.00  0.00
ATOM     67 2H5* DA      3      29.564  38.872  31.599  1.00  0.00
ATOM     68  C4* DA      3      29.689  37.639  29.932  1.00  0.00
ATOM     69  H4* DA      3      29.515  38.482  29.326  1.00  0.00
ATOM     70  O4* DA      3      30.690  36.859  29.299  1.00  0.00
ATOM     71  C1* DA      3      30.173  35.617  28.838  1.00  0.00
ATOM     72  H1* DA      3      29.871  35.727  27.814  1.00  0.00
ATOM     73  N9  DA      3      31.203  34.568  28.970  1.00  0.00
ATOM     74  C8  DA      3      31.678  33.994  30.119  1.00  0.00
ATOM     75  H8  DA      3      31.326  34.301  31.081  1.00  0.00
ATOM     76  N7  DA      3      32.577  33.066  29.960  1.00  0.00
ATOM     77  C5  DA      3      32.725  33.021  28.578  1.00  0.00
ATOM     78  C6  DA      3      33.491  32.280  27.613  1.00  0.00
ATOM     79  N6  DA      3      34.373  31.315  28.001  1.00  0.00
ATOM     80 1H6  DA      3      34.935  30.764  27.353  1.00  0.00
ATOM     81 2H6  DA      3      34.546  31.064  28.972  1.00  0.00
ATOM     82  N1  DA      3      33.363  32.485  26.317  1.00  0.00
ATOM     83  C2  DA      3      32.524  33.408  25.901  1.00  0.00
ATOM     84  H2  DA      3      32.461  33.539  24.844  1.00  0.00
ATOM     85  N3  DA      3      31.743  34.191  26.619  1.00  0.00
ATOM     86  C4  DA      3      31.879  33.960  27.945  1.00  0.00
ATOM     87  C3* DA      3      28.425  36.779  30.009  1.00  0.00
ATOM     88  H3* DA      3      27.962  36.824  30.976  1.00  0.00
ATOM     89  C2* DA      3      28.934  35.394  29.692  1.00  0.00
ATOM     90 1H2* DA      3      29.183  34.909  30.601  1.00  0.00
ATOM     91 2H2* DA      3      28.214  34.800  29.190  1.00  0.00
ATOM     92  O3* DA      3      27.500  37.228  29.028  1.00  0.00
ATOM     93  P   DT      4      26.031  36.708  28.833  1.00  0.00
ATOM     94  O1P DT      4      25.209  37.850  28.517  1.00  0.00
ATOM     95  O2P DT      4      25.653  35.872  29.964  1.00  0.00
ATOM     96  O5* DT      4      26.128  35.803  27.551  1.00  0.00
ATOM     97  C5* DT      4      26.425  36.362  26.262  1.00  0.00
ATOM     98 1H5* DT      4      27.268  37.023  26.347  1.00  0.00
```

```
ATOM     99 2H5* DT      4      25.584  36.939  25.932  1.00  0.00
ATOM    100  C4* DT      4      26.737  35.290  25.219  1.00  0.00
ATOM    101  H4* DT      4      26.972  35.777  24.285  1.00  0.00
ATOM    102  O4* DT      4      27.889  34.542  25.652  1.00  0.00
ATOM    103  C1* DT      4      27.662  33.137  25.469  1.00  0.00
ATOM    104  H1* DT      4      27.897  32.925  24.436  1.00  0.00
ATOM    105  N1  DT      4      28.495  32.322  26.388  1.00  0.00
ATOM    106  C6  DT      4      28.235  32.269  27.732  1.00  0.00
ATOM    107  H6  DT      4      27.408  32.837  28.132  1.00  0.00
ATOM    108  C5  DT      4      29.055  31.487  28.609  1.00  0.00
ATOM    109  C7  DT      4      28.757  31.442  30.088  1.00  0.00
ATOM    110 1H7  DT      4      27.874  32.045  30.314  1.00  0.00
ATOM    111 2H7  DT      4      28.579  30.411  30.397  1.00  0.00
ATOM    112 3H7  DT      4      29.612  31.825  30.649  1.00  0.00
ATOM    113  C4  DT      4      30.117  30.727  28.073  1.00  0.00
ATOM    114  O4  DT      4      30.812  30.005  28.740  1.00  0.00
ATOM    115  N3  DT      4      30.316  30.862  26.722  1.00  0.00
ATOM    116  H3  DT      4      31.059  30.331  26.317  1.00  0.00
ATOM    117  C2  DT      4      29.579  31.635  25.854  1.00  0.00
ATOM    118  O2  DT      4      29.874  31.660  24.689  1.00  0.00
ATOM    119  C3* DT      4      25.618  34.265  25.008  1.00  0.00
ATOM    120  H3* DT      4      24.679  34.587  25.446  1.00  0.00
ATOM    121  C2* DT      4      26.148  32.976  25.620  1.00  0.00
ATOM    122 1H2* DT      4      25.858  32.951  26.660  1.00  0.00
ATOM    123 2H2* DT      4      25.768  32.086  25.113  1.00  0.00
ATOM    124  O3* DT      4      25.488  34.064  23.592  1.00  0.00
ATOM    125  P   FB      5      24.073  33.704  22.917  1.00  0.00
ATOM    126  O1P FB      5      24.264  33.733  21.466  1.00  0.00
ATOM    127  O2P FB      5      23.051  34.587  23.469  1.00  0.00
ATOM    128  O5* FB      5      23.708  32.208  23.378  1.00  0.00
ATOM    129  C5* FB      5      24.282  31.078  22.715  1.00  0.00
ATOM    130 1H5* FB      5      25.353  31.112  22.807  1.00  0.00
ATOM    131 2H5* FB      5      24.049  31.103  21.663  1.00  0.00
ATOM    132  C4* FB      5      23.748  29.790  23.324  1.00  0.00
ATOM    133  H4* FB      5      24.088  28.949  22.754  1.00  0.00
ATOM    134  O4* FB      5      24.271  29.636  24.645  1.00  0.00
ATOM    135  C1* FB      5      23.290  29.133  25.532  1.00  0.00
ATOM    136  N6  FB      5      23.256  27.673  25.511  1.00  0.00
ATOM    137  N7  FB      5      23.691  26.563  28.260  1.00  0.00
ATOM    138  C5  FB      5      24.559  26.351  27.206  1.00  0.00
ATOM    139  C4  FB      5      25.565  25.354  27.324  1.00  0.00
ATOM    140  O4A FB      5      25.758  24.708  28.325  1.00  0.00
ATOM    141  N3  FB      5      26.321  25.096  26.212  1.00  0.00
ATOM    142  H3  FB      5      27.039  24.386  26.268  1.00  0.00
ATOM    143  C2  FB      5      26.053  25.694  25.027  1.00  0.00
ATOM    144  N2  FB      5      26.841  25.278  23.994  1.00  0.00
ATOM    145 1HN2 FB      5      26.571  25.479  23.033  1.00  0.00
ATOM    146 2HN2 FB      5      27.263  24.352  24.014  1.00  0.00
ATOM    147  N1  FB      5      25.072  26.540  24.787  1.00  0.00
ATOM    148  C6  FB      5      24.322  26.920  25.890  1.00  0.00
ATOM    149  C3* FB      5      22.230  29.744  23.475  1.00  0.00
ATOM    150  H3* FB      5      21.744  30.597  23.030  1.00  0.00
ATOM    151  C2* FB      5      22.011  29.716  24.963  1.00  0.00
ATOM    152 1H2* FB      5      21.887  30.718  25.315  1.00  0.00
ATOM    153 2H2* FB      5      21.151  29.153  25.242  1.00  0.00
ATOM    154  O3* FB      5      21.722  28.513  22.959  1.00  0.00
ATOM    155  H1* FB      5      23.466  29.505  26.520  1.00  0.00
```

```
ATOM    156  C8  FB      5      22.567  25.833  28.324  1.00  0.00
ATOM    157  O8  FB      5      22.304  24.875  27.766  1.00  0.00
ATOM    158  H8  FB      5      21.902  26.171  28.904  1.00  0.00
ATOM    159  H6  FB      5      22.827  27.420  24.665  1.00  0.00
ATOM    160  P   DA      6      20.930  28.359  21.656  1.00  0.00
ATOM    161  O1P DA      6      20.866  29.633  20.981  1.00  0.00
ATOM    162  O2P DA      6      19.725  27.692  21.985  1.00  0.00
ATOM    163  O5* DA      6      21.762  27.339  20.817  1.00  0.00
ATOM    164  C5* DA      6      22.854  27.716  19.966  1.00  0.00
ATOM    165 1H5* DA      6      23.510  28.394  20.476  1.00  0.00
ATOM    166 2H5* DA      6      22.461  28.235  19.117  1.00  0.00
ATOM    167  C4* DA      6      23.649  26.489  19.503  1.00  0.00
ATOM    168  H4* DA      6      24.454  26.820  18.865  1.00  0.00
ATOM    169  O4* DA      6      24.211  25.823  20.655  1.00  0.00
ATOM    170  C1* DA      6      23.717  24.485  20.807  1.00  0.00
ATOM    171  H1* DA      6      24.472  23.792  20.471  1.00  0.00
ATOM    172  N9  DA      6      23.373  24.226  22.224  1.00  0.00
ATOM    173  C8  DA      6      22.427  24.846  22.999  1.00  0.00
ATOM    174  H8  DA      6      21.805  25.629  22.617  1.00  0.00
ATOM    175  N7  DA      6      22.330  24.425  24.231  1.00  0.00
ATOM    176  C5  DA      6      23.307  23.432  24.283  1.00  0.00
ATOM    177  C6  DA      6      23.828  22.501  25.251  1.00  0.00
ATOM    178  N6  DA      6      23.334  22.452  26.522  1.00  0.00
ATOM    179 1H6  DA      6      23.581  21.723  27.188  1.00  0.00
ATOM    180 2H6  DA      6      22.514  22.976  26.821  1.00  0.00
ATOM    181  N1  DA      6      24.779  21.635  24.941  1.00  0.00
ATOM    182  C2  DA      6      25.273  21.637  23.719  1.00  0.00
ATOM    183  H2  DA      6      26.018  20.900  23.499  1.00  0.00
ATOM    184  N3  DA      6      24.938  22.414  22.705  1.00  0.00
ATOM    185  C4  DA      6      23.957  23.288  23.033  1.00  0.00
ATOM    186  C3* DA      6      22.801  25.436  18.785  1.00  0.00
ATOM    187  H3* DA      6      21.893  25.858  18.388  1.00  0.00
ATOM    188  C2* DA      6      22.521  24.400  19.865  1.00  0.00
ATOM    189 1H2* DA      6      21.610  24.670  20.375  1.00  0.00
ATOM    190 2H2* DA      6      22.403  23.404  19.468  1.00  0.00
ATOM    191  O3* DA      6      23.587  24.886  17.712  1.00  0.00
ATOM    192  P   DT      7      23.077  23.761  16.720  1.00  0.00
ATOM    193  O1P DT      7      23.745  23.966  15.441  1.00  0.00
ATOM    194  O2P DT      7      21.622  23.716  16.752  1.00  0.00
ATOM    195  O5* DT      7      23.626  22.424  17.355  1.00  0.00
ATOM    196  C5* DT      7      25.031  22.171  17.447  1.00  0.00
ATOM    197 1H5* DT      7      25.497  22.962  17.993  1.00  0.00
ATOM    198 2H5* DT      7      25.447  22.164  16.463  1.00  0.00
ATOM    199  C4* DT      7      25.337  20.843  18.131  1.00  0.00
ATOM    200  H4* DT      7      26.400  20.695  18.131  1.00  0.00
ATOM    201  O4* DT      7      24.912  20.878  19.500  1.00  0.00
ATOM    202  C1* DT      7      24.457  19.588  19.913  1.00  0.00
ATOM    203  H1* DT      7      25.313  18.978  20.159  1.00  0.00
ATOM    204  N1  DT      7      23.524  19.710  21.057  1.00  0.00
ATOM    205  C6  DT      7      22.391  20.479  20.957  1.00  0.00
ATOM    206  H6  DT      7      22.167  20.962  20.031  1.00  0.00
ATOM    207  C5  DT      7      21.533  20.657  22.084  1.00  0.00
ATOM    208  C7  DT      7      20.285  21.499  21.967  1.00  0.00
ATOM    209 1H7  DT      7      20.212  21.933  20.967  1.00  0.00
ATOM    210 2H7  DT      7      19.404  20.879  22.153  1.00  0.00
ATOM    211 3H7  DT      7      20.305  22.301  22.708  1.00  0.00
ATOM    212  C4  DT      7      21.865  20.059  23.317  1.00  0.00
```

```
ATOM    213  O4  DT      7      21.223  20.244  24.317  1.00  0.00
ATOM    214  N3  DT      7      22.994  19.270  23.316  1.00  0.00
ATOM    215  H3  DT      7      23.255  18.836  24.179  1.00  0.00
ATOM    216  C2  DT      7      23.846  19.073  22.256  1.00  0.00
ATOM    217  O2  DT      7      24.821  18.388  22.406  1.00  0.00
ATOM    218  C3* DT      7      24.657  19.623  17.503  1.00  0.00
ATOM    219  H3* DT      7      24.029  19.900  16.674  1.00  0.00
ATOM    220  C2* DT      7      23.840  19.022  18.645  1.00  0.00
ATOM    221 1H2* DT      7      22.804  19.357  18.570  1.00  0.00
ATOM    222 2H2* DT      7      23.866  17.946  18.665  1.00  0.00
ATOM    223  O3* DT      7      25.687  18.723  17.053  1.00  0.00
ATOM    224  P   DT      8      25.422  17.306  16.397  1.00  0.00
ATOM    225  O1P DT      8      26.567  16.986  15.563  1.00  0.00
ATOM    226  O2P DT      8      24.065  17.319  15.806  1.00  0.00
ATOM    227  O5* DT      8      25.417  16.313  17.623  1.00  0.00
ATOM    228  C5* DT      8      26.557  16.195  18.481  1.00  0.00
ATOM    229 1H5* DT      8      26.830  17.173  18.827  1.00  0.00
ATOM    230 2H5* DT      8      27.379  15.795  17.926  1.00  0.00
ATOM    231  C4* DT      8      26.294  15.300  19.687  1.00  0.00
ATOM    232  H4* DT      8      27.209  15.181  20.234  1.00  0.00
ATOM    233  O4* DT      8      25.348  15.925  20.555  1.00  0.00
ATOM    234  C1* DT      8      24.532  14.951  21.189  1.00  0.00
ATOM    235  H1* DT      8      25.062  14.553  22.034  1.00  0.00
ATOM    236  N1  DT      8      23.236  15.544  21.581  1.00  0.00
ATOM    237  C6  DT      8      22.482  16.184  20.637  1.00  0.00
ATOM    238  H6  DT      8      22.831  16.204  19.628  1.00  0.00
ATOM    239  C5  DT      8      21.258  16.821  20.997  1.00  0.00
ATOM    240  C7  DT      8      20.428  17.512  19.944  1.00  0.00
ATOM    241 1H7  DT      8      20.218  18.539  20.249  1.00  0.00
ATOM    242 2H7  DT      8      20.960  17.524  18.990  1.00  0.00
ATOM    243 3H7  DT      8      19.479  16.987  19.820  1.00  0.00
ATOM    244  C4  DT      8      20.841  16.825  22.338  1.00  0.00
ATOM    245  O4  DT      8      19.817  17.393  22.660  1.00  0.00
ATOM    246  N3  DT      8      21.657  16.144  23.219  1.00  0.00
ATOM    247  H3  DT      8      21.382  16.137  24.179  1.00  0.00
ATOM    248  C2  DT      8      22.807  15.461  22.902  1.00  0.00
ATOM    249  O2  DT      8      23.387  14.833  23.749  1.00  0.00
ATOM    250  C3* DT      8      25.716  13.925  19.356  1.00  0.00
ATOM    251  H3* DT      8      25.533  13.825  18.303  1.00  0.00
ATOM    252  C2* DT      8      24.417  13.871  20.134  1.00  0.00
ATOM    253 1H2* DT      8      23.607  14.075  19.468  1.00  0.00
ATOM    254 2H2* DT      8      24.244  12.920  20.591  1.00  0.00
ATOM    255  O3* DT      8      26.636  12.914  19.814  1.00  0.00
ATOM    256  P   DC      9      26.453  11.367  19.539  1.00  0.00
ATOM    257  O1P DC      9      27.750  10.745  19.707  1.00  0.00
ATOM    258  O2P DC      9      25.762  11.205  18.273  1.00  0.00
ATOM    259  O5* DC      9      25.519  10.830  20.687  1.00  0.00
ATOM    260  C5* DC      9      25.918  10.880  22.063  1.00  0.00
ATOM    261 1H5* DC      9      26.266  11.867  22.299  1.00  0.00
ATOM    262 2H5* DC      9      26.728  10.202  22.214  1.00  0.00
ATOM    263  C4* DC      9      24.778  10.518  23.002  1.00  0.00
ATOM    264  H4* DC      9      25.153  10.493  24.007  1.00  0.00
ATOM    265  O4* DC      9      23.764  11.531  22.946  1.00  0.00
ATOM    266  C1* DC      9      22.481  10.947  23.103  1.00  0.00
ATOM    267  H1* DC      9      22.337  10.763  24.151  1.00  0.00
ATOM    268  N1  DC      9      21.423  11.828  22.547  1.00  0.00
ATOM    269  C6  DC      9      21.387  12.135  21.222  1.00  0.00
```

```
ATOM    270  H6  DC      9      22.116  11.723  20.568  1.00  0.00
ATOM    271  C5  DC      9      20.386  13.007  20.723  1.00  0.00
ATOM    272  H5  DC      9      20.345  13.257  19.688  1.00  0.00
ATOM    273  C4  DC      9      19.463  13.492  21.655  1.00  0.00
ATOM    274  N4  DC      9      18.472  14.326  21.227  1.00  0.00
ATOM    275 1H4  DC      9      18.310  14.547  20.247  1.00  0.00
ATOM    276 2H4  DC      9      17.711  14.645  21.823  1.00  0.00
ATOM    277  N3  DC      9      19.491  13.172  22.916  1.00  0.00
ATOM    278  C2  DC      9      20.450  12.356  23.386  1.00  0.00
ATOM    279  O2  DC      9      20.415  12.103  24.562  1.00  0.00
ATOM    280  C3* DC      9      24.077   9.196  22.688  1.00  0.00
ATOM    281  H3* DC      9      24.511   8.702  21.846  1.00  0.00
ATOM    282  C2* DC      9      22.647   9.607  22.409  1.00  0.00
ATOM    283 1H2* DC      9      22.526   9.706  21.355  1.00  0.00
ATOM    284 2H2* DC      9      21.941   8.891  22.757  1.00  0.00
ATOM    285  O3* DC      9      24.152   8.346  23.837  1.00  0.00
ATOM    286  P   DA3    10      23.686   6.813  23.813  1.00  0.00
ATOM    287  O1P DA3    10      24.710   5.996  24.501  1.00  0.00
ATOM    288  O2P DA3    10      23.261   6.470  22.437  1.00  0.00
ATOM    289  O5* DA3    10      22.386   6.883  24.747  1.00  0.00
ATOM    290  C5* DA3    10      22.484   7.096  26.161  1.00  0.00
ATOM    291 1H5* DA3    10      22.973   8.054  26.349  1.00  0.00
ATOM    292 2H5* DA3    10      23.098   6.302  26.592  1.00  0.00
ATOM    293  C4* DA3    10      21.082   7.078  26.782  1.00  0.00
ATOM    294  H4* DA3    10      21.189   7.155  27.831  1.00  0.00
ATOM    295  O4* DA3    10      20.337   8.200  26.353  1.00  0.00
ATOM    296  C1* DA3    10      19.005   7.887  26.027  1.00  0.00
ATOM    297  H1* DA3    10      18.376   8.167  26.839  1.00  0.00
ATOM    298  N9  DA3    10      18.646   8.625  24.818  1.00  0.00
ATOM    299  C8  DA3    10      19.218   8.524  23.588  1.00  0.00
ATOM    300  H8  DA3    10      20.021   7.843  23.422  1.00  0.00
ATOM    301  N7  DA3    10      18.720   9.304  22.677  1.00  0.00
ATOM    302  C5  DA3    10      17.716   9.969  23.362  1.00  0.00
ATOM    303  C6  DA3    10      16.723  10.964  23.090  1.00  0.00
ATOM    304  N6  DA3    10      16.620  11.547  21.861  1.00  0.00
ATOM    305 1H6  DA3    10      16.041  12.367  21.689  1.00  0.00
ATOM    306 2H6  DA3    10      17.330  11.440  21.140  1.00  0.00
ATOM    307  N1  DA3    10      15.874  11.366  24.000  1.00  0.00
ATOM    308  C2  DA3    10      15.957  10.853  25.200  1.00  0.00
ATOM    309  H2  DA3    10      15.250  11.216  25.898  1.00  0.00
ATOM    310  N3  DA3    10      16.798   9.951  25.651  1.00  0.00
ATOM    311  C4  DA3    10      17.656   9.546  24.699  1.00  0.00
ATOM    312  C3* DA3    10      20.265   5.855  26.420  1.00  0.00
ATOM    313  H3* DA3    10      20.775   5.279  25.691  1.00  0.00
ATOM    314  C2* DA3    10      18.985   6.386  25.844  1.00  0.00
ATOM    315 1H2* DA3    10      18.937   6.127  24.824  1.00  0.00
ATOM    316 2H2* DA3    10      18.148   5.961  26.323  1.00  0.00
ATOM    317  O3* DA3    10      19.980   5.096  27.592  1.00  0.00
ATOM    318  H3T DA3    10      20.824   4.796  27.971  1.00  0.00
TER
ATOM    319  H5T DT5    11       7.126  16.818  24.719  1.00  0.00
ATOM    320  O5* DT5    11       7.872  17.211  24.233  1.00  0.00
ATOM    321  C5* DT5    11       8.819  17.691  25.185  1.00  0.00
ATOM    322 1H5* DT5    11       8.335  18.385  25.875  1.00  0.00
ATOM    323 2H5* DT5    11       9.638  18.198  24.682  1.00  0.00
ATOM    324  C4* DT5    11       9.428  16.518  25.941  1.00  0.00
ATOM    325  H4* DT5    11       8.672  15.917  26.449  1.00  0.00
```

```
ATOM    326  O4* DT5    11      10.138  15.707  25.010  1.00  0.00
ATOM    327  C1* DT5    11      11.310  15.147  25.588  1.00  0.00
ATOM    328  H1* DT5    11      11.023  14.260  26.109  1.00  0.00
ATOM    329  N1  DT5    11      12.349  14.880  24.571  1.00  0.00
ATOM    330  C6  DT5    11      12.519  15.735  23.529  1.00  0.00
ATOM    331  H6  DT5    11      11.947  16.627  23.491  1.00  0.00
ATOM    332  C5  DT5    11      13.450  15.444  22.488  1.00  0.00
ATOM    333  C7  DT5    11      13.630  16.400  21.334  1.00  0.00
ATOM    334 1H7  DT5    11      14.666  16.744  21.299  1.00  0.00
ATOM    335 2H7  DT5    11      13.406  15.895  20.392  1.00  0.00
ATOM    336 3H7  DT5    11      12.971  17.263  21.450  1.00  0.00
ATOM    337  C4  DT5    11      14.200  14.260  22.547  1.00  0.00
ATOM    338  O4  DT5    11      14.943  13.923  21.669  1.00  0.00
ATOM    339  N3  DT5    11      14.029  13.507  23.675  1.00  0.00
ATOM    340  H3  DT5    11      14.604  12.702  23.762  1.00  0.00
ATOM    341  C2  DT5    11      13.150  13.759  24.694  1.00  0.00
ATOM    342  O2  DT5    11      13.103  13.027  25.635  1.00  0.00
ATOM    343  C3* DT5    11      10.508  16.970  26.916  1.00  0.00
ATOM    344  H3* DT5    11      10.714  17.974  26.837  1.00  0.00
ATOM    345  C2* DT5    11      11.752  16.186  26.578  1.00  0.00
ATOM    346 1H2* DT5    11      12.453  16.851  26.117  1.00  0.00
ATOM    347 2H2* DT5    11      12.198  15.726  27.434  1.00  0.00
ATOM    348  O3* DT5    11      10.017  16.619  28.226  1.00  0.00
ATOM    349  P   DG     12      10.425  17.459  29.533  1.00  0.00
ATOM    350  O1P DG     12       9.532  17.060  30.645  1.00  0.00
ATOM    351  O2P DG     12      10.540  18.885  29.152  1.00  0.00
ATOM    352  O5* DG     12      11.891  16.900  29.843  1.00  0.00
ATOM    353  C5* DG     12      12.093  15.693  30.544  1.00  0.00
ATOM    354 1H5* DG     12      11.592  14.883  30.008  1.00  0.00
ATOM    355 2H5* DG     12      11.621  15.791  31.525  1.00  0.00
ATOM    356  C4* DG     12      13.558  15.366  30.722  1.00  0.00
ATOM    357  H4* DG     12      13.648  14.524  31.372  1.00  0.00
ATOM    358  O4* DG     12      14.109  15.008  29.448  1.00  0.00
ATOM    359  C1* DG     12      15.331  15.702  29.232  1.00  0.00
ATOM    360  H1* DG     12      16.127  15.128  29.675  1.00  0.00
ATOM    361  N9  DG     12      15.584  15.876  27.794  1.00  0.00
ATOM    362  C8  DG     12      15.077  16.810  26.939  1.00  0.00
ATOM    363  H8  DG     12      14.394  17.556  27.274  1.00  0.00
ATOM    364  N7  DG     12      15.457  16.703  25.703  1.00  0.00
ATOM    365  C5  DG     12      16.323  15.620  25.721  1.00  0.00
ATOM    366  C6  DG     12      17.094  14.991  24.723  1.00  0.00
ATOM    367  O6  DG     12      17.079  15.384  23.574  1.00  0.00
ATOM    368  N1  DG     12      17.865  13.945  25.189  1.00  0.00
ATOM    369  H1  DG     12      18.454  13.464  24.522  1.00  0.00
ATOM    370  C2  DG     12      17.882  13.549  26.476  1.00  0.00
ATOM    371  N2  DG     12      18.718  12.501  26.729  1.00  0.00
ATOM    372 1H2  DG     12      19.095  12.367  27.665  1.00  0.00
ATOM    373 2H2  DG     12      19.444  12.255  26.060  1.00  0.00
ATOM    374  N3  DG     12      17.183  14.067  27.459  1.00  0.00
ATOM    375  C4  DG     12      16.415  15.106  27.044  1.00  0.00
ATOM    376  C3* DG     12      14.416  16.510  31.259  1.00  0.00
ATOM    377  H3* DG     12      13.834  17.282  31.718  1.00  0.00
ATOM    378  C2* DG     12      15.145  16.983  30.017  1.00  0.00
ATOM    379 1H2* DG     12      14.519  17.672  29.481  1.00  0.00
ATOM    380 2H2* DG     12      16.075  17.466  30.235  1.00  0.00
ATOM    381  O3* DG     12      15.344  15.961  32.206  1.00  0.00
ATOM    382  P   DA     13      16.234  16.775  33.187  1.00  0.00
```

```
ATOM    383  O1P DA     13      15.792  16.463  34.520  1.00  0.00
ATOM    384  O2P DA     13      16.285  18.168  32.771  1.00  0.00
ATOM    385  O5* DA     13      17.652  16.139  32.980  1.00  0.00
ATOM    386  C5* DA     13      17.940  14.797  33.395  1.00  0.00
ATOM    387 1H5* DA     13      17.169  14.151  33.044  1.00  0.00
ATOM    388 2H5* DA     13      17.928  14.766  34.462  1.00  0.00
ATOM    389  C4* DA     13      19.282  14.273  32.878  1.00  0.00
ATOM    390  H4* DA     13      19.397  13.254  33.214  1.00  0.00
ATOM    391  O4* DA     13      19.267  14.294  31.436  1.00  0.00
ATOM    392  C1* DA     13      20.282  15.164  30.918  1.00  0.00
ATOM    393  H1* DA     13      21.115  14.557  30.602  1.00  0.00
ATOM    394  N9  DA     13      19.763  15.955  29.793  1.00  0.00
ATOM    395  C8  DA     13      18.841  16.960  29.827  1.00  0.00
ATOM    396  H8  DA     13      18.378  17.248  30.742  1.00  0.00
ATOM    397  N7  DA     13      18.572  17.525  28.683  1.00  0.00
ATOM    398  C5  DA     13      19.399  16.837  27.810  1.00  0.00
ATOM    399  C6  DA     13      19.713  16.855  26.409  1.00  0.00
ATOM    400  N6  DA     13      19.098  17.724  25.556  1.00  0.00
ATOM    401 1H6  DA     13      19.356  17.828  24.577  1.00  0.00
ATOM    402 2H6  DA     13      18.456  18.452  25.863  1.00  0.00
ATOM    403  N1  DA     13      20.634  16.056  25.888  1.00  0.00
ATOM    404  C2  DA     13      21.242  15.217  26.684  1.00  0.00
ATOM    405  H2  DA     13      21.990  14.624  26.218  1.00  0.00
ATOM    406  N3  DA     13      21.080  15.021  27.976  1.00  0.00
ATOM    407  C4  DA     13      20.154  15.857  28.492  1.00  0.00
ATOM    408  C3* DA     13      20.501  15.099  33.294  1.00  0.00
ATOM    409  H3* DA     13      20.332  15.633  34.211  1.00  0.00
ATOM    410  C2* DA     13      20.705  16.020  32.105  1.00  0.00
ATOM    411 1H2* DA     13      20.073  16.882  32.220  1.00  0.00
ATOM    412 2H2* DA     13      21.718  16.369  32.015  1.00  0.00
ATOM    413  O3* DA     13      21.624  14.209  33.432  1.00  0.00
ATOM    414  P   DA     14      23.038  14.635  34.003  1.00  0.00
ATOM    415  O1P DA     14      23.382  13.681  35.035  1.00  0.00
ATOM    416  O2P DA     14      23.032  16.044  34.338  1.00  0.00
ATOM    417  O5* DA     14      24.030  14.444  32.799  1.00  0.00
ATOM    418  C5* DA     14      24.360  13.145  32.307  1.00  0.00
ATOM    419 1H5* DA     14      23.460  12.658  31.996  1.00  0.00
ATOM    420 2H5* DA     14      24.778  12.574  33.102  1.00  0.00
ATOM    421  C4* DA     14      25.353  13.172  31.148  1.00  0.00
ATOM    422  H4* DA     14      25.596  12.158  30.899  1.00  0.00
ATOM    423  O4* DA     14      24.737  13.772  29.996  1.00  0.00
ATOM    424  C1* DA     14      25.481  14.898  29.508  1.00  0.00
ATOM    425  H1* DA     14      26.052  14.603  28.639  1.00  0.00
ATOM    426  N9  DA     14      24.544  15.987  29.169  1.00  0.00
ATOM    427  C8  DA     14      23.718  16.678  30.016  1.00  0.00
ATOM    428  H8  DA     14      23.657  16.437  31.057  1.00  0.00
ATOM    429  N7  DA     14      23.020  17.634  29.474  1.00  0.00
ATOM    430  C5  DA     14      23.400  17.559  28.139  1.00  0.00
ATOM    431  C6  DA     14      23.108  18.251  26.916  1.00  0.00
ATOM    432  N6  DA     14      22.259  19.264  26.811  1.00  0.00
ATOM    433 1H6  DA     14      22.060  19.674  25.926  1.00  0.00
ATOM    434 2H6  DA     14      21.571  19.423  27.458  1.00  0.00
ATOM    435  N1  DA     14      23.670  17.921  25.770  1.00  0.00
ATOM    436  C2  DA     14      24.523  16.921  25.754  1.00  0.00
ATOM    437  H2  DA     14      24.957  16.700  24.805  1.00  0.00
ATOM    438  N3  DA     14      24.923  16.166  26.761  1.00  0.00
ATOM    439  C4  DA     14      24.340  16.524  27.929  1.00  0.00
```

```
ATOM    440  C3* DA     14      26.643  13.963  31.408  1.00  0.00
ATOM    441  H3* DA     14      26.804  14.174  32.453  1.00  0.00
ATOM    442  C2* DA     14      26.445  15.259  30.639  1.00  0.00
ATOM    443 1H2* DA     14      26.016  15.984  31.315  1.00  0.00
ATOM    444 2H2* DA     14      27.369  15.671  30.262  1.00  0.00
ATOM    445  O3* DA     14      27.742  13.192  30.881  1.00  0.00
ATOM    446  P   DT     15      29.261  13.616  31.007  1.00  0.00
ATOM    447  O1P DT     15      30.065  12.408  30.873  1.00  0.00
ATOM    448  O2P DT     15      29.402  14.459  32.215  1.00  0.00
ATOM    449  O5* DT     15      29.525  14.524  29.740  1.00  0.00
ATOM    450  C5* DT     15      29.439  13.984  28.412  1.00  0.00
ATOM    451 1H5* DT     15      28.541  13.401  28.319  1.00  0.00
ATOM    452 2H5* DT     15      30.277  13.337  28.242  1.00  0.00
ATOM    453  C4* DT     15      29.416  15.086  27.352  1.00  0.00
ATOM    454  H4* DT     15      29.390  14.629  26.378  1.00  0.00
ATOM    455  O4* DT     15      28.235  15.894  27.482  1.00  0.00
ATOM    456  C1* DT     15      28.465  17.155  26.853  1.00  0.00
ATOM    457  H1* DT     15      28.354  17.011  25.795  1.00  0.00
ATOM    458  N1  DT     15      27.551  18.216  27.351  1.00  0.00
ATOM    459  C6  DT     15      27.351  18.378  28.701  1.00  0.00
ATOM    460  H6  DT     15      27.873  17.735  29.382  1.00  0.00
ATOM    461  C5  DT     15      26.461  19.394  29.184  1.00  0.00
ATOM    462  C7  DT     15      26.267  19.589  30.669  1.00  0.00
ATOM    463 1H7  DT     15      26.429  20.638  30.927  1.00  0.00
ATOM    464 2H7  DT     15      25.248  19.318  30.952  1.00  0.00
ATOM    465 3H7  DT     15      26.974  18.971  31.226  1.00  0.00
ATOM    466  C4  DT     15      25.752  20.201  28.265  1.00  0.00
ATOM    467  O4  DT     15      24.941  21.029  28.605  1.00  0.00
ATOM    468  N3  DT     15      26.028  19.979  26.935  1.00  0.00
ATOM    469  H3  DT     15      25.547  20.562  26.265  1.00  0.00
ATOM    470  C2  DT     15      26.928  19.063  26.433  1.00  0.00
ATOM    471  O2  DT     15      27.137  19.017  25.247  1.00  0.00
ATOM    472  C3* DT     15      30.581  16.080  27.404  1.00  0.00
ATOM    473  H3* DT     15      31.056  16.058  28.367  1.00  0.00
ATOM    474  C2* DT     15      29.924  17.433  27.163  1.00  0.00
ATOM    475 1H2* DT     15      30.012  18.017  28.046  1.00  0.00
ATOM    476 2H2* DT     15      30.368  17.971  26.366  1.00  0.00
ATOM    477  O3* DT     15      31.532  15.768  26.373  1.00  0.00
ATOM    478  P   DC     16      32.912  16.531  26.199  1.00  0.00
ATOM    479  O1P DC     16      33.794  15.659  25.433  1.00  0.00
ATOM    480  O2P DC     16      33.354  17.000  27.505  1.00  0.00
ATOM    481  O5* DC     16      32.588  17.805  25.312  1.00  0.00
ATOM    482  C5* DC     16      32.181  17.683  23.940  1.00  0.00
ATOM    483 1H5* DC     16      31.398  16.955  23.869  1.00  0.00
ATOM    484 2H5* DC     16      33.009  17.333  23.358  1.00  0.00
ATOM    485  C4* DC     16      31.679  19.014  23.364  1.00  0.00
ATOM    486  H4* DC     16      31.494  18.898  22.310  1.00  0.00
ATOM    487  O4* DC     16      30.446  19.410  23.982  1.00  0.00
ATOM    488  C1* DC     16      30.254  20.832  23.882  1.00  0.00
ATOM    489  H1* DC     16      29.592  21.017  23.050  1.00  0.00
ATOM    490  N1  DC     16      29.679  21.359  25.148  1.00  0.00
ATOM    491  C6  DC     16      30.176  20.955  26.352  1.00  0.00
ATOM    492  H6  DC     16      30.975  20.244  26.388  1.00  0.00
ATOM    493  C5  DC     16      29.624  21.490  27.550  1.00  0.00
ATOM    494  H5  DC     16      30.014  21.213  28.507  1.00  0.00
ATOM    495  C4  DC     16      28.561  22.403  27.401  1.00  0.00
ATOM    496  N4  DC     16      27.985  22.935  28.518  1.00  0.00
```

```
ATOM    497 1H4  DC     16      28.220  22.645  29.465  1.00  0.00
ATOM    498 2H4  DC     16      27.167  23.541  28.495  1.00  0.00
ATOM    499  N3  DC     16      28.118  22.803  26.234  1.00  0.00
ATOM    500  C2  DC     16      28.665  22.322  25.111  1.00  0.00
ATOM    501  O2  DC     16      28.258  22.786  24.075  1.00  0.00
ATOM    502  C3* DC     16      32.599  20.212  23.582  1.00  0.00
ATOM    503  H3* DC     16      33.075  20.146  24.543  1.00  0.00
ATOM    504  C2* DC     16      31.647  21.398  23.564  1.00  0.00
ATOM    505 1H2* DC     16      31.957  22.147  24.259  1.00  0.00
ATOM    506 2H2* DC     16      31.630  21.843  22.595  1.00  0.00
ATOM    507  O3* DC     16      33.557  20.280  22.518  1.00  0.00
ATOM    508  P   DA     17      34.922  21.053  22.608  1.00  0.00
ATOM    509  O1P DA     17      35.792  20.490  21.624  1.00  0.00
ATOM    510  O2P DA     17      35.347  21.048  23.966  1.00  0.00
ATOM    511  O5* DA     17      34.670  22.527  22.183  1.00  0.00
ATOM    512  C5* DA     17      34.075  22.857  20.949  1.00  0.00
ATOM    513 1H5* DA     17      33.048  22.571  20.993  1.00  0.00
ATOM    514 2H5* DA     17      34.536  22.313  20.169  1.00  0.00
ATOM    515  C4* DA     17      34.192  24.326  20.627  1.00  0.00
ATOM    516  H4* DA     17      33.622  24.519  19.754  1.00  0.00
ATOM    517  O4* DA     17      33.645  25.101  21.699  1.00  0.00
ATOM    518  C1* DA     17      34.530  26.159  22.071  1.00  0.00
ATOM    519  H1* DA     17      34.285  27.029  21.485  1.00  0.00
ATOM    520  N9  DA     17      34.404  26.465  23.509  1.00  0.00
ATOM    521  C8  DA     17      35.027  25.876  24.581  1.00  0.00
ATOM    522  H8  DA     17      35.738  25.087  24.466  1.00  0.00
ATOM    523  N7  DA     17      34.702  26.343  25.752  1.00  0.00
ATOM    524  C5  DA     17      33.798  27.348  25.433  1.00  0.00
ATOM    525  C6  DA     17      32.975  28.291  26.138  1.00  0.00
ATOM    526  N6  DA     17      32.975  28.358  27.501  1.00  0.00
ATOM    527 1H6  DA     17      32.360  28.967  28.038  1.00  0.00
ATOM    528 2H6  DA     17      33.497  27.718  28.095  1.00  0.00
ATOM    529  N1  DA     17      32.197  29.138  25.495  1.00  0.00
ATOM    530  C2  DA     17      32.159  29.110  24.182  1.00  0.00
ATOM    531  H2  DA     17      31.514  29.823  23.718  1.00  0.00
ATOM    532  N3  DA     17      32.817  28.315  23.365  1.00  0.00
ATOM    533  C4  DA     17      33.609  27.446  24.036  1.00  0.00
ATOM    534  C3* DA     17      35.626  24.817  20.408  1.00  0.00
ATOM    535  H3* DA     17      36.328  24.014  20.299  1.00  0.00
ATOM    536  C2* DA     17      35.896  25.649  21.647  1.00  0.00
ATOM    537 1H2* DA     17      36.313  25.005  22.390  1.00  0.00
ATOM    538 2H2* DA     17      36.580  26.453  21.480  1.00  0.00
ATOM    539  O3* DA     17      35.624  25.615  19.227  1.00  0.00
ATOM    540  P   DT     18      36.836  26.377  18.591  1.00  0.00
ATOM    541  O1P DT     18      36.851  26.062  17.185  1.00  0.00
ATOM    542  O2P DT     18      38.038  26.152  19.371  1.00  0.00
ATOM    543  O5* DT     18      36.397  27.872  18.762  1.00  0.00
ATOM    544  C5* DT     18      35.229  28.366  18.097  1.00  0.00
ATOM    545 1H5* DT     18      34.411  27.696  18.270  1.00  0.00
ATOM    546 2H5* DT     18      35.416  28.384  17.048  1.00  0.00
ATOM    547  C4* DT     18      34.817  29.754  18.570  1.00  0.00
ATOM    548  H4* DT     18      33.924  30.036  18.046  1.00  0.00
ATOM    549  O4* DT     18      34.490  29.719  19.966  1.00  0.00
ATOM    550  C1* DT     18      34.965  30.898  20.616  1.00  0.00
ATOM    551  H1* DT     18      34.235  31.681  20.490  1.00  0.00
ATOM    552  N1  DT     18      35.232  30.615  22.045  1.00  0.00
ATOM    553  C6  DT     18      36.155  29.662  22.383  1.00  0.00
```

```
ATOM    554  H6  DT     18      36.632  29.107  21.601  1.00  0.00
ATOM    555  C5  DT     18      36.468  29.403  23.752  1.00  0.00
ATOM    556  C7  DT     18      37.505  28.366  24.111  1.00  0.00
ATOM    557 1H7  DT     18      38.219  28.788  24.821  1.00  0.00
ATOM    558 2H7  DT     18      37.024  27.503  24.578  1.00  0.00
ATOM    559 3H7  DT     18      38.040  28.043  23.215  1.00  0.00
ATOM    560  C4  DT     18      35.808  30.129  24.768  1.00  0.00
ATOM    561  O4  DT     18      36.024  29.949  25.939  1.00  0.00
ATOM    562  N3  DT     18      34.881  31.054  24.339  1.00  0.00
ATOM    563  H3  DT     18      34.377  31.558  25.040  1.00  0.00
ATOM    564  C2  DT     18      34.550  31.329  23.029  1.00  0.00
ATOM    565  O2  DT     18      33.699  32.144  22.788  1.00  0.00
ATOM    566  C3* DT     18      35.878  30.849  18.391  1.00  0.00
ATOM    567  H3* DT     18      36.747  30.491  17.871  1.00  0.00
ATOM    568  C2* DT     18      36.215  31.254  19.816  1.00  0.00
ATOM    569 1H2* DT     18      37.066  30.675  20.130  1.00  0.00
ATOM    570 2H2* DT     18      36.465  32.297  19.914  1.00  0.00
ATOM    571  O3* DT     18      35.282  31.934  17.666  1.00  0.00
ATOM    572  P   DA     19      36.032  33.268  17.282  1.00  0.00
ATOM    573  O1P DA     19      35.502  33.707  16.000  1.00  0.00
ATOM    574  O2P DA     19      37.468  33.082  17.420  1.00  0.00
ATOM    575  O5* DA     19      35.567  34.293  18.388  1.00  0.00
ATOM    576  C5* DA     19      34.202  34.724  18.480  1.00  0.00
ATOM    577 1H5* DA     19      33.576  33.884  18.689  1.00  0.00
ATOM    578 2H5* DA     19      33.901  35.126  17.541  1.00  0.00
ATOM    579  C4* DA     19      34.013  35.792  19.553  1.00  0.00
ATOM    580  H4* DA     19      33.014  36.175  19.482  1.00  0.00
ATOM    581  O4* DA     19      34.173  35.218  20.852  1.00  0.00
ATOM    582  C1* DA     19      35.061  35.974  21.662  1.00  0.00
ATOM    583  H1* DA     19      34.505  36.700  22.219  1.00  0.00
ATOM    584  N9  DA     19      35.789  35.071  22.563  1.00  0.00
ATOM    585  C8  DA     19      36.669  34.086  22.224  1.00  0.00
ATOM    586  H8  DA     19      36.904  33.871  21.208  1.00  0.00
ATOM    587  N7  DA     19      37.186  33.426  23.215  1.00  0.00
ATOM    588  C5  DA     19      36.594  34.035  24.317  1.00  0.00
ATOM    589  C6  DA     19      36.626  33.903  25.742  1.00  0.00
ATOM    590  N6  DA     19      37.413  32.968  26.348  1.00  0.00
ATOM    591 1H6  DA     19      37.399  32.786  27.350  1.00  0.00
ATOM    592 2H6  DA     19      37.942  32.264  25.838  1.00  0.00
ATOM    593  N1  DA     19      35.923  34.686  26.532  1.00  0.00
ATOM    594  C2  DA     19      35.147  35.592  25.992  1.00  0.00
ATOM    595  H2  DA     19      34.597  36.190  26.679  1.00  0.00
ATOM    596  N3  DA     19      34.963  35.855  24.718  1.00  0.00
ATOM    597  C4  DA     19      35.709  35.056  23.925  1.00  0.00
ATOM    598  C3* DA     19      35.024  36.940  19.476  1.00  0.00
ATOM    599  H3* DA     19      35.561  36.927  18.549  1.00  0.00
ATOM    600  C2* DA     19      35.935  36.699  20.662  1.00  0.00
ATOM    601 1H2* DA     19      36.741  36.071  20.350  1.00  0.00
ATOM    602 2H2* DA     19      36.348  37.599  21.067  1.00  0.00
ATOM    603  O3* DA     19      34.320  38.182  19.608  1.00  0.00
ATOM    604  P   DG3    20      34.999  39.595  19.425  1.00  0.00
ATOM    605  O1P DG3    20      33.974  40.507  18.957  1.00  0.00
ATOM    606  O2P DG3    20      36.204  39.443  18.638  1.00  0.00
ATOM    607  O5* DG3    20      35.418  40.029  20.875  1.00  0.00
ATOM    608  C5* DG3    20      34.456  40.423  21.837  1.00  0.00
ATOM    609 1H5* DG3    20      33.705  39.676  21.898  1.00  0.00
ATOM    610 2H5* DG3    20      33.998  41.322  21.512  1.00  0.00
```

```
ATOM    611  C4* DG3    20      35.066  40.642  23.216  1.00  0.00
ATOM    612  H4* DG3    20      34.329  41.044  23.863  1.00  0.00
ATOM    613  O4* DG3    20      35.489  39.412  23.780  1.00  0.00
ATOM    614  C1* DG3    20      36.486  39.672  24.757  1.00  0.00
ATOM    615  H1* DG3    20      36.015  39.977  25.664  1.00  0.00
ATOM    616  N9  DG3    20      37.311  38.496  25.002  1.00  0.00
ATOM    617  C8  DG3    20      37.979  37.739  24.091  1.00  0.00
ATOM    618  H8  DG3    20      37.895  37.932  23.047  1.00  0.00
ATOM    619  N7  DG3    20      38.704  36.788  24.584  1.00  0.00
ATOM    620  C5  DG3    20      38.504  36.902  25.955  1.00  0.00
ATOM    621  C6  DG3    20      38.999  36.191  27.065  1.00  0.00
ATOM    622  O6  DG3    20      39.769  35.274  26.972  1.00  0.00
ATOM    623  N1  DG3    20      38.538  36.635  28.284  1.00  0.00
ATOM    624  H1  DG3    20      38.855  36.169  29.102  1.00  0.00
ATOM    625  C2  DG3    20      37.692  37.701  28.451  1.00  0.00
ATOM    626  N2  DG3    20      37.384  38.024  29.740  1.00  0.00
ATOM    627 1H2  DG3    20      36.541  38.555  29.944  1.00  0.00
ATOM    628 2H2  DG3    20      37.518  37.342  30.484  1.00  0.00
ATOM    629  N3  DG3    20      37.213  38.404  27.440  1.00  0.00
ATOM    630  C4  DG3    20      37.624  37.978  26.221  1.00  0.00
ATOM    631  C3* DG3    20      36.298  41.524  23.254  1.00  0.00
ATOM    632  H3* DG3    20      36.711  41.609  22.306  1.00  0.00
ATOM    633  C2* DG3    20      37.299  40.814  24.161  1.00  0.00
ATOM    634 1H2* DG3    20      38.120  40.412  23.563  1.00  0.00
ATOM    635 2H2* DG3    20      37.699  41.475  24.933  1.00  0.00
ATOM    636  O3* DG3    20      36.032  42.770  23.761  1.00  0.00
ATOM    637  H3T DG3    20      36.166  43.205  23.534  1.00  0.00
TER
ATOM    638  O11 FA     21      29.919  25.331  29.922  1.00  0.00
ATOM    639  C11 FA     21      29.476  25.916  28.975  1.00  0.00
ATOM    640  O10 FA     21      28.447  26.741  29.139  1.00  0.00
ATOM    641  CA1 FA     21      28.074  27.499  28.095  1.00  0.00
ATOM    642  C9B FA     21      27.008  28.383  28.362  1.00  0.00
ATOM    643  C9A FA     21      26.197  28.620  29.596  1.00  0.00
ATOM    644  C9  FA     21      25.358  27.406  30.017  1.00  0.00
ATOM    645  O9  FA     21      25.159  27.489  31.427  1.00  0.00
ATOM    646  HO9 FA     21      25.364  26.623  31.821  1.00  0.00
ATOM    647  C8A FA     21      23.991  27.572  29.321  1.00  0.00
ATOM    648  O7  FA     21      23.969  28.902  28.794  1.00  0.00
ATOM    649  H8A FA     21      23.198  27.532  30.054  1.00  0.00
ATOM    650  H9  FA     21      25.844  26.462  29.762  1.00  0.00
ATOM    651  H9a FA     21      26.796  29.037  30.395  1.00  0.00
ATOM    652  C6A FA     21      25.176  29.639  29.053  1.00  0.00
ATOM    653  H6a FA     21      24.979  30.447  29.745  1.00  0.00
ATOM    654  O6A FA     21      25.676  30.140  27.789  1.00  0.00
ATOM    655  C5M FA     21      26.611  29.273  27.385  1.00  0.00
ATOM    656  C5B FA     21      27.215  29.203  26.129  1.00  0.00
ATOM    657  H5B FA     21      26.859  29.907  25.394  1.00  0.00
ATOM    658  C4B FA     21      28.161  28.259  25.822  1.00  0.00
ATOM    659  O4  FA     21      28.705  28.182  24.575  1.00  0.00
ATOM    660  CM  FA     21      27.943  28.890  23.583  1.00  0.00
ATOM    661 1HM  FA     21      27.024  29.275  24.029  1.00  0.00
ATOM    662 2HM  FA     21      28.534  29.717  23.186  1.00  0.00
ATOM    663 3HM  FA     21      27.690  28.206  22.769  1.00  0.00
ATOM    664  C4A FA     21      28.596  27.410  26.825  1.00  0.00
ATOM    665  C3A FA     21      29.583  26.413  26.576  1.00  0.00
ATOM    666  CA1 FA     21      29.996  25.722  27.707  1.00  0.00
```

```
ATOM    667  C1  FA     21      30.989  24.822  27.365  1.00  0.00
ATOM    668  O1  FA     21      31.588  24.121  28.132  1.00  0.00
ATOM    669  C2A FA     21      31.229  24.853  25.892  1.00  0.00
ATOM    670 1H2A FA     21      32.269  25.044  25.697  1.00  0.00
ATOM    671 2H2A FA     21      30.980  23.893  25.479  1.00  0.00
ATOM    672  C3  FA     21      30.306  25.957  25.343  1.00  0.00
ATOM    673 1H3  FA     21      30.874  26.752  24.886  1.00  0.00
ATOM    674 2H3  FA     21      29.620  25.560  24.616  1.00  0.00
TER
END
```

File F-2: Average structure of rMD refined α-AFB_1-FAPY modified tetramer 5'-$C^1T^2\underline{X}^3A^4$-3'

```
ATOM      1  H5T DC5     1      24.449  20.542   4.634  1.00  0.00
ATOM      2  O5* DC5     1      23.579  20.238   4.319  1.00  0.00
ATOM      3  C5* DC5     1      22.931  21.336   3.679  1.00  0.00
ATOM      4 1H5* DC5     1      21.986  21.001   3.249  1.00  0.00
ATOM      5 2H5* DC5     1      23.564  21.712   2.873  1.00  0.00
ATOM      6  C4* DC5     1      22.679  22.445   4.710  1.00  0.00
ATOM      7  H4* DC5     1      23.621  22.783   5.147  1.00  0.00
ATOM      8  O4* DC5     1      22.022  23.541   4.053  1.00  0.00
ATOM      9  C1* DC5     1      20.653  23.611   4.531  1.00  0.00
ATOM     10  H1* DC5     1      20.525  24.526   5.112  1.00  0.00
ATOM     11  N1  DC5     1      19.677  23.543   3.448  1.00  0.00
ATOM     12  C6  DC5     1      19.644  22.510   2.559  1.00  0.00
ATOM     13  H6  DC5     1      20.414  21.752   2.592  1.00  0.00
ATOM     14  C5  DC5     1      18.602  22.430   1.589  1.00  0.00
ATOM     15  H5  DC5     1      18.579  21.615   0.881  1.00  0.00
ATOM     16  C4  DC5     1      17.622  23.471   1.636  1.00  0.00
ATOM     17  N4  DC5     1      16.581  23.467   0.753  1.00  0.00
ATOM     18 1H4  DC5     1      16.392  22.697   0.116  1.00  0.00
ATOM     19 2H4  DC5     1      15.793  24.108   0.817  1.00  0.00
ATOM     20  N3  DC5     1      17.688  24.451   2.543  1.00  0.00
ATOM     21  C2  DC5     1      18.692  24.518   3.460  1.00  0.00
ATOM     22  O2  DC5     1      18.725  25.397   4.296  1.00  0.00
ATOM     23  C3* DC5     1      21.753  21.920   5.808  1.00  0.00
ATOM     24  H3* DC5     1      21.819  20.835   5.907  1.00  0.00
ATOM     25  C2* DC5     1      20.376  22.380   5.392  1.00  0.00
ATOM     26 1H2* DC5     1      19.900  21.611   4.782  1.00  0.00
ATOM     27 2H2* DC5     1      19.738  22.609   6.249  1.00  0.00
ATOM     28  O3* DC5     1      22.078  22.527   7.065  1.00  0.00
ATOM     29  P   DT      2      21.833  21.679   8.412  1.00  0.00
ATOM     30  O1P DT      2      22.741  22.196   9.460  1.00  0.00
ATOM     31  O2P DT      2      21.847  20.240   8.068  1.00  0.00
ATOM     32  O5* DT      2      20.336  22.094   8.799  1.00  0.00
ATOM     33  C5* DT      2      20.084  23.427   9.258  1.00  0.00
ATOM     34 1H5* DT      2      20.336  24.141   8.471  1.00  0.00
ATOM     35 2H5* DT      2      20.724  23.620  10.122  1.00  0.00
ATOM     36  C4* DT      2      18.616  23.564   9.666  1.00  0.00
ATOM     37  H4* DT      2      18.467  24.545  10.122  1.00  0.00
ATOM     38  O4* DT      2      17.761  23.450   8.511  1.00  0.00
ATOM     39  C1* DT      2      16.845  22.353   8.681  1.00  0.00
ATOM     40  H1* DT      2      15.911  22.749   9.085  1.00  0.00
ATOM     41  N1  DT      2      16.595  21.690   7.396  1.00  0.00
ATOM     42  C6  DT      2      17.441  20.750   6.883  1.00  0.00
ATOM     43  H6  DT      2      18.367  20.537   7.396  1.00  0.00
ATOM     44  C5  DT      2      17.129  20.066   5.669  1.00  0.00
ATOM     45  C7  DT      2      18.096  19.055   5.102  1.00  0.00
ATOM     46 1H7  DT      2      18.400  19.355   4.097  1.00  0.00
ATOM     47 2H7  DT      2      18.982  18.980   5.737  1.00  0.00
ATOM     48 3H7  DT      2      17.616  18.076   5.037  1.00  0.00
ATOM     49  C4  DT      2      15.927  20.340   4.993  1.00  0.00
ATOM     50  O4  DT      2      15.653  19.761   3.962  1.00  0.00
ATOM     51  N3  DT      2      15.108  21.299   5.600  1.00  0.00
ATOM     52  H3  DT      2      14.233  21.513   5.142  1.00  0.00
ATOM     53  C2  DT      2      15.393  21.996   6.782  1.00  0.00
ATOM     54  O2  DT      2      14.641  22.821   7.259  1.00  0.00
```

```
ATOM     55  C3* DT      2      18.169  22.461  10.629  1.00  0.00
ATOM     56  H3* DT      2      18.990  22.043  11.215  1.00  0.00
ATOM     57  C2* DT      2      17.501  21.440   9.717  1.00  0.00
ATOM     58 1H2* DT      2      18.252  20.805   9.241  1.00  0.00
ATOM     59 2H2* DT      2      16.774  20.820  10.247  1.00  0.00
ATOM     60  O3* DT      2      17.163  23.017  11.503  1.00  0.00
ATOM     61  P   FB      3      17.324  22.863  13.101  1.00  0.00
ATOM     62  O1P FB      3      16.135  23.465  13.745  1.00  0.00
ATOM     63  O2P FB      3      18.676  23.336  13.471  1.00  0.00
ATOM     64  O5* FB      3      17.270  21.268  13.345  1.00  0.00
ATOM     65  C5* FB      3      16.028  20.540  13.275  1.00  0.00
ATOM     66 1H5* FB      3      15.597  20.690  12.283  1.00  0.00
ATOM     67 2H5* FB      3      15.332  20.917  14.026  1.00  0.00
ATOM     68  C4* FB      3      16.330  19.050  13.500  1.00  0.00
ATOM     69  H4* FB      3      16.469  18.823  14.559  1.00  0.00
ATOM     70  O4* FB      3      15.249  18.279  12.925  1.00  0.00
ATOM     71  C1* FB      3      15.443  18.165  11.496  1.00  0.00
ATOM     72  N6  FB      3      15.038  16.846  10.995  1.00  0.00
ATOM     73  N7  FB      3      14.414  14.472   9.685  1.00  0.00
ATOM     74  C5  FB      3      13.443  15.414   9.847  1.00  0.00
ATOM     75  C4  FB      3      12.150  15.153   9.373  1.00  0.00
ATOM     76  O4A FB      3      11.884  14.121   8.792  1.00  0.00
ATOM     77  N3  FB      3      11.247  16.177   9.638  1.00  0.00
ATOM     78  H3  FB      3      10.291  16.063   9.327  1.00  0.00
ATOM     79  C2  FB      3      11.609  17.330  10.321  1.00  0.00
ATOM     80  N2  FB      3      10.586  18.209  10.526  1.00  0.00
ATOM     81 1HN2 FB      3       9.644  17.963  10.229  1.00  0.00
ATOM     82 2HN2 FB      3      10.513  18.699  11.416  1.00  0.00
ATOM     83  N1  FB      3      12.834  17.606  10.788  1.00  0.00
ATOM     84  C6  FB      3      13.773  16.619  10.543  1.00  0.00
ATOM     85  C3* FB      3      17.544  18.661  12.655  1.00  0.00
ATOM     86  H3* FB      3      18.271  19.474  12.584  1.00  0.00
ATOM     87  C2* FB      3      16.954  18.334  11.300  1.00  0.00
ATOM     88 1H2* FB      3      17.148  19.193  10.689  1.00  0.00
ATOM     89 2H2* FB      3      17.418  17.451  10.856  1.00  0.00
ATOM     90  O3* FB      3      18.208  17.485  13.135  1.00  0.00
ATOM     91  H1* FB      3      14.904  18.958  10.962  1.00  0.00
ATOM     92  C8  FB      3      14.337  13.267  10.313  1.00  0.00
ATOM     93  O8  FB      3      13.836  13.155  11.413  1.00  0.00
ATOM     94  H8  FB      3      14.257  12.439   9.610  1.00  0.00
ATOM     95  H6  FB      3      15.555  16.105  11.464  1.00  0.00
ATOM     96  P   DA3     4      19.548  17.647  13.999  1.00  0.00
ATOM     97  O1P DA3     4      20.272  18.849  13.529  1.00  0.00
ATOM     98  O2P DA3     4      20.234  16.337  14.063  1.00  0.00
ATOM     99  O5* DA3     4      18.911  17.952  15.429  1.00  0.00
ATOM    100  C5* DA3     4      18.474  16.850  16.226  1.00  0.00
ATOM    101 1H5* DA3     4      19.355  16.332  16.612  1.00  0.00
ATOM    102 2H5* DA3     4      17.900  16.155  15.610  1.00  0.00
ATOM    103  C4* DA3     4      17.612  17.361  17.375  1.00  0.00
ATOM    104  H4* DA3     4      17.925  18.333  17.762  1.00  0.00
ATOM    105  O4* DA3     4      16.229  17.438  16.953  1.00  0.00
ATOM    106  C1* DA3     4      15.373  16.863  17.962  1.00  0.00
ATOM    107  H1* DA3     4      15.144  17.627  18.708  1.00  0.00
ATOM    108  N9  DA3     4      14.136  16.300  17.388  1.00  0.00
ATOM    109  C8  DA3     4      13.663  15.011  17.489  1.00  0.00
ATOM    110  H8  DA3     4      14.209  14.233  18.001  1.00  0.00
ATOM    111  N7  DA3     4      12.513  14.794  16.907  1.00  0.00
```

```
ATOM    112  C5  DA3     4      12.170  16.037  16.394  1.00  0.00
ATOM    113  C6  DA3     4      11.059  16.563  15.656  1.00  0.00
ATOM    114  N6  DA3     4       9.999  15.801  15.260  1.00  0.00
ATOM    115 1H6  DA3     4       9.269  16.141  14.637  1.00  0.00
ATOM    116 2H6  DA3     4       9.976  14.787  15.354  1.00  0.00
ATOM    117  N1  DA3     4      11.066  17.864  15.344  1.00  0.00
ATOM    118  C2  DA3     4      12.084  18.630  15.712  1.00  0.00
ATOM    119  H2  DA3     4      12.006  19.667  15.422  1.00  0.00
ATOM    120  N3  DA3     4      13.181  18.311  16.381  1.00  0.00
ATOM    121  C4  DA3     4      13.173  16.997  16.697  1.00  0.00
ATOM    122  C3* DA3     4      17.650  16.323  18.509  1.00  0.00
ATOM    123  H3* DA3     4      18.385  15.538  18.319  1.00  0.00
ATOM    124  C2* DA3     4      16.232  15.759  18.578  1.00  0.00
ATOM    125 1H2* DA3     4      16.162  14.858  17.966  1.00  0.00
ATOM    126 2H2* DA3     4      15.930  15.524  19.601  1.00  0.00
ATOM    127  O3* DA3     4      17.945  16.987  19.736  1.00  0.00
ATOM    128  H3T DA3     4      17.291  17.697  19.864  1.00  0.00
TER
ATOM    129  O11 FA      5      10.886  15.413   4.777  1.00  0.00
ATOM    130  C11 FA      5      11.288  16.301   5.501  1.00  0.00
ATOM    131  O10 FA      5      12.565  16.348   5.954  1.00  0.00
ATOM    132  CA1 FA      5      12.922  17.389   6.774  1.00  0.00
ATOM    133  C9B FA      5      14.274  17.401   7.253  1.00  0.00
ATOM    134  C9A FA      5      15.393  16.452   7.002  1.00  0.00
ATOM    135  C9  FA      5      15.046  15.006   7.349  1.00  0.00
ATOM    136  O9  FA      5      15.775  14.139   6.483  1.00  0.00
ATOM    137  HO9 FA      5      15.472  13.227   6.641  1.00  0.00
ATOM    138  C8A FA      5      15.533  14.802   8.790  1.00  0.00
ATOM    139  O7  FA      5      16.166  16.037   9.197  1.00  0.00
ATOM    140  H8A FA      5      16.274  14.001   8.818  1.00  0.00
ATOM    141  H9  FA      5      13.975  14.820   7.252  1.00  0.00
ATOM    142  H9a FA      5      15.792  16.559   5.991  1.00  0.00
ATOM    143  C6A FA      5      16.387  16.931   8.086  1.00  0.00
ATOM    144  H6a FA      5      17.427  16.874   7.758  1.00  0.00
ATOM    145  O6A FA      5      16.010  18.294   8.476  1.00  0.00
ATOM    146  C5M FA      5      14.712  18.433   8.106  1.00  0.00
ATOM    147  C5B FA      5      13.790  19.466   8.470  1.00  0.00
ATOM    148  H5B FA      5      14.114  20.271   9.126  1.00  0.00
ATOM    149  C4B FA      5      12.487  19.442   7.999  1.00  0.00
ATOM    150  O4  FA      5      11.614  20.419   8.354  1.00  0.00
ATOM    151  CM  FA      5      12.200  21.381   9.249  1.00  0.00
ATOM    152 1HM  FA      5      13.061  21.851   8.770  1.00  0.00
ATOM    153 2HM  FA      5      11.460  22.146   9.495  1.00  0.00
ATOM    154 3HM  FA      5      12.521  20.881  10.164  1.00  0.00
ATOM    155  C4A FA      5      12.068  18.415   7.167  1.00  0.00
ATOM    156  C3A FA      5      10.724  18.409   6.706  1.00  0.00
ATOM    157  CA1 FA      5      10.407  17.325   5.882  1.00  0.00
ATOM    158  C1  FA      5       9.064  17.424   5.499  1.00  0.00
ATOM    159  O1  FA      5       8.514  16.600   4.797  1.00  0.00
ATOM    160  C2A FA      5       8.443  18.667   6.123  1.00  0.00
ATOM    161 1H2A FA      5       8.083  19.342   5.344  1.00  0.00
ATOM    162 2H2A FA      5       7.614  18.390   6.777  1.00  0.00
ATOM    163  C3  FA      5       9.576  19.344   6.944  1.00  0.00
ATOM    164 1H3  FA      5       9.791  20.346   6.566  1.00  0.00
ATOM    165 2H3  FA      5       9.316  19.398   8.003  1.00  0.00
TER
END
```

File F-3: Average structure of rMD refined *cis*-5*R*,6*S*-thymine glycol modified duplex 5'-G[1]T[2]G[3]C[4]G[5]Tg[6]G[7]T[8]T[9]T[10]G[11]T[12]-3'•5'-A[13]C[14]A[15]A[16]A[17]C[18]G[19]C[20]G[21]C[22]A[23]C[24]-3'

```
ATOM      1  H5T DG5     1      47.096  36.175  24.762  1.00  0.00
ATOM      2  O5* DG5     1      46.864  36.120  25.706  1.00  0.00
ATOM      3  C5* DG5     1      45.952  35.038  25.881  1.00  0.00
ATOM      4 1H5* DG5     1      46.401  34.126  25.481  1.00  0.00
ATOM      5 2H5* DG5     1      45.028  35.235  25.333  1.00  0.00
ATOM      6  C4* DG5     1      45.660  34.858  27.377  1.00  0.00
ATOM      7  H4* DG5     1      46.587  34.807  27.951  1.00  0.00
ATOM      8  O4* DG5     1      44.839  35.947  27.829  1.00  0.00
ATOM      9  C1* DG5     1      43.547  35.573  28.198  1.00  0.00
ATOM     10  H1* DG5     1      43.512  35.520  29.253  1.00  0.00
ATOM     11  N9  DG5     1      42.572  36.546  27.703  1.00  0.00
ATOM     12  C8  DG5     1      42.152  36.756  26.432  1.00  0.00
ATOM     13  H8  DG5     1      42.510  36.190  25.616  1.00  0.00
ATOM     14  N7  DG5     1      41.302  37.711  26.291  1.00  0.00
ATOM     15  C5  DG5     1      41.135  38.175  27.577  1.00  0.00
ATOM     16  C6  DG5     1      40.342  39.230  28.095  1.00  0.00
ATOM     17  O6  DG5     1      39.621  39.992  27.515  1.00  0.00
ATOM     18  N1  DG5     1      40.447  39.373  29.441  1.00  0.00
ATOM     19  H1  DG5     1      39.910  40.083  29.864  1.00  0.00
ATOM     20  C2  DG5     1      41.232  38.621  30.213  1.00  0.00
ATOM     21  N2  DG5     1      41.219  38.911  31.521  1.00  0.00
ATOM     22 1H2  DG5     1      41.792  38.380  32.162  1.00  0.00
ATOM     23 2H2  DG5     1      40.638  39.661  31.867  1.00  0.00
ATOM     24  N3  DG5     1      41.995  37.649  29.784  1.00  0.00
ATOM     25  C4  DG5     1      41.904  37.463  28.449  1.00  0.00
ATOM     26  C3* DG5     1      44.792  33.640  27.627  1.00  0.00
ATOM     27  H3* DG5     1      44.931  32.907  26.873  1.00  0.00
ATOM     28  C2* DG5     1      43.396  34.186  27.643  1.00  0.00
ATOM     29 1H2* DG5     1      43.028  34.219  26.648  1.00  0.00
ATOM     30 2H2* DG5     1      42.737  33.601  28.218  1.00  0.00
ATOM     31  O3* DG5     1      45.163  33.096  28.886  1.00  0.00
ATOM     32  P   DT      2      44.550  31.740  29.414  1.00  0.00
ATOM     33  O1P DT      2      45.574  30.944  30.073  1.00  0.00
ATOM     34  O2P DT      2      43.934  30.941  28.363  1.00  0.00
ATOM     35  O5* DT      2      43.494  32.329  30.439  1.00  0.00
ATOM     36  C5* DT      2      43.960  32.858  31.672  1.00  0.00
ATOM     37 1H5* DT      2      44.625  33.650  31.475  1.00  0.00
ATOM     38 2H5* DT      2      44.491  32.110  32.203  1.00  0.00
ATOM     39  C4* DT      2      42.839  33.377  32.526  1.00  0.00
ATOM     40  H4* DT      2      43.253  33.816  33.407  1.00  0.00
ATOM     41  O4* DT      2      42.115  34.375  31.806  1.00  0.00
ATOM     42  C1* DT      2      40.729  34.207  31.972  1.00  0.00
ATOM     43  H1* DT      2      40.462  34.648  32.902  1.00  0.00
ATOM     44  N1  DT      2      39.970  34.812  30.855  1.00  0.00
ATOM     45  C6  DT      2      39.997  34.274  29.598  1.00  0.00
ATOM     46  H6  DT      2      40.558  33.392  29.432  1.00  0.00
ATOM     47  C5  DT      2      39.332  34.846  28.574  1.00  0.00
ATOM     48  C7  DT      2      39.397  34.224  27.199  1.00  0.00
ATOM     49 1H7  DT      2      38.389  33.986  26.852  1.00  0.00
ATOM     50 2H7  DT      2      39.847  34.926  26.494  1.00  0.00
ATOM     51 3H7  DT      2      39.993  33.310  27.227  1.00  0.00
ATOM     52  C4  DT      2      38.577  36.048  28.774  1.00  0.00
```

```
ATOM     53  O4  DT      2      37.987  36.672  27.936  1.00  0.00
ATOM     54  N3  DT      2      38.563  36.496  30.053  1.00  0.00
ATOM     55  H3  DT      2      38.017  37.296  30.244  1.00  0.00
ATOM     56  C2  DT      2      39.210  35.938  31.111  1.00  0.00
ATOM     57  O2  DT      2      39.086  36.434  32.198  1.00  0.00
ATOM     58  C3* DT      2      41.826  32.317  32.907  1.00  0.00
ATOM     59  H3* DT      2      42.155  31.342  32.644  1.00  0.00
ATOM     60  C2* DT      2      40.596  32.719  32.134  1.00  0.00
ATOM     61 1H2* DT      2      40.611  32.231  31.195  1.00  0.00
ATOM     62 2H2* DT      2      39.706  32.454  32.624  1.00  0.00
ATOM     63  O3* DT      2      41.616  32.392  34.316  1.00  0.00
ATOM     64  P   DG      3      40.715  31.341  35.082  1.00  0.00
ATOM     65  O1P DG      3      41.455  30.810  36.220  1.00  0.00
ATOM     66  O2P DG      3      40.250  30.219  34.254  1.00  0.00
ATOM     67  O5* DG      3      39.519  32.283  35.528  1.00  0.00
ATOM     68  C5* DG      3      39.713  33.163  36.633  1.00  0.00
ATOM     69 1H5* DG      3      40.471  33.861  36.386  1.00  0.00
ATOM     70 2H5* DG      3      40.054  32.611  37.469  1.00  0.00
ATOM     71  C4* DG      3      38.448  33.913  37.016  1.00  0.00
ATOM     72  H4* DG      3      38.685  34.590  37.817  1.00  0.00
ATOM     73  O4* DG      3      37.967  34.662  35.895  1.00  0.00
ATOM     74  C1* DG      3      36.659  34.261  35.494  1.00  0.00
ATOM     75  H1* DG      3      35.948  34.945  35.913  1.00  0.00
ATOM     76  N9  DG      3      36.555  34.246  34.022  1.00  0.00
ATOM     77  C8  DG      3      37.128  33.382  33.144  1.00  0.00
ATOM     78  H8  DG      3      37.753  32.587  33.474  1.00  0.00
ATOM     79  N7  DG      3      36.871  33.620  31.895  1.00  0.00
ATOM     80  C5  DG      3      36.035  34.723  31.947  1.00  0.00
ATOM     81  C6  DG      3      35.382  35.466  30.913  1.00  0.00
ATOM     82  O6  DG      3      35.432  35.325  29.717  1.00  0.00
ATOM     83  N1  DG      3      34.589  36.468  31.385  1.00  0.00
ATOM     84  H1  DG      3      34.088  37.010  30.716  1.00  0.00
ATOM     85  C2  DG      3      34.445  36.756  32.687  1.00  0.00
ATOM     86  N2  DG      3      33.627  37.779  32.971  1.00  0.00
ATOM     87 1H2  DG      3      33.479  38.056  33.930  1.00  0.00
ATOM     88 2H2  DG      3      33.158  38.273  32.225  1.00  0.00
ATOM     89  N3  DG      3      35.056  36.132  33.678  1.00  0.00
ATOM     90  C4  DG      3      35.832  35.109  33.246  1.00  0.00
ATOM     91  C3* DG      3      37.288  33.010  37.419  1.00  0.00
ATOM     92  H3* DG      3      37.620  32.053  37.770  1.00  0.00
ATOM     93  C2* DG      3      36.474  32.905  36.139  1.00  0.00
ATOM     94 1H2* DG      3      36.884  32.125  35.524  1.00  0.00
ATOM     95 2H2* DG      3      35.445  32.674  36.312  1.00  0.00
ATOM     96  O3* DG      3      36.551  33.686  38.454  1.00  0.00
ATOM     97  P   DC      4      35.282  33.007  39.158  1.00  0.00
ATOM     98  O1P DC      4      35.313  33.295  40.604  1.00  0.00
ATOM     99  O2P DC      4      35.197  31.550  38.967  1.00  0.00
ATOM    100  O5* DC      4      34.098  33.775  38.396  1.00  0.00
ATOM    101  C5* DC      4      33.836  35.139  38.737  1.00  0.00
ATOM    102 1H5* DC      4      34.687  35.728  38.472  1.00  0.00
ATOM    103 2H5* DC      4      33.695  35.225  39.793  1.00  0.00
ATOM    104  C4* DC      4      32.605  35.685  38.019  1.00  0.00
ATOM    105  H4* DC      4      32.476  36.715  38.300  1.00  0.00
ATOM    106  O4* DC      4      32.800  35.614  36.599  1.00  0.00
ATOM    107  C1* DC      4      31.668  35.068  35.935  1.00  0.00
ATOM    108  H1* DC      4      30.994  35.870  35.716  1.00  0.00
ATOM    109  N1  DC      4      32.095  34.346  34.710  1.00  0.00
```

```
ATOM    110  C6  DC      4      32.987  33.323  34.788  1.00  0.00
ATOM    111  H6  DC      4      33.344  33.028  35.750  1.00  0.00
ATOM    112  C5  DC      4      33.418  32.709  33.668  1.00  0.00
ATOM    113  H5  DC      4      34.126  31.915  33.729  1.00  0.00
ATOM    114  C4  DC      4      32.915  33.173  32.437  1.00  0.00
ATOM    115  N4  DC      4      33.336  32.593  31.305  1.00  0.00
ATOM    116 1H4  DC      4      34.007  31.839  31.342  1.00  0.00
ATOM    117 2H4  DC      4      32.983  32.911  30.413  1.00  0.00
ATOM    118  N3  DC      4      32.039  34.142  32.341  1.00  0.00
ATOM    119  C2  DC      4      31.613  34.730  33.461  1.00  0.00
ATOM    120  O2  DC      4      30.788  35.596  33.338  1.00  0.00
ATOM    121  C3* DC      4      31.322  34.899  38.305  1.00  0.00
ATOM    122  H3* DC      4      31.458  34.206  39.109  1.00  0.00
ATOM    123  C2* DC      4      31.025  34.195  36.997  1.00  0.00
ATOM    124 1H2* DC      4      31.470  33.225  37.017  1.00  0.00
ATOM    125 2H2* DC      4      29.978  34.079  36.828  1.00  0.00
ATOM    126  O3* DC      4      30.270  35.823  38.618  1.00  0.00
ATOM    127  P   DG      5      28.904  35.314  39.294  1.00  0.00
ATOM    128  O1P DG      5      28.355  36.366  40.152  1.00  0.00
ATOM    129  O2P DG      5      29.068  34.096  40.090  1.00  0.00
ATOM    130  O5* DG      5      27.997  35.043  37.999  1.00  0.00
ATOM    131  C5* DG      5      27.350  36.134  37.349  1.00  0.00
ATOM    132 1H5* DG      5      28.089  36.842  37.034  1.00  0.00
ATOM    133 2H5* DG      5      26.698  36.618  38.040  1.00  0.00
ATOM    134  C4* DG      5      26.551  35.676  36.137  1.00  0.00
ATOM    135  H4* DG      5      26.035  36.521  35.729  1.00  0.00
ATOM    136  O4* DG      5      27.444  35.154  35.139  1.00  0.00
ATOM    137  C1* DG      5      26.986  33.919  34.626  1.00  0.00
ATOM    138  H1* DG      5      26.294  34.111  33.839  1.00  0.00
ATOM    139  N9  DG      5      28.096  33.096  34.108  1.00  0.00
ATOM    140  C8  DG      5      28.960  32.291  34.789  1.00  0.00
ATOM    141  H8  DG      5      28.964  32.217  35.857  1.00  0.00
ATOM    142  N7  DG      5      29.759  31.602  34.034  1.00  0.00
ATOM    143  C5  DG      5      29.398  31.980  32.743  1.00  0.00
ATOM    144  C6  DG      5      29.860  31.549  31.458  1.00  0.00
ATOM    145  O6  DG      5      30.704  30.738  31.183  1.00  0.00
ATOM    146  N1  DG      5      29.223  32.174  30.402  1.00  0.00
ATOM    147  H1  DG      5      29.514  31.913  29.471  1.00  0.00
ATOM    148  C2  DG      5      28.226  33.098  30.533  1.00  0.00
ATOM    149  N2  DG      5      27.701  33.600  29.406  1.00  0.00
ATOM    150 1H2  DG      5      28.048  33.294  28.508  1.00  0.00
ATOM    151 2H2  DG      5      26.958  34.282  29.456  1.00  0.00
ATOM    152  N3  DG      5      27.774  33.486  31.721  1.00  0.00
ATOM    153  C4  DG      5      28.393  32.907  32.784  1.00  0.00
ATOM    154  C3* DG      5      25.561  34.547  36.438  1.00  0.00
ATOM    155  H3* DG      5      25.409  34.414  37.493  1.00  0.00
ATOM    156  C2* DG      5      26.218  33.339  35.795  1.00  0.00
ATOM    157 1H2* DG      5      26.883  32.894  36.496  1.00  0.00
ATOM    158 2H2* DG      5      25.514  32.604  35.486  1.00  0.00
ATOM    159  O3* DG      5      24.314  34.844  35.799  1.00  0.00
ATOM    160  P   TG      6      22.984  34.103  36.398  1.00  0.00
ATOM    161  O2P TG      6      23.203  33.729  37.837  1.00  0.00
ATOM    162  O1P TG      6      21.744  34.928  36.196  1.00  0.00
ATOM    163  O5* TG      6      22.963  32.775  35.443  1.00  0.00
ATOM    164  C5* TG      6      21.955  32.693  34.417  1.00  0.00
ATOM    165 1H5* TG      6      21.263  33.534  34.512  1.00  0.00
ATOM    166 2H5* TG      6      21.394  31.767  34.559  1.00  0.00
```

```
ATOM    167  C4* TG      6      22.618  32.694  33.033  1.00  0.00
ATOM    168  H4* TG      6      22.865  33.706  32.706  1.00  0.00
ATOM    169  O4* TG      6      23.827  31.913  33.080  1.00  0.00
ATOM    170  C1* TG      6      24.022  31.317  31.790  1.00  0.00
ATOM    171  H1* TG      6      24.353  32.101  31.106  1.00  0.00
ATOM    172  C2* TG      6      22.612  30.872  31.405  1.00  0.00
ATOM    173 1H2* TG      6      22.380  29.920  31.887  1.00  0.00
ATOM    174 2H2* TG      6      22.490  30.772  30.324  1.00  0.00
ATOM    175  N1  TG      6      24.995  30.236  31.797  1.00  0.00
ATOM    176  C2  TG      6      25.930  30.128  30.836  1.00  0.00
ATOM    177  O2  TG      6      25.984  30.908  29.907  1.00  0.00
ATOM    178  N3  TG      6      26.737  29.048  30.907  1.00  0.00
ATOM    179  H3  TG      6      27.335  28.881  30.117  1.00  0.00
ATOM    180  C4  TG      6      26.822  28.207  31.960  1.00  0.00
ATOM    181  O4  TG      6      27.571  27.252  31.941  1.00  0.00
ATOM    182  C5  TG      6      26.120  28.645  33.216  1.00  0.00
ATOM    183  O5  TG      6      25.925  27.535  34.089  1.00  0.00
ATOM    184  HO5 TG      6      25.431  26.850  33.605  1.00  0.00
ATOM    185  CM  TG      6      26.963  29.705  33.938  1.00  0.00
ATOM    186 1HM  TG      6      27.121  30.569  33.289  1.00  0.00
ATOM    187 2HM  TG      6      26.455  30.039  34.845  1.00  0.00
ATOM    188 3HM  TG      6      27.937  29.293  34.212  1.00  0.00
ATOM    189  C6  TG      6      24.773  29.236  32.829  1.00  0.00
ATOM    190  H6  TG      6      24.305  29.701  33.698  1.00  0.00
ATOM    191  O6  TG      6      23.936  28.186  32.348  1.00  0.00
ATOM    192  HO6 TG      6      24.351  27.808  31.552  1.00  0.00
ATOM    193  C3* TG      6      21.736  31.981  31.989  1.00  0.00
ATOM    194  H3* TG      6      20.823  31.583  32.438  1.00  0.00
ATOM    195  O3* TG      6      21.395  32.887  30.927  1.00  0.00
ATOM    196  P   DG      7      20.109  32.470  30.004  1.00  0.00
ATOM    197  O1P DG      7      19.423  33.638  29.501  1.00  0.00
ATOM    198  O2P DG      7      19.149  31.729  30.773  1.00  0.00
ATOM    199  O5* DG      7      20.781  31.577  28.847  1.00  0.00
ATOM    200  C5* DG      7      21.404  32.229  27.744  1.00  0.00
ATOM    201 1H5* DG      7      22.146  32.896  28.112  1.00  0.00
ATOM    202 2H5* DG      7      20.678  32.801  27.229  1.00  0.00
ATOM    203  C4* DG      7      22.056  31.263  26.770  1.00  0.00
ATOM    204  H4* DG      7      22.572  31.834  26.024  1.00  0.00
ATOM    205  O4* DG      7      23.014  30.455  27.441  1.00  0.00
ATOM    206  C1* DG      7      23.077  29.151  26.890  1.00  0.00
ATOM    207  H1* DG      7      23.731  29.154  26.052  1.00  0.00
ATOM    208  N9  DG      7      23.553  28.207  27.909  1.00  0.00
ATOM    209  C8  DG      7      23.134  28.121  29.198  1.00  0.00
ATOM    210  H8  DG      7      22.332  28.724  29.566  1.00  0.00
ATOM    211  N7  DG      7      23.817  27.293  29.931  1.00  0.00
ATOM    212  C5  DG      7      24.736  26.753  29.034  1.00  0.00
ATOM    213  C6  DG      7      25.777  25.789  29.213  1.00  0.00
ATOM    214  O6  DG      7      26.123  25.225  30.219  1.00  0.00
ATOM    215  N1  DG      7      26.458  25.530  28.033  1.00  0.00
ATOM    216  H1  DG      7      27.209  24.856  28.080  1.00  0.00
ATOM    217  C2  DG      7      26.201  26.123  26.818  1.00  0.00
ATOM    218  N2  DG      7      26.970  25.747  25.787  1.00  0.00
ATOM    219 1H2  DG      7      26.823  26.153  24.874  1.00  0.00
ATOM    220 2H2  DG      7      27.694  25.057  25.924  1.00  0.00
ATOM    221  N3  DG      7      25.244  27.037  26.642  1.00  0.00
ATOM    222  C4  DG      7      24.561  27.290  27.784  1.00  0.00
ATOM    223  C3* DG      7      21.096  30.276  26.096  1.00  0.00
```

```
ATOM    224  H3* DG      7      20.097  30.392  26.455  1.00  0.00
ATOM    225  C2* DG      7      21.656  28.910  26.452  1.00  0.00
ATOM    226 1H2* DG      7      21.101  28.506  27.250  1.00  0.00
ATOM    227 2H2* DG      7      21.615  28.235  25.652  1.00  0.00
ATOM    228  O3* DG      7      21.145  30.506  24.688  1.00  0.00
ATOM    229  P   DT      8      20.101  29.835  23.674  1.00  0.00
ATOM    230  O1P DT      8      19.783  30.786  22.611  1.00  0.00
ATOM    231  O2P DT      8      18.849  29.374  24.292  1.00  0.00
ATOM    232  O5* DT      8      21.039  28.674  23.088  1.00  0.00
ATOM    233  C5* DT      8      20.610  27.313  23.100  1.00  0.00
ATOM    234 1H5* DT      8      19.859  27.184  22.351  1.00  0.00
ATOM    235 2H5* DT      8      20.177  27.075  24.046  1.00  0.00
ATOM    236  C4* DT      8      21.788  26.390  22.817  1.00  0.00
ATOM    237  H4* DT      8      22.351  26.752  21.980  1.00  0.00
ATOM    238  O4* DT      8      22.652  26.310  23.961  1.00  0.00
ATOM    239  C1* DT      8      23.259  25.011  24.082  1.00  0.00
ATOM    240  H1* DT      8      24.280  25.082  23.782  1.00  0.00
ATOM    241  N1  DT      8      23.187  24.472  25.458  1.00  0.00
ATOM    242  C6  DT      8      22.230  24.913  26.334  1.00  0.00
ATOM    243  H6  DT      8      21.492  25.605  25.987  1.00  0.00
ATOM    244  C5  DT      8      22.216  24.500  27.619  1.00  0.00
ATOM    245  C7  DT      8      21.163  25.027  28.565  1.00  0.00
ATOM    246 1H7  DT      8      20.556  24.201  28.941  1.00  0.00
ATOM    247 2H7  DT      8      21.640  25.519  29.414  1.00  0.00
ATOM    248 3H7  DT      8      20.518  25.742  28.051  1.00  0.00
ATOM    249  C4  DT      8      23.217  23.562  28.102  1.00  0.00
ATOM    250  O4  DT      8      23.336  23.156  29.226  1.00  0.00
ATOM    251  N3  DT      8      24.099  23.128  27.156  1.00  0.00
ATOM    252  H3  DT      8      24.805  22.487  27.448  1.00  0.00
ATOM    253  C2  DT      8      24.125  23.517  25.849  1.00  0.00
ATOM    254  O2  DT      8      24.934  23.024  25.111  1.00  0.00
ATOM    255  C3* DT      8      21.367  24.952  22.565  1.00  0.00
ATOM    256  H3* DT      8      20.488  24.738  23.090  1.00  0.00
ATOM    257  C2* DT      8      22.497  24.153  23.105  1.00  0.00
ATOM    258 1H2* DT      8      22.159  23.301  23.545  1.00  0.00
ATOM    259 2H2* DT      8      23.114  23.887  22.340  1.00  0.00
ATOM    260  O3* DT      8      21.206  24.714  21.183  1.00  0.00
ATOM    261  P   DT      9      20.393  23.442  20.650  1.00  0.00
ATOM    262  O1P DT      9      20.003  23.675  19.257  1.00  0.00
ATOM    263  O2P DT      9      19.153  23.220  21.384  1.00  0.00
ATOM    264  O5* DT      9      21.448  22.244  20.824  1.00  0.00
ATOM    265  C5* DT      9      22.581  22.178  19.957  1.00  0.00
ATOM    266 1H5* DT      9      23.135  23.091  20.039  1.00  0.00
ATOM    267 2H5* DT      9      22.259  22.087  18.946  1.00  0.00
ATOM    268  C4* DT      9      23.488  21.000  20.313  1.00  0.00
ATOM    269  H4* DT      9      24.398  21.092  19.756  1.00  0.00
ATOM    270  O4* DT      9      23.791  21.024  21.718  1.00  0.00
ATOM    271  C1* DT      9      23.540  19.760  22.342  1.00  0.00
ATOM    272  H1* DT      9      24.444  19.179  22.310  1.00  0.00
ATOM    273  N1  DT      9      23.057  19.944  23.734  1.00  0.00
ATOM    274  C6  DT      9      21.922  20.673  23.970  1.00  0.00
ATOM    275  H6  DT      9      21.416  21.120  23.143  1.00  0.00
ATOM    276  C5  DT      9      21.454  20.856  25.220  1.00  0.00
ATOM    277  C7  DT      9      20.220  21.698  25.443  1.00  0.00
ATOM    278 1H7  DT      9      19.932  22.197  24.515  1.00  0.00
ATOM    279 2H7  DT      9      19.397  21.069  25.785  1.00  0.00
ATOM    280 3H7  DT      9      20.419  22.451  26.207  1.00  0.00
```

```
ATOM    281  C4  DT      9      22.147  20.284  26.353  1.00  0.00
ATOM    282  O4  DT      9      21.839  20.392  27.505  1.00  0.00
ATOM    283  N3  DT      9      23.280  19.554  26.025  1.00  0.00
ATOM    284  H3  DT      9      23.794  19.137  26.789  1.00  0.00
ATOM    285  C2  DT      9      23.775  19.352  24.751  1.00  0.00
ATOM    286  O2  DT      9      24.780  18.702  24.544  1.00  0.00
ATOM    287  C3* DT      9      22.862  19.633  20.042  1.00  0.00
ATOM    288  H3* DT      9      21.988  19.717  19.425  1.00  0.00
ATOM    289  C2* DT      9      22.520  19.104  21.424  1.00  0.00
ATOM    290 1H2* DT      9      21.521  19.411  21.665  1.00  0.00
ATOM    291 2H2* DT      9      22.564  18.032  21.487  1.00  0.00
ATOM    292  O3* DT      9      23.846  18.797  19.413  1.00  0.00
ATOM    293  P   DT     10      23.428  17.447  18.654  1.00  0.00
ATOM    294  O1P DT     10      24.391  17.128  17.594  1.00  0.00
ATOM    295  O2P DT     10      22.083  17.500  18.075  1.00  0.00
ATOM    296  O5* DT     10      23.494  16.403  19.869  1.00  0.00
ATOM    297  C5* DT     10      24.759  15.917  20.308  1.00  0.00
ATOM    298 1H5* DT     10      25.422  16.740  20.430  1.00  0.00
ATOM    299 2H5* DT     10      25.169  15.270  19.566  1.00  0.00
ATOM    300  C4* DT     10      24.640  15.170  21.631  1.00  0.00
ATOM    301  H4* DT     10      25.613  14.853  21.949  1.00  0.00
ATOM    302  O4* DT     10      24.080  16.011  22.646  1.00  0.00
ATOM    303  C1* DT     10      23.596  15.188  23.704  1.00  0.00
ATOM    304  H1* DT     10      24.432  14.959  24.330  1.00  0.00
ATOM    305  N1  DT     10      22.521  15.840  24.489  1.00  0.00
ATOM    306  C6  DT     10      21.512  16.522  23.860  1.00  0.00
ATOM    307  H6  DT     10      21.505  16.560  22.795  1.00  0.00
ATOM    308  C5  DT     10      20.535  17.126  24.565  1.00  0.00
ATOM    309  C7  DT     10      19.456  17.898  23.843  1.00  0.00
ATOM    310 1H7  DT     10      19.720  18.020  22.791  1.00  0.00
ATOM    311 2H7  DT     10      18.505  17.367  23.918  1.00  0.00
ATOM    312 3H7  DT     10      19.338  18.883  24.299  1.00  0.00
ATOM    313  C4  DT     10      20.536  17.074  26.011  1.00  0.00
ATOM    314  O4  DT     10      19.760  17.614  26.751  1.00  0.00
ATOM    315  N3  DT     10      21.549  16.344  26.556  1.00  0.00
ATOM    316  H3  DT     10      21.564  16.255  27.545  1.00  0.00
ATOM    317  C2  DT     10      22.532  15.707  25.870  1.00  0.00
ATOM    318  O2  DT     10      23.333  15.052  26.476  1.00  0.00
ATOM    319  C3* DT     10      23.701  13.977  21.607  1.00  0.00
ATOM    320  H3* DT     10      22.922  14.159  20.911  1.00  0.00
ATOM    321  C2* DT     10      23.146  13.918  23.007  1.00  0.00
ATOM    322 1H2* DT     10      22.099  13.851  22.968  1.00  0.00
ATOM    323 2H2* DT     10      23.513  13.073  23.520  1.00  0.00
ATOM    324  O3* DT     10      24.404  12.779  21.277  1.00  0.00
ATOM    325  P   DG     11      23.605  11.448  20.877  1.00  0.00
ATOM    326  O1P DG     11      24.388  10.638  19.958  1.00  0.00
ATOM    327  O2P DG     11      22.323  11.711  20.254  1.00  0.00
ATOM    328  O5* DG     11      23.411  10.750  22.290  1.00  0.00
ATOM    329  C5* DG     11      24.467  10.007  22.850  1.00  0.00
ATOM    330 1H5* DG     11      25.333  10.605  22.837  1.00  0.00
ATOM    331 2H5* DG     11      24.649   9.151  22.272  1.00  0.00
ATOM    332  C4* DG     11      24.171   9.590  24.266  1.00  0.00
ATOM    333  H4* DG     11      25.033   9.153  24.702  1.00  0.00
ATOM    334  O4* DG     11      23.811  10.731  25.019  1.00  0.00
ATOM    335  C1* DG     11      22.819  10.444  25.974  1.00  0.00
ATOM    336  H1* DG     11      23.264  10.420  26.918  1.00  0.00
ATOM    337  N9  DG     11      21.741  11.432  25.977  1.00  0.00
```

```
ATOM    338  C8  DG     11      21.044  11.924  24.925  1.00  0.00
ATOM    339  H8  DG     11      21.245  11.612  23.928  1.00  0.00
ATOM    340  N7  DG     11      20.147  12.803  25.234  1.00  0.00
ATOM    341  C5  DG     11      20.230  12.871  26.613  1.00  0.00
ATOM    342  C6  DG     11      19.498  13.633  27.569  1.00  0.00
ATOM    343  O6  DG     11      18.629  14.449  27.400  1.00  0.00
ATOM    344  N1  DG     11      19.854  13.369  28.857  1.00  0.00
ATOM    345  H1  DG     11      19.379  13.871  29.570  1.00  0.00
ATOM    346  C2  DG     11      20.806  12.494  29.206  1.00  0.00
ATOM    347  N2  DG     11      21.027  12.364  30.521  1.00  0.00
ATOM    348 1H2  DG     11      21.733  11.723  30.854  1.00  0.00
ATOM    349 2H2  DG     11      20.487  12.907  31.181  1.00  0.00
ATOM    350  N3  DG     11      21.521  11.783  28.368  1.00  0.00
ATOM    351  C4  DG     11      21.188  12.015  27.077  1.00  0.00
ATOM    352  C3* DG     11      23.014   8.631  24.414  1.00  0.00
ATOM    353  H3* DG     11      22.405   8.719  23.612  1.00  0.00
ATOM    354  C2* DG     11      22.359   9.085  25.602  1.00  0.00
ATOM    355 1H2* DG     11      21.416   9.098  25.482  1.00  0.00
ATOM    356 2H2* DG     11      22.567   8.447  26.291  1.00  0.00
ATOM    357  O3* DG     11      23.413   7.298  24.582  1.00  0.00
ATOM    358  P   DT3    12      22.379   6.085  24.498  1.00  0.00
ATOM    359  O1P DT3    12      23.082   4.846  24.183  1.00  0.00
ATOM    360  O2P DT3    12      21.347   6.291  23.496  1.00  0.00
ATOM    361  O5* DT3    12      21.756   6.075  25.967  1.00  0.00
ATOM    362  C5* DT3    12      22.473   5.492  27.038  1.00  0.00
ATOM    363 1H5* DT3    12      23.432   5.927  27.071  1.00  0.00
ATOM    364 2H5* DT3    12      22.593   4.456  26.866  1.00  0.00
ATOM    365  C4* DT3    12      21.775   5.717  28.364  1.00  0.00
ATOM    366  H4* DT3    12      22.386   5.366  29.155  1.00  0.00
ATOM    367  O4* DT3    12      21.528   7.100  28.568  1.00  0.00
ATOM    368  C1* DT3    12      20.335   7.319  29.309  1.00  0.00
ATOM    369  H1* DT3    12      20.606   7.624  30.287  1.00  0.00
ATOM    370  N1  DT3    12      19.490   8.338  28.657  1.00  0.00
ATOM    371  C6  DT3    12      19.238   8.273  27.321  1.00  0.00
ATOM    372  H6  DT3    12      19.670   7.487  26.756  1.00  0.00
ATOM    373  C5  DT3    12      18.465   9.186  26.709  1.00  0.00
ATOM    374  C7  DT3    12      18.234   9.094  25.220  1.00  0.00
ATOM    375 1H7  DT3    12      17.492   8.323  25.003  1.00  0.00
ATOM    376 2H7  DT3    12      17.859  10.047  24.845  1.00  0.00
ATOM    377 3H7  DT3    12      19.168   8.856  24.707  1.00  0.00
ATOM    378  C4  DT3    12      17.872  10.257  27.456  1.00  0.00
ATOM    379  O4  DT3    12      17.176  11.127  27.021  1.00  0.00
ATOM    380  N3  DT3    12      18.141  10.243  28.783  1.00  0.00
ATOM    381  H3  DT3    12      17.699  10.931  29.338  1.00  0.00
ATOM    382  C2  DT3    12      18.914   9.331  29.426  1.00  0.00
ATOM    383  O2  DT3    12      19.042   9.416  30.606  1.00  0.00
ATOM    384  C3* DT3    12      20.419   5.080  28.474  1.00  0.00
ATOM    385  H3* DT3    12      19.975   5.072  27.555  1.00  0.00
ATOM    386  C2* DT3    12      19.676   5.967  29.365  1.00  0.00
ATOM    387 1H2* DT3    12      18.626   6.028  29.071  1.00  0.00
ATOM    388 2H2* DT3    12      19.723   5.581  30.302  1.00  0.00
ATOM    389  O3* DT3    12      20.503   3.772  29.037  1.00  0.00
ATOM    390  H3T DT3    12      20.216   3.130  28.364  1.00  0.00
TER
ATOM    391  H5T DA5    13      10.685  14.890  33.017  1.00  0.00
ATOM    392  O5* DA5    13      10.435  15.477  33.754  1.00  0.00
ATOM    393  C5* DA5    13      10.429  14.715  34.938  1.00  0.00
```

```
ATOM    394 1H5* DA5    13       9.872  13.790  34.774  1.00  0.00
ATOM    395 2H5* DA5    13       9.940  15.276  35.737  1.00  0.00
ATOM    396  C4* DA5    13      11.872  14.402  35.327  1.00  0.00
ATOM    397  H4* DA5    13      11.820  13.724  36.126  1.00  0.00
ATOM    398  O4* DA5    13      12.562  13.800  34.257  1.00  0.00
ATOM    399  C1* DA5    13      13.797  14.397  33.998  1.00  0.00
ATOM    400  H1* DA5    13      14.536  13.882  34.545  1.00  0.00
ATOM    401  N9  DA5    13      14.090  14.352  32.557  1.00  0.00
ATOM    402  C8  DA5    13      13.553  15.093  31.566  1.00  0.00
ATOM    403  H8  DA5    13      12.831  15.835  31.753  1.00  0.00
ATOM    404  N7  DA5    13      13.975  14.802  30.378  1.00  0.00
ATOM    405  C5  DA5    13      14.876  13.793  30.613  1.00  0.00
ATOM    406  C6  DA5    13      15.698  13.006  29.807  1.00  0.00
ATOM    407  N6  DA5    13      15.773  13.086  28.511  1.00  0.00
ATOM    408 1H6  DA5    13      16.405  12.499  28.040  1.00  0.00
ATOM    409 2H6  DA5    13      15.261  13.770  28.043  1.00  0.00
ATOM    410  N1  DA5    13      16.464  12.086  30.321  1.00  0.00
ATOM    411  C2  DA5    13      16.447  11.940  31.620  1.00  0.00
ATOM    412  H2  DA5    13      17.081  11.197  32.004  1.00  0.00
ATOM    413  N3  DA5    13      15.736  12.596  32.510  1.00  0.00
ATOM    414  C4  DA5    13      14.958  13.517  31.934  1.00  0.00
ATOM    415  C3* DA5    13      12.693  15.599  35.731  1.00  0.00
ATOM    416  H3* DA5    13      12.090  16.455  35.875  1.00  0.00
ATOM    417  C2* DA5    13      13.646  15.764  34.577  1.00  0.00
ATOM    418 1H2* DA5    13      13.220  16.413  33.867  1.00  0.00
ATOM    419 2H2* DA5    13      14.567  16.166  34.872  1.00  0.00
ATOM    420  O3* DA5    13      13.381  15.289  36.942  1.00  0.00
ATOM    421  P   DC     14      14.192  16.404  37.746  1.00  0.00
ATOM    422  O1P DC     14      14.085  16.161  39.186  1.00  0.00
ATOM    423  O2P DC     14      13.751  17.764  37.467  1.00  0.00
ATOM    424  O5* DC     14      15.655  16.125  37.179  1.00  0.00
ATOM    425  C5* DC     14      16.383  15.015  37.681  1.00  0.00
ATOM    426 1H5* DC     14      15.762  14.160  37.660  1.00  0.00
ATOM    427 2H5* DC     14      16.663  15.193  38.688  1.00  0.00
ATOM    428  C4* DC     14      17.618  14.735  36.856  1.00  0.00
ATOM    429  H4* DC     14      18.068  13.833  37.204  1.00  0.00
ATOM    430  O4* DC     14      17.293  14.578  35.473  1.00  0.00
ATOM    431  C1* DC     14      18.449  14.789  34.701  1.00  0.00
ATOM    432  H1* DC     14      19.007  13.900  34.728  1.00  0.00
ATOM    433  N1  DC     14      18.146  15.178  33.304  1.00  0.00
ATOM    434  C6  DC     14      17.257  16.159  33.032  1.00  0.00
ATOM    435  H6  DC     14      16.771  16.642  33.840  1.00  0.00
ATOM    436  C5  DC     14      16.993  16.500  31.757  1.00  0.00
ATOM    437  H5  DC     14      16.289  17.259  31.542  1.00  0.00
ATOM    438  C4  DC     14      17.672  15.814  30.744  1.00  0.00
ATOM    439  N4  DC     14      17.432  16.101  29.477  1.00  0.00
ATOM    440 1H4  DC     14      16.769  16.811  29.203  1.00  0.00
ATOM    441 2H4  DC     14      17.936  15.545  28.824  1.00  0.00
ATOM    442  N3  DC     14      18.528  14.874  30.980  1.00  0.00
ATOM    443  C2  DC     14      18.788  14.550  32.254  1.00  0.00
ATOM    444  O2  DC     14      19.592  13.687  32.440  1.00  0.00
ATOM    445  C3* DC     14      18.652  15.835  36.867  1.00  0.00
ATOM    446  H3* DC     14      18.203  16.748  37.075  1.00  0.00
ATOM    447  C2* DC     14      19.187  15.826  35.483  1.00  0.00
ATOM    448 1H2* DC     14      19.059  16.731  35.059  1.00  0.00
ATOM    449 2H2* DC     14      20.174  15.617  35.495  1.00  0.00
ATOM    450  O3* DC     14      19.653  15.553  37.819  1.00  0.00
```

```
ATOM    451  P   DA     15      20.806  16.606  38.158  1.00  0.00
ATOM    452  O1P DA     15      21.206  16.453  39.555  1.00  0.00
ATOM    453  O2P DA     15      20.410  17.991  37.934  1.00  0.00
ATOM    454  O5* DA     15      21.951  16.144  37.145  1.00  0.00
ATOM    455  C5* DA     15      22.603  14.898  37.359  1.00  0.00
ATOM    456 1H5* DA     15      21.871  14.128  37.388  1.00  0.00
ATOM    457 2H5* DA     15      23.090  14.912  38.298  1.00  0.00
ATOM    458  C4* DA     15      23.602  14.576  36.266  1.00  0.00
ATOM    459  H4* DA     15      23.938  13.569  36.399  1.00  0.00
ATOM    460  O4* DA     15      22.977  14.683  34.986  1.00  0.00
ATOM    461  C1* DA     15      23.674  15.531  34.107  1.00  0.00
ATOM    462  H1* DA     15      24.285  14.938  33.492  1.00  0.00
ATOM    463  N9  DA     15      22.744  16.313  33.278  1.00  0.00
ATOM    464  C8  DA     15      21.910  17.321  33.648  1.00  0.00
ATOM    465  H8  DA     15      21.861  17.676  34.649  1.00  0.00
ATOM    466  N7  DA     15      21.175  17.798  32.690  1.00  0.00
ATOM    467  C5  DA     15      21.576  17.045  31.593  1.00  0.00
ATOM    468  C6  DA     15      21.230  17.012  30.234  1.00  0.00
ATOM    469  N6  DA     15      20.299  17.828  29.723  1.00  0.00
ATOM    470 1H6  DA     15      20.073  17.782  28.739  1.00  0.00
ATOM    471 2H6  DA     15      19.822  18.490  30.319  1.00  0.00
ATOM    472  N1  DA     15      21.823  16.167  29.396  1.00  0.00
ATOM    473  C2  DA     15      22.737  15.373  29.891  1.00  0.00
ATOM    474  H2  DA     15      23.199  14.699  29.185  1.00  0.00
ATOM    475  N3  DA     15      23.169  15.279  31.131  1.00  0.00
ATOM    476  C4  DA     15      22.540  16.154  31.942  1.00  0.00
ATOM    477  C3* DA     15      24.788  15.528  36.220  1.00  0.00
ATOM    478  H3* DA     15      24.832  16.111  37.071  1.00  0.00
ATOM    479  C2* DA     15      24.549  16.337  35.017  1.00  0.00
ATOM    480 1H2* DA     15      24.077  17.183  35.279  1.00  0.00
ATOM    481 2H2* DA     15      25.425  16.587  34.585  1.00  0.00
ATOM    482  O3* DA     15      25.971  14.795  36.085  1.00  0.00
ATOM    483  P   DA     16      27.406  15.480  36.161  1.00  0.00
ATOM    484  O1P DA     16      28.271  14.675  37.002  1.00  0.00
ATOM    485  O2P DA     16      27.396  16.842  36.684  1.00  0.00
ATOM    486  O5* DA     16      27.834  15.450  34.628  1.00  0.00
ATOM    487  C5* DA     16      28.216  14.214  34.047  1.00  0.00
ATOM    488 1H5* DA     16      27.367  13.579  34.024  1.00  0.00
ATOM    489 2H5* DA     16      28.948  13.749  34.648  1.00  0.00
ATOM    490  C4* DA     16      28.779  14.377  32.641  1.00  0.00
ATOM    491  H4* DA     16      29.036  13.405  32.268  1.00  0.00
ATOM    492  O4* DA     16      27.790  14.964  31.788  1.00  0.00
ATOM    493  C1* DA     16      28.227  16.198  31.222  1.00  0.00
ATOM    494  H1* DA     16      28.585  16.011  30.230  1.00  0.00
ATOM    495  N9  DA     16      27.099  17.143  31.194  1.00  0.00
ATOM    496  C8  DA     16      26.512  17.799  32.232  1.00  0.00
ATOM    497  H8  DA     16      26.885  17.723  33.227  1.00  0.00
ATOM    498  N7  DA     16      25.454  18.487  31.918  1.00  0.00
ATOM    499  C5  DA     16      25.366  18.291  30.542  1.00  0.00
ATOM    500  C6  DA     16      24.499  18.733  29.528  1.00  0.00
ATOM    501  N6  DA     16      23.451  19.524  29.794  1.00  0.00
ATOM    502 1H6  DA     16      22.846  19.829  29.045  1.00  0.00
ATOM    503 2H6  DA     16      23.264  19.815  30.743  1.00  0.00
ATOM    504  N1  DA     16      24.691  18.394  28.267  1.00  0.00
ATOM    505  C2  DA     16      25.717  17.626  27.985  1.00  0.00
ATOM    506  H2  DA     16      25.856  17.387  26.958  1.00  0.00
ATOM    507  N3  DA     16      26.609  17.122  28.818  1.00  0.00
```

```
ATOM    508  C4  DA     16      26.373  17.495  30.095  1.00  0.00
ATOM    509  C3* DA     16      29.990  15.308  32.574  1.00  0.00
ATOM    510  H3* DA     16      30.466  15.401  33.530  1.00  0.00
ATOM    511  C2* DA     16      29.403  16.622  32.088  1.00  0.00
ATOM    512 1H2* DA     16      29.071  17.190  32.934  1.00  0.00
ATOM    513 2H2* DA     16      30.109  17.219  31.546  1.00  0.00
ATOM    514  O3* DA     16      30.908  14.772  31.611  1.00  0.00
ATOM    515  P   DA     17      32.404  15.345  31.502  1.00  0.00
ATOM    516  O1P DA     17      33.329  14.263  31.137  1.00  0.00
ATOM    517  O2P DA     17      32.886  15.979  32.737  1.00  0.00
ATOM    518  O5* DA     17      32.213  16.431  30.341  1.00  0.00
ATOM    519  C5* DA     17      32.183  16.010  28.976  1.00  0.00
ATOM    520 1H5* DA     17      31.354  15.349  28.827  1.00  0.00
ATOM    521 2H5* DA     17      33.079  15.478  28.748  1.00  0.00
ATOM    522  C4* DA     17      32.060  17.204  28.040  1.00  0.00
ATOM    523  H4* DA     17      32.130  16.863  27.025  1.00  0.00
ATOM    524  O4* DA     17      30.781  17.836  28.231  1.00  0.00
ATOM    525  C1* DA     17      30.923  19.247  28.360  1.00  0.00
ATOM    526  H1* DA     17      30.928  19.680  27.377  1.00  0.00
ATOM    527  N9  DA     17      29.811  19.820  29.145  1.00  0.00
ATOM    528  C8  DA     17      29.707  19.991  30.494  1.00  0.00
ATOM    529  H8  DA     17      30.464  19.669  31.173  1.00  0.00
ATOM    530  N7  DA     17      28.610  20.562  30.890  1.00  0.00
ATOM    531  C5  DA     17      27.941  20.809  29.693  1.00  0.00
ATOM    532  C6  DA     17      26.722  21.415  29.334  1.00  0.00
ATOM    533  N6  DA     17      25.858  21.927  30.167  1.00  0.00
ATOM    534 1H6  DA     17      25.026  22.328  29.802  1.00  0.00
ATOM    535 2H6  DA     17      26.057  21.917  31.134  1.00  0.00
ATOM    536  N1  DA     17      26.341  21.500  28.071  1.00  0.00
ATOM    537  C2  DA     17      27.144  21.001  27.155  1.00  0.00
ATOM    538  H2  DA     17      26.804  21.083  26.150  1.00  0.00
ATOM    539  N3  DA     17      28.314  20.415  27.324  1.00  0.00
ATOM    540  C4  DA     17      28.659  20.349  28.632  1.00  0.00
ATOM    541  C3* DA     17      33.112  18.286  28.316  1.00  0.00
ATOM    542  H3* DA     17      33.891  17.918  28.965  1.00  0.00
ATOM    543  C2* DA     17      32.311  19.400  28.971  1.00  0.00
ATOM    544 1H2* DA     17      32.286  19.227  30.028  1.00  0.00
ATOM    545 2H2* DA     17      32.728  20.369  28.796  1.00  0.00
ATOM    546  O3* DA     17      33.663  18.748  27.067  1.00  0.00
ATOM    547  P   DC     18      35.161  19.306  27.018  1.00  0.00
ATOM    548  O1P DC     18      35.712  19.224  25.660  1.00  0.00
ATOM    549  O2P DC     18      36.061  18.567  27.903  1.00  0.00
ATOM    550  O5* DC     18      34.951  20.801  27.537  1.00  0.00
ATOM    551  C5* DC     18      34.922  21.912  26.629  1.00  0.00
ATOM    552 1H5* DC     18      35.851  21.959  26.100  1.00  0.00
ATOM    553 2H5* DC     18      34.846  22.803  27.210  1.00  0.00
ATOM    554  C4* DC     18      33.765  21.849  25.639  1.00  0.00
ATOM    555  H4* DC     18      33.895  21.026  24.967  1.00  0.00
ATOM    556  O4* DC     18      32.532  21.699  26.356  1.00  0.00
ATOM    557  C1* DC     18      31.622  22.780  26.116  1.00  0.00
ATOM    558  H1* DC     18      30.935  22.465  25.349  1.00  0.00
ATOM    559  N1  DC     18      30.903  23.192  27.346  1.00  0.00
ATOM    560  C6  DC     18      31.515  23.127  28.561  1.00  0.00
ATOM    561  H6  DC     18      32.528  22.793  28.616  1.00  0.00
ATOM    562  C5  DC     18      30.896  23.473  29.709  1.00  0.00
ATOM    563  H5  DC     18      31.424  23.397  30.648  1.00  0.00
ATOM    564  C4  DC     18      29.549  23.931  29.601  1.00  0.00
```

```
ATOM    565  N4  DC     18      28.866  24.280  30.699  1.00  0.00
ATOM    566 1H4  DC     18      29.303  24.211  31.608  1.00  0.00
ATOM    567 2H4  DC     18      27.915  24.609  30.619  1.00  0.00
ATOM    568  N3  DC     18      28.960  24.033  28.407  1.00  0.00
ATOM    569  C2  DC     18      29.606  23.689  27.262  1.00  0.00
ATOM    570  O2  DC     18      29.041  23.826  26.206  1.00  0.00
ATOM    571  C3* DC     18      33.618  23.155  24.853  1.00  0.00
ATOM    572  H3* DC     18      34.532  23.715  24.853  1.00  0.00
ATOM    573  C2* DC     18      32.508  23.888  25.581  1.00  0.00
ATOM    574 1H2* DC     18      32.922  24.453  26.394  1.00  0.00
ATOM    575 2H2* DC     18      31.960  24.553  24.941  1.00  0.00
ATOM    576  O3* DC     18      33.197  22.874  23.515  1.00  0.00
ATOM    577  P   DA     19      33.634  23.851  22.327  1.00  0.00
ATOM    578  O1P DA     19      33.496  23.178  21.040  1.00  0.00
ATOM    579  O2P DA     19      35.015  24.287  22.463  1.00  0.00
ATOM    580  O5* DA     19      32.628  25.075  22.508  1.00  0.00
ATOM    581  C5* DA     19      31.351  25.040  21.886  1.00  0.00
ATOM    582 1H5* DA     19      30.780  24.241  22.303  1.00  0.00
ATOM    583 2H5* DA     19      31.483  24.842  20.853  1.00  0.00
ATOM    584  C4* DA     19      30.602  26.355  22.030  1.00  0.00
ATOM    585  H4* DA     19      29.747  26.333  21.391  1.00  0.00
ATOM    586  O4* DA     19      30.141  26.518  23.378  1.00  0.00
ATOM    587  C1* DA     19      30.416  27.819  23.892  1.00  0.00
ATOM    588  H1* DA     19      29.582  28.467  23.701  1.00  0.00
ATOM    589  N9  DA     19      30.660  27.710  25.345  1.00  0.00
ATOM    590  C8  DA     19      31.736  27.170  25.985  1.00  0.00
ATOM    591  H8  DA     19      32.596  26.821  25.458  1.00  0.00
ATOM    592  N7  DA     19      31.612  27.064  27.271  1.00  0.00
ATOM    593  C5  DA     19      30.359  27.624  27.502  1.00  0.00
ATOM    594  C6  DA     19      29.592  27.834  28.670  1.00  0.00
ATOM    595  N6  DA     19      30.002  27.460  29.890  1.00  0.00
ATOM    596 1H6  DA     19      29.415  27.633  30.693  1.00  0.00
ATOM    597 2H6  DA     19      30.896  27.005  30.006  1.00  0.00
ATOM    598  N1  DA     19      28.393  28.442  28.554  1.00  0.00
ATOM    599  C2  DA     19      27.990  28.830  27.364  1.00  0.00
ATOM    600  H2  DA     19      27.043  29.314  27.331  1.00  0.00
ATOM    601  N3  DA     19      28.597  28.675  26.203  1.00  0.00
ATOM    602  C4  DA     19      29.790  28.051  26.342  1.00  0.00
ATOM    603  C3* DA     19      31.456  27.589  21.722  1.00  0.00
ATOM    604  H3* DA     19      32.410  27.319  21.313  1.00  0.00
ATOM    605  C2* DA     19      31.607  28.278  23.068  1.00  0.00
ATOM    606 1H2* DA     19      32.516  27.949  23.511  1.00  0.00
ATOM    607 2H2* DA     19      31.644  29.336  22.982  1.00  0.00
ATOM    608  O3* DA     19      30.753  28.412  20.782  1.00  0.00
ATOM    609  P   DC     20      31.457  29.684  20.100  1.00  0.00
ATOM    610  O1P DC     20      31.067  29.787  18.689  1.00  0.00
ATOM    611  O2P DC     20      32.921  29.669  20.188  1.00  0.00
ATOM    612  O5* DC     20      30.816  30.841  21.005  1.00  0.00
ATOM    613  C5* DC     20      29.471  31.234  20.763  1.00  0.00
ATOM    614 1H5* DC     20      28.880  30.366  20.612  1.00  0.00
ATOM    615 2H5* DC     20      29.426  31.822  19.883  1.00  0.00
ATOM    616  C4* DC     20      28.891  32.010  21.931  1.00  0.00
ATOM    617  H4* DC     20      27.846  32.172  21.782  1.00  0.00
ATOM    618  O4* DC     20      29.072  31.290  23.149  1.00  0.00
ATOM    619  C1* DC     20      28.992  32.203  24.253  1.00  0.00
ATOM    620  H1* DC     20      27.968  32.219  24.578  1.00  0.00
ATOM    621  N1  DC     20      29.887  31.821  25.371  1.00  0.00
```

```
ATOM    622  C6  DC     20      31.118  31.292  25.141  1.00  0.00
ATOM    623  H6  DC     20      31.457  31.195  24.135  1.00  0.00
ATOM    624  C5  DC     20      31.880  30.877  26.172  1.00  0.00
ATOM    625  H5  DC     20      32.833  30.436  25.991  1.00  0.00
ATOM    626  C4  DC     20      31.363  31.042  27.471  1.00  0.00
ATOM    627  N4  DC     20      32.091  30.626  28.517  1.00  0.00
ATOM    628 1H4  DC     20      32.991  30.194  28.367  1.00  0.00
ATOM    629 2H4  DC     20      31.739  30.746  29.456  1.00  0.00
ATOM    630  N3  DC     20      30.208  31.610  27.714  1.00  0.00
ATOM    631  C2  DC     20      29.472  32.022  26.681  1.00  0.00
ATOM    632  O2  DC     20      28.440  32.576  26.939  1.00  0.00
ATOM    633  C3* DC     20      29.557  33.345  22.195  1.00  0.00
ATOM    634  H3* DC     20      30.577  33.268  21.971  1.00  0.00
ATOM    635  C2* DC     20      29.366  33.552  23.654  1.00  0.00
ATOM    636 1H2* DC     20      30.233  33.914  24.071  1.00  0.00
ATOM    637 2H2* DC     20      28.616  34.246  23.820  1.00  0.00
ATOM    638  O3* DC     20      28.950  34.389  21.458  1.00  0.00
ATOM    639  P   DG     21      29.667  35.811  21.319  1.00  0.00
ATOM    640  O1P DG     21      29.240  36.479  20.092  1.00  0.00
ATOM    641  O2P DG     21      31.127  35.731  21.279  1.00  0.00
ATOM    642  O5* DG     21      29.171  36.574  22.645  1.00  0.00
ATOM    643  C5* DG     21      27.816  37.027  22.749  1.00  0.00
ATOM    644 1H5* DG     21      27.171  36.174  22.760  1.00  0.00
ATOM    645 2H5* DG     21      27.567  37.614  21.892  1.00  0.00
ATOM    646  C4* DG     21      27.584  37.854  24.019  1.00  0.00
ATOM    647  H4* DG     21      26.535  38.074  24.103  1.00  0.00
ATOM    648  O4* DG     21      28.004  37.105  25.175  1.00  0.00
ATOM    649  C1* DG     21      29.104  37.725  25.848  1.00  0.00
ATOM    650  H1* DG     21      28.724  38.242  26.714  1.00  0.00
ATOM    651  N9  DG     21      30.094  36.712  26.253  1.00  0.00
ATOM    652  C8  DG     21      30.972  36.025  25.475  1.00  0.00
ATOM    653  H8  DG     21      31.015  36.160  24.415  1.00  0.00
ATOM    654  N7  DG     21      31.715  35.180  26.117  1.00  0.00
ATOM    655  C5  DG     21      31.318  35.339  27.441  1.00  0.00
ATOM    656  C6  DG     21      31.770  34.725  28.655  1.00  0.00
ATOM    657  O6  DG     21      32.610  33.883  28.820  1.00  0.00
ATOM    658  N1  DG     21      31.144  35.193  29.774  1.00  0.00
ATOM    659  H1  DG     21      31.426  34.807  30.664  1.00  0.00
ATOM    660  C2  DG     21      30.186  36.125  29.751  1.00  0.00
ATOM    661  N2  DG     21      29.676  36.467  30.942  1.00  0.00
ATOM    662 1H2  DG     21      28.946  37.162  30.998  1.00  0.00
ATOM    663 2H2  DG     21      30.024  36.031  31.785  1.00  0.00
ATOM    664  N3  DG     21      29.715  36.706  28.663  1.00  0.00
ATOM    665  C4  DG     21      30.328  36.276  27.530  1.00  0.00
ATOM    666  C3* DG     21      28.395  39.147  24.067  1.00  0.00
ATOM    667  H3* DG     21      28.637  39.504  23.082  1.00  0.00
ATOM    668  C2* DG     21      29.626  38.755  24.860  1.00  0.00
ATOM    669 1H2* DG     21      30.352  38.324  24.193  1.00  0.00
ATOM    670 2H2* DG     21      30.090  39.587  25.357  1.00  0.00
ATOM    671  O3* DG     21      27.627  40.134  24.779  1.00  0.00
ATOM    672  P   DC     22      28.142  41.651  24.916  1.00  0.00
ATOM    673  O1P DC     22      27.010  42.574  24.761  1.00  0.00
ATOM    674  O2P DC     22      29.179  42.027  23.947  1.00  0.00
ATOM    675  O5* DC     22      28.718  41.620  26.411  1.00  0.00
ATOM    676  C5* DC     22      27.809  41.661  27.512  1.00  0.00
ATOM    677 1H5* DC     22      27.172  40.807  27.476  1.00  0.00
ATOM    678 2H5* DC     22      27.200  42.524  27.429  1.00  0.00
```

```
ATOM    679  C4* DC     22      28.529  41.676  28.843  1.00  0.00
ATOM    680  H4* DC     22      27.804  41.669  29.629  1.00  0.00
ATOM    681  O4* DC     22      29.343  40.514  28.968  1.00  0.00
ATOM    682  C1* DC     22      30.630  40.810  29.444  1.00  0.00
ATOM    683  H1* DC     22      30.593  40.815  30.498  1.00  0.00
ATOM    684  N1  DC     22      31.625  39.832  28.958  1.00  0.00
ATOM    685  C6  DC     22      31.892  39.709  27.630  1.00  0.00
ATOM    686  H6  DC     22      31.388  40.337  26.937  1.00  0.00
ATOM    687  C5  DC     22      32.772  38.799  27.196  1.00  0.00
ATOM    688  H5  DC     22      32.975  38.701  26.155  1.00  0.00
ATOM    689  C4  DC     22      33.391  37.995  28.165  1.00  0.00
ATOM    690  N4  DC     22      34.270  37.067  27.762  1.00  0.00
ATOM    691 1H4  DC     22      34.474  36.957  26.779  1.00  0.00
ATOM    692 2H4  DC     22      34.729  36.478  28.442  1.00  0.00
ATOM    693  N3  DC     22      33.166  38.098  29.444  1.00  0.00
ATOM    694  C2  DC     22      32.292  39.022  29.863  1.00  0.00
ATOM    695  O2  DC     22      32.126  39.115  31.051  1.00  0.00
ATOM    696  C3* DC     22      29.478  42.855  29.010  1.00  0.00
ATOM    697  H3* DC     22      29.360  43.555  28.223  1.00  0.00
ATOM    698  C2* DC     22      30.832  42.222  28.994  1.00  0.00
ATOM    699 1H2* DC     22      31.204  42.253  28.026  1.00  0.00
ATOM    700 2H2* DC     22      31.490  42.720  29.608  1.00  0.00
ATOM    701  O3* DC     22      29.247  43.482  30.269  1.00  0.00
ATOM    702  P   DA     23      29.906  44.895  30.604  1.00  0.00
ATOM    703  O1P DA     23      28.986  45.712  31.394  1.00  0.00
ATOM    704  O2P DA     23      30.285  45.649  29.415  1.00  0.00
ATOM    705  O5* DA     23      31.188  44.443  31.434  1.00  0.00
ATOM    706  C5* DA     23      31.098  44.284  32.843  1.00  0.00
ATOM    707 1H5* DA     23      30.306  43.633  33.057  1.00  0.00
ATOM    708 2H5* DA     23      30.886  45.214  33.281  1.00  0.00
ATOM    709  C4* DA     23      32.356  43.729  33.459  1.00  0.00
ATOM    710  H4* DA     23      32.220  43.582  34.508  1.00  0.00
ATOM    711  O4* DA     23      32.674  42.485  32.873  1.00  0.00
ATOM    712  C1* DA     23      34.053  42.269  32.876  1.00  0.00
ATOM    713  H1* DA     23      34.291  41.590  33.625  1.00  0.00
ATOM    714  N9  DA     23      34.508  41.765  31.586  1.00  0.00
ATOM    715  C8  DA     23      34.151  42.182  30.343  1.00  0.00
ATOM    716  H8  DA     23      33.474  42.987  30.192  1.00  0.00
ATOM    717  N7  DA     23      34.680  41.504  29.380  1.00  0.00
ATOM    718  C5  DA     23      35.488  40.594  30.054  1.00  0.00
ATOM    719  C6  DA     23      36.353  39.557  29.675  1.00  0.00
ATOM    720  N6  DA     23      36.542  39.234  28.388  1.00  0.00
ATOM    721 1H6  DA     23      37.172  38.484  28.144  1.00  0.00
ATOM    722 2H6  DA     23      36.055  39.741  27.662  1.00  0.00
ATOM    723  N1  DA     23      37.023  38.861  30.572  1.00  0.00
ATOM    724  C2  DA     23      36.868  39.168  31.834  1.00  0.00
ATOM    725  H2  DA     23      37.428  38.592  32.530  1.00  0.00
ATOM    726  N3  DA     23      36.096  40.097  32.352  1.00  0.00
ATOM    727  C4  DA     23      35.417  40.771  31.397  1.00  0.00
ATOM    728  C3* DA     23      33.574  44.566  33.237  1.00  0.00
ATOM    729  H3* DA     23      33.492  45.044  32.343  1.00  0.00
ATOM    730  C2* DA     23      34.619  43.588  33.195  1.00  0.00
ATOM    731 1H2* DA     23      35.283  43.808  32.537  1.00  0.00
ATOM    732 2H2* DA     23      35.035  43.594  34.078  1.00  0.00
ATOM    733  O3* DA     23      33.832  45.476  34.275  1.00  0.00
ATOM    734  P   DC3    24      34.849  46.673  34.087  1.00  0.00
ATOM    735  O1P DC3    24      34.648  47.670  35.123  1.00  0.00
```

```
ATOM    736  O2P DC3    24      34.690  47.311  32.788  1.00  0.00
ATOM    737  O5* DC3    24      36.258  45.947  34.222  1.00  0.00
ATOM    738  C5* DC3    24      36.759  45.600  35.503  1.00  0.00
ATOM    739 1H5* DC3    24      36.036  45.017  36.003  1.00  0.00
ATOM    740 2H5* DC3    24      36.928  46.477  36.072  1.00  0.00
ATOM    741  C4* DC3    24      38.049  44.808  35.405  1.00  0.00
ATOM    742  H4* DC3    24      38.351  44.483  36.367  1.00  0.00
ATOM    743  O4* DC3    24      37.871  43.673  34.578  1.00  0.00
ATOM    744  C1* DC3    24      39.059  43.356  33.874  1.00  0.00
ATOM    745  H1* DC3    24      39.523  42.529  34.358  1.00  0.00
ATOM    746  N1  DC3    24      38.753  43.056  32.462  1.00  0.00
ATOM    747  C6  DC3    24      38.015  43.910  31.716  1.00  0.00
ATOM    748  H6  DC3    24      37.612  44.774  32.173  1.00  0.00
ATOM    749  C5  DC3    24      37.799  43.663  30.416  1.00  0.00
ATOM    750  H5  DC3    24      37.221  44.331  29.834  1.00  0.00
ATOM    751  C4  DC3    24      38.375  42.518  29.873  1.00  0.00
ATOM    752  N4  DC3    24      38.187  42.253  28.573  1.00  0.00
ATOM    753 1H4  DC3    24      37.636  42.879  28.003  1.00  0.00
ATOM    754 2H4  DC3    24      38.598  41.427  28.162  1.00  0.00
ATOM    755  N3  DC3    24      39.087  41.690  30.562  1.00  0.00
ATOM    756  C2  DC3    24      39.282  41.935  31.864  1.00  0.00
ATOM    757  O2  DC3    24      39.938  41.149  32.473  1.00  0.00
ATOM    758  C3* DC3    24      39.134  45.632  34.747  1.00  0.00
ATOM    759  H3* DC3    24      38.699  46.333  34.032  1.00  0.00
ATOM    760  C2* DC3    24      39.941  44.569  34.030  1.00  0.00
ATOM    761 1H2* DC3    24      40.220  44.934  33.139  1.00  0.00
ATOM    762 2H2* DC3    24      40.786  44.347  34.551  1.00  0.00
ATOM    763  O3* DC3    24      39.932  46.310  35.715  1.00  0.00
ATOM    764  H3T DC3    24      40.577  46.866  35.242  1.00  0.00
TER
END
```

APPENDIX G

SCRIPTS

File G-1: This script takes a MARDIGRAS intensity file and assigns error values from NOESY spectra files and generates AMBER restraints segregated by error.

```
#!/usr/bin/perl -w

#errtrans
#		A program to combine two intensity files and one restraint file in the format
#			% errtrans first.INT.1 second.INT.1 restraint.rst
#
#Subroutines
# &pullerr
#		Takes an intensity file and creates a hash containing only errors using atom names and numbers as pointers
# &filesep
#		Removes first element from final combined files
# &newfile
#		Stores an array with given file name
#
#Variables
# @ARGV
#		Array containg arguments passed to the function in sequential order [Intensity file, Intensity file, Restraint file]
# $#ARGV
#		Scalar containing the number of arguments passed to the program
# INT
#		Filehandle for an intensity files
# $FirstRes
#		Scalar containg first residue name and number
# $SecondRes
#		Scalar containing second residue name and number
# $combo1
#		Scalar containg one option for hash indexing for information
# $combo2
#		Scalar containing second option for hash indexing for information
# $hashindex
#		Scalar containing final decision for indexing hash in intensity
# %INT1
#		Hash containing all the errors from first intensity file indexed via residue name and number and sorted
alphanumerically
# %INT2
#		Hash containing all the errors from the second intensity file indexed via residue name and number and sorted
alphanumerically
# @LArray
#		Array containing data from intensity files, [First Residue Name, First Residue Number, Second Residue Name,
Second Residue Number, Intensity, Error]
# %@$OUTPUT
#		Hash, Array, or Scalar containing subroutine calculation
# $RSTfile
#		Name of restraint file
# $jcount,$icount,$tencount,$twentycount,$thirtycount,$fortycount,$fiftycount,$nocount,$jj
#		Counter variables intitialized at zero
# @lineArray
```

```
#		Array containing data from restraint file in format [First Residue Name, First Residue Number, Second Residue
Name, Second Residue Number, Lower Bound, Upper Bound, Center, Width, steps, steps, steps, steps]
# RST
#		Filehandle for restraint file
# @newarray
#		Array containing combined restraint data and intensity errors with errors at both begining and end of array
element
# @sortednew
#		A version of @newarray sorted via errors
# @tens,@twentys,@thirtys,@fortys,@fiftys,@nos
#		Arrays containing only elements of indicated errors
# @aten,@atwenty,@athirty,@aforty,@afifty,@anos
#		Arrays with data finally formatted correctly
# @final
#		Array containing formatted data in subroutine
# $line
#		Scalar for each line to be formatted in subroutine
# $pieces
#		Scalar containing one piece of information that needs to be reorganized in an array

#					CODE

# Test to insure correct number of inputs
if($#ARGV != 3) {
    print " You Suck, do it right!\n just kidden to use this correctly try\n
        %\MAR2AMBER primary.INT secondary.INT totat_rst_list pdbfile \n\n\n";
    exit;
}

%INT1 = &pullerr($ARGV[0]);
%INT2 = &pullerr($ARGV[1]);

#			The following code reads the data from the restraint file and stores it into a hash, RST
$RSTfile = $ARGV[2];
print "\tReading $RSTfile...\n";
open(RST, $RSTfile)||die("Cannot open file!");

$jcount = 0;
while(<RST>){
    if(/H/ && substr($_, 0, 6) ne "HEADER") #only reads lines that start w/ H and are > 1 character deep
    {
        @lineArray = split;		# Separates information on each line
                                # Create possible indexing paramters
            $FirstRes = "$lineArray[0] \t $lineArray[1]";
            $SecondRes = "$lineArray[2] \t $lineArray[3]";
        $combo1 = "$FirstRes \t $SecondRes";
        $combo2 = "$SecondRes \t $FirstRes";

        # Assign indexing paramters based upon a sorting algorithm
        if($combo1 lt $combo2) {
            $hashindex[$jcount] = "$combo1";
                    }else{
                        $hashindex[$jcount] = "$combo2";
                    }
            $RST{$hashindex[$jcount]} = "$lineArray[4] \t $lineArray[5]"; # Store information for the restraint file in a hash,
%RST
            $jcount++;
}}

#The following code creates a hash containing the combined data from the intensity and restraint files with the errors
being at both beginning and end for sorting purposes
$icount = 0;
while ($icount < $jcount){
            if (exists $INT1{$hashindex[$icount]}) {
                    $newarray[$icount] = "$INT1{$hashindex[$icount]} \t $hashindex[$icount] \t
$RST{$hashindex[$icount]} \t $INT1{$hashindex[$icount]}\n";
                    } elsif (exists $INT2{$hashindex[$icount]}){
                        $newarray[$icount] = "$INT2{$hashindex[$icount]} \t $hashindex[$icount] \t
$RST{$hashindex[$icount]} \t $INT2{$hashindex[$icount]}\n";
                    }
```

```
                        else{
                        $newarray[$icount] = "no_error \t $hashindex[$icount] \t $RST{$hashindex[$icount]} \t no_error\n";
                        }
                $icount++;
}
@sortednew = sort @newarray;

# Initialize counter variables
$zcount = 0;$tencount = 0;$twentycount = 0;$thirtycount = 0;$fortycount = 0;$fiftycount = 0;$nocount = 0;

# Following code sorts new restraint file by error and places each in a new variable
while ($zcount < $jcount) {
        $a = substr($sortednew[$zcount],0,2);

        if ($a eq "10") {
                $tens[$tencount] = $sortednew[$zcount];
                $tencount++;
        }elsif ($a eq "20") {
                $twentys[$twentycount] = $sortednew[$zcount];
                $twentycount++;
        }elsif ($a eq "30") {
                $thirtys[$thirtycount] = $sortednew[$zcount];
                $thirtycount++;
        }elsif ($a eq "40") {
                $fortys[$fortycount] = $sortednew[$zcount];
                $fortycount++;
        }elsif ($a eq "50") {
                $fiftys[$fiftycount] = $sortednew[$zcount];
                $fiftycount++;
        }else {
                $nos[$nocount] = $sortednew[$zcount];
                $nocount++;
        }
$zcount++;
}

# Following code removes the first element (error) from the arrays
$jj = 0;
foreach $line (@sortednew) {
        @pieces = split(' ',$line);
        $atotal[$jj] = "$pieces[1] \t $pieces[2] \t $pieces[3] \t $pieces[4] \t $pieces[5] \t $pieces[6] \t $pieces[7]\n";
        $jj++;
}
@aten = &filesep(@tens);
@atwenty = &filesep(@twentys);
@athirty = &filesep(@thirtys);
@aforty = &filesep(@fortys);
@afifty = &filesep(@fiftys);
@anos = &filesep(@nos);


#                               SUBROUTINES

# The following subroutine creates a hash containing all the errors from the intensity files stored by the sorted residue
names and numbers with name of intensity file passed to subroutine
sub pullerr
{
        local(@LArray,$FirstRes, $SecondRes, $combo1, $combo2, $hashindex, %INT, %OUTPUT, $_);
        print "\tReading $_[0]...\n";
        open(INT,$_[0]);
        while(<INT>) {
                if(/H/ && substr($_, 0, 6) ne "HEADER")  #only reads lines that start w/ H and are > 1 character deep
                {
                        @LArray = split(' ',$_);            # Separates information on each line
                        # Create possible indexing paramters
                        $FirstRes = "$LArray[0] \t $LArray[1]";
                        $SecondRes = "$LArray[2] \t $LArray[3]";
                    $combo1 = "$FirstRes \t $SecondRes";
                    $combo2 = "$SecondRes \t $FirstRes";
```

```
                                        # Assign indexing paramters based upon a sorting algorithm
                                if($combo1 lt $combo2) {
                                        $hashindex = "$combo1";
                                        }else{
                                                $hashindex = "$combo2";
                                        }
                                                $INT{$hashindex} = "$LArray[5]";          # Store information
for the intensity file in a hash, INTT
                        }
                        }
                close(INT);
                %OUTPUT = %INT;
}

# The following subroutine removes the first space-delimited block of information in an array with array passed to
subroutine
sub filesep
{
                local(@final,@pieces,@OUTPUT);
                $jj = 0;
                foreach $line (@_) {
                        @pieces = split(' ',$line);
                        $final[$jj] = "$pieces[1] \t $pieces[2] \t $pieces[3] \t $pieces[4] \t $pieces[5] \t $pieces[6]\n";
                        $jj++;
                }
                @OUTPUT = @final;
}

#                               MORE CODE

$pdbfile = $ARGV[3];

open(PDB, $pdbfile)||die("Cannot open file!");

print "\tReading $pdbfile...\n";

$count = 1;


while(<PDB>){  #Makes a Hash with Res Variable and increments though sequence starting at 1

    if(/ O3\* /){
        chomp($_);
                if(/DC5/ || /C5/ || /DC3/ || /C3/ || / C/ || /DC/)
                {
                   $res = "CYT";
                }
                if(/DG5/ || /G5/ || /DG3/ || /G3/ || / G/ || /DG/)
                {
                   $res = "GUA";
                }
                if(/DA5/ || /A5/ || /DA3/ || /A3/ || / A/ || /DA/)
                {
                   $res = "ADE";
                }
                if(/DT5/ || /T5/ || /DT3/ || /T3/ || / T/ || /DT/)
                {
                   $res = "THY";
                }

                #adducts
                if(/tg/ || /TG/)
                {
                    $res = "TG";
                }
        if(/fa/ || /FA/)
                {
                    $res = "FA";
                }
        if(/fb/ || /FB/)
```

```
            {
                $res = "FB";
            }
        $seq[$count] = $res;
        print "\t$count\t$seq[$count]\n";

            $count++;
    }
}
close(PDB);
print "\n\tSequence stored.\n";

@master = &restyp(\@atotal,\@seq);
@ten = &restyp(\@aten,\@seq);
@twenty = &restyp(\@atwenty,\@seq);
@thirty = &restyp(\@athirty,\@seq);
@forty = &restyp(\@aforty,\@seq);
@fifty = &restyp(\@afifty,\@seq);
@noerror = &restyp(\@anos,\@seq);


# "MASTER" file containing all errors , everything.rst
&newfile(@master,"master.list");
# File containing errors of 10, ten.rst
&newfile(@ten,"ten.temp");
# File containng errors of 20. twenty.rst
&newfile(@twenty,"twenty.temp");
# File containing errors of 30, thirty.rst
&newfile(@thirty,"thirty.temp");
# File containing errors of 40, forty.rst
&newfile(@forty,"forty.temp");
# File containing errors of 50, fifty.rst
&newfile(@fifty,"fifty.temp");
# File containing the peaks with no errors, noerror.rst
&newfile(@noerror,"noerror.temp");

print "\n\tConverting files into AMBER format.\n\n\n";
system "/sb/apps/amber9/x86_64/exe/makeDIST_RST -ual ten.temp -pdb $pdbfile -map
/home/brownkl/MyScripts/map.AMBER -rst 10.RST";
system "rm ten.temp";

print "\n\tConverting files into AMBER format.\n\n\n";
system "/sb/apps/amber9/x86_64/exe/makeDIST_RST -ual twenty.temp -pdb $pdbfile -map
/home/brownkl/MyScripts/map.AMBER -rst 20.RST";
system "rm twenty.temp";

print "\n\tConverting files into AMBER format.\n\n\n";
system "/sb/apps/amber9/x86_64/exe/makeDIST_RST -ual thirty.temp -pdb $pdbfile -map
/home/brownkl/MyScripts/map.AMBER -rst 30.RST";
system "rm thirty.temp";

print "\n\tConverting files into AMBER format.\n\n\n";
system "/sb/apps/amber9/x86_64/exe/makeDIST_RST -ual forty.temp -pdb $pdbfile -map
/home/brownkl/MyScripts/map.AMBER -rst 40.RST";
system "rm forty.temp";

print "\n\tConverting files into AMBER format.\n\n\n";
system "/sb/apps/amber9/x86_64/exe/makeDIST_RST -ual fifty.temp -pdb $pdbfile -map
/home/brownkl/MyScripts/map.AMBER -rst 50.RST";
system "rm fifty.temp";

print "\n\tConverting files into AMBER format.\n\n\n";
system "/sb/apps/amber9/x86_64/exe/makeDIST_RST -ual noerror.temp -pdb $pdbfile -map
/home/brownkl/MyScripts/map.AMBER -rst noerror.RST";
system "rm noerror.temp";


print "\n\tEnjoy your new restraints... Be-Yotch!!!.\n\n";
```

```
#                                   MORE SUBROUTINES

sub restyp
{
local($ii, @pieces, @resdat);
$ii = 0;
foreach $dat (@{$_[0]}){
            @pieces = split(" ",$dat);
            if (defined($pieces[6])){
            $resdat[$ii] = "$pieces[1] \t $_[1]->[$pieces[1]] \t $pieces[0] \t $pieces[3] \t $_[1]->[$pieces[3]] \t $pieces[2] \t
$pieces[4] \t $pieces[5] \t $pieces[6]\n";
            $ii++;
            }else{
            $resdat[$ii] = "$pieces[1]\t$_[1]->[$pieces[1]]\t$pieces[0]\t$pieces[3]\t$_[1]-
>[$pieces[3]]\t$pieces[2]\t$pieces[4]\t$pieces[5]\n";
            $ii++;
            }
            }
            @OUTPUT = @resdat;
}

# The following subroutine stores the passed matrix in a .rst file with given file name with information passed to subroutine
in format: filename(@arraydata, "filename")
sub newfile
{
            local($name);
            $name = pop @_;
            open (FHANDLE, "> $name");
            print FHANDLE @_;
            close (FHANDLE);
}
```

File G-2: This script is used for running multiple MARDIGRAS jobs from a single input file and index runs at multiple correlation times to standard distances for A form and B form DNA with *syn* or *anti* base configuration.

```
#!/bin/csh

#file usage mardirun <project name> <correlation time> <mixing time> <noise>
#input files must have very specific formats for this to work (see below)
#

set mardigras=/sb/apps/Linux/bin/mardigras
set sugdist = /home/brownkl/MyScripts/sugdist.range
set sugdist1 = /home/brownkl/MyScripts/sugdist1.range
set long = /home/brownkl/MyScripts/long.range
set short = /home/brownkl/MyScripts/short.range

cat<<eof>constrain.dat
eof

cat<<eof>1.PARM
PDB FILE $1_$2ns.pdb
INT FILE $1_$3.INT.1
OUT FILE $1_$3ms_$2ns
FREQUENCY 800.0
RANDMARDI 50
MINITN 2
MAXITN 10
NOISE ABSOLUTE UNNORMALIZED $4
NORMALIZE ALL
METHYL JUMP 3
PRINT DISTANCES
eof

$mardigras 1.PARM

\rm 1.PARM

rand-restr<<rand-restr_end>rand-restr.junk
1
50
$1_$3ms_$2ns
10
$1_$3ms_$2ns.rrst
rand-restr_end
\rm rand-restr.junk

rrange<<rrange_end>rrange.junk
$sugdist
$1_$3ms_$2ns.rrst
$1_$3ms_$2ns.ianti
rrange_end
\rm rrange.junk

rrange<<rrange_end1>rrange.junk
```

```
$sugdist1
$1_$3ms_$2ns.rrst
$1_$3ms_$2ns.isyn
rrange_end1
\rm rrange.junk

rrange<<rrange_end1>rrange.junk
$long
$1_$3ms_$2ns.rrst
$1_$3ms_$2ns.ilong
rrange_end1
\rm rrange.junk

rrange<<rrange_end1>rrange.junk
$short
$1_$3ms_$2ns.rrst
$1_$3ms_$2ns.ishort
rrange_end1
\rm rrange.junk

echo $1_$3ms_$2ns.ianti >> INDEX
grep r-center $1_$3ms_$2ns.ianti >> INDEX
grep average $1_$3ms_$2ns.ianti >> INDEX

echo $1_$3ms_$2ns.isyn >> INDEX
grep r-center $1_$3ms_$2ns.isyn >> INDEX
grep average $1_$3ms_$2ns.isyn >> INDEX

echo $1_$3ms_$2ns.ilong >> INDEX
grep r-center $1_$3ms_$2ns.ilong >> INDEX
grep average $1_$3ms_$2ns.ilong >> INDEX

echo $1_$3ms_$2ns.ishort >> INDEX
grep r-center $1_$3ms_$2ns.ishort >> INDEX
grep average $1_$3ms_$2ns.ishort >> INDEX

\rm constrain.dat

\rm $1_$3ms_$2ns.ianti
\rm $1_$3ms_$2ns.isyn
\rm $1_$3ms_$2ns.ilong
\rm $1_$3ms_$2ns.ishort
```

File G-3: This script is used for preparing pdb files with a common star and prime nomenclature.

```
#!/usr/bin/perl

#star2prime

#This will replace the "*" in insight PDBs with "'" and """
#where appropriate

$infile = $ARGV[0];
$outfile = $ARGV[1];

open(INPUT, $infile)||die("Cannot Open File");


print "\n\tChanging atom names...";

while(<INPUT>){ #move line by line thru the file

    #if there is a filename in the header, change it
    s/$infile/$outfile/g;

    #change the atom names
    s/ DT /  T /g;
    s/ DT3/  T /g;
    s/ 1HM / HM1 /g;
    s/ 2HM / HM2 /g;
    s/ 3HM / HM3 /g;
    s/H1\*/H1'/g;
    s/H3\*/H3'/g;
    s/H4\*/H4'/g;

    #H2', H2", H5', H5" are also changed
    #a bug in Insight switchs H5" and H5', this fixes it
    s/1H2\*/H2'1/g;
    s/2H2\*/H2'2/g;
    s/2H5\*/H5'2/g;
    s/1H5\*/H5'1/g;

    $replace .= $_;
}
close(INPUT);

open(OUTPUT, ">$outfile");
print OUTPUT $replace;
close(OUTPUT);

print "\tDONE!\n";
```

File G-3: This script is used for quickly analyzing pdb files with CORMA

```
#!/bin/csh

#
#usage %corma.com pdbfilenum INTfile outfile

set corma=/u/6.2/mardigras5.2_6.2/corma
set cormain=/u/6.2/mardigras5.2_6.2/corma.in
set prep=/home/kyleb/MyScripts/tgpdb.prep

cat<<eof>constrain.dat
eof

$cormain<<eof>corma.junk2

$1.pdb
file.pdb
y
i
3
3
;

eof
\rm corma.junk2
$prep file.pdb $1.PDB
$corma<<eof>corma.junk

800;     #Spectrometer Frequence
n;       #ENTER EXPERIMENT TYPE: [N(OESY)/R(OESY),
\r       #Enter RELAXATION DELAY (in sec), OR RETURN FOR FULL RELAXED
SIMULATION:
n;       #INCLUDE KINETIC EXCHANGE? [def=n]:
n;       #ENSEMBLE (MULTIPLE FAST EXCHANGE)? [def=n]
$1.PDB;          #ENTER PDB FILE-NAME :
y;       #COMPARE WITH EXPERIMENTAL INTENSITIES?
$2;      #ENTER NAME OF EXPERIMENTAL INTENSITY FILE:
a;       #NORMALIZE USING ONLY FIXED-DISTANCE INTENSITIES [def=f],
$3;      #NAME FOR INTENSITY FILE TO BE CREATED:
3;       #CUTOFF LEVEL FOR INTENSITIES?
y;       #DISPLAY INTENSITIES IN EXTENDED PRECISION? [def=y]:
n;       #ADD RANDOM NOISE? [def=n]
3;       # SELECT METHYL JUMP MODEL:
n;       #GENERATE POSTSCRIPT FILES FOR PLOTTING? [def=n]
y;       #WRITE SUBMATRICES FILE IN EXTENDED FORMAT? [def=n]:
25;      # NUMBER OF RESIDUES PRINTED IN A ROW? [def=25]
eof
\rm corma.junk
\rm file.pdb
\rm constrain.dat
```

REFERENCES

1. DePinho, R. A. (2000) The age of cancer. *Nature 408*, 248-254.

2. Hoeijmakers, J. H. (2001) Genome maintenance mechanisms for preventing cancer. *Nature 411*, 366-374.

3. Kawanishi, S., Hiraku, Y. and Oikawa, S. (2001) Mechanism of guanine-specific DNA damage by oxidative stress and its role in carcinogenesis and aging. *Mutat. Res. 488*, 65-76.

4. Finkel, T. and Holbrook, N. J. (2000) Oxidants, oxidative stress and the biology of ageing. *Nature 408*, 239-247.

5. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature 362*, 709-715.

6. Ames, B. N., Shigenaga, M. K. and Hagen, T. M. (1993) Oxidants, antioxidants, and the degenerative diseases of aging. *Proc. Natl. Acad. Sci. USA 90*, 7915-7922.

7. Halliwell, B. and Gutteridge, J. M. (1989) *Free Radicals in Biology and Medicine*, Claredon Press, Oxford.

8. Weinberg, R. A. (1989) Oncogenes, antioncogenes, and the molecular bases of multistep carcinogenesis. *Cancer Res. 49*, 3713-3721.

9. Miller, E. C. (1978) Some current perspectives on chemical carcinogenesis in humans and experimental animals: Presidential Address. *Cancer Res. 38*, 1479-1496.

10. Singer, B. and Grungerger, D. (1983) *Molecular Biology of Mutagens and Carcinogens*, Plenum Press, New York.

11. Miller, J. A. (1970) Carcinogenesis by Chemicals: An Overview—G. H. A. Clowes Memorial Lecture *Cancer Res. 30*, 559-576.

12. (1976) *Free Radicals in Biology*, Vol. 1-5, Academic, New York.

13. Huttermann, J., Kohnlein, W., Teoule, R. and Bertinchamps, A. J. (1978) *Effects of Ionizing Radiation of DNA*, Springer, New York.

14. Wang, D., Kreutzer, D. A. and Essigmann, J. M. (1998) Mutagenicity and repair of oxidative DNA damage: insights from studies using defined lesions. *Mutat. Res. 400*, 99-115.

15. Friedberg, E. C. (1985) *DNA Repair*, W.H. Freeman and Co., New York.

16. Laval, J., Jurado, J., Saparbaev, M. and Sidorkina, O. (1998) Antimutagenic role of base-excision repair enzymes upon free radical-induced DNA damage. *Mutat. Res. 402*, 93-102.

17. Delaney, J. C. and Essigmann, J. M. (1999) Context-dependent mutagenesis by DNA lesions. *Chem. Biol. 6*, 743-753.

18. Lindahl, T. and Wood, R. D. (1999) Quality control by DNA repair. *Science 286*, 1897-1905.

19. Essigmann, J. M. and Wood, M. L. (1993) The relationship between the chemical structures and mutagenic specificities of the DNA lesions formed by chemical and physical mutagens. *Toxicol. Lett. 67*, 29-39.

20. Singer, B. and Essigmann, J. M. (1991) Site-specific mutagenesis: retrospective and prospective. *Carcinogenesis 12*, 949-955.

21. Basu, A. K. and Essigmann, J. M. (1988) Site-specifically modified oligodeoxynucleotides as probes for the structural and biological effects of DNA-damaging agents. *Chem. Res. Toxicol. 1*, 1-18.

22. Morrison, R. T. and Boyd, R. N. (1992) *Organic Chemistry*, 6th ed., Prentice Hall, Englewood Cliffs, New Jersey.

23. Carey, F. A. (2008) *Organic Chemistry*, 7th ed., McGraw-Hill, Boston.

24. Voet, D., Voet, J. G. and Pratt, C. W. (2002) *Fundamentals of Biochemistry*, Upgrade ed., John Wiley & Sons, Inc., New York.

25. Tomasz, M., Lipman, R., Lee, M. S., Verdine, G. L. and Nakanishi, K. (1987) Reaction of acid-activated mitomycin C with calf thymus DNA and model guanines: elucidation of the base-catalyzed degradation of N7-alkylguanine nucleosides. *Biochemistry 26*, 2010-2027.

26. Burgdorf, L. T. and Carell, T. (2002) Synthesis, stability, and conformation of the formamidopyrimidine G DNA lesion. *Chem. Eur. J. 8*, 293-301.

27. Haraguchi, K., Delaney, M. O., Wiederholt, C. J., Sambandam, A., Hantosi, Z. and Greenberg, M. M. (2002) Synthesis and characterization of oligodeoxynucleotides containing formamidopyrimidine lesions and nonhydrolyzable analogues. *J. Am. Chem. Soc. 124*, 3263-3269.

28. Aramini, J. M., Cleaver, S. H., Pon, R. T., Cunningham, R. P. and Germann, M. W. (2004) Solution structure of a DNA duplex containing an α-anomeric adenosine: insights into substrate recognition by endonuclease IV. *J. Mol. Biol. 338*, 77-91.

29. Greenberg, M. M., Hantosi, Z., Wiederholt, C. J. and Rithner, C. D. (2001) Studies on N4-(2-deoxy-D-pentofuranosyl)-4,6-diamino-5-formamidopyrimidine (Fapy.dA) and N6-(2-deoxy-D-pentofuranosyl)-6-diamino-5-formamido-4-hydroxypyrimidine (Fapy.dG). *Biochemistry 40*, 15856-15861.

30. Patro, J. N., Wiederholt, C. J., Jiang, Y. L., Delaney, J. C., Essigmann, J. M. and Greenberg, M. M. (2007) Studies on the replication of the ring opened formamidopyrimidine, Fapy.dG in *Escherichia coli*. *Biochemistry 46*, 10202-10212.

31. Brown, K. L., Deng, J. Z., Iyer, R. S., Iyer, L. G., Voehler, M. W., Stone, M. P., Harris, C. M. and Harris, T. M. (2006) Unraveling the aflatoxin-FAPY conundrum: Structural basis for differential replicative processing of isomeric forms of the formamidopyrimidine-type DNA adduct of aflatoxin B_1. *J. Am. Chem. Soc. 128*, 15188-15199.

32. Tudek, B., Boiteux, S. and Laval, J. (1992) Biological properties of imidazole ring-opened N^7-methylguanine in M13mp18 phage DNA. *Nucleic Acids Res. 20*, 3079-3084.

33. Boiteux, S. and Laval, J. (1983) Imidazole open ring 7-methylguanine: an inhibitor of DNA synthesis. *Biochem. Biophys. Res. Commun. 110*, 552-558.

34. Asagoshi, K., Terato, H., Ohyama, Y. and Ide, H. (2002) Effects of a guanine-derived formamidopyrimidine lesion on DNA replication: translesion DNA synthesis, nucleotide insertion, and extension kinetics. *J. Biol. Chem. 277*, 14589-14597.

35. Saladino, R., Crestini, C., Mincione, E., Costanzo, G., Di Mauro, E. and Negri, R. (1997) Mechanism of degradation of 2'-deoxycytidine by formamide: implications for chemical DNA sequencing procedures. *Bioorg. Med. Chem. 5*, 2041-2048.

36. Busch, F., Pieck, J. C., Ober, M., Gierlich, J., Hsu, G. W., Beese, L. S. and Carell, T. (2008) Dissecting the differences between the α and β anomers of the oxidative DNA lesion FaPydG. *Chem. Eur. J. 14*, 2125-2132.

37. Lustig, M. J., Cadet, J., Boorstein, R. J. and Teebor, G. W. (1992) Synthesis of the diastereomers of thymidine glycol, determination of concentrations and rates of interconversion of their *cis-trans* epimers at equilibrium and demonstration of differential alkali lability within DNA. *Nucleic Acids Res. 20*, 4839-4845.

38. Ide, H., Yamaoka, T. and Kimura, Y. (1994) Replication of DNA templates containing the α-anomer of deoxyadenosine, a major adenine lesion produced by hydroxyl radicals. *Biochemistry 33*, 7127-7133.

39. Shimizu, H., Yagi, R., Kimura, Y., Makino, K., Terato, H., Ohyama, Y. and Ide, H. (1997) Replication bypass and mutagenic effect of α-deoxyadenosine site-specifically incorporated into single-stranded vectors. *Nucleic Acids Res. 25*, 597-603.

40. Haines, J. A., Reese, C. B. and Todd, L. (1962) The methylation of guanosine and related compounds with diazomethane. *J. Chem. Soc.*, 5281-5288.

41. O'Connor, T. R., Boiteux, S. and Laval, J. (1988) Ring-opened 7-methylguanine residues in DNA are a block to *in vitro* DNA synthesis. *Nucleic Acids Res. 16*, 5879-5894.

42. Boiteux, S., Belleney, J., Roques, B. P. and Laval, J. (1984) Two rotameric forms of open ring 7-methylguanine are present in alkylated polynucleotides. *Nucleic Acids Res. 12*, 5429-5439.

43. Barrows, L. R. and Magee, P. N. (1982) Nonenzymatic methylation of DNA by S-adenosylmethionine *in vitro*. *Carcinogenesis 3*, 349-351.

44. Rydberg, B. and Lindahl, T. (1982) Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *Embo. J. 1*, 211-216.

45. Holliday, R. and Ho, T. (1998) Gene silencing and endogenous DNA methylation in mammalian cells. *Mutat. Res. 400*, 361-368.

46. Kodama, M. and Saito, H. (1980) Formation of methylurea from methylamine and carbamyl phosphate: a possible environmental hazard. *Cancer Lett. 10*, 319-324.

47. Vaughan, P. and Sedqwick, B. (1991) A weak adaptive response to alkylation damage of nitrosation. *Environ. Mol. Mutagen. 15*, 69-70.

48. Bartsch, H., Ohsima, H., Shuker, D. E. G., Pignatelli, B. and Calmes, S. (1990) Exposure of humans to endogenous N-nitroso compounds; implications in cancer etiology. *Mutat. Res. 238*, 255-267.

49. Hecht, S. S. (1999) Tobacco smoke carcinogens and lung cancer. *J. Natl. Cancer Inst. 91*, 1194-1210.

50. Wogan, G. N., Hecht, S. S., Felton, J. S., Conney, A. H. and Loeb, L. A. (2004) Environmental and chemical carcinogenesis. *Semin. Cancer Biol. 14*, 473-486.

51. Hecht, S. S. (1998) Biochemistry, biology, and carcinogenicity of tobacco-specific N-nitrosamines. *Chem. Res. Toxicol. 11*, 559-603.

52. Smith, G. B., Castonguay, A., Donnelly, P. J., Reid, K. R., Petsikas, D. and Massey, T. E. (1999) Biotransformation of the tobacco-specific carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) in freshly isolated human lung cells. *Carcinogenesis 20*, 1809-1818.

53. Vodicka, P. and Hemminki, K. (1988) Depurination and imidazole ring-opening in nucleosides and DNA alkylated by styrene oxide. *Chem. Biol. Interact. 68*, 117-126.

54. Hendler, S., Furer, E. and Srinivasan, P. R. (1970) Synthesis and chemical properties of monomers and polymers containing 7-methylguanine and an investigation of their substrate or template properties for bacterial deoxyribonucleic acid or ribonucleic acid polymerases. *Biochemistry 9*, 4141-4153.

55. Barbarella, G., Tugnoli, V. and Zambianchi, M. (1991) Imidazole ring-opening of 7-methylguanosine at physiological pH. *Nucleosides Nucleotides 10*, 1759-1769.

56. Christov, P. (2007) Synthesis and *In Vitro* Replication Studies of N^5-Alkylated Formamidopyrimidine Adducts in DNA, Ph.D In *Chemistry*, Vanderbilt University, Nashville, TN.

57. Chetsanga, C. J. and Lindahl, T. (1979) Release of 7-methylguanine residues whose imidazole rings have been opened from damaged DNA by a DNA glycosylase from *Escherichia coli*. *Nucleic Acids Res. 6*, 3673-3684.

58. Laval, J., Pierre, J. and Laval, F. (1981) Release of 7-methylguanine residues from alkylated DNA by extracts of *Micrococcus luteus* and *Escherichia coli*. *Proc. Natl. Acad. Sci. USA 78*, 852-855.

59. Chetsanga, C. J., Bearie, B. and Makaroff, C. (1982) Alkaline opening of imidazole ring of 7-methylguanosine. Analysis of the resulting pyrimidine derivatives. *Chem. Biol. Interact. 41*, 217-233.

60. Hemminki, K., Peltonen, K. and Vodicka, P. (1989) Depurination from DNA of 7-methylguanine, 7-(2-amonoethyl)-guanine and ring-opened 7-methylguanines. *Chem. Biol. Interact. 70*, 289-303.

61. Beranek, D. T., Weiss, C. C., Evans, F. E., Chetsanga, C. J. and Kadlubar, F. F. (1983) Identification of N5-methyl-N5-formyl-2,5,6-triamino-4-

hydroxypyrimidine as a major adduct in rat liver DNA after treatment with the carcinogens, N,N-dimethylnitrosamine or 1,2-dimethylhydrazine. *Biochem. Biophys. Res. Commun. 110*, 625-631.

62. Kadlubar, F. F., Beranek, D. T., Weis, C. C., Evans, F. E., Cox, R. and Irving, C. C. (1984) Characterization of the purine ring-opened 7-methylguanine and its persistence in rat bladder epithelial DNA after treatment with the carcinogen N-methylnitrosourea. *Carcinogenesis 5*, 587-592.

63. Humphreys, W. G. and Guengerich, F. P. (1991) Structure of formamidopyrimidine adducts as determined by NMR using specifically ^{15}N-labeled guanosine. *Chem. Res. Toxicol. 4*, 632-636.

64. Box, H. C., Ligam, K. T., French, J. B., Potienko, G. and Alderfer, J. L. (1981) ^{13}C NMR characterization of alkyalted derivatives of guanosine, adenosine, and cytidine. *Carbohydr., Nucleosides, Nucleotides 8*, 189-195.

65. Raoul, S., Bardet, M. and Cadet, J. (1995) γ Irradiation of 2'-deoxyadenosine in oxygen-free aqueous solutions: Identification and conformational features of formamidopyrimidine nucleoside derivatives. *Chem. Res. Toxicol. 8*, 924-933.

66. Berger, M. and Cadet, J. (1985) Isolation and characterization of the radiation-induced degradation products of 2'-deoxyguanisine in oxygen-free aqeouos solutions. *Z. Naturforsch. 40B*, 1519-1531.

67. Haraguchi, K. and Greenberg, M. M. (2001) Synthesis of oligonucleotides containing Fapy.dG (N6-(2-deoxy-α,β-D-erythro-pentofuranosyl)-2,6-diamino-4-hydroxy-5- formamidopyrimidine). *J. Am. Chem. Soc. 123*, 8636-8637.

68. Chetsanga, C. J., Lozon, M., Makaroff, C. and Savage, L. (1981) Purification and characterization of *Escherichia coli* formamidopyrimidine-DNA glycosylase that excises damaged 7-methylguanine from deoxyribonucleic acid. *Biochemistry 20*, 5201-5207.

69. van der Kemp, P. A., Thomas, D., Barbey, R., De Oliveira, R. and Boiteux, S. (1996) Cloning and expression in *Escherichia coli* of the OGG1 gene of *Saccharomyces cerevisiae*, which codes for a DNA glycosylase that excises 7,8-dihydro-8-oxoguanine and 2,6-diamino-4-hydroxy-5-N-methylformamidopyrimidine. *Proc. Natl. Acad. Sci. USA 93*, 5197-5202.

70. Asagoshi, K., Yamada, T., Terato, H., Ohyama, Y., Monden, Y., Arai, T., Nishimura, S., Aburatani, H., Lindahl, T. and Ide, H. (2000) Distinct repair activities of human 7,8-dihydro-8-oxoguanine DNA glycosylase and formamidopyrimidine DNA glycosylase for formamidopyrimidine and 7,8-dihydro-8-oxoguanine. *J. Biol. Chem. 275*, 4956-4964.

71. Graves, R., Laval, J. and Pegg, A. E. (1992) Sequence specificity of DNA repair by *Escherichia coli* Fpg protein. *Carcinogenesis 13*, 1455-1459.

72. Busby, W. F. and Wogan, G. N. (1984) Aflatoxins, In *Chemical Carcinogens* (Searle, C., Ed.) pp 945-1136, American Chemical Society, Washington, DC.

73. Smela, M. E., Currier, S. S., Bailey, E. A. and Essigmann, J. M. (2001) The chemistry and biology of aflatoxin B_1: from mutational spectrometry to carcinogenesis. *Carcinogenesis 22*, 535-545.

74. Bennett, J. W. and Klich, M. (2003) Mycotoxins. *Clin. Microbiol. Rev. 16*, 497-516.

75. Diener, U. L., Cole, R. L., Sanders, G. A., Payne, G. A., Lee, L. S. and Klich, M. A. (1987) Epidemiology of aflatoxin formation by *Aspergillus flavus*. *Annu. Rev. Phytopathol. 25*, 249-270.

76. Detroy, R. W., Lillehoj, E. B. and Ciegler, A. (1971) Aflatoxin and related compounds, In *Microbial Toxins* (Ciegler, A., Kadis, S. and Ajl, S. J., Eds.) pp 3-178, Academic Press, New York.

77. Blout, W. P. (1961) Turkey "X" disease. *J. Br. Turkey Fed. 9*, 52-58.

78. (1969) *Aflatoxin, Scientific Background, Control, and Implications*, Academic Press, New York.

79. Henry, S. H., Bosch, F. X., Troxell, T. C. and Bolger, P. M. (1999) Policy forum: public health. Reducing liver cancer--global control of aflatoxin. *Science 286*, 2453-2454.

80. Wilson, D. M. and Payne, G. A. (1994) Factors Affecting *Aspergillus flavus* Group Infection and Aflatoxin Contamination of Crops, In *The toxicology of aflatoxins. Human heatlh, veterinary and agricultural significance* (Eaton, D. L. and Groopman, J. D., Eds.) pp 309-325, Academic Press, San Diego, CA.

81. Abbas, H. K., Williams, W. P., Windham, G. L., Pringle, H. C., 3rd, Xie, W. and Shier, W. T. (2002) Aflatoxin and fumonisin contamination of commercial corn (Zea mays) hybrids in Mississippi. *J. Agric. Food Chem. 50*, 5246-5254.

82. Heathcote, J. G. and Hibbert, J. R. (1978) *Aflatoxins: Chemical and Biological Aspects.*, Elsevier Scientific Publishing Company, Amsterdam, The Netherlands.

83. Hartley, R. D., Nesbitt, B. F. and O'Kelly, J. (1963) Toxic metabolites of *Aspergillus flavus*. *Nature 198*, 1056-1058.

84. Squire, R. A. (1981) Ranking animal carcinogens: a proposed regulatory approach. *Science 214*, 877-880.

85. Garner, R. C., Miller, E. C. and Miller, J. A. (1972) Liver microsomal metabolism of aflatoxin B_1 to a reactive derivative toxic to *Salmonella typhimurium* TA 1530. *Cancer Res. 32*, 2058-2066.

86. Shimada, T. and Guengerich, F. P. (1989) Evidence for cytochrome P-450_{NF}, the nifedipine oxidase, being the principal enzyme involved in the bioactivation of aflatoxins in human liver. *Proc. Natl. Acad. Sci. USA 86*, 462-465.

87. Raney, K. D., Shimada, T., Kim, D. H., Groopman, J. D., Harris, T. M. and Guengerich, F. P. (1992) Oxidation of aflatoxins and sterigmatocystin by human liver microsomes: significance of aflatoxin Q_1 as a detoxication product of aflatoxin B_1. *Chem. Res. Toxicol. 5*, 202-210.

88. Ueng, Y. F., Shimada, T., Yamazaki, H. and Guengerich, F. P. (1995) Oxidation of aflatoxin B_1 by bacterial recombinant human cytochrome P450 enzymes. *Chem. Res. Toxicol. 8*, 218-225.

89. Swenson, D. H., Lin, J. K., Miller, E. C. and Miller, J. A. (1977) Aflatoxin B_1-2,3-oxide as a probable intermediate in the covalent binding of aflatoxins B_1 and B_2 to rat liver DNA and ribosomal RNA *in vivo*. *Cancer Res. 37*, 172-181.

90. Shimada, T., Iwasaki, M., Martin, M. V. and Guengerich, F. P. (1989) Human liver microsomal cytochrome P-450 enzymes involved in the bioactivation of procarcinogens detected by *umu* gene response in *Salmonella typhimurium* TA 1535/pSK1002. *Cancer Res. 49*, 3218-3228.

91. Yu, F. L., Huang, J. X., Bender, W., Wu, Z. R. and Chang, J. C. (1991) Evidence for the covalent binding of aflatoxin B_1-dichloride to cytosine in DNA. *Carcinogenesis 12*, 997-1002.

92. Iyer, R. S., Coles, B., Raney, K., Thier, R., Guengerich, F. P. and Harris, T. M. (1994) DNA adduction by the potent carcinogen aflatoxin B_1: mechanistic studies. *J. Am. Chem. Soc. 116*, 1603-1609.

93. Essigmann, J. M., Croy, R. G., Nadzan, A. M., Busby, W. F., Jr., Reinhold, V. N., Buchi, G. and Wogan, G. N. (1977) Structural identification of the major DNA adduct formed by aflatoxin B_1 *in vitro*. *Proc. Natl. Acad. Sci. USA 74*, 1870-1874.

94. Croy, R. G., Essigmann, J. M., Reinhold, V. N. and Wogan, G. N. (1978) Identification of the principal aflatoxin B_1-DNA adduct formed *in vivo* in rat liver. *Proc. Natl. Acad. Sci. USA 75*, 1745-1749.

95. Autrup, H., Essigmann, J. M., Croy, R. G., Trump, B. F., Wogan, G. N. and Harris, C. C. (1979) Metabolism of aflatoxin B_1 and identification of the major aflatoxin B_1-DNA adducts formed in cultured human bronchus and colon. *Cancer Res. 39*, 694-698.

96. Johnson, W. W. and Guengerich, F. P. (1997) Reaction of aflatoxin B_1 *exo*-8,9-epoxide with DNA: kinetic analysis of covalent binding and DNA-induced hydrolysis. *Proc. Natl. Acad. Sci. USA 94*, 6121-6125.

97. Raney, K. D., Coles, B., Guengerich, F. P. and Harris, T. M. (1992) The *endo*-8,9-epoxide of aflatoxin B_1: a new metabolite. *Chem. Res. Toxicol. 5*, 333-335.

98. Gopalakrishnan, S., Harris, T. M. and Stone, M. P. (1990) Intercalation of aflatoxin B_1 in two oligodeoxynucleotide adducts: comparative ^{1}H NMR analysis of d(ATCAFBGAT).d(ATCGAT) and d(ATAFBGCAT)2. *Biochemistry 29*, 10438-10448.

99. Raney, K. D., Gopalakrishnan, S., Byrd, S., Stone, M. P. and Harris, T. M. (1990) Alteration of the aflatoxin cyclopentenone ring to a δ-lactone reduces intercalation with DNA and decreases formation of guanine N7 adducts by aflatoxin epoxides. *Chem. Res. Toxicol. 3*, 254-261.

100. Johnson, W. W., Harris, T. M. and Guengerich, F. P. (1996) Kinetics and mechanism of hydrolysis of aflatoxin B_1 *exo*-8,9-oxide and rearrangement of the dihydrodiol. *J. Am. Chem. Soc. 118*, 8213-82220.

101. Guengerich, F. P., Voehler, M., Williams, K. M., Deng, Z. and Harris, T. M. (2002) Structure of the aflatoxin B_1 dialdehyde adduct formed from reaction with methylamine. *Chem. Res. Toxicol. 15*, 793-798.

102. Guengerich, F. P., Arneson, K. O., Williams, K. M., Deng, Z. and Harris, T. M. (2002) Reaction of aflatoxin B_1 oxidation products with lysine. *Chem. Res. Toxicol. 15*, 780-792.

103. Sabbioni, G., Skipper, P. L., Buchi, G. and Tannenbaum, S. R. (1987) Isolation and characterization of the major serum albumin adduct formed by aflatoxin B_1 *in vivo* in rats. *Carcinogenesis 8*.

104. Iyer, R. S. and Harris, T. M. (1993) Preparation of aflatoxin B_1 8,9-epoxide using m-chloroperbenzoic acid. *Chem. Res. Toxicol. 6*, 313-316.

105. Baertschi, S. W., Kevin D. Raney, Michael P. Stone, and Thomas M. Harris. (1988) Preparation of the 8,9-epoxide of the mycotoxin aflatoxin B_1: The ultimate carcinogenic species. *J. Am. Chem. Soc. 110*, 7929-7931.

106. Raney, V. M., Harris, T. M. and Stone, M. P. (1993) DNA conformation mediates aflatoxin B_1-DNA binding and the formation of guanine N7 adducts by aflatoxin B_1 8,9-*exo*-epoxide. *Chem. Res. Toxicol. 6*, 64-68.

107. Kobertz, W. R., Wang, D., Wogan, G. N. and Essigmann, J. M. (1997) An intercalation inhibitor altering the target selectivity of DNA damaging agents: synthesis of site-specific aflatoxin B_1 adducts in a p53 mutational hotspot. *Proc. Natl. Acad. Sci. USA 94*, 9579-9584.

108. Baertschi, S. W., Raney, K. D., Shimada, T., Harris, T. M. and Guengerich, F. P. (1989) Comparison of rates of enzymatic oxidation of aflatoxin B_1, aflatoxin G_1, and sterigmatocystin and activities of the epoxides in forming guanyl-N7 adducts and inducing different genetic responses. *Chem. Res. Toxicol. 2*, 114-112.

109. Asao, T., Buechi, G., Abdel-Kader, M. M., Chang, S. B., Wick, E. L. and Wogan, G. N. (1965) The structures of aflatoxins B and G. *J. Am. Chem. Soc. 87*, 882-886.

110. Deng, Z. (1998) DNA adducts of aflatoxin B_1, Masters In *Chemistry*, Vanderbilt University, Nashville, TN.

111. Ojha, R. P., Roychoudhury, M. and Sanyal, N. K. (1990) Mode of action of intercalators: a theoretical study. *Indian J. Biochem. Biophys. 27*, 228-239.

112. Lamm, G. and Pack, G. R. (1990) Acidic domains around nucleic acids. *Proc. Natl. Acad. Sci. USA 87*, 9030-9033.

113. Lamm, G., Wong, G. R. and Pack, G. R. (1996) DNA-mediated acid catalysis: calculations of the rates of DNA-catalyzed hydrolysis of diol epoxides. *J. Am. Chem. Soc. 118*, 3325-3331.

114. Gopalakrishnan, S., Byrd, S., Stone, M. P. and Harris, T. M. (1989) Carcinogen-nucleic acid interactions: equilibrium binding studies of aflatoxin B_1 with the oligodeoxynucleotide d(ATGCAT)2 and with plasmid pBR322 support intercalative association with the B-DNA helix. *Biochemistry 28*, 726-734.

115. Mao, H., Deng, Z., Wang, F., Harris, T. M. and Stone, M. P. (1998) An intercalated and thermally stable FAPY adduct of aflatoxin B_1 in a DNA duplex: structural refinement from ^{1}H NMR. *Biochemistry 37*, 4374-4387.

116. Giri, I. and Stone, M. P. (2002) Thermal stabilization of the DNA duplex by adducts of aflatoxin B_1. *Biopolymers 65*, 190-201.

117. Croy, R. G. and Wogan, G. N. (1981) Temporal patterns of covalent DNA adducts in rat liver after single and multiple doses of aflatoxin B_1. *Cancer Res. 41*, 197-203.

118. Muench, K. F., Misra, R. P. and Humayun, M. Z. (1983) Sequence specificity in aflatoxin B_1--DNA interactions. *Proc. Natl. Acad. Sci. USA 80*, 6-10.

119. Benasutti, M., Ejadi, S., Whitlow, M. D. and Loechler, E. L. (1988) Mapping the binding site of aflatoxin B_1 in DNA: systematic analysis of the reactivity of aflatoxin B_1 with guanines in different DNA sequences. *Biochemistry 27*, 472-481.

120. Levy, D. D., Groopman, J. D., Lim, S. E., Seidman, M. M. and Kraemer, K. H. (1992) Sequence specificity of aflatoxin B_1-induced mutations in a plasmid replicated in xeroderma pigmentosum and DNA repair proficient human cells. *Cancer Res. 52*, 5668-5673.

121. Denissenko, M. F., Cahill, J., Koudriakova, T. B., Gerber, N. and Pfeifer, G. P. (1999) Quantitation and mapping of aflatoxin B_1-induced DNA damage in genomic DNA using aflatoxin B_1-8,9-epoxide and microsomal activation systems. *Mutat. Res. 425*, 205-211.

122. Misra, R. P., Muench, K. F. and Humayun, M. Z. (1983) Covalent and noncovalent interactions of aflatoxin with defined deoxyribonucleic acid sequences. *Biochemistry 22*, 3351-3359.

123. Refolo, L. M., Conley, M. P., Sambamurti, K., Jacobsen, J. S. and Humayun, M. Z. (1985) Sequence context effects in DNA replication blocks induced by aflatoxin B_1. *Proc. Natl. Acad. Sci. USA 82*, 3096-3100.

124. Puisieux, A., Lim, S., Groopman, J. and Ozturk, M. (1991) Selective targeting of p53 gene mutational hotspots in human cancers by etiologically defined carcinogens. *Cancer Res. 51*, 6185-6189.

125. Yu, F. L. (1983) Preferential binding of aflatoxin B_1 to the transcriptionally active regions of rat liver nucleolar chromatin *in vivo* and *in vitro*. *Carcinogenesis 4*, 889-893.

126. Marien, K., Moyer, R., Loveland, P., Van Holde, K. and Bailey, G. (1987) Comparative binding and sequence interaction specificities of aflatoxin B_1, aflatoxicol, aflatoxin M_1, and aflatoxicol M_1 with purified DNA. *J. Biol. Chem. 262*, 7455-7462.

127. Yu, F. L., Bender, W. and Geronimo, I. H. (1990) Base and sequence specificities of aflatoxin B_1 binding to single- and double-stranded DNAs. *Carcinogenesis 11*, 475-478.

128. Denissenko, M. F., Koudriakova, T. B., Smith, L., O'Connor, T. R., Riggs, A. D. and Pfeifer, G. P. (1998) The p53 codon 249 mutational hotspot in hepatocellular

carcinoma is not related to selective formation or persistence of aflatoxin B_1 adducts. *Oncogene 17*, 3007-3014.

129. Trottier, Y., Waithe, W. I. and Anderson, A. (1992) Kinds of mutations induced by aflatoxin B_1 in a shuttle vector replicating in human cells transiently expressing cytochrome P4501A2 cDNA. *Mol. Carcinog. 6*, 140-147.

130. Preito-Alamo, M. J., Jurado, J., Abril, N., Diaz-Pohl, C., Bolcsfoldi, G. and Pueyo, C. (1996) Mutational specificity of aflatoxin B_1. Comparison of *in vivo* host-mediated assay with *in vitro* S9 metabolic activation. *Carcinogenesis 17*, 1997-2002.

131. Sambamurti, K., Callahan, J., Luo, X., Perkins, C. P., Jacobsen, J. S. and Humayun, M. Z. (1988) Mechanisms of mutagenesis by a bulky DNA lesion at the guanine N7 position. *Genetics 120*, 863-873.

132. Parken, D. M., Bray, F., Ferlay, P. and Pisani, P. (2001) Estimating the world cancer burden: globocan 2000. *J. Cancer 94*, 153-156.

133. Ming, L., Thorgeirsson, S. S., Gail, M. H., Lu, P., Harris, C. C., Wang, N., Shao, Y., Wu, Z., Liu, G., Wang, X. and Sun, Z. (2002) Dominant role of hepatitis B virus and cofactor role of aflatoxin in hepatocarcinogenesis in Qidong, China. *Hepatology 36*, 1214-1220.

134. Gorina, S. and Pavletich, N. P. (1996) Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2. *Science 274*, 1001-1005.

135. Montalto, G., Cervello, M., Giannitrapani, L., Dantona, F., Terranova, A. and Castagnetta, L. A. (2002) Epidemiology, risk factors, and natural history of hepatocellular carcinoma. *Ann. N. Y. Acad. Sci. 963*, 13-20.

136. Henry, S. H., Bosch, F. X. and Bowers, J. C. (2002) Aflatoxin, hepatitis and worldwide liver cancer risks. *Adv. Exp. Med. Biol. 504*, 229-233.

137. Dominguez-Malagon, H. and Gaytan-Graham, S. (2001) Hepatocellular carcinoma: an update. *Ultrastruct. Pathol. 25*, 497-516.

138. Wogan, G. N. (2000) Impacts of chemicals on liver cancer risk. *Semin. Cancer Biol. 10*, 201-210.

139. Wogan, G. N. (1992) Aflatoxins as risk factors for hepatocellular carcinoma in humans. *Cancer Res. 52*, 2114s-2118s.

140. Shen, H. M. and Ong, C. N. (1996) Mutations of the p53 tumor suppressor gene and *ras* oncogenes in aflatoxin hepatocarcinogenesis. *Mutat. Res. 366*, 23-44.

141. Ross, R. K., Yuan, J. M., Yu, M. C., Wogan, G. N., Qian, G. S., Tu, J. T., Groopman, J. D., Gao, Y. T. and Henderson, B. E. (1992) Urinary aflatoxin biomarkers and risk of hepatocellular carcinoma. *Lancet. 339*, 943-946.

142. Bannasch, P., Khoshkhou, N. I., Hacker, H. J., Radaeva, S., Mrozek, M., Zillmann, U., Kopp-Schneider, A., Haberkorn, U., Elgas, M., Tolle, T. and et al. (1995) Synergistic hepatocarcinogenic effect of hepadnaviral infection and dietary aflatoxin B_1 in woodchucks. *Cancer Res. 55*, 3318-3330.

143. Chen, C. J., Wang, L. Y., Lu, S. N., Wu, M. H., You, S. L., Zhang, Y. J., Wang, L. W. and Santella, R. M. (1996) Elevated aflatoxin exposure and increased risk of hepatocellular carcinoma. *Hepatology 24*, 38-42.

144. Wang, X. W., Forrester, K., Yeh, H., Feitelson, M. A., Gu, J. R. and Harris, C. C. (1994) Hepatitis B virus X protein inhibits p53 sequence-specific DNA binding, transcriptional activity, and association with transcription factor ERCC3. *Proc. Natl. Acad. Sci. USA 91*, 2230-2234.

145. Truant, R., Antunovic, J., Greenblatt, J., Prives, C. and Cromlish, J. A. (1995) Direct interaction of the hepatitis B virus HBx protein with p53 leads to inhibition by HBx of p53 response element-directed transactivation. *J. Virol. 69*, 1851-1859.

146. Buchhop, S., Gibson, M. K., Wang, X. W., Wagner, P., Sturzbecher, H. W. and Harris, C. C. (1997) Interaction of p53 with the human Rad51 protein. *Nucleic Acids Res. 25*, 3868-3874.

147. Schaefer, S., Seifer, M., Grimmsmann, T., Fink, L., Wenderhold, S., Hohne, M. W. and Gerlich, W. H. (1998) Properties of tumour suppressor p53 in murine hepatocyte lines transformed by hepatitis B virus X protein. *J. Gen. Virol. 79 (Pt 4)*, 767-777.

148. Terradillos, O., Pollicino, T., Lecoeur, H., Tripodi, M., Gougeon, M. L., Tiollais, P. and Buendia, M. A. (1998) p53-independent apoptotic effects of the hepatitis B virus HBx protein *in vivo* and *in vitro*. *Oncogene 17*, 2115-2123.

149. Lee, S. G. and Rho, H. M. (2000) Transcriptional repression of the human p53 gene by hepatitis B viral X protein. *Oncogene 19*, 468-471.

150. Su, Q., Schroder, C. H., Otto, G. and Bannasch, P. (2000) Overexpression of p53 protein is not directly related to hepatitis B x protein expression and is associated with neoplastic progression in hepatocellular carcinomas rather than hepatic preneoplasia. *Mutat. Res. 462*, 365-380.

151. Wild, C. P., Yin, F., Turner, P. C., Chemin, I., Chapot, B., Mendy, M., Whittle, H., Kirk, G. D. and Hall, A. J. (2000) Environmental and genetic determinants of aflatoxin-albumin adducts in the Gambia. *Int. J. Cancer 86*, 1-7.

152. Bos, J. L., Fearon, E. R., Hamilton, S. R., Verlaan-de Vries, M., van Boom, J. H., van der Eb, A. J. and Vogelstein, B. (1987) Prevalence of *ras* gene mutations in human colorectal cancers. *Nature 327*, 293-297.

153. Vogelstein, B., Fearon, E. R., Hamilton, S. R., Kern, S. E., Preisinger, A. C., Leppert, M., Nakamura, Y., White, R., Smits, A. M. and Bos, J. L. (1988) Genetic alterations during colorectal-tumor development. *N. Engl. J. Med. 319*, 525-532.

154. McMahon, G., Hanson, L., Lee, J. J. and Wogan, G. N. (1986) Identification of an activated *c-Ki-ras* oncogene in rat liver tumores induced by aflatoxin B_1. *Proc. Natl. Acad. Sci. USA 83*, 9418-9422.

155. McMahon, G., Davis, E. F., Huber, L. J., Kim, Y. and Wogan, G. N. (1990) Characterization of *c-Ki-ras* and *N-ras* oncogenes in aflatoxin B_1-induced rat liver tumors. *Proc. Natl. Acad. Sci. USA 87*, 1104-1108.

156. Soman, N. R. and Wogan, G. N. (1993) Activation of the *c-Ki-ras* oncogene in aflatoxin B_1-induced hepatocellular carcinoma and adenoma in the rat: detection by denaturing gradient gel electrophoresis. *Proc. Natl. Acad. Sci. USA 90*, 2045-2049.

157. Chang, Y. J., Mathews, C., Mangold, K., Marien, K., Hendricks, J. and Bailey, G. (1991) Analysis of *ras* gene mutations in rainbow trout liver tumors initiated by aflatoxin B_1. *Mol. Carcinog. 4*, 112-119.

158. Riley, J., Mandel, H. G., Sinha, S., Judah, D. J. and Neal, G. E. (1997) *In vitro* activation of the human Harvey-*ras* proto-oncogene by aflatoxin B_1. *Carcinogenesis 18*, 905-910.

159. Bressac, B., Kew, M., Wands, J. and Ozturk, M. (1991) Selective G to T mutations of p53 gene in hepatocellular carcinoma from southern Africa. *Nature 350*, 429-431.

160. Hsu, I. C., Metcalf, R. A., Sun, T., Welsh, J. A., Wang, N. J. and Harris, C. C. (1991) Mutational hotspot in the p53 gene in human hepatocellular carcinomas. *Nature 350*, 427-428.

161. Gentil, A., Margot, A. and Sarasin, A. (1984) Apurinic sites cause mutations in simian virus 40. *Mutat. Res. 129*, 141-147.

162. Bailey, E. A., Iyer, R. S., Stone, M. P., Harris, T. M. and Essigmann, J. M. (1996) Mutational properties of the primary aflatoxin B_1-DNA adduct. *Proc. Natl. Acad. Sci. USA 93*, 1535-1539.

163. Loeb, L. A. and Preston, B. D. (1986) Mutagenesis by apurinic/apyrimidinic sites. *Annu. Rev. Genet. 20*, 201-230.

164. Gentil, A., Renault, G., Madzak, C., Margot, A., Cabral-Neto, J. B., Vasseur, J. J., Rayner, B., Imbach, J. L. and Sarasin, A. (1990) Mutagenic properties of a unique abasic site in mammalian cells. *Biochem. Biophys. Res. Commun. 173*, 704-710.

165. Gentil, A., Cabral-Neto, J. B., Mariage-Samson, R., Margot, A., Imbach, J. L., Rayner, B. and Sarasin, A. (1992) Mutagenicity of a unique apurinic/apyrimidinic site in mammalian cells. *J. Mol. Biol. 227*, 981-984.

166. Kamiya, H., Suzuki, M., Komatsu, Y., Miura, H., Kikuchi, K., Sakaguchi, T., Murata, N., Masutani, C., Hanaoka, F. and Ohtsuka, E. (1992) An abasic site analogue activates a *c-Ha-ras* gene by a point mutation at modified and adjacent positions. *Nucleic Acids Res. 20*, 4409-4415.

167. Lawrence, C. W., Borden, A., Banerjee, S. K. and LeClerc, J. E. (1990) Mutation frequency and spectrum resulting from a single abasic site in a single-stranded vector. *Nucleic Acids Res. 18*, 2153-2157.

168. Foster, P. L., Eisenstadt, E. and Miller, J. H. (1983) Base substitution mutations induced by metabolically activated aflatoxin B_1. *Proc. Natl. Acad. Sci. USA 80*, 2695-2698.

169. Smela, M. E., Hamm, M. L., Henderson, P. T., Harris, C. M., Harris, T. M. and Essigmann, J. M. (2002) The aflatoxin B_1 formamidopyrimidine adduct plays a major role in causing the types of mutations observed in human hepatocellular carcinoma. *Proc. Natl. Acad. Sci. USA 99*, 6655-6660.

170. Johnston, D. S. and Stone, M. P. (1995) Refined solution structure of 8,9-dihydro-8-(N7-guanyl)-9-hydroxyaflatoxin B_1 opposite CpA in the complementary strand of an oligodeoxynucleotide duplex as determined by ^{1}H NMR. *Biochemistry 34*, 14037-14050.

171. Jones, W. R., Johnston, D. S. and Stone, M. P. (1998) Refined structure of the doubly intercalated d(TATAFBGCATA)2 aflatoxin B_1 adduct. *Chem. Res. Toxicol. 11*, 873-881.

172. Giri, I., Jenkins, M. D., Schnetz-Boutaud, N. C. and Stone, M. P. (2002) Structural refinement of the 8,9-dihydro-8-(N7-guanyl)-9-hydroxy-aflatoxin B_1 adduct in a 5'-Cp(AFB)G-3' sequence. *Chem. Res. Toxicol. 15*, 638-647.

173. Giri, I., Johnston, D. S. and Stone, M. P. (2002) Mispairing of the 8,9-dihydro-8-(N7-guanyl)-9-hydroxy-aflatoxin B_1 adduct with deoxyadenosine results in extrusion of the mismatched dA toward the major groove. *Biochemistry 41*, 5462-5472.

174. Courtemanche, C. and Anderson, A. (1999) Multiple mutations in a shuttle vector modified by ultraviolet irradiation, (+/-)-7 β, 8 α-dihydroxy-9 α, 10 α-epoxy 7,8,9,10-tetrahydrobenzo[a]pyrine, and aflatoxin B_1 have different properties than single mutations and may be generated during translesion synthesis. *Mutat. Res. 430*, 23-36.

175. Lin, J. K., Miller, J. A. and Miller, E. C. (1977) 2,3-Dihydro-2-(guan-7-yl)-3-hydroxy-aflatoxin B_1, a major acid hydrolysis product of aflatoxin B_1-DNA or -ribosomal RNA adducts formed in hepatic microsome-mediated reactions and in rat liver *in vivo*. *Cancer Res. 37*, 4430-4438.

176. Hertzog, P. J., Smith, J. R. and Garner, R. C. (1982) Characterisation of the imidazole ring-opened forms of *trans*-8,9-dihydro-8,9-dihydro-8-(7-guanyl)9-hydroxy aflatoxin B_1. *Carcinogenesis 3*, 723-725.

177. Hertzog, P. J., Lindsay Smith, J. R. and Garner, R. C. (1980) A high pressure liquid chromatography study on the removal of DNA-bound aflatoxin B_1 in rat liver and *in vitro*. *Carcinogenesis 1*, 787-793.

178. Groopman, J. D., Croy, R. G. and Wogan, G. N. (1981) *In vitro* reactions of aflatoxin B_1-adducted DNA. *Proc. Natl. Acad. Sci. USA 78*, 5445-5449.

179. Essigmann, J. M., Croy, R. G., Bennett, R. A. and Wogan, G. N. (1982) Metabolic activation of aflatoxin B_1: patterns of DNA adduct formation, removal, and excretion in relation to carcinogenesis. *Drug Metab. Rev. 13*, 581-602.

180. Kaden, D. A., Call, K. M., Leong, P. M., Komives, E. A. and Thilly, W. G. (1987) Killing and mutation of human lymphoblast cells by aflatoxin B_1: evidence for an inducible repair response. *Cancer Res. 47*, 1993-2001.

181. Leadon, S. A. and Snowden, M. M. (1988) Differential repair of DNA damage in the human metallothionein gene family. *Mol. Cell Biol. 8*, 5331-5338.

182. Ball, R. W., Wilson, D. M. and Coulombe, R. A., Jr. (1990) Comparative formation and removal of aflatoxin B_1-DNA adducts in cultured mammalian tracheal epithelium. *Cancer Res. 50*, 4918-4922.

183. Alekseyev, Y. O., Hamm, M. L. and Essigmann, J. M. (2004) Aflatoxin B_1 formamidopyrimidine adducts are preferentially repaired by the nucleotide excision repair pathway *in vivo*. *Carcinogenesis 25*, 1045-1051.

184. Garner, R. C. and Wright, C. M. (1973) Induction of mutations in DNA-repair deficient bacteria by a liver microsomal metabolite of aflatoxin B_1. *Br. J. Cancer 28*, 544-551.

185. Sarasin, A. R., Smith, C. A. and Hanawalt, P. C. (1977) Repair of DNA in human cells after treatment with activated aflatoxin B_1. *Cancer Res. 37*, 1786-1793.

186. Leadon, S. A., Tyrrell, R. M. and Cerutti, P. A. (1981) Excision repair of aflatoxin B_1-DNA adducts in human fibroblasts. *Cancer Res. 41*, 5125-5129.

187. Takahashi, Y., Nakatsuru, Y., Zhang, S., Shimizu, Y., Kume, H., Tanaka, K., Ide, F. and Ishikawa, T. (2002) Enhanced spontaneous and aflatoxin-induced liver tumorigenesis in xeroderma pigmentosum group A gene-deficient mice. *Carcinogenesis 23*, 627-633.

188. Friedberg, E. C., Walker, G. C. and Siede, W. (1995) *DNA Repair and Mutagenesis*, American Chemical Society for Microbiology, Washington, DC.

189. Oleykowski, C. A., Mayernik, J. A., Lim, S. E., Groopman, J. D., Grossman, L., Wogan, G. N. and Yeung, A. T. (1993) Repair of aflatoxin B_1 DNA adducts by the UvrABC endonuclease of *Escherichia coli*. *J. Biol. Chem. 268*, 7990-8002.

190. Chetsanga, C. J. and Frenette, G. P. (1983) Excision of aflatoxin B_1-imidazole ring opened guanine adducts from DNA by formamidopyrimidine-DNA glycosylase. *Carcinogenesis 4*, 997-1000.

191. Tchou, J., Bodepudi, V., Shibutani, S., Antoshechkin, I., Miller, J., Grollman, A. P. and Johnson, F. (1994) Substrate specificity of Fpg protein. Recognition and cleavage of oxidatively damaged DNA. *J. Biol. Chem. 269*, 15318-15324.

192. Croy, R. G. and Wogan, G. N. (1981) Quantitative comparison of covalent aflatoxin-DNA adducts formed in rat and mouse livers and kidneys. *J. Natl. Cancer. Inst. 66*, 761-768.

193. Hutchinson, F. (1985) Chemical changes induced in DNA by ionizing radiation. *Prog. Nucleic Acid Res. Mol. Biol. 32*, 115-154.

194. Ames, B. N. (1983) Dietary carcinogens and anticarcinogens. Oxygen radicals and degenerative diseases. *Science 221*, 1256-1264.

195. Frenkel, K., Chrzan, K., Troll, W., Teebor, G. W. and Steinberg, J. J. (1986) Radiation-like modification of bases in DNA exposed to tumor promoter-activated polymorphonuclear leukocytes. *Cancer Res. 46*, 5533-5540.

196. Beer, M., Stern, S., Carmalt, D. and Mohlhenrich, K. H. (1966) Determination of base sequence in nucleic acids with the electron microscope. V. The thymine-specific reactions of osmium tetroxide with deoxyribonucleic acid and its components. *Biochemistry 5*, 2283-2288.

197. Teoule, R., Bonicel, A., Bert, C., Cadet, J. and Polverelli, M. (1974) Identification of radioproducts resulting from the breakage of thymine moiety by γ irradiation of *E. coli* DNA in an aerated aqueous solution. *Radiat. Res. 57*, 46-58.

198. Cathcart, R., Schwiers, E., Saul, R. L. and Ames, B. N. (1984) Thymine glycol and thymidine glycol in human and rat urine: a possible assay for oxidative DNA damage. *Proc. Natl. Acad. Sci. USA 81*, 5633-5637.

199. Frenkel, K., Goldstein, M. S., Duker, N. J. and Teebor, G. W. (1981) Identification of the *cis*-thymine glycol moiety in oxidized deoxyribonucleic acid. *Biochemistry 20*, 750-754.

200. Frenkel, K., Goldstein, M. S. and Teebor, G. W. (1981) Identification of the *cis*-thymine glycol moiety in chemically oxidized and γ-irradiated deoxyribonucleic acid by high-pressure liquid chromatography analysis. *Biochemistry 20*, 7566-7571.

201. Vaishnav, Y., Holwitt, E., Swenberg, C., Lee, H. C. and Kan, L. S. (1991) Synthesis and characterization of stereoisomers of 5,6-dihydro-5,6-dihydroxy-thymidine. *J. Biomol. Struct. Dyn. 8*, 935-951.

202. Wang, Y. (2002) HPLC isolation and mass spectrometric characterization of two isomers of thymine glycols in oligodeoxynucleotides. *Chem. Res. Toxicol. 15*, 671-676.

203. Fujita, S. and Steenken, S. (1981) Pattern of hydroxyl radical addition to uracil and methyl- and carboxyl-substituted uracils. Electron transfer of hydroxyl adducts with N,N,N',N'-tetramethyl-p-phenylenediamine and tetranitromethane. *J. Am. Chem. Soc. 103*, 2540-2545.

204. Schulte-Frohlinde, D. and von Sonntag, C. (1985) Oxidative Stress, (Sies, H., Ed.) pp 11-40, Academic Press, London.

205. Leadon, S. A. (1987) Production of thymine glycols in DNA by radiation and chemical carcinogens as detected by a monoclonal antibody. *Br. J. Cancer Suppl. 8*, 113-117.

206. Teebor, G., Cummings, A., Frenkel, K., Shaw, A., Voituriez, L. and Cadet, J. (1987) Quantitative measurement of the diastereoisomers of *cis* thymidine glycol in γ-irradiated DNA. *Free Radic. Res. Commun. 2*, 303-309.

207. Flippen, J. L. (1973) The crystal and molecular structures of reaction products from -irradiation of thymine and cytosine: *cis*-thymine glycol, $C_5H_8N_2O_4$, and *trans*-1-carbamoylimidazolidone-4,5-diol, $C_4H_7N_3O_4$. *Acta Crystallogr. Sect. B 29*, 1756-1762.

208. Adelman, R., Saul, R. L. and Ames, B. N. (1988) Oxidative damage to DNA: relation to species metabolic rate and life span. *Proc. Natl. Acad. Sci. USA 85*, 2706-2708.

209. Saul, R. L. and Ames, B. N. (1986) *Mechanisms of DNA Damage and Repair*, Plenum Press, New York.

210. Ames, B. N. (1984) The detection of environmental mutagens and potential carcinogens. *Cancer 53*, 2034-2040.

211. Ide, H., Kow, Y. W. and Wallace, S. S. (1985) Thymine glycols and urea residues in M13 DNA constitute replicative blocks *in vitro*. *Nucleic Acids Res. 13*, 8035-8052.

212. Jensen, W. E., Jones, A. S. and Ross, G. W. (1965) The synthesis and properties of 2-deoxy-D-ribosylureas. *J. Chem. Soc. Part II*, 2463-2465.

213. Rouet, P. and Essigmann, J. M. (1985) Possible role for thymine glycol in the selective inhibition of DNA synthesis on oxidized DNA templates. *Cancer Res. 45*, 6113-6118.

214. Clark, J. M. and Beardsley, G. P. (1986) Thymine glycol lesions terminate chain elongation by DNA polymerase I *in vitro*. *Nucleic Acids Res. 14*, 737-749.

215. Hayes, R. C. and LeClerc, J. E. (1986) Sequence dependence for bypass of thymine glycols in DNA by DNA polymerase I. *Nucleic Acids Res. 14*, 1045-1061.

216. Hayes, R. C., Petrullo, L. A., Huang, H. M., Wallace, S. S. and LeClerc, J. E. (1988) Oxidative damage in DNA. Lack of mutagenicity by thymine glycol lesions. *J. Mol. Biol. 201*, 239-246.

217. Clark, J. M. and Beardsley, G. P. (1989) Template length, sequence context, and 3'-5' exonuclease activity modulate replicative bypass of thymine glycol lesions *in vitro*. *Biochemistry 28*, 775-779.

218. Lida, S. and Hayatsu, H. (1970) The permanganate oxidation of thymine. *Biochim. Biophys. Acta. 213*, 1-13.

219. Basu, A. K., Loechler, E. L., Leadon, S. A. and Essigmann, J. M. (1989) Genetic effects of thymine glycol: site-specific mutagenesis and molecular modeling studies. *Proc. Natl. Acad. Sci. USA 86*, 7677-7681.

220. Kusumoto, R., Masutani, C., Iwai, S. and Hanaoka, F. (2002) Translesion synthesis by human DNA polymerase η across thymine glycol lesions. *Biochemistry 41*, 6090-6099.

221. Fischhaber, P. L., Gerlach, V. L., Feaver, W. J., Hatahet, Z., Wallace, S. S. and Friedberg, E. C. (2002) Human DNA polymerase κ bypasses and extends beyond thymine glycols during translesion synthesis *in vitro*, preferentially incorporating correct nucleotides. *J. Biol. Chem. 277*, 37604-37611.

222. Zuo, S., Boorstein, R. J. and Teebor, G. W. (1995) Oxidative damage to 5-methylcytosine in DNA. *Nucleic Acids Res. 23*, 3239-3243.

223. Pfeifer, G. P. (2000) p53 mutational spectra and the role of methylated CpG sequences. *Mutat. Res. 450*, 155-166.

224. Ocampo-Hafalla, M. T., Altamirano, A., Basu, A. K., Chan, M. K., Ocampo, J. E., Cummings, A., Jr., Boorstein, R. J., Cunningham, R. P. and Teebor, G. W. (2006) Repair of thymine glycol by hNth1 and hNeil1 is modulated by base pairing and *cis-trans* epimerization. *DNA Repair (Amst) 5*, 444-454.

225. Ikeda, S., Biswas, T., Roy, R., Izumi, T., Boldogh, I., Kurosky, A., Sarker, A. H., Seki, S. and Mitra, S. (1998) Purification and characterization of human NTH1, a homolog of *Escherichia coli* endonuclease III. Direct identification of Lys-212 as the active nucleophilic residue. *J. Biol. Chem. 273*, 21585-21593.

226. Aspinwall, R., Rothwell, D. G., Roldan-Arjona, T., Anselmino, C., Ward, C. J., Cheadle, J. P., Sampson, J. R., Lindahl, T., Harris, P. C. and Hickson, I. D. (1997) Cloning and characterization of a functional human homolog of *Escherichia coli* endonuclease III. *Proc. Natl. Acad. Sci. USA 94*, 109-114.

227. Marenstein, D. R., Chan, M. K., Altamirano, A., Basu, A. K., Boorstein, R. J., Cunningham, R. P. and Teebor, G. W. (2003) Substrate specificity of human endonuclease III (hNTH1). Effect of human APE1 on hNTH1 activity. *J. Biol. Chem. 278*, 9005-9012.

228. Katafuchi, A., Nakano, T., Masaoka, A., Terato, H., Iwai, S., Hanaoka, F. and Ide, H. (2004) Differential specificity of human and *Escherichia coli* endonuclease III and VIII homologues for oxidative base lesions. *J. Biol. Chem. 279*, 14464-14471.

229. Hazra, T. K., Izumi, T., Boldogh, I., Imhoff, B., Kow, Y. W., Jaruga, P., Dizdaroglu, M. and Mitra, S. (2002) Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. *Proc. Natl. Acad. Sci. USA 99*, 3523-3528.

230. Miller, H., Fernandes, A. S., Zaika, E., McTigue, M. M., Torres, M. C., Wente, M., Iden, C. R. and Grollman, A. P. (2004) Stereoselective excision of thymine glycol from oxidatively damaged DNA. *Nucleic Acids Res. 32*, 338-345.

231. Clark, J. M., Pattabiraman, N., Jarvis, W. and Beardsley, G. P. (1987) Modeling and molecular mechanical studies of the *cis*-thymine glycol radiation damage lesion in DNA. *Biochemistry 26*, 5404-5409.

232. Kao, J. Y., Goljer, I., Phan, T. A. and Bolton, P. H. (1993) Characterization of the effects of a thymine glycol residue on the structure, dynamics, and stability of duplex DNA by NMR. *J. Biol. Chem. 268*, 17787-17793.

233. Kung, H. C. and Bolton, P. H. (1997) Structure of a duplex DNA containing a thymine glycol residue in solution. *J. Biol. Chem. 272*, 9227-9236.

234. Aller, P., Rould, M. A., Hogg, M., Wallace, S. S. and Doublie, S. (2007) A structural rationale for stalling of a replicative DNA polymerase at the most common oxidative thymine lesion, thymine glycol. *Proc. Natl. Acad. Sci. USA 104*, 814-818.

235. Robert W. Murray, M. S. (2002) Synthesis of epoxides using dimethyl dioxirane, www.orgsyn.org.

236. Adam, W., Chan, Y. Y., Cremer, D., Gauss, J., Scheutzow, D. and Schindler, M. (1987) Spectral and chemical properties of dimethyldioxirane as determined by experiment and *ab initio* calculations. *J. Org. Chem. 52*, 2800-2803.

237. Jones, W. R. and Stone, M. P. (1998) Site-specific targeting of aflatoxin adduction directed by triple helix formation in the major groove of oligodeoxyribonucleotides. *Nucleic Acids Res. 26*, 1070-1075.

238. Koshlap, K. M., Gillespie, P., Dervan, P. B. and Feigon, J. (1993) Nonnatural deoxyribonucleoside D3 incorporated in an intramolecular DNA triplex binds sequence-specifically by intercalation. *J. Am. Chem. Soc. 115*, 7908-7909.

239. Radhakrishnan, I., Patel, D. J., Priestly, E. S., Nash, H. M. and Dervan, P. B. (1993) NMR structural studies on a nonnatural deoxyribonucleoside which mediates recognition of GC base pairs in pyrimidine-purine-pyrimidine DNA triplexes. *Biochemistry 32*, 11228-11234.

240. Veselkov, A. N., Baranovskii, S. F., Dymant, L. N., Petrenko, N. V., Osetrov, S. G., Veselkov, D. A., Parkes, X. and Davies, D. (1998) ^{1}H-NMR study of complex formation between the aromatic dye ethidium bromide and the single-stranded noncomplementary deoxytetranucleotide 5'-d(CpTpGpA) in aqueous solution. *Biofizika 43*, 205-214.

241. Jones, W. R., Johnston, D. S. and Stone, M. P. (1999) Site-specific synthesis of aflatoxin B_1 adducts within an oligodeoxyribonucleotide containing the human p53 codon 249 sequence. *Chem. Res. Toxicol. 12*, 707-714.

242. Zang, H., Goodenough, A. K., Choi, J. Y., Irimia, A., Loukachevitch, L. V., Kozekov, I. D., Angel, K. C., Rizzo, C. J., Egli, M. and Guengerich, F. P. (2005) DNA adduct bypass polymerization by *Sulfolobus solfataricus* DNA polymerase Dpo4: analysis and crystal structures of multiple base pair substitution and frameshift products with the adduct 1,N^2-ethenoguanine. *J. Biol. Chem. 280*, 29750-29764.

243. Ling, H., Boudsocq, F., Woodgate, R. and Yang, W. (2004) Snapshots of replication through an abasic lesion: structural basis for base substitution and frameshift. *Mol. Cell 13*, 751-762.

244. Ling, H., Boudsocq, F., Polsky, B. S., Woodgate, R. and Yang, W. (2003) Replication of a *cis–syn* thymine dimer at atomic resolution. *Nature 424*, 1083-1087.

245. Iwai, S. (2000) Synthesis of a thymine glycol building block and its incorporation into oligonucleotides. *Nucleic Acids Symp. Ser.*, 121-122.

246. Marky, L. A. and Breslauer, K. J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers 26*, 1601-1620.

247. Longfellow, C. E., Kierzek, R. and Turner, D. H. (1990) Thermodynamic and spectroscopic study of bulge loops in oligoribonucleotides. *Biochemistry 29*, 278-285.

248. Petersheim, M. and Turner, D. H. (1983) Base-stacking and base-pairing contributions to helix stability: thermodynamics of double-helix formation with CCGG, CCGGp, CCGGAp, ACCGGp, CCGGUp, and ACCGGUp. *Biochemistry 22*, 256-263.

249. Gralla, J. and Crothers, D. M. (1973) Free energy of imperfect nucleic acid helices. Small internal loops resulting from mismatches. *J. Mol. Biol. 78*, 301-319.

250. Allawi, H. T. and SantaLucia, J., Jr. (1997) Thermodynamics and NMR of internal G.T mismatches in DNA. *Biochemistry 36*, 10581-10594.

251. Bohon, J. and de los Santos, C. R. (2005) Effect of 6-thioguanine on the stability of duplex DNA. *Nucleic Acids Res. 33*, 2880-2886.

252. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, J., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai,

M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, A. D. Daniels, O. Farkas, A. D. Rabuck, K. Raghavachari and Ortiz., J. V. (2004) Gaussian 03, Gaussian, Inc, Wallingford CT.

253. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., 3rd, DeBolt, S., Ferguson, D., Seibel, G. and Knollman, P. (1995) AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Commun. 91*, 1-41.

254. Schlegel, H. B. (1984) Estimating the Hessian for gradient-type geometry optimizations. *Theor. Chim. Acta. 66*, 333-340.

255. Schlegel, H. B. (1982) Optimization of equilibrium geometries and transition structures. *J. Comput. Chem. 3*, 214.

256. Schlegel, H. B. (1989) *New Theoretical Concepts for Understanding Organic Reactions*, Vol. 33, Kluwer Academic, The Netherlands.

257. Schlegel, H. B. (1995) Geometry optimization on potential energy surfaces, In *Modern Electronic Structure Theory* (Yarkony, D. R., Ed.), World Scientific Publishing, Singapore.

258. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277-293.

259. Jeener, J., Meier, B. H., Bachmann, P. and Ernst, R. R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys. 71*, 4546-4553.

260. Wagner, R. and Berger, S. (1996) Gradient-Selected NOESY—A Fourfold Reduction of the Measurement Time for the NOESY Experiment. *J. Magn. Reson. 123*, 119-121.

261. Piotto, M., Saudek, V. and Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR 2*, 661-665.

262. Derome, A. E. and Williamson, M. P. (1990) Rapid-pulsing artifacts in double-quantum-filtered COSY. *J. Magn. Reson. 88*, 177-185.

263. Delaglio, F., Wu, Z. and Bax, A. (2001) Measurement of homonuclear proton couplings from regular 2D COSY spectra. *J. Magn. Reson. 149*, 276-281.

264. Wang, H., Zuiderweg, E. R. P. and Glick, G. D. (1995) Solution structure of a disulfide cross-linked DNA hairpin. *J. Am. Chem. Soc. 117*, 2981-2991.

265. Blondet, P., Albrand, J. P., Von Kienlin, M., Decorps, M. and Lavanchy, N. (1987) Use of rotating-phase DANTE pulses for *in vivo* proton NMR spectral editing with a single irradiation facility. *J. Mag. Res. 71*, 342-346.

266. Bax, A. and Davis, D. G. (1985) MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J. Mag. Res.*, 355-360.

267. Willker, W., Leibfritz, D., Kerssebaum, R. and Bermel, W. (1993) Gradient selection in inverse heteronuclear correlation spectroscopy. *Magn. Resn. Chem. 31*, 287-292.

268. Carr, H. Y. and Purcell, E. M. (1954) Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev. 94*, 630.

269. Vold, R. L., Waugh, J. S., Klein, M. P. and Phelps, D. E. (1968) Measurement of spin relaxation in complex systems. *J. Chem. Phys. 48*, 3831-3832.

270. Macke, T. and Case, D. A. (1998) Modeling unusual nucleic acid structures., In *Molecular Modeling of Nucleic Acids.* (Leontes, N. B. and SantaLucia, J., Eds.) pp 379-393, American Chemical Society, Washington, DC.

271. Goddard, T. D. and Kneller, D. G. SPARKY 3.11, University of California, San Francisco.

272. Keepers, J. W. and James, T. L. (1984) A theoretical study of distance determination from NMR. Two-dimensional nuclear Overhauser effect spectra. *J. Magn. Reson. 57*, 404-426.

273. James, T. L. (1991) Relaxation matrix analysis of two-dimensional nuclear Overhauser effect spectra. *Curr. Opin. Struct. Biol. 1*, 1042-1053.

274. Liu, H., Spielmann, H. P., Ulyanov, N. B., Wemmer, D. E. and James, T. L. (1995) Interproton distance bounds from 2D NOE intensities: effect of experimental noise and peak integration errors. *J. Biomol. NMR 6*, 390-402.

275. Borgias, B. A. and James, T. L. (1990) MARDIGRAS - a procedure for matrix analysis of relaxation for discerning geometry of an aqueous structure. *J. Magn. Reson. 87*, 475-487.

276. Borgias, B. A. and James, T. L. (1989) Two-dimensional nuclear Overhauser effect: complete relaxation matrix analysis. *Methods Enzymol. 176*, 169-183.

277. Van de Ven, F. J. and Hilbers, C. W. (1988) Nucleic acids and nuclear magnetic resonance. *Eur. J. Biochem. 178*, 1-38.

278. van Wijk, J., Huckriede, B. D., Ippel, J. H. and Altona, C. (1992) Furanose sugar conformations in DNA from NMR coupling constants. *Methods Enzymol. 211*, 286-306.

279. Clore, G. M., Murphy, E. C., Gronenborn, A. M. and Bax, A. (1998) Determination of three-bond $^{1}H3'-^{31}P$ couplings in nucleic acids and protein-nucleic acid complexes by quantitative *J* correlation spectroscopy. *J. Magn. Reson. 134*, 164-167.

280. Wu, Z., Tjandra, N. and Bax, A. (2001) Measurement of $^{1}H3'-^{31}P$ dipolar couplings in a DNA oligonucleotide by constant-time NOESY difference spectroscopy. *J. Biomol. NMR 19*, 367-370.

281. Lankhorst, P. P., Haasnoot, A. G., Erkelens, C. and Altona, C. (1984) Carbon-13 NMR in conformational analysis of nucleic acid fragments. The magnitude of torsional angle in d(TpA) from CCOP and HCOP NMR coupling constants. *Nucleic Acids Res. 12*, 5419-5428.

282. Kim, S. G., Lin, L. J. and Reid, B. R. (1992) Determination of nucleic acid backbone conformation by ^{1}H NMR. *Biochemistry 31*, 3564-3574.

283. Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B. and Woods, R. J. (2005) The AMBER biomolecular simulation programs. *J. Comput. Chem. 26*, 1668-1688.

284. Wang, J., Cieplak, P. and Kollman, P. A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem. 21*, 1049-1074.

285. Clore, G. M., Brunger, A. T., Karplus, M. and Gronenborn, A. M. (1986) Application of molecular dynamics with interproton distance restraints to three-dimensional protein structure determination. *J. Mol. Biol. 191*, 523-551.

286. Tsui, V. and Case, D. A. (2000) Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers 56*, 275-291.

287. Bashford, D. and Case, D. A. (2000) Generalized Born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem. 51*, 129-152.

288. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys. 81*, 3684-3690.

289. Izaguirre, J. A., Catarello, D. P., Wozniak, J. M. and Skeel, R. D. (2001) Langevin stabilization of molecular dynamics. *J. Chem. Phys 114*, 2090-2098.

290. Loncharich, R. J., Brooks, B. R. and Pastor, R. W. (1992) Langevin dynamics of peptides: The frictional dependence of isomerization rates of N-actylananyl-N'-methylamide. *Biopolymers 32*, 523-535.

291. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H. and Pedersen, L. G. (1995) A smooth particle mesh Ewald method. *J. Chem. Phys. 103*, 8577-8593.

292. Rychaert, J. P., Ciccotti, G. and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys. 23*, 327-341.

293. Lavery, R. and Sklenar, H. (1996) CURVES 5.1 Helical Analysis of Irregular Nucleic Acids,, Laboratoire de Biochimie Theorique CNRS, Paris, France.

294. Thomas, P. D., Basus, V. J. and James, T. L. (1991) Protein solution structure determination using distances from two-dimensional nuclear Overhauser effect experiments: Effect of approximations on the accuracy of derived structures. *Proc. Natl. Acad. Sci. USA 88*, 1237-1241.

295. Jiang, Y. L., Wiederholt, C. J., Patro, J. N., Haraguchi, K. and Greenberg, M. M. (2005) Synthesis of oligonucleotides containing Fapy.dG (N^6-(2-deoxy-α,β-D-erythropentofuranosyl)-2,6-diamino-4-hydroxy-5 -formamidopyrimidine) using a 5'-dimethoxytrityl dinucleotide phosphoramidite. *J. Org. Chem. 70*, 141-149.

296. Haraguchi, K., Delaney, M. O., Wiederholt, C. J., Sambandam, A., Hantosi, Z. and Greenberg, M. M. (2001) Synthesis and characterization of oligonucleotides containing formamidopyrimidine lesions (Fapy.dA, Fapy.dG) at defined sites. *Nucleic Acids Res. Suppl.*, 129-130.

297. Patro, J. N., Haraguchi, K., Delaney, M. O. and Greenberg, M. M. (2004) Probing the configurations of formamidopyrimidine lesions Fapy.dA and Fapy.dG in DNA using endonuclease IV. *Biochemistry 43*, 13397-13403.

298. Birk, C., Vosa, J. and Wirsching, J. (1997) Preparation, structural elucidation and reactions of benzyl and synthesis of the corresponding 2'-deoxy-4'-thionucleosides. *Carbohydr. Res. 304*, 239-247.

299. Mastihubova, M. and Biely, P. (2004) Deoxy and deoxyfluoro analogues of acetylated methyl β-xylopyranoside--substrates for acetylxylan esterases. *Carbohydr Res. 339*, 2101-2110.

300. Asagoshi, K., Yamada, T., Okada, Y., Terato, H., Ohyama, Y., Seki, S. and Ide, H. (2000) Recognition of formamidopyrimidine by *Escherichia coli* and mammalian thymine glycol glycosylases. Distinctive paired base effects and biological and mechanistic implications. *J. Biol. Chem. 275*, 24781-24786.

301. Ward, D. I., Jeffs, S. M., Coe, P. L. and Walker, R. T. (1993) A mild procedure for the anomerization of 2'-deoxynucleosides. *Tetrahedron Letters 34*, 6779-6782.

302. Gopalakrishnan, S., Liu, X. and Patel, D. J. (1992) Solution structure of the covalent sterigmatocystin-DNA adduct. *Biochemistry 31*, 10790-10801.

303. Gopalakrishnan, S. and Patel, D. J. (1993) Formation and structural features of a sterigmatocystin-formamidopyrimidine adduct at the DNA duplex level. *J. Am. Chem. Soc. 115*, 9321-9322.

304. Reid, B. R. (1987) Sequence-specific assignments and their use in NMR studies of DNA structure. *Q. Rev. Biophys. 20*, 1-34.

305. Patel, D. J., Shapiro, L. and Hare, D. (1987) DNA and RNA: NMR studies of conformations and dynamics in solution. *Q. Rev. Biophys. 20*, 35-112.

306. Boelens, R., Scheek, R. M., Dijkstra, K. and Kaptein, R. (1985) Sequential assignment of imino- and amino-proton resonances in ^{1}H NMR spectra of oligonucleotides by two-dimensional NMR spectroscopy. Application to a *lac* operator fragment. *J. Mag. Res. 62*, 378-386.

307. Hazra, T. K., Das, A., Das, S., Choudhury, S., Kow, Y. W. and Roy, R. (2007) Oxidative DNA damage repair in mammalian cells: a new perspective. *DNA Repair (Amst) 6*, 470-480.

308. Weiss, B. and Cunningham, R. P. (1985) Genetic mapping of nth, a gene affecting endonuclease III (thymine glycol-DNA glycosylase) in *Escherichia coli* K-12. *J. Bacteriol. 162*, 607-610.

309. Jiang, D., Hatahet, Z., Melamede, R. J., Kow, Y. W. and Wallace, S. S. (1997) Characterization of *Escherichia coli* endonuclease VIII. *J. Biol. Chem. 272*, 32230-32239.

310. Lin, J. J. and Sancar, A. (1989) A new mechanism for repairing oxidative damage to DNA: (A)BC excinuclease removes AP sites and thymine glycols from DNA. *Biochemistry 28*, 7979-7984.

311. Reardon, J. T., Bessho, T., Kung, H. C., Bolton, P. H. and Sancar, A. (1997) *In vitro* repair of oxidative DNA damage by human nucleotide excision repair system: possible explanation for neurodegeneration in xeroderma pigmentosum patients. *Proc. Natl. Acad. Sci. USA 94*, 9463-9468.

312. Kow, Y. W., Wallace, S. S. and Van Houten, B. (1990) UvrABC nuclease complex repairs thymine glycol, an oxidative DNA base damage. *Mutat. Res. 235*, 147-156.

313. Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

314. SantaLucia, J., Jr. and Hicks, D. (2004) The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct. 33*, 415-440.

315. Williams, M. C. and Rouzina, I. (2002) Force spectroscopy of single DNA and RNA molecules. *Curr. Opin. Struct. Biol. 12*, 330-336.

316. Marmur, J. and Doty, P. (1962) Determination of the base composition of dioxyribonucleic acid from its thermal melting temperature. *J. Mol. Biol. 5*, 109-118.

317. Wartell, R. M. and Benight, A. S. (1985) Thermal denaturation of DNA molecules. *Phys. Rep. 126*, 67-107.

318. Luo, R., Gilson, H. S., Potter, M. J. and Gilson, M. K. (2001) The physical basis of nucleic acid base stacking in water. *Biophys. J. 80*, 140-148.

319. Sponer, J., Jurecka, P. and Hobza, P. (2004) Accurate interaction energies of hydrogen-bonded nucleic acid base pairs. *J. Am. Chem. Soc. 126*, 10142-10151.

320. Hunter, C. A. (1993) Sequence-dependent DNA structure. The role of base stacking interactions. *J. Mol. Biol. 230*, 1025-1054.

321. Danilov, V. I. and Tolokh, I. S. (1984) Nature of the stacking of nucleic acid bases in water: a Monte Carlo simulation. *J. Biomol. Struct. Dyn. 2*, 119-130.

322. Cieplak, P. and Kollman, P. A. (1988) Calculation of the free energy of association of nucleic acid bases *in vacuo* and water solutions. *J. Am. Chem. Soc. 110*, 3734-3739.

323. Yakovchuk, P., Protozanova, E. and Frank-Kamenetskii, M. D. (2006) Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res. 34*, 564-574.

324. Protozanova, E., Yakovchuk, P. and Frank-Kamenetskii, M. D. (2004) Stacked-unstacked equilibrium at the nick site of DNA. *J. Mol. Biol. 342*, 775-785.

325. Yan, S., Shapiro, R., Geacintov, N. E. and Broyde, S. (2001) Stereochemical, structural, and thermodynamic origins of stability differences between stereoisomeric benzo[a]pyrene diol epoxide deoxyadenosine adducts in a DNA mutational hot spot sequence. *J. Am. Chem. Soc. 123*, 7054-7066.

326. Rodger, A. and Nordén, B. (1997) *Circular Dichroism and Linear Dichroism*, Oxford University Press, Oxford; New York.

327. Norden, B. and Tjerneld, F. (1982) Structure of methylene blue-DNA complexes studied by linear and circular dichroism spectroscopy. *Biopolymers 21*, 1713-1734.

328. Lyng, R., Hard, T. and Norden, B. (1987) Induced CD of DNA intercalators: electric dipole allowed transitions. *Biopolymers 26*, 1327-1345.

329. Schipper, P. E., Norden, B. and Tjerneld, F. (1980) Determination of binding geometry of DNA-adduct systems through induced circular dichroism. *Chem. Phys. Lett. 70*, 17-21.

330. Gray, D. M., Ratliff, R. L. and Vaughan, M. R. (1992) Circular Dichroism Spectroscopy of DNA, In *Methods in Enzymology, Vol 211* (Lilley, D. M. J. and Dahlberg, J. E., Eds.) pp 389-406, Academic Press, San Diego.

331. Newberne, P. M. and Butler, W. H. (1969) Acute and chronic effects of aflatoxin on the liver of domestic and laboratory animals: a review. *Cancer Res. 29*, 236-250.

332. Miller, E. C. and Miller, J. A. (1981) Searches for ultimate chemical carcinogens and their reactions with cellular macromolecules. *Cancer 47*, 2327-2345.

333. Clifford, J. I. and Rees, K. R. (1967) The action of aflatoxin B_1 on the rat liver. *Biochem. J. 102*, 65-75.

334. Sarasin, A. and Moule, Y. (1975) Translational step inhibited *in vivo* by aflatoxin B_1 in rat-liver polysomes. *Eur. J. Biochem. 54*, 329-340.

335. Sarasin, A. and Moule, Y. (1976) Helical polysomes induced by aflatoxin B_1 *in vivo*. A new hypothesis for helix formation by chemicals and carcinogens. *Exp. Cell Res. 97*, 346-358.

336. Yu, F. L., Cahill, J. M., Lipinski, L. J. and Dipple, A. (1996) Effect of aflatoxin B_1-8,9-epoxide-DNA adducts on transcription of a supF gene fragment. *Cancer Lett. 109*, 77-83.

337. Woody, R. W. (1995) Circular Dichroism, In *Methods in Enzymology, Vol 246* (Kenneth, S., Ed.) pp 34-71, Academic Press, San Diego.

338. Laspia, M. F. and Wallace, S. S. (1988) Excision repair of thymine glycols, urea residues, and apurinic sites in *Escherichia coli*. *J. Bacteriol. 170*, 3359-3366.

339. Krosky, D. J., Schwarz, F. P. and Stivers, J. T. (2004) Linear free energy correlations for enzymatic base flipping: how do damaged base pairs facilitate specific recognition? *Biochemistry 43*, 4188-4195.

340. Gorenstein, D. G. (1994) Conformation and dynamcis of DNA and protein-DNA complexes by ^{31}P NMR. *Chem. Rev. 94*, 1315-1336.

341. Wu, Z., Tjandra, N. and Bax, A. (2001) ^{31}P chemical shift anisotropy as an aid in determining nucleic acid structure in liquid crystals. *J. Am. Chem. Soc. 123*, 3617-3618.

342. Sanner, M. F. (1996) MSMS, The Scripps Research Institute, La Jolla, CA.

343. Bax, A. (2003) Weak alignment offers new NMR opportunities to study protein structure and dynamics. *Protein Sci. 12*, 1-16.

344. Wu, Z., Delaglio, F., Tjandra, N., Zhurkin, V. B. and Bax, A. (2003) Overall structure and sugar dynamics of a DNA dodecamer from homo- and heteronuclear dipolar couplings and ^{31}P chemical shift anisotropy. *J. Biomol. NMR 26*, 297-315.

345. Clark, F. H. and Cahoon, N. M. (1987) Ionization constants by curve fitting: determination of partition and distribution coefficients of acids and bases and their ions. *J. Pharm. Sci. 76*, 611-620.

346. Dixon, S. L. and Jurs, P. C. (1993) Estimation of pK_a for organic oxyacids using calculated atomic charges. *J. Comp. Chem. 14*, 1460-1467.

347. Csizmadia, F., Tsantili-Kakoulidou, A., Panderi, I. and Darvas, F. (1997) Prediction of distribution coefficient from structure. 1. Estimation method. *J. Pharm. Sci. 86*, 865-871.