2015

Computational studies of damaged DNA : an investigation of DNA O-linked adducts formed due to exposure to phenolic carcinogens

Majdi Yazdi, Mohadeseh

Lethbridge, Alta. : University of Lethbridge, Dept. of Chemistry and Biochemistry

COMPUTATIONAL STUDIES OF DAMAGED DNA:
AN INVESTIGATION OF DNA O-LINKED ADDUCTS FORMED DUE TO
EXPOSURE TO PHENOLIC CARCINOGENS

MOHADESEH MAJDI YAZDI
B.Sc., Sharif University of Technology, 2004
M.Sc., Tarbiat Modares University, 2008

A Thesis

Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfillment of the
Requirements for the Degree

MASTER OF SCIENCE

Department of Chemistry and Biochemistry
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

© Mohadeseh Majdi Yazdi, 2015

COMPUTATIONAL STUDIES OF DAMAGED DNA:
AN INVESTIGATION OF DNA O-LINKED ADDUCTS FORMED DUE TO
EXPOSURE TO PHENOLIC CARCINOGENS

MOHADESEH MAJDI YAZDI

Date of Defence: September, 12, 2014

Dr. P. Dibble
Supervisor   Professor   Ph.D.

Dr. M. Gerken
Thesis Examination Committee Member   Professor   Ph.D.

Dr. S. Wetmore
Thesis Examination Committee Member   Professor   Ph.D.

Dr. H. Muchall
External Examiner
University Concordia
Montreal, Quebec   Associate Professor   Ph.D.

Dr. M. Roussel
Chair, Thesis Examination Committee   Professor   Ph.D.

Dedications

This thesis is dedicated to my family, my mother, Zahra, and my father Ali Akbar, who did not only raise and nurture me but also dedicated themselves unconditionally over the years for my education, and intellectual development. Their paternal care and support have been shown in incredible ways during my life and they always, were a source of motivation and strength during moments of despair and discouragement, my brother Mehdi and my sister Motahareh, for their endless love, support and encouragement. My sister in law, Mahboubeh, and my dear nephew, Mehrad, who bring happiness to my family and they have been always hoping in me and have never left my side. I would be forever indebted to such a lovely family.

Abstract

This thesis systematically develops a computational model to identify the conformational and base-pairing preferences of ^{PhO}dG, $^{4-Cl-PhO}$dG, $^{DCP-O}$dG, $^{TCP-O}$dG, and $^{PCP-O}$dG by gradually increasing the size of the system also structural properties of unsubstituted O-linked. All adducts at nucleoside level adopted *syn* conformation. Moreover, effect of protonation at N^3 and N^7 site on the structural properties and deglycosilation barrier of adducted guanosine was probed. It was highly desirable to include O-linked phenolic as well as C8-dG adducts into a DNA strand in order to understand the detrimental effect of them and the conformational distortion of double helix duplex the desired modified base into *Nar*I DNA duplex through the employment of molecular dynamic simulation (MD) was assessed. The *anti*-conformation against cytosine is preferred with this model for all adducts and *syn* conformer for all unsubstituted O-linked and *ortho* and *para* C–linked structures against guanine mismatch is the lowest energy structure.

Acknowledgement

I would like to express my deepest appreciation to my supervisor Dr. Peter Dibble for his invaluable assistance, support and guidance which this work would not have been possible without his help. His supportive experience, encouragement and collaboration during the project promoted the work results and directed the research in a fruitful approach. His vast knowledge and experience have played an important role in success of this project. I am sincerely grateful for his help.

I would like to extend my most sincere gratitude to the dean of graduate studies at the University of Lethbrige, Professor Robert Wood, for his infinite supports, wise decisions, and kindness. Words cannot describe how much I appreciate the values and lessons that he has taught me as my dean and a very responsible person at the University of Lethbridge.

I wish to express my appreciation for uncountable understanding, care, kindness as well as camaraderie as a friend to Margaret Miles behavioral health consultant and Dr. Julie Smith at the University of Lethbridge Health Centre.

Special thanks go to Kathleen Schrage, manager of School of Graduate Studies for her unassuming supports during my education at the University of Lethbridge and also Lorie Peter administrative support for all her punctuality and care.

I would like to thank to the members of my dissertation committee, Dr. Michel Gerken and Dr. Stacey Wetmore for taking time to review my dissertation. I would

teaching and making chemistry enjoyable for the students. I would like to thank my coordinators during my TA ship at the University of Lethbridge Dr. Ying Zheng and Dr. Greg Patenaude, and John Eng.

Finally, I would like to thank to my family for their infinite confidence and encouragement.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

3,4-EQ, 3,4-estronequinone

4-OHEN, 4-hydroxyequilenin

A, adenine

AA, aromatic amine

AAF, 2-acetylaminofluorene

ABP, aminobiphenyl

AF, 2-aminofluorene

AFB, aflatoxin Bi

AFB-FAPY, aflatoxin Bi formamidopyrimidine

AN, aniline

A_N, nucleophilic addition

AP, aminopyrene

aq., aqueous

B, major groove binding conformation-B-type

B[a]P, benzo[a]pyrene

BPQ, benzo[a]pyrene-7,8-dione

C, cytosine

C8-(4-CH20CH3-Ph)-dG, C8-(4-methoxymethylphenyl)-2'-deoxyguanosine adduct

C8-(4-Me-Ph)-dG, C8-(4-methylphenyl)-2'-deoxyguanosine adduct

C8-Ph-dG, C8-phenyl-2'-deoxyguanosine adduct

dA, 2'-deoxyadenosine

dC, 2'-deoxycytidine

DCP-O-dG, 8-(2,4-dichlorophenoxy)-2'-deoxyguanosine

deg.; degree

DFT, density functional theory

dG, 2'-deoxyguanosine

dGMP, 2'-deoxyguanosine 5'-monophosphate

D_N, leaving group dissociation

DNA, deoxyribonucleic acid

E.Coli, Escherichia Coli

EPA, Environmental Protection Agency

G, guanine

IARC, International Agency for Research on Cancer

1Q, 2-amino-3-methylimidazo[4,5-/]quinolone

J, Joules

K, Kelvin

K_1, rate constant for the glycosidic bond cleavage of the monoprotonated species

K_2, rate constant for the glycosidic bond cleavage of the monoprotonated species

K_a, acid dissociation constant or association equilibrium constant

K_B, Boltzmann's constant (1.381×10^{-23} J•K^{-1})

MD, molecular dynamics

mer, denotes length of an oligonucleotide

*Nar*I, 5'-CTCGGCGCCATC-3'

N2-(4-PhOH)-dG, N2-(4-hydroxyphenyl)-2'-deoxyguanosine adduct

NI, 5-guanidino-4-nitroimidazole

NMR, nuclear magnetic resonance

o-, ortho

ODN, oligodeoxynucleotide

OTA, ochratoxin A

p-, para

PAH, polycyclic aromatic hydrocarbons

PCM, polarisable continuum model

PCP, pentachlorophenol

PES, Potential Energy Surface

Ph, phenyl

PhOH, phenoxyl

Ph-OdG, 8-phenoxy-2'-deoxyguanosine

RESP, restrained electrostatic potential

S, base displaced stacked conformation - stacked

S_NAr, nucleophilic aromatic substitution

TCP, trichlorophenol

TCP-OdG, 8-(2,4,6-trichlorophenoxy)-2'-deoxyguanosine

T, thymine

T_m, thermal melting

ZPVE, zero-point vibrational energy

W, minor groove binding conformation - wedge

Chapter 1. Introduction

1.1. Overview

A significant focus of nucleic acids research is on the damage to DNA, the formation mechanism of DNA lesions by electrophilic species. The assessment of structural and genotoxic features of mutated complexes is of particular interest. Phenols are known to form O-linked and *ortho* and *para* C-linked adducts at the C8 site of deoxyguanosine (dG). This dissertation studies the O-linked class of phenolic dG adducts using a bottom-up approach (small nucleobase to large DNA duplex) by using quantum mechanical modeling and molecular dynamics simulations for small and large models respectively. Their proton affinity, hydrolytic stability, as well as their structural impact on the DNA duplex, has been considered. Furthermore, the effect of functionalization with chlorine on the conformational preference and glycosidic bond stability will be probed. To assess the modified duplex structure, as well as distortions induced to the helix, the O-linked dG adduct was incorporated into a DNA duplex. Theoretical results indicate a destabilizing effect on the natural duplex. Likewise, in order to elucidate the effects of the C-linked lesions and expand our knowledge of the relative effect of sequence dependence, the stability of the associated hydrogen bonds, base pair parameters and sugar puckering, the *ortho* and *para* C-linked adducts will be compared and contrasted with the O-linked analogues.

1.2. General Background

1.2.1. History of DNA

The history of deoxyribonucleic acid (DNA) began in the late 1860s via structural identification by Swiss chemist Friedrich Miescher via isolation of "nuclein" from leukocytes obtained from pus.[1] Following Miescher's discovery, other scientists such as Phoebus Levene and Erwin Chargaff obtained more specific information about the chemistry of the building blocks and their topological constructive interactions of the DNA molecule.[2] Based on their work, American biologist James Watson, English physicists Francis Crick, Maurice Wilkins and Rosalind Franklin, postulated their noteworthy conclusion which is the three-dimensional structure of the double helix. Later on in 1944, Avery, MacLeod, and McCarty proposed the genetical characteristics of DNA.[3-7] These and other discoveries, over many decades, ultimately resulted in more precise information about the importance of DNA to life.

1.2.2. Components of DNA

DNA is an integral biopolymer selected by nature to store the information required to build organisms (Figure 1.1). These organisms in turn replicate the information in DNA. The chemical structure of DNA includes two classes of bases: purines and pyrimidines. The purines, having a double-ringed structure, are adenine and guanine. Likewise, the pyrimidines bases are cytosine and thymine and are single-ringed structures. N1 of the pyrimidines or N9 of the purines are attached to deoxyribose by a glycosidic bond to form nucleosides (Figure 1.2).[8, 9]
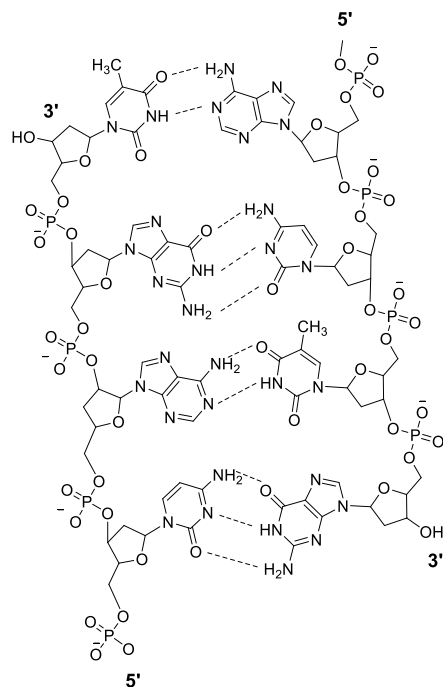
Figure 1.1. Schematic model of the DNA double helix.
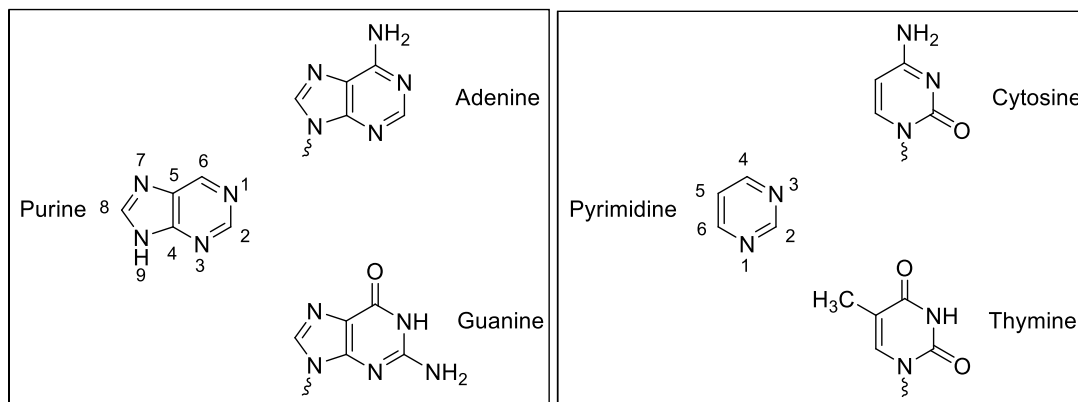
Figure 1.2. The chemical structures of purines and pyrimidines. The sugar moiety can be attached to the bases via dashed lines.

The only difference between ribose, the sugar component of RNA and the deoxyribose sugar in DNA is the absence of a hydroxyl group at the 2′-position. Each nucleoside base has two different conformations: syn and anti. There is reversible equilibrium between these two conformers (Figure 1.3).
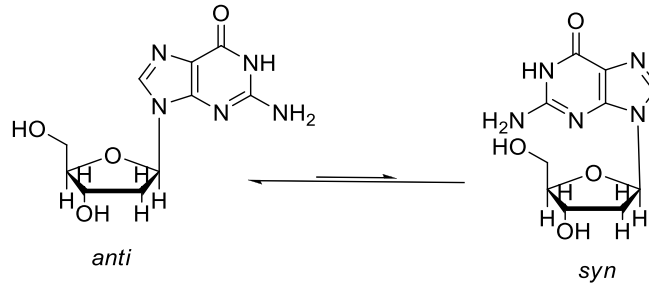
Figure 1.3. *Anti* and *syn* conformation of deoxyguanosine.[10]

DNA nucleotides are composed of a nitrogenous base, deoxyribose and at least one phosphate group, and are the subunits or building blocks of nucleic acids (DNA and RNA). Genes are formed by arrangements of the four different nucleotides. The nucleosides which are involved in the formation of genes are 2'-deoxyadenosine (dA), 2'-deoxyguanosine (dG), 2'-deoxythymidine (dT) and 2'-deoxycytidine (dC) (Figure 1.4).
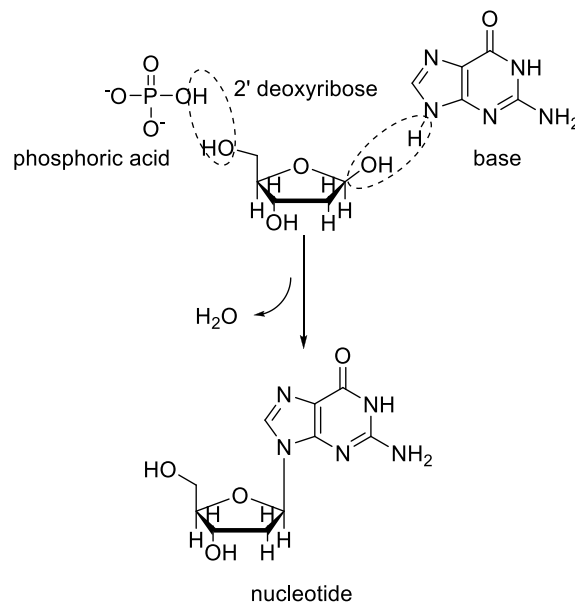
Figure 1.4. Nucleotide can form by reaction of base, sugar and phosphate group and releasing water.[11]

4

The negative charge on the phosphate groups allows positively charged histone proteins to interact with them to form chromatin. Also, this property makes it possible to separate DNA using gel electrophoresis. To follow conventions, these natural negative charges should be neutralize during the process of simulation.

1.2.3. The DNA Double Helix

Base pairing two polynucleotide chains by hydrogen bonds in an anti-parallel alignment forms the DNA double helix (Figure 1.1).[12]. The pairing occurs between adenine and thymine, guanine and cytosine, which results in the self-encoding character of DNA.[13] Most DNA double helices are right-handed except Z-DNA which is left-handed. One of the unique features of the DNA double helix is pairing of the 5′ end of one strand with the 3′ end of its complementary strand via the phosphate groups which attach the 3' end of one sugar to the 5′ end of the other sugar of nucleotide component. On the other hand, intramolecular hydrogen bonds and the availability of the outer edges of the nitrogen bases keep the complementary base pairs together. These also provide the possibility of bonding to other molecules like the proteins with the helices. In fact, these hydrogen bonds play imperative roles in replication and expression. A second important contribution in the stability of a duplex is stacking interactions between the flanking bases. The flat and relatively water-insoluble bases tend to stack above each other due to presence of some weak interactions.[14, 15]

1.2.3.1. Conformations of the Double Helix

Different geometries and dimensions have been identified for the double helix. Early X-ray diffraction studies revealed the B and the A forms of DNA as the two major types of duplex structures.[14, 16] B-DNA is the most common conformation of DNA. DNA also adopts the A conformation which is shorter and wider which cannot easily exist under normal physiological conditions. Z-DNA was first discovered in 1979 and is a left-handed conformation. The Z-DNA helix is left-handed and has a structure that repeats every 2 base pairs. The major and minor grooves, unlike A- and B-DNA, show little difference in width.[17]
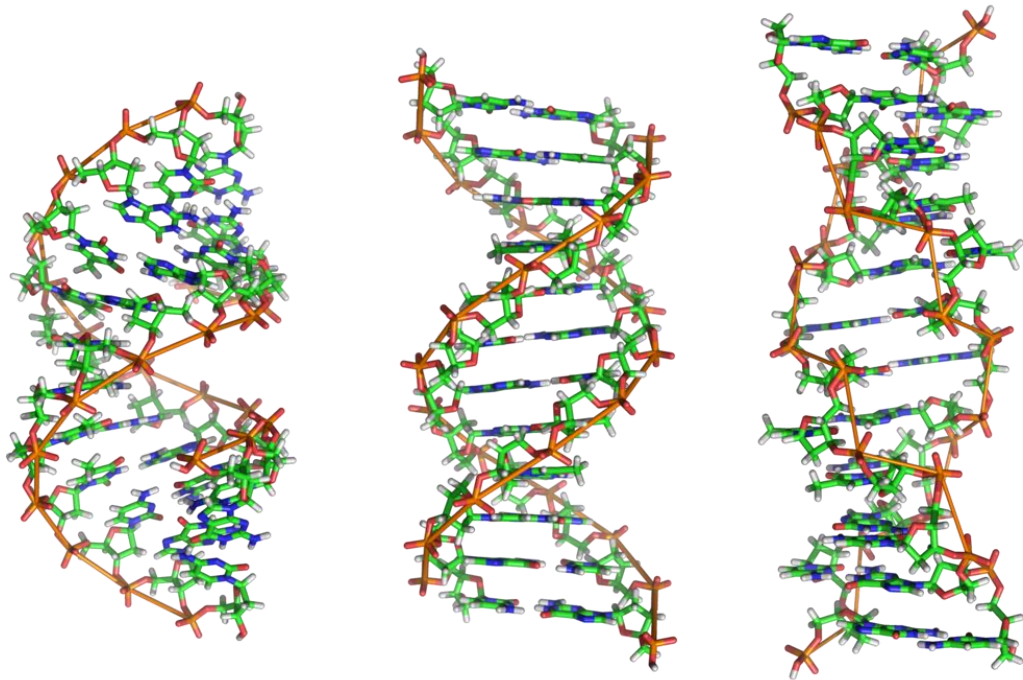
Figure 1.5. Models of the A, B, and Z Forms of DNA. The sugar-phosphate backbone of each chain is located on the outside and the bases oriented inward. (Left) The more compact A form of DNA, (middle) The B form of DNA, the usual form found in cells (right) Z DNA is a left-handed helix and has a zig zag (hence "Z") appearance.

1.2.3.2. The Minor Groove and Major Groove in DNA Structure

The DNA molecule is a long extended polymer and the major and minor grooves of DNA occur where the backbones are far apart and close together, respectively, (Figure 1.6) making the size of two grooves different. The bases are oriented inside the helices (inward), nonetheless they are accessible through the major and minor grooves. The geometry of the base pairs are determining factors in specifying different grooves. In particular, the angle between the glycosidic bonds (the angle at which the two sugars protrude from the base pairs) is about 120° for the narrow angle or 240° for the wide angle. Greater stacking of the flanking base leads to more narrow angles between the sugars on the backbone of

the strand, generating a minor groove, and the larger angle on the other edge (sugar backbone) causes to the formation of the major groove.[11, 18, 19]
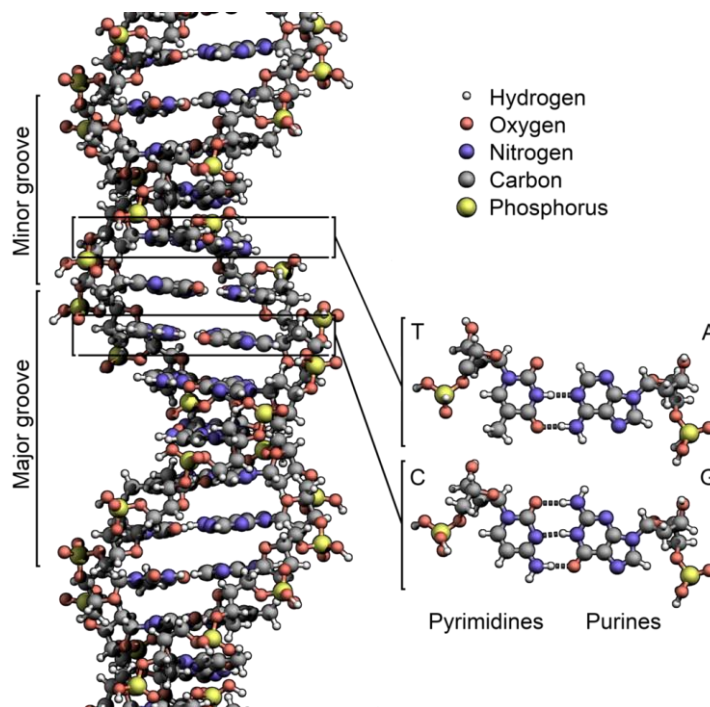
Figure 1.6. Space-filling model of the double helix. The typical model of the major and minor grooves. The major groove width is around 22 Å whereas this width is reduced to 12 Å in minor.

The bases which present in the major groove are more accessible because the backbones are not in the way. Hence, proteins that regulate biochemical reactions in the cell and are involved in transcription (copying DNA to RNA) or replication (copying DNA to DNA) more favorably react with bases on the major groove side.

1.2.3.3. Conformations of Mutated DNA Complexes

In the context of fully paired duplexes, there are three prototypical conformational patterns which have been observed for C8 substituted dG (with aromatic amines) which can be extended to other C8 substituted modified complexes.[20] They are classified broadly as the major groove B-type (B), base

displaced stacked (S), and minor groove wedge (W) conformations (**Error! eference source not found.**).[21]
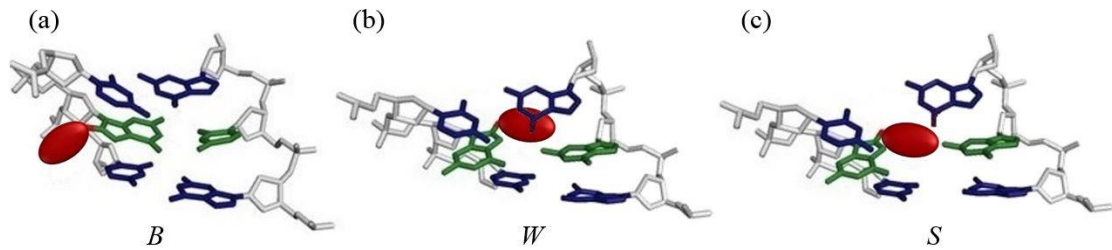
Figure 1.7. Conformational motifs in duplex DNA (a) B, (b) S (stacked), and (c) W (wedge) conformers.[10]

1.2.4. Bases Can Flip Out from the Double Helix

The pairings of each base on one polynucleotide strand with the complementary base on the other strand are energetically favorable. However, exogenous and endogenous agents can induce DNA lesions that alter the chemical structure of the nucleobases. This sometimes can cause bases to protrude from the double helix, a phenomenon known as base flipping. As DNA is flexible, it is assumed that the process of flipping out one base is not energetically expensive.[22] In addition, distortion of double helix and formation of new hydrogen bonds and stacking interactions in some modified complexes are energetically disfavored and preferred, respectively.

1.3. DNA Damage and Mutation

DNA damage and mutation are fundamentally different concepts. Damage is the physical and chemical alteration of DNA structure which causes abnormalities in the DNA features. These abnormalities can be recognized by enzymes and repaired depending on information that is available in the undamaged sequence of

the complementary DNA strand as well as presence of some enzymes that may detect/accommodate damage. On the other hand, mutation is an alteration in the arrangement of bases which affect the sequence of the DNA and cannot be repaired by enzymes, so these mutations are replicated when the cell replicates. Protein function and regulation can be altered by mutations. DNA damage can induce errors in DNA synthesis during the process of replication or repair. Because these errors often mainly lead to mutation, they are considered to be related topics.[23]

1.3.1. DNA Damage

The chemistry of DNA damage is complex and there are various types of DNA lesions that can form. Although biological systems are surrounded by a medium which they need to survive, there are unfriendly components that are inevitable. Reactive oxygen species in the atmosphere, exposure to damaging solar and cosmic radiations, and a variety of natural and cooking-induced carcinogens in food are other basic sources of damage. In addition to these natural agents, technological developments have resulted in man-made hazards from pollution. Unfortunately, there is also the possibility that hazards created by humans induce damage that may inflict mutations, or cannot be repaired by existing systems.[24] Therefore, both internal and external sources inflict damage to DNA. In addition, certain types of damage can be induced by classical chemical reactions. Together, they can produce more than 70 distinct chemical modifications of native nucleotides. Indeed, about 15 of these have been distinguished in the cell. Structural DNA damages can be cytotoxic. It is estimated that DNA damage events in a single human cell range from 10^4–10^6 per day, requiring about 10^{16}-10^{18} repair

events per day in an adult human, having an average of 10^{12} cells. DNA damage associated with endogenous species is more extensive (greater than 75%) than damage caused by environmental factors.[23, 25]

1.3.1.1. Damage to the DNA Backbone and Double-Strand Breaks

The backbone of DNA is exposed to various types of damage. Deoxyribose sugars have several reactive positions toward oxygen radical species.[26-28] Single strand breaks (SSBs) in the DNA backbone can occur by exposure to ionizing radiation or radiomimetic anti-cancer drugs such as bleomycin or neocarzinostatin. These modes of damage often induce multiple damaged sites in close proximity in the double helix and thus lead to the formation of double strand breaks (DSBs). These are severe lesions because it has been shown that they can induce gross chromosomal aberrations (translocations that are known as genetic exchange between nonhomologous chromosomes).[29-31] These types of mutations change the structure of the DNA backbone to phosphate, 3-phosphoglycolate, and 3-phosphoglycaldehyde end groups.

1.3.1.2. DNA Interstrand Crosslinks

A highly cytotoxic lesion which results in a covalent connection between two DNA bases on opposite strands of DNA is known as an interstrand crosslink (ICL).ICLs block DNA replication and transcription. There is evidence that ICLs occur naturally through the reaction of DNA with bifunctional electrophiles which are generated by peroxidation of lipids but these lesions are expected to be rare

(Figure 1.8).[32] Interestingly, a large number of antitumor drugs, such as the chloroethyl nitroso ureas, nitrogen mustards, mitomycins, and cisplatin, form ICLs.[33]
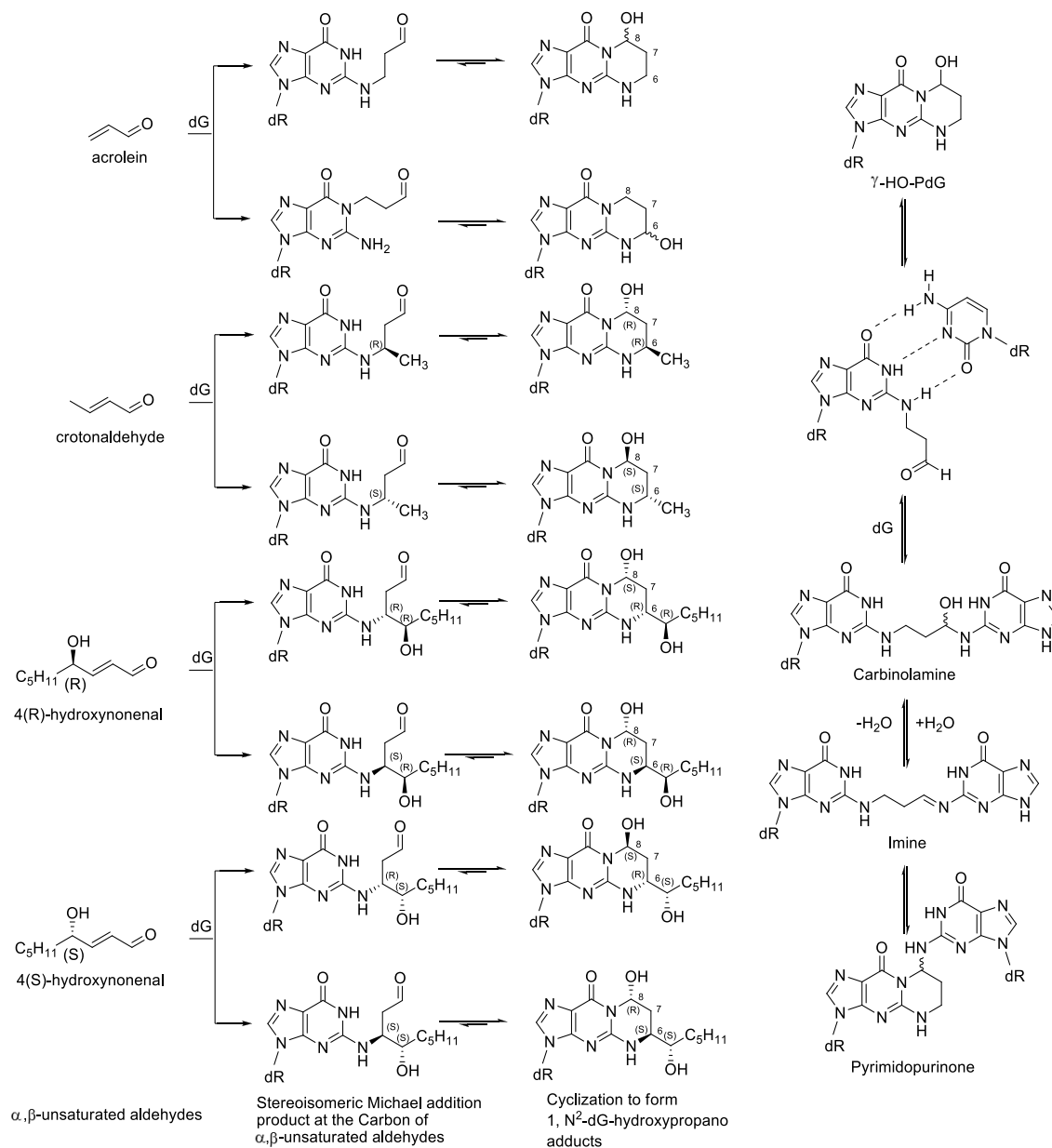
Figure 1.8. N 2-dG cyclic adducts arising from Michael addition of acrolein, crotonaldehyde, and 4-HNE to dG and formation of ICL by the generated intermediates.[10]

Figure 1.8 gives examples of the formation of interstrand crosslinks, the 1, N^2-dG adducts of acrolein, crotonaldehyde, and 4-HNE yield interstrand cross-links in a specific sequence of DNA. These take place via the opening of the 8–hydroxypropano ring to the corresponding aldehydes followed by an attack with the N^2-amino group of dG in the complementry strand in the sequence. The cross-links arising from acrolein and crotonaldehyde exist in duplex DNA as carbinolamine linkages, which enable the cross-linked CG base pairs to maintain Watson-Crick hydrogen bonding with minimal distortion of the duplex. The crosslinking chemistry of crotonaldehyde and 4-HNE depends upon the stereochemistry of the C6 carbon, which orients the reactive aldehydes within the minor groove in the sequence.[34]

1.3.1.3. Damage to the DNA Nucleobases

1.3.1.3.1. General Overview

The DNA nucleobases are particularly susceptible to DNA damage (Figure 1.9). Hydrolysis is the simplest reaction that is potentially harmful to DNA (1). Indeed, the glycosidic bond of the purine nucleotides is prone to acid-catalyzed hydrolysis.[35-37] A second hydrolytic reaction (deamination) can occur at the exocyclic amine groups of C, 5-MeC, A, and G, forming uracil, thymine, hypoxanthine, and xanthine nucleobases, respectively (2).[38] Reactive oxygen species (ROS) such as superoxide radical anions, hydrogen peroxide, or hydroxyl radicals (by-products of oxygen metabolism) can react with DNA to give rise to over 100 oxidative DNA modifications found to date (3,4).[26, 39] A common example of methylation of DNA by some cofactors of enzymatic reactions, such as S-

adenosylmethionine (SAM), is the formation of 7-methylguanine, a relatively harmless lesion, and 3-methyladenine. The latter is highly cytotoxic due to its ability to block DNA replication (7, 8).[40] In addition, there are numerous unavoidable exogenous sources leading to mutation, among them cyclobutane dimers which are photoadducts between flanking pyrimidine residues in DNA formed by UV radiation from sunlight.[41] Other radiation frequencies, like X-rays and γ-rays, may also cause DNA damage (9, 10). Genotoxic compounds, such as aromatic amines and aromatic hydrocarbons that can be found in cigarette smoke, and manmade chemical compounds such as hair dye can cause DNA damage through the formation of bulky DNA adducts (addition products) (11, 12, 13).[42]

It has been demonstrated that among the various DNA damage pathways, base modification is the most common. DNA base modification generally refers to changes in the structures and chemical properties of the bases causing their loss of functionality. Guanine has the smallest oxidation potential among all the DNA bases guanine (1.29V), adenine (1.42V), cytosine (1.6V), thymine (1.7V). Guanine is frequently modified by one-electron oxidants.[26, 43]
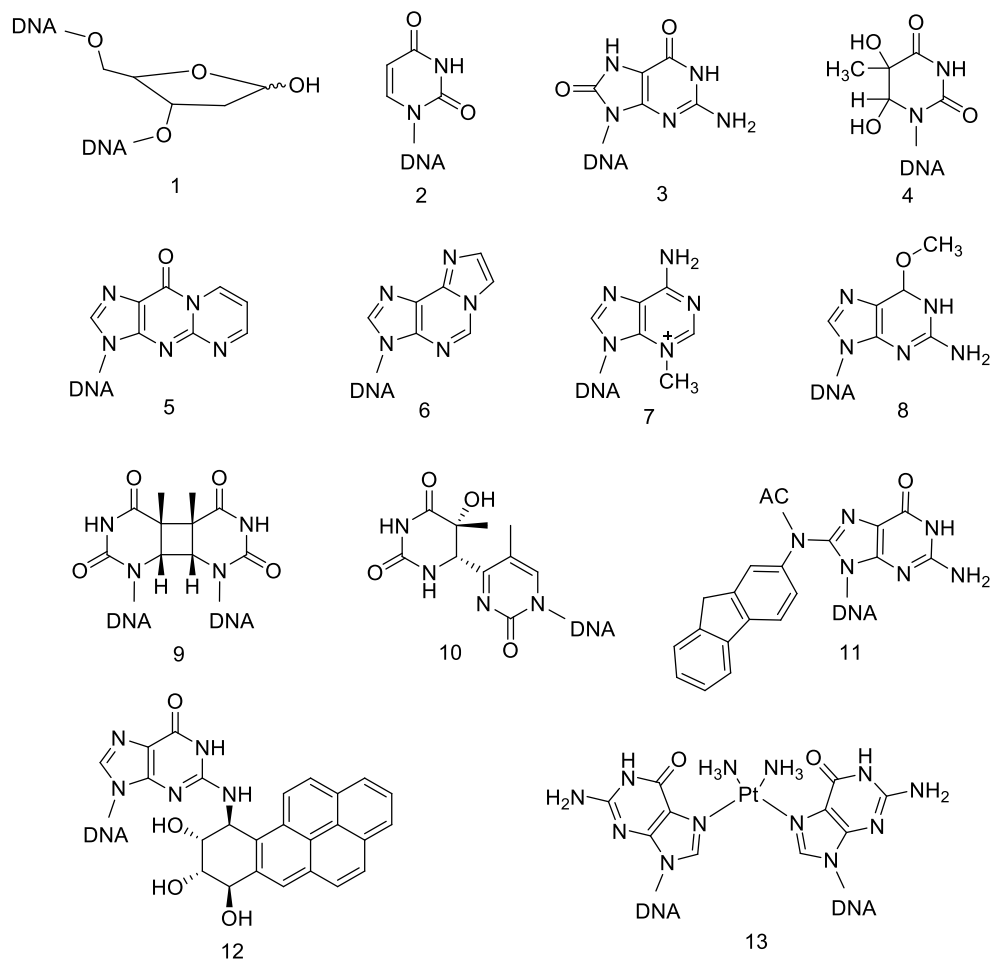
Figure 1.9. Various types of damage to the nucleobases.[24]

1.3.1.3.2. Reactive Oxygen (ROS) Species and Nitrogen Species (RNS)

The greatest variety of damage is induced by free radicals. Many damaging agents, including some beneficial drugs, are capable of generating free radicals. ROS and RNS are byproducts of oxygen and nitrogen metabolism, and

consequently the response to these types of reagents can lead to DNA lesion (Figure 1.10).
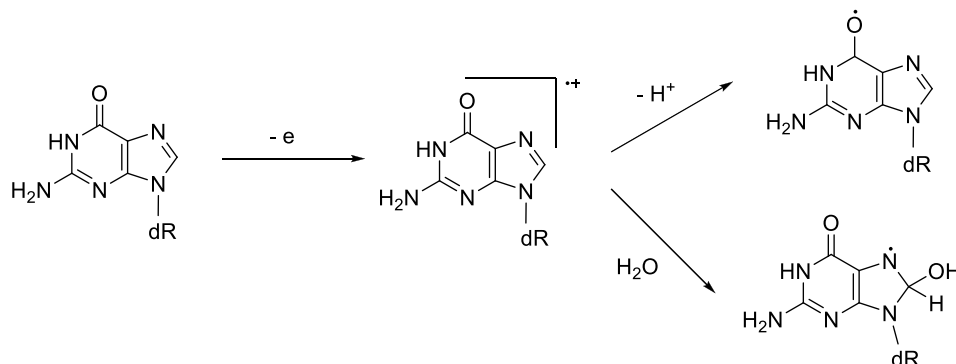
Figure 1.10. One electron oxidation reactions of the guanine moiety of DNA and hydration reactions of the initially formed guanine radical cations leading to the transient generation of the reducing 8-hydroxy-7,8 - dihydroguanyl.[10]

Free radicals also have been implicated in the etiology of neurological diseases such as Alzheimer's[44] and Parkinson's.[45] It has been shown that oxidatively generated DNA damage contributes to aging.[46, 47] Cytotoxic lesions may result in abnormal cell physiology, apoptosis, and cell death if DNA replication or transcription is inhibited. Endogenously generated products of oxidative stress can trigger DNA damage. Peroxidation of polyunsaturated fatty acyl moieties on phospholipids generates many electrophilic products, including a series of bifunctional aldehydes capable of reacting with DNA to form adducts (Figure 1.11).
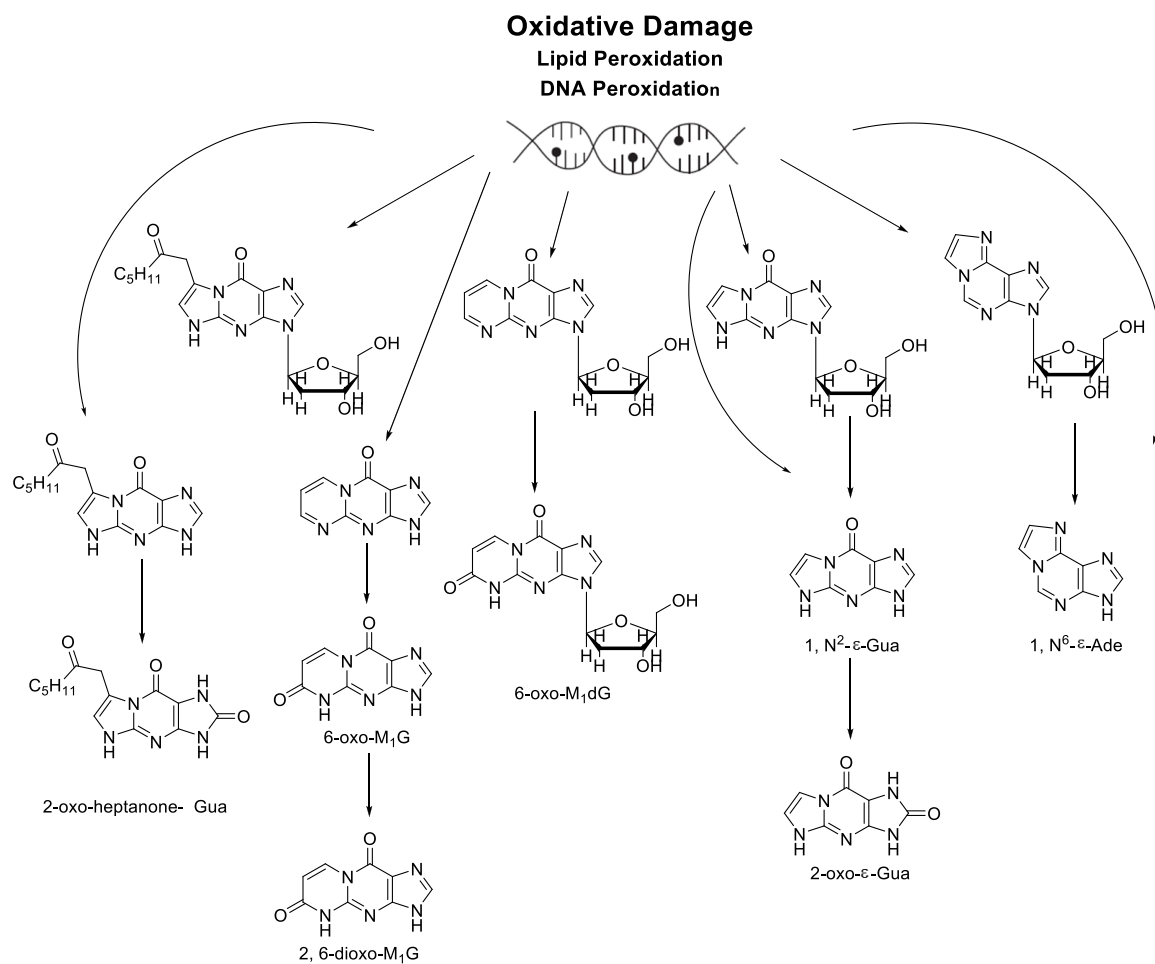
Figure 1.11. Several types of oxidative damage.[10]

Most adducts from this class contain exocyclic rings that block Watson-Crick base pairing. They are highly mutagenic, and can induce a range of base pair substitutions and, in some cases, frameshift mutations. In addition to oxidation, some nucleoside adducts are subject to glycolytic cleavage, which yields the free base.

1.3.1.3.3. Polycyclic Aromatic Hydrocarbons

Many environmental chemical substances have been implicated by the World Health Organization's International Agency for Research on Cancer (IARC). Among

the different carcinogens are polycyclic aromatic hydrocarbons (PAHs), compound that contain two or more fused benzene rings. One example is benzo[*a*]pyrene (B[*a*]P) which is the product of fossil fuel combustion and is therefore ubiquitous in our environment. Fifty carcinogens exist in cigarette smoke, including PAHs, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone, 1,3-butadiene,[48-50] and aromatic amines and heterocyclic aromatic amines. All of these compounds are metabolically activated to reactive intermediates that form premutagenic covalent adducts with DNA. During the twentieth century, a combination of epidemiological and animal experiments has provided convincing evidence that polycyclic aromatic hydrocarbons (PAHs) and the representative compound benzo[*a*]pyrene (B[*a*]P), are carcinogenic (Figure 1.12).[10]

Bay-Reigon PAH
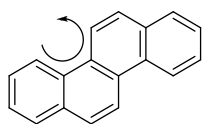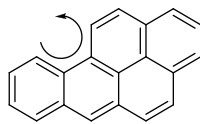
Chrysene   Benzo[a]pyrene   Benz[a]anthracene
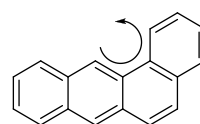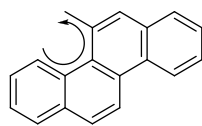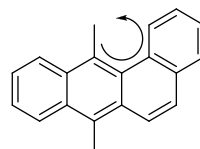
5-Methylchrysene   7,12-Dimethylbenz[a]anthracene

Fjord-Reigon PAH
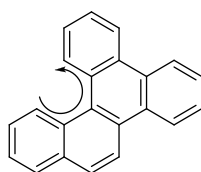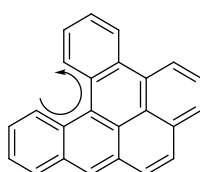
Benzo[g]cherysene   Dibenzo[a,f] pyrene

Figure 1.12. Structures of PAH carcinogens.[10]

There are different types of PAH-DNA adducts formed depending on the pathways of PAH activation that exist in the target organ (Figure 1.13).
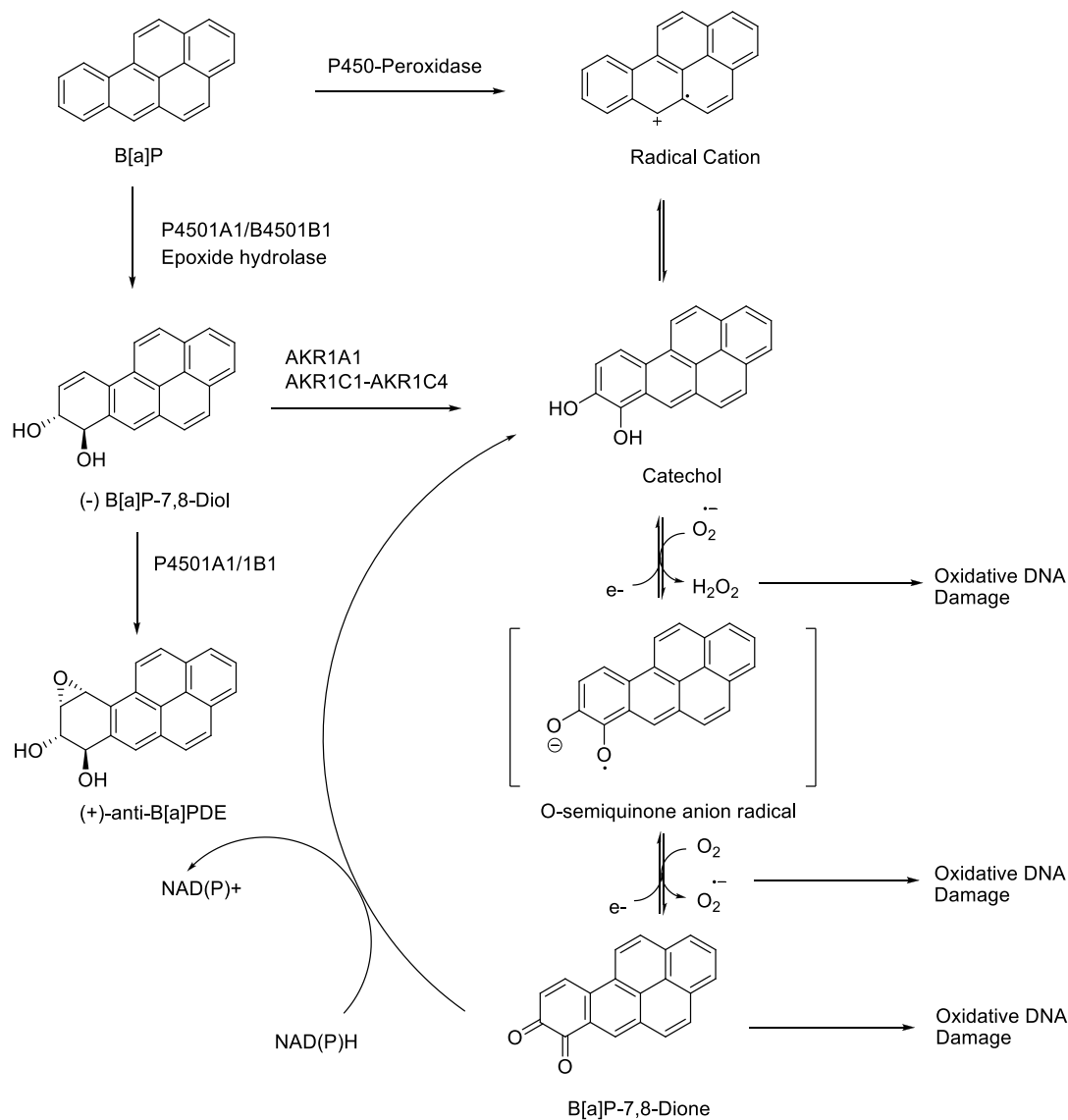
Figure 1.13. Different pathways involved in the metabolic activation of PAHs.[10]

Not only the aforementioned C8 linked dG adducts can damage DNA but also addition of PAHs to the different sites of the nucleobases leads to other lesions. For instance, bulky B[*a*]P-6-N7-dG and B[*a*]P-6-N7-dA adducts result from Michael addition of benzo[*a*]pyrene-7,8-dione (BPQ) to the N2 site of dG and dA, respectively. Metabolic activation of PAHs results in the formation of diol epoxides. These soluble intermediates react with N2 in dG and N6 in dA. The N2-dG PAH

adduct acquires an *anti* conformation, and consequently forms Watson-Crick pairing with complementary bases in the DNA helix.[51, 52]

It is difficult to study the carcinogenicity of these species experimentally due to the variety of lesions formed by even a single PAH compound. Therefore, computational approaches may clarify the structural aspect and biological implication of these species.

The binding of the highly reactive metabolic diol epoxide intermediates of fjord and bay region PAHs to the exocyclic amino groups of adenine and guanine in DNA lead to a variety of structurally different lesions that are locally distorted and cause destabilization of the DNA structure. The acquired conformations depend directly on the topology of the polycyclic aromatic ring systems and the stereochemical properties of the DNA adduct formed. The variable conformations that have been observed can be classified as: intercalation with displacement of the modified and partner bases from the interior of the double helix, intercalation without base displacement, flipping out the adducts from the duplex and formation of external minor groove conformations. The surrounding base sequence context is one of the most important limiting factors in local structural DNA perturbations and acquiring the different conformations of the PAH diol epoxide DNA.[53] Correlations between the structural properties of the series of the bay region B[*a*]P-N2-dG and N6-dA adducts, as well as selected fjord region PAH diol epoxide N6-dA adducts, have provided new insights into the factors that govern the recognition of the structural and biological characteristics of this family of DNA lesions.[54, 55]

1.3.1.3.4. Estrogens

IARC has classified endogenous human estrogens as carcinogens. Risks of estrogen exposure have been determined by experimental and epidemiological data, which strongly suggest that exposure to estrogen hormone increases the risk of cancer, especially breast and endometrial cancer.[56] Reactive intermediates derived from estrogen metabolism cause DNA damage by electrophilic and oxidative reactions leading to genotoxicity.[57] For example, equine estrogens and endogenous estrogens are oxidized to o-quinones, which are electrophiles as well as potent redox active compounds (Figure 1.14).[58]

The mechanism of action involves a genotoxic mechanism that passes through the formation of o-quinone intermediates and involves the metabolic activation of human estrogens that bind to cellular DNA causing mutations.[59, 60] Equilin and equilenin, which are known as equine estrogen, and o-quinone metabolites have hormone replacement therapy applications and can form covalent adducts with DNA.[61] Furthermore, a mechanism involving ROS derived from redox cycling between the catechol and o-quinone derivatives of endogenous human and equine estrogens also leads to oxidatively damaged DNA.[62]
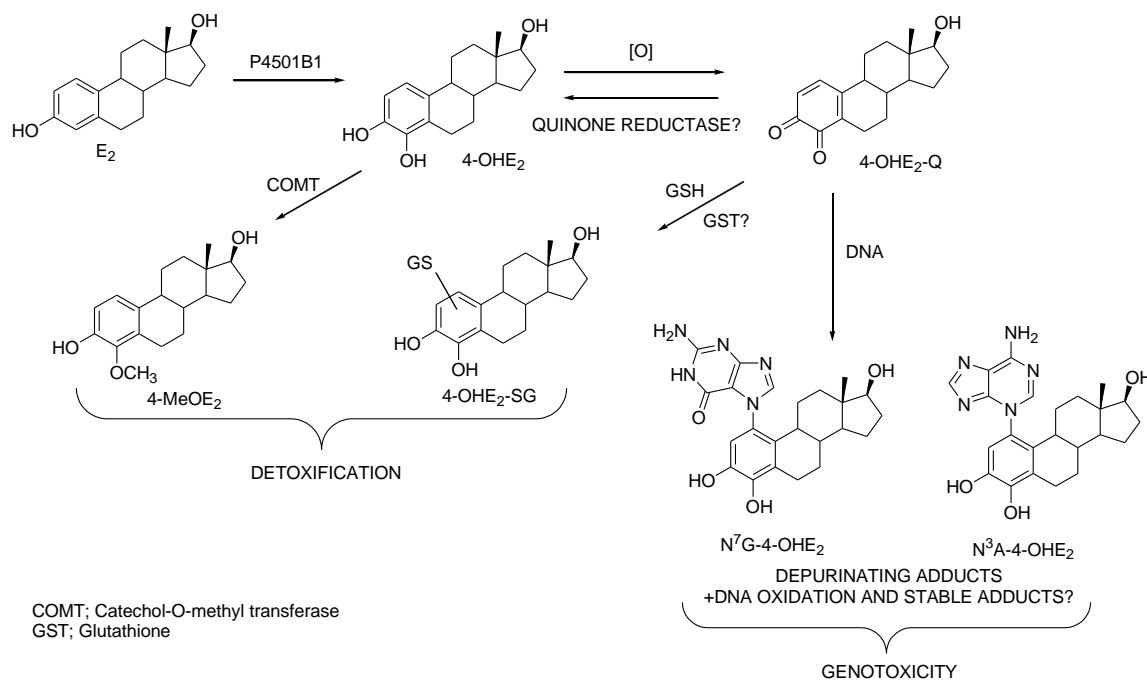
Figure 1.14. Mechanism of reaction of Estrogen with DNA bases.[10]

1.3.1.3.5. Aflatoxin

The aflatoxins are mycotoxins that are among the most naturally toxic and cancer-causing substances in animals and humans.[63] The members of the aflatoxin family are produced by a fungal *Aspergillus* species that grows in stored grains and other crops (e.g. maize) particularly when the storage environments are humid.[64, 65] Aflatoxin B1 is metabolized in the liver to an exo-8,9-epoxide that is a highly reactive intermediate that binds covalently to DNA. Among different types of Ochratoxines, ochratoxin A (OTA) can be found in cereals, coffee, wine and fruit juice making OTA the most abundant ochratoxin. It has been determined to be the most toxic species.[66] On the other hand, ochratoxin B (OTB) is not as toxic and is not found in foodstuffs. Among the different toxic effects of OTA, nephrotoxicity[67] and urothelial tract carcinogenesis in rodents and chickens are the most

prevalent[68, 69] OTA has been suggested to cause testicular cancer in young men.[69] Since OTA is carcinogenic in animals, but there is a lack of evidence for its carcinogenicity in humans, OTA is classified as a group 2B carcinogen (possibly carcinogenic to humans) by the International Agency for Research on Cancer (Figure 1.15).[65]

Ochratoxin A; R^1= Cl, R^2=H
Ochratoxin B; R^1= H, R^2=H
Ochratoxin C; R^1= Cl, R^2=CH$_2$CH$_3$

Ochratoxin α; R= Cl
Ochratoxin β; R= H
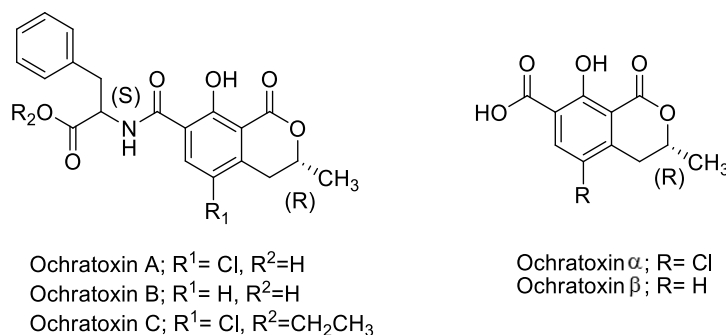
Figure 1.15. Structures of different natural ochratoxins.[65]

Experimental studies have suggested that the C8 position of 2′-deoxyguanosine (dG) is more reactive with OTA.[70-72] Specifically, the photoreaction of OTA in the presence of dG forms C-linked OTB-dG adducts.[73] The possibility of the formation of an O-linked OTA-dG adduct is less likely than C-linked OTB-dG, since there are not enough electron withdrawing groups on OTA and its derivatives for the formation of radical oxygen (Figure 1.16).[74]

24
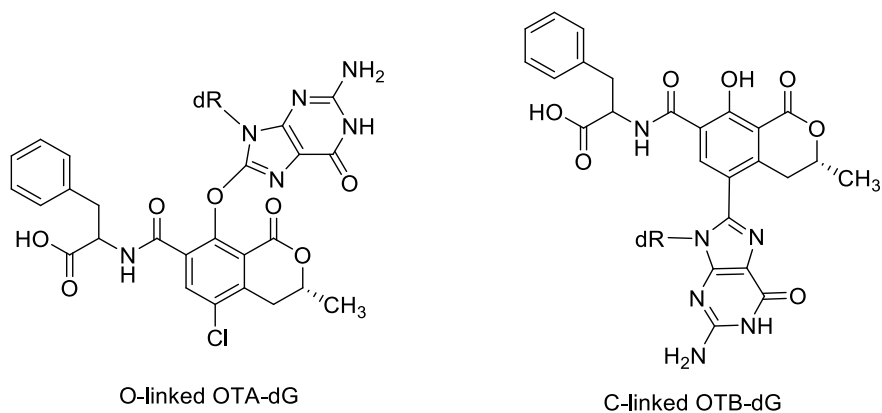
Figure 1.16. Two types of adducts formed by photochemical reaction of OTA from oxygen atom and carbon atom with C8 site of dG.[52, 65]

Computational studies on the nucleobase of the OTA adducted DNA helix of the C8-modification of 2′-deoxyguanosine (dG) have been performed.[65] The perpendicular orientation of OTA with respect to the nucleobases is observed in OTA. At the smallest level, addition of the deoxyribose sugar moiety could form a more twisted structure as well as restricting rotation about the C–C linkage at the nucleoside level. This can be rationalized by steric hindrance between OTA and the sugar. Therefore, the nucleoside preferentially adopts a *syn* conformation (by 10−20 kJ mol⁻¹), which is stabilized by an O5′–H•••N3 hydrogen bond. However, elimination of this hydrogen bond provides a better idea about the DNA environment, and the computational data show an increase in the energy level of *anti/syn* conformers (<5 kJ mol⁻¹) after removal of the non-native hydrogen bonds. According to the theoretical results at the nucleotide level,[65] the presence of the 5′-monophosphate group leads to formation of stable *syn* conformer. This can be related to stabilizing interactions between the amino group of base and the phosphate, leading to energy differences of up to 20 kJ mol⁻¹ between the

nucleotide *syn* and *anti* conformers. However, MD simulation and more specific free energy calculations elucidate the comparable stability of the *syn* and *anti* conformers of OTB-dG adduct in a specific sequence of DNA paired with cytosine.[65, 75, 76]

1.3.1.3.6. Aromatic Amines and Heterocyclic Aromatic Amines

An important class of aromatic mutagens is the aromatic amines and heterocyclic amines, which contribute to the etiology of gastrointestinal cancers.[52, 77] These complexes are produced by broiling or barbecuing meats at high temperatures. Aromatic amines (AAs) and heterocyclic aromatic amines (HAAs) are common environmental and dietary contaminants and carcinogens. The dye, chemical, and rubber manufacturing industries were major sources cancers, such as bladder tumors, until the first half of the twentieth century.[78-82] Numerous epidemiological studies have demonstrated that AAs, such as 4-aminobiphenyl (ABP), 2-naphthylamine (2-NA), and benzidine (Bz), all contaminants in aniline dyes, were found to lead to DNA damage. 2-Aminofluorene (AF) and N2-acetylaminofluorene (AAF), which were originally developed as pesticides, attracted the attention of scientists more than all other AAs. The biochemistry and biological effects of AF and AAF and the genotoxic properties of their DNA adducts have shown that they lead to lesions. According to Figure 1.17, mechanisms of the formation of dG C8-AA and HAA adducts pass though the formation of N7-dG as intermediate.[83, 84]
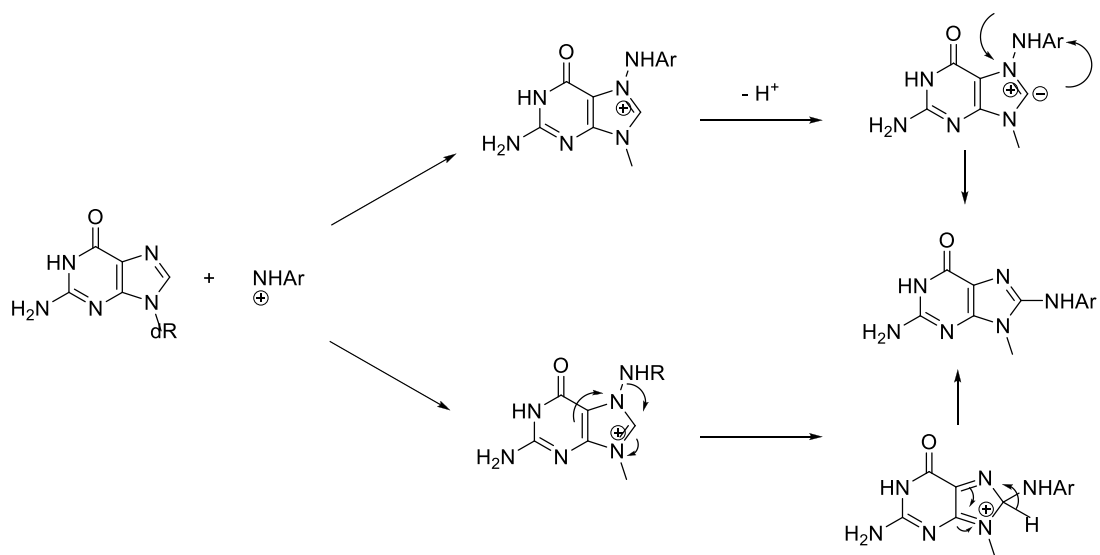
Figure 1.17. Two mechanisms for the formation of AA and HAA adducts at the C8 site of dG via an N7-dG intermediate.[10]

The causal role of HAAs in human cancer remains controversial, however, because the amounts of HAAs present in the diet are generally low. Examples of common AAs and HAAs and aromatic amine adducts are shown in Figure 1.18.[85-87]
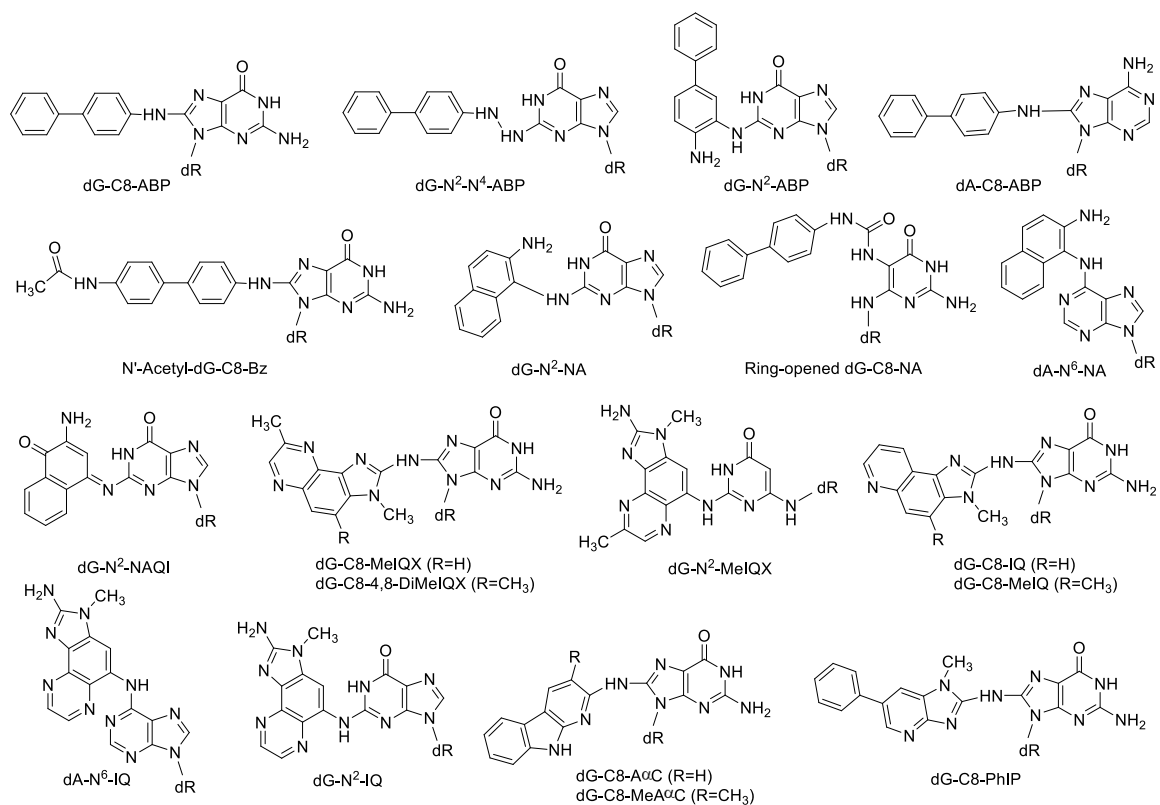
Figure 1.18. Structures of the principal AA and HAA DNA adducts[10].

Theoretical and experimental studies have revealed that there are several factors which determine the structures of aromatic amine DNA adducts. The bulkiness of adducts, their coplanarity and the type of atoms which are involved in the formation of covalent linkages between the bulky group and the nucleobase (C8, N2, etc.), are among the most significant ones. It has been shown that aromatic amine lesions can be categorized in three conformational motifs(S, B, W). However, the population balance of each conformer depends directly on the sequences surrounding the lesion site. The structural changes at the lesion site (e.g., rotamers, C8 versus N2 linkage, *etc.*) are also distinct factors that can influence the S/B/W ratios of aromatic amine adducts. These differences also

determine the nature of the conformation-specific repair and mutational outcomes.[82, 88]

1.3.1.3.7. Aristolochic Acid

Aristolochic acids are nitrophenanthrene carboxylic acid derivatives commonly found in the plants of the *Aristolochiaceae* and *Asarum* family. These compounds are also found in Chinese herbal medicines used for clinical purposes. These compounds mainly include Aristolochic Acid I AAI and Aristolochic Acid II (AAII), the first being the most abundant carcinogen.[89,90] More than 100 cases of Aristocholic acid nephropay (AAN) have been reported in Belgium. Identification of many cases of urothelial cancer in AAN patients in Belgium and the U.K highlights the potential carcinogenicity of aristolcholic acids in human beings.[89]

Aristolcholic acids attack the purine bases in DNA (guanine and adenine, Figure 1.19). The nitro group is reduced to form aristolactam, which in turn forms an intermediate nitrenium ion that is the ultimate carcinogen and leads to adducts.[91] While the aristolocholic acid-N2-dG is found to be non-mutagenic, the N6-dA adduct is highly mutagenic, predominantly leading to transversions of A to T that are rarely observed in other human tumors.[92] The dA adduct can tautomerise between the imino and amino forms, and molecular modelling, as well as NMR studies, have shown that the amino form is more dominant.[93] The thermodynamic stabilities of adducted DNA show that the aristolocholic acid-dA lesion destabilizes the helix and this destabilization is sequence dependent.[94] However, the reason for the sequence dependence is unknown. Within the DNA strand, the aristolactam moiety was found to intercalate, pushing the opposing

base (dT) out and stacking between the flanking bases of the opposing base. This leads to widening in the groove dimension and destabilization of the helix.[93, 94]
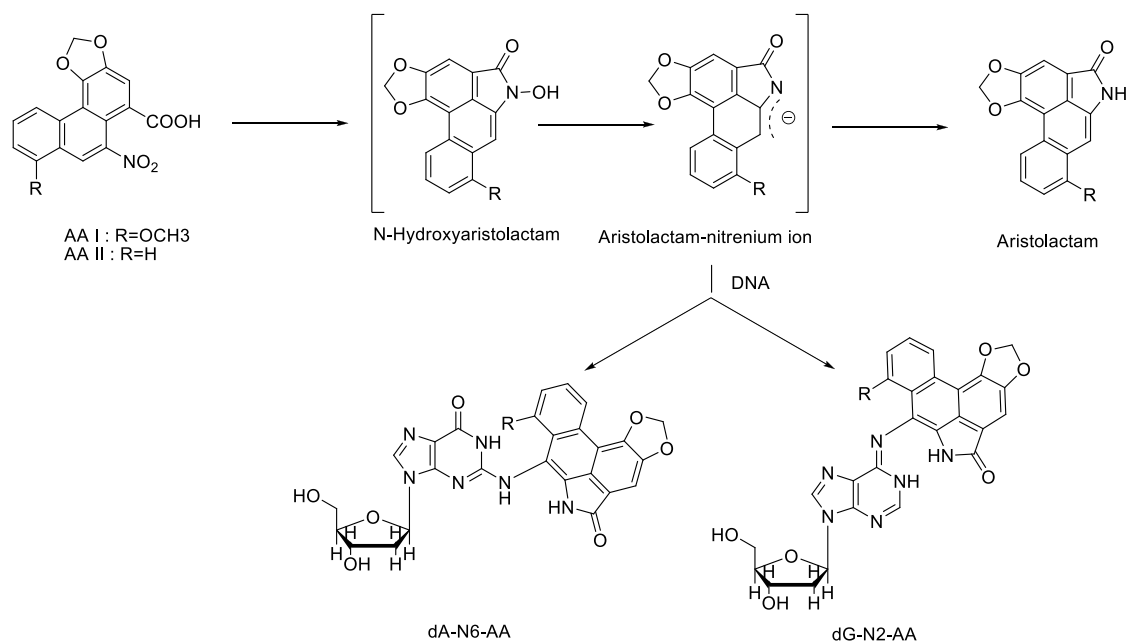
Figure 1.19. Mechanism for adduct formation by the Aristolochic Acids.[94]

As mentioned above, previous studies of different adducts have shown that the conformational preference of the adduct depends on various factors such as the site of attachment, flanking bases, opposing bases, etc. Computational studies of more adducts (nucleobases modified by phenolic, heteroaromatic as well as bi- and polycyclic moieties) can help us predict their carcinogenicity and conformational preference based on such factors. This can be done without actually synthesizing modified DNA, saving time, cost and chemicals. With such a large and increasing number of DNA lesions, it has not been possible for scientists to focus on all equally. Chemically modified purines, and in particular phenoxyl DNA adducts and their propensity to mispair, have garnered a great deal of attention. [95-98] This thesis

covers descriptions of the computational details of DNA O-linked adducts formed due to exposure to phenols. Understanding the reasons for the detrimental effects of phenolic carcinogens is the first step to understanding how these compounds lead to disease as well as to the development of strategies to combat such damage.

1.3.1.3.8. O-linked and C-linked Adducts

Phenols are organic compounds that not only contain antioxidant properties, but also have toxic pro-oxidant properties associated with aging and disease.[99] This pro-oxidant activity is interpreted by initial one-electron oxidation of phenols into reactive phenoxyl radical intermediates with peroxidase enzymes or with transition metals with redox-active properties. The generated radicals may damage lipids, proteins, and DNA.[26, 100] Chlorophenols (CPs) are persistent environmental toxic chemicals found in pesticides, disinfectants, wood preservatives, personal care formulations. Some are substantial by-products of wood pulp bleaching with chlorine.[101, 102] Experimental studies imply that metabolic activation in the case of chlorosubstituted phenol passes through a predominant pathway which involves formation of an o-quinone species ultimately forming adducts at the N2/N1 site of dG (Figure 1.20).[103]
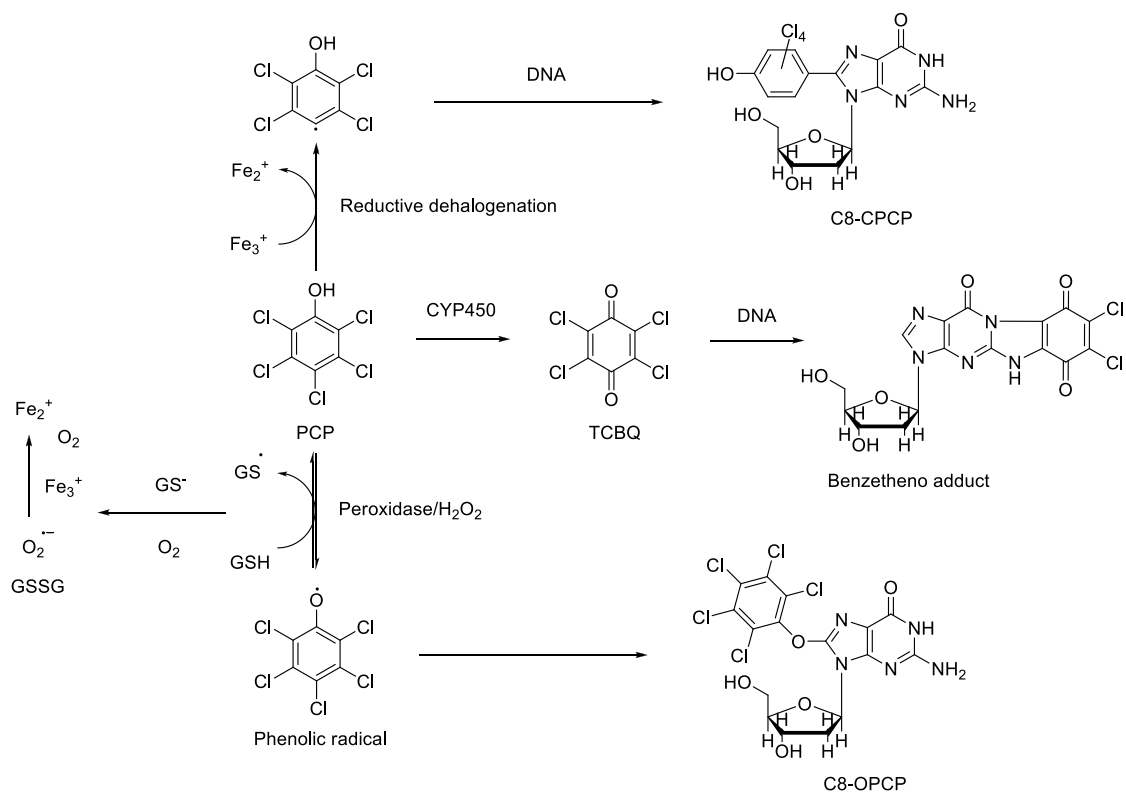
Figure 1.20. Proposed mechanisms for DNA adduct formation by PCP.[104]

Studies have suggested that the possibility of formation of phenoxyl radical intermediates in highly chlorinated species results in stable, carbon-bonded or oxygen-bonded C8-dG adducts making pentachlorophenol mutated complexes very probable (Figure 1.20 and Scheme 1.1).[104] Thus, substituted phenolic compounds display ambident reactivity at the C8 site of dG,[73] which complicates the elucidation of their mutagenic profile (Scheme 1.1).
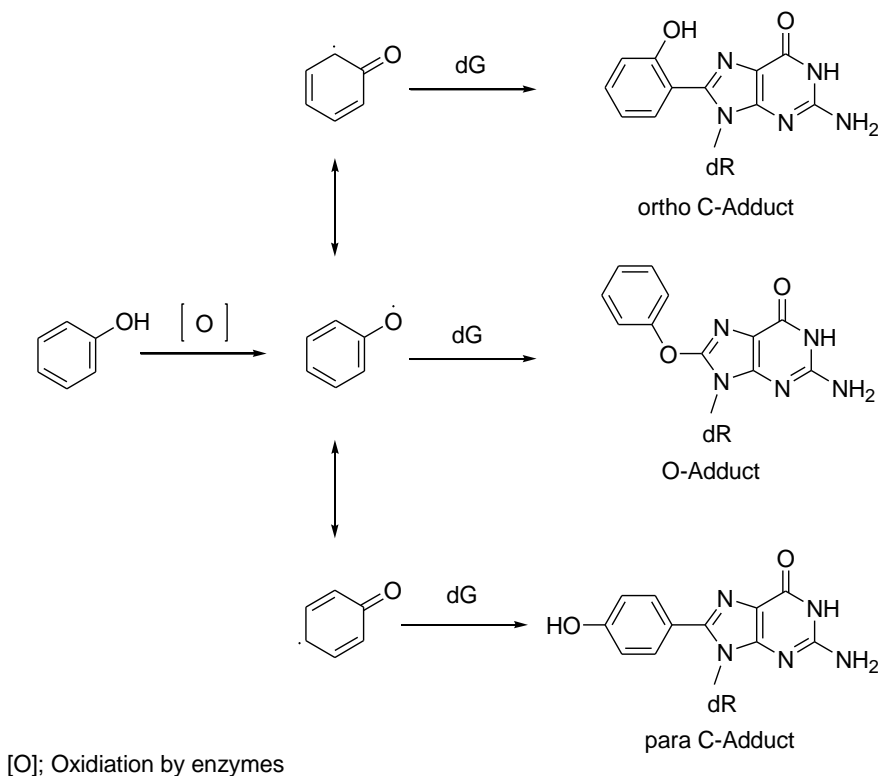
Scheme 1.1. Ambident reactivity of phenolic radicals forms O-linked and C-linked adducts with the C8 site of 2′-deoxyguanosine.[95]

Previous research has strived to uncover the structural properties of the *ortho* and *para* C-bonded dG phenolic adducts,[95] to provide greater insight into the conformational and base-pairing preferences of the adducts in biologically relevant systems. Experimental and theoretical calculations on small (nucleobase and nucleoside) models of *ortho* and *para* C-linked structures ([o-PhOH]dG and [p-PhOH]dG) elucidated a twisted *syn* conformation as the lowest energy structure of the nucleobase adducts.[95] The *ortho* adduct is planar, and contains an O–H•••N7 hydrogen bond.[95, 96, 98, 105] Furthermore, the degree of twist about the C–C linkage was determined to be dependent on the adduct (*para* > *ortho*) and solvent (polar > aprotic as modeled using implicit (PCM) solvation). The thermodynamic stability

of the C-bonded phenoxyl DNA adducts and their propensities to mispair have also been investigated.[97, 98] The conformational preference of the [o-PhOH]dG and [p-PhOH]dG lesions in two decanucleotide sequences (ODN1=5´-CCATXCTACC-3´ and ODN2=5´-GGTAGXATGG-3´ where X denotes the adduct) were determined by MD simulations. The experimentally measured melting temperatures of the duplexes indicate that both [PhOH]dG adducts destabilize the duplex when base-paired with C (the normal partner of G), but show a sequence dependent increase in duplex stability when mismatched with G, which may lead to mutagenic hotspots.[52, 106]

A combined experimental and computational approach can give us a greater understanding of the structural preference of the phenolic C8 purine adducts. For example, it is important to understand the possibility of *anti* to *syn* conversion since the presence of the *syn* (unnatural) orientation could create the potential for mispairing in the DNA strand. Additionally, the possibility of enhanced abasic site formation post-damage is a cause for concern since the effects of the phenolic adduction on the stability of the glycosidic bond are uncertain.

1.3.1.3.9. O-linked benzylic Adducts

DNA benzylation can take place by generation of benzylic radicals from N-nitrobenzylmethylamine in tobacco[107] as well as other organic species such as alcohol and some chemical compounds in plants. Since the alkylation of guanine at O6 is common and also cytotoxic,[108] it would be worthwhile to understand the mutagenicity and toxicity properties of these adducts.[109] Computational modeling of benzylic adducts by Wetmore and coworkers[110] indicate the significant flexibility of the benzylic moiety due to the presence of methylene group which

reduces the steric clashes between the ring and guanosine. In addition the CH_2 linkage increases the number of dihedrals which should be considered. On the other hand, the small size of the aromatic ring decreases the bulkiness of this adduct, ultimately resulting in a lower barrier to rotation of the aromatic ring with respect to the nucleobase. Hence, different conformational preferences has been induced in this mutated complex. T-shaped conformations and some intercalated structures are among the imperative ones. This diversity makes the mutagenic properties of this modified adduct complex. However due to the less notable bulky effect of the benzylic moiety, other factors such as the type of complementary and flanking bases and the solvation of the benzyl moiety can be considered as determining factors (Figure 1.21).[110]
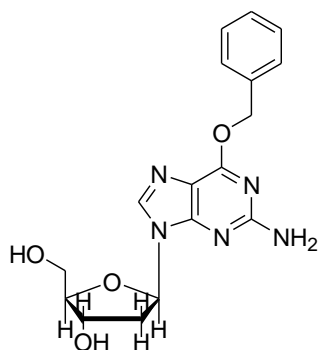
Figure 1.21. The benzylic adduct by attachment to the O6 site of Guanosine.

1.4. DNA Damage and Chemotherapeutic Applications

Assessment of mutation has different aspects. At first glance, DNA damage by genotoxic exogenous and endogenous agents implies that such damage must be avoided for maintaining the integrity of the genome while probing the cancer therapy applications of DNA lesions is also worthwhile. In other words, we may

35

wish to damage DNA in order to diminish the survival of tumor cells. Chemotherapy methods are commonly employed based on ionizing radiation which mainly causes double strand breaks. Other intracellular and intercellular signaling pathways are induced that also play a critical role in destroying tumor tissue.[111] Platinum-based compounds are commonly used in cancer chemotherapy.[112]

1.5. Experimental verses Computational Techniques

Different techniques can help scientists gain structural information on site-specific DNA lesions. The most important ones are high resolution nuclear magnetic resonance (NMR) methods, X-ray crystallography, and molecular dynamic simulation methods (MD). These techniques can provide valuable insights into the relationships between the structural features of DNA lesions and their impact on DNA replication. MD simulation, which studies the evolution in time of dynamic ensembles, provides insights into the dynamics of structural characteristics of DNA lesions. Over the past two decades, NMR methods have elucidated new insights into the structural properties of bulky PAH derived DNA adducts, adducts derived from aromatic amines and many other DNA lesions.[113, 114] These techniques have some pros and cons. Lesions in solution can be highly mobile and heterogeneous which can make studying and defining the structures of DNA lesions difficult and even impossible in some cases. Crystallography can provide outstanding resolutions of the structural properties of DNA lesions in proteins. Growing crystals of sufficient quality to yield high resolution structures (resolutions below 2 Å) is challenging. DNA-containing lesions exhibit

conformational flexibility and heterogeneity that diminish resolution.[115] Molecular dynamic simulations are an effective and complementary method that can provide essential information for the interpretation of NMR data at the atomic level. Indeed, modeling techniques can expand our knowledge about structural properties where experimental data has not yet been acquired. Dynamic trajectories provide unique information on lesion mobility and ensemble analysis using statistical mechanical methods can give us insights relating to the thermodynamics.[116]

Since very little is known about the corresponding O-linked adduct, this thesis represents a milestone to gain deeper insight into understanding the structural properties, conformational preference, and stability of O-linked adducts. Likewise, improving our perception into biological processes provides valuable information for evaluating the interaction between nucleotides in a duplex by applying MD simulations to the study of adducted DNA strands.

1.6. Thesis Approach and Summary

The purpose of this thesis is to provide a comprehensive computational overview of the structural properties and conformational preferences of oxygen-bonded C8-dG phenoxyl adducts by initial modeling using a small model and progressively increasing the size of system to the DNA helix. New understandings of the geometrical properties and conformational flexibility of phenoxyl adducts in a DNA environment is distinguished. In addition, the favored conformation of various chlorosubstituted adducts is compared and contrasted with that of unsubstituted species. The effect of protonation on adduct degradation by

destabilization of the glycosidic bond is determined. However, the direct dependence of protonation and functionalization of phenoxyl moiety with chlorine on the heterolytic cleavage of glycosidic bond is demystified. Another aspect of this project was to determine the influence of this damage when it is incorporated into a duplex, and how phenoxyl adducts may contribute to genotoxicity in cells. These theoretical results on adducted DNA strands can inspire new research developments on how to combat the detrimental effects of these adducts.

The aim of Chapter 2 is to provide information regarding the geometrical properties of unsubstituted adducts using small models such as nucleobase and nucleoside and nucleotide models by implementing density functional theory (DFT) calculations. The structural properties of unsubstituted O-linked adducts are compared and contrasted with *ortho*[95] and *para*[95] C-linked mutated complexes which were performed by Andrea Millen in the Wetmore research group.

Chapter 3 focuses on implementing the structural properties of the small models to draft a DNA model. The adduct structure is inserted into a specific site (G^3) of a 12mer *Nar*I recognition sequence (5′–CTCG^{1}G^{2}CG^{3}CCATC–3′). This will clarify different aspects of the impact of duplex environment on the conformational heterogeneity and structural favorability of unsubstituted O-linked adducts and *ortho* and *para* C-linked mutated species. Base-pairing preferences are considered in order to rationalize the results from free energy calculations.

Finally, Chapter 4 will address the stability of chloro substituted O-linked adducts at the nucleobase, nucleoside and nucleotide level. The influence of

chlorine substitution on the hydrolytic stability of protonated biaryl ether nucleoside adducts is investigated by determining the deglycosylation barrier associated with the formation of abasic sites. In both Chapters 2 and 4, the further expansion of the computational model to a deoxynucleoside monophosphate lead us to a better understanding of the *anti/syn* conformational preferences and energy barriers between conformers of adducts at this level.

Modeling the structure of modified nucleobases with phenoxyl adducts and improving the size of the system from a small nucleobase to a large DNA duplex perspective allows for comparison between small models and the double helix. This clarifies the importance of small computational models and what these results specifically tell us about conformation that can be relevant when they are applied to larger models such as the DNA double helix. Furthermore, one of the main aims with the smaller computational models is the availability of experimental evidence (provided by the Manderville research group) with which we can establish the accuracy of theoretical results, as well as facilitating the interpretation of experimental results. The Future Work section, Chapter 5, outlines other studies that can be done to determine the mutagenicity of O-linked adducts, and how we might improve the efficiency of computational systems in order to develop new methodologies which may be useful in answering important questions about related systems.

1.7. References

1.   Dahm, R., Friedrich Miescher and the Discovery of DNA. *Developmental Biology* **2005,** *278* (2), 274-288.

2.   Dahm, R., The First Discovery of DNA. *American Scientist* **2008,** *96* (4), 320-327.

3.   Avery, O. T.; Macleod, C. M.; McCarty, M., Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. *The Journal of Experimental Medicine* **1944,** *79* (2), 137-58.

4.   Watson, J. D.; Crick, F. H. C., Molecular Structure of Nucleic Acids - a Structure for Deoxyribose Nucleic Acid. *Nature* **1953,** *171* (4356), 737-738.

5.   Wilkins, M. H.; Stokes, A. R.; Wilson, H. R., Molecular Structure of Deoxypentose Nucleic Acids. *Nature* **1953,** *171* (4356), 738-40.

6.   Franklin, R. E.; Gosling, R. G., Molecular Configuration in Sodium Thymonucleate. *Nature* **1953,** *171* (4356), 740-1.

7.   Strausberg, R. L.; Levy, S.; Rogers, Y. H., Emerging DNA Sequencing Technologies for Human Genomic Medicine. *Drug Discovery Today* **2008,** *13* (13-14), 569-77.

8.   Liu, B.; Baskin, R. J.; Kowalczykowski, S. C., DNA Unwinding Heterogeneity by Recbcd Results from Static Molecules Able to Equilibrate. *Nature* **2013,** *500* (7463), 482-485.

9.   Bauer, W. R.; Crick, F. H. C.; White, J. H., Supercoiled DNA. *Scientific American* **1980,** *243* (1), 118-&.

10.   Broyde, N. E. G. a. S., *The Chemical Biology of DNA Damage.* © 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim: New York University.

11.   Genomics,   S.   o.   B.   M.   B.   a. Http://Biology.Kenyon.Edu/Bmb/Biomolecules.Htm the Structures of DNA and Rna.

12.   Watson, J. D.; Crick, F. H. C., Molecular Structure of Nucleic Acids - a Structure for Deoxyribose Nucleic Acid *Clinical Orthopaedics and Related Research* **2007,** (462), 3-5.

13.   Elson, D.; Chargaff, E., On the Desoxyribonucleic Acid Content of Sea Urchin Gametes. *Experientia* **1952,** *8* (4), 143-145.

14.   Franklin, R. E.; Gosling, R. G., Molecular Configuration in Sodium Thymonucleate. *Nature* **1953,** *171* (4356), 740-741.

15.   Crick, F. H. C., Linking Numbers and Nucleosomes. *Proceedings of the National Academy of Sciences of the United States of America* **1976,** *73* (8), 2639-2643.

16.   Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R., Molecular Structure of Nucleic Acids: Molecular Structure of Deoxypentose Nucleic Acids. *Nature* **1953,** *171* (4356), 738-740.

17.   Rich, A.; Zhang, S., Z-DNA: The Long Road to Biological Function. *Nature Reviews Genetics* **2003,** *4* (7), 566-572.

18.   Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E., Crystal-Structure Analysis of a Complete Turn of B-DNA. *Nature* **1980,** *287* (5784), 755-758.

19.   Pabo, C. O.; Sauer, R. T., Protein-DNA Recognition. *Annual Review of Biochemistry* **1984,** *53*, 293-321.

20.   Vaidyanathan, V. G.; Liang, F. T.; Beard, W. A.; Shock, D. D.; Wilson, S. H.; Cho, B. P., Insights into the Conformation of Aminofluorene-Deoxyguanine Adduct in a DNA Polymerase Active Site. *Journal of Biological Chemistry* **2013,** *288* (32), 23573-23585.

21.   Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E., Crystal Structure Analysis of a Complete Turn of B-DNA. *Nature* **1980,** *287* (5784), 755-758.

22.   Spies, M. A.; Schowen, R. L., The Trapping of a Spontaneously "Flipped-out" Base from Double Helical Nucleic Acids by Host–Guest Complexation with B-Cyclodextrin: The Intrinsic Base-Flipping Rate Constant for DNA and Rna. *Journal of the American Chemical Society* **2002,** *124* (47), 14049-14053.

23.   Best, B. P., Nuclear DNA Damage as a Direct Cause of Aging. *Rejuvenation Research* **2009,** *12* (3), 199-208.

24.   Scharer, O. D., Chemistry and Biology of DNA Repair. *Angewandte Chemie International Edition* **2003,** *42* (26), 2946-74.

25.   Greenberg, M. M., The Formamidopyrimidines: Purine Lesions Formed in Competition with 8-Oxopurines from Oxidative Stress. *Accounts of Chemical Research* **2012,** *45* (4), 588-597.

26.   Burrows, C. J.; Muller, J. G., Oxidative Nucleobase Modifications Leading to Strand Scission. *Chemical Reviews* **1998,** *98* (3), 1109-1152.

27.   Greenberg, M. M., Investigating Nucleic Acid Damage Processes Via Independent Generation of Reactive Intermediates. *Chemical Research in Toxicology* **1998,** *11* (11), 1235-48.

28.   Pogozelski, W. K.; Tullius, T. D., Oxidative Strand Scission of Nucleic Acids: Routes Initiated by Hydrogen Abstraction from the Sugar Moiety. *Chemical Reviews* **1998,** *98* (3), 1089-1108.

29.   Wallace, S. S., Enzymatic Processing of Radiation-Induced Free Radical Damage in DNA. *Radiation Research* **1998,** *150* (5), S60-79.

30.   Wang, X.; Ira, G.; Tercero, J. A.; Holmes, A. M.; Diffley, J. F.; Haber, J. E., Role of DNA Replication Proteins in Double-Strand Break-Induced Recombination in Saccharomyces Cerevisiae. *Molecular and Cellular Biology* **2004,** *24* (16), 6891-9.

31.   Piazza, A.; Serero, A.; Boule, J. B.; Legoix-Ne, P.; Lopes, J.; Nicolas, A., Stimulation of Gross Chromosomal Rearrangements by the Human Ceb1 and Ceb25 Minisatellites in Saccharomyces Cerevisiae Depends on G-Quadruplexes or Cdc13. *Plos Genetics* **2012,** *8* (11), 17.

32.   Niedernhofer, L. J.; Riley, M.; Schnetz-Boutaud, N.; Sanduwaran, G.; Chaudhary, A. K.; Reddy, G. R.; Marnett, L. J., Temperature-Dependent Formation of a Conjugate between Tris(Hydroxymethyl)Aminomethane Buffer and the Malondialdehyde-DNA Adduct Pyrimidopurinone. *Chemical Research in Toxicology* **1997,** *10* (5), 556-61.

33.   Lawley, P. D.; Phillips, D. H., DNA Adducts from Chemotherapeutic Agents. *Mutation Research* **1996,** *355* (1-2), 13-40.

34.   Kozekov, I. D.; Nechev, L. V.; Sanchez, A.; Harris, C. M.; Lloyd, R. S.; Harris, T. M., Interchain Cross-Linking of DNA Mediated by the Principal Adduct of Acrolein. *Chemical Research in Toxicology* **2001,** *14* (11), 1482-1485.

35.   Gates, K. S., An Overview of Chemical Processes That Damage Cellular DNA: Spontaneous Hydrolysis, Alkylation, and Reactions with Radicals. *Chemical Research in Toxicology* **2009,** *22* (11), 1747-1760.

36.   Stivers, J. T.; Jiang, Y. L., A Mechanistic Perspective on the Chemistry of DNA Repair Glycosylases. *Chemical Reviews* **2003,** *103* (7), 2729-2759.

37.   Zoltewicz, J. A.; Clark, D. F.; Sharpless, T. W.; Grahe, G., Kinetics and Mechanism of the Acid-Catalyzed Hydrolysis of Some Purine Nucleosides. *Journal of the American Chemical Society* **1970,** *92* (6), 1741-1750.

38.   Lindahl, T., Instability and Decay of the Primary Structure of DNA. *Nature* **1993,** *362* (6422), 709-15.

39.   Beckman, K. B.; Ames, B. N., Oxidative Decay of DNA. *The Journal of Biological Chemistry* **1997,** *272* (32), 19633-6.

40.   McCarty, M., Discovering Genes Are Made of DNA. *Nature* **2003,** *421* (6921), 406.

41.   Ravanat, J. L.; Douki, T.; Cadet, J., Direct and Indirect Effects of Uv Radiation on DNA and Its Components. *Journal of Photochemistry Photobiology B* **2001,** *63* (1-3), 88-102.

42.   Jamieson, E. R.; Lippard, S. J., Structure, Recognition, and Processing of Cisplatin-DNA Adducts. *Chemical Reviews* **1999,** *99* (9), 2467-2498.

43.   Jena, N. R.; Mishra, P. C., Formation of Ring-Opened and Rearranged Products of Guanine: Mechanisms and Biological Significance. *Free Radical Biology and Medicine* **2012,** *53* (1), 81-94.

44.   Ravanat, J. L.; Douki, T.; Duez, P.; Gremaud, E.; Herbert, K.; Hofer, T.; Lasserre, L.; Saint-Pierre, C.; Favier, A.; Cadet, J., Cellular Background Level of 8-Oxo-7,8-Dihydro-2 '-Deoxyguanosine: An Isotope Based Method to Evaluate Artefactual Oxidation of DNA During Its Extraction and Subsequent Work-Up. *Carcinogenesis* **2002,** *23* (11), 1911-1918.

45.   Chao, M.-R.; Yen, C.-C.; Hu, C.-W., Prevention of Artifactual Oxidation in Determination of Cellular 8-Oxo-7,8-Dihydro-2'-Deoxyguanosine by Isotope-Dilution Lc-Ms/Ms with Automated Solid-Phase Extraction. *Free Radical Biology and Medicine* **2008,** *44* (3), 464-473.

46.   Collins, A. R.; Cadet, J.; Moller, L.; Poulsen, H. E.; Vina, J., Are We Sure We Know How to Measure 8-Oxo-7,8-Dihydroguanine in DNA from Human Cells? *Archives of Biochemistry and Biophysics* **2004,** *423* (1), 57-65.

47.   Floyd, R. A.; Watson, J. J.; Wong, P. K.; Altmiller, D. H.; Rickard, R. C., Hydroxyl Free Radical Adduct of Deoxyguanosine: Sensitive Detection and Mechanisms of Formation. *Free radical research communications* **1986,** *1* (3), 163-72.

48.   Eot-Houllier, G.; Gonera, M.; Gasparutto, D.; Giustranti, C.; Sage, E., Interplay between DNA N-Glycosylases/Ap Lyases at Multiply Damaged Sites and Biological Consequences. *Nucleic Acids Research* **2007,** *35* (10), 3355-3366.

49.   Chakravarti, D.; Pelling, J. C.; Cavalieri, E. L.; Rogan, E. G., Relating Aromatic Hydrocarbon-Induced DNA-Adducts and C-H-Ras Mutations in Mouse Skin Papillomas - the Role of Apurinic Sites. *Proceedings of the National Academy of Sciences of the United States of America* **1995,** *92* (22), 10422-10426.

50.   Casale, G. P.; Singhal, M.; Bhattacharya, S.; RamaNathan, R.; Roberts, K. P.; Barbacci, D. C.; Zhao, J.; Jankowiak, R.; Gross, M. L.; Cavalieri, E. L.; Small, G. J.; Rennard, S. I.; Mumford, J. L.; Shen, M., Detection and Quantification of Depurinated Benzo[a]Pyrene-Adducted DNA Bases in the Urine of Cigarette Smokers and Women Exposed to Household Coal Smoke. *Chemical Research in Toxicology* **2001,** *14* (2), 192-201.

51.   Cai, Y.; Wang, L. H.; Ding, S. A.; Schwaid, A.; Geacintov, N. E.; Broyde, S., A Bulky DNA Lesion Derived from a Highly Potent Polycyclic Aromatic Tumorigen Stabilizes Nucleosome Core Particle Structure. *Biochemistry* **2010,** *49* (46), 9943-9945.

52.   Millen, A. L.; Sharma, P.; Wetmore, S. D., C8-Linked Bulky Guanosine DNA Adducts: Experimental and Computational Insights into Adduct Conformational Preferences and Resulting Mutagenicity. *Future Medicinal Chemistry* **2012,** *4* (15), 1981-2007.

53.   Donny-Clark, K.; Broyde, S., Influence of Local Sequence Context on Damaged Base Conformation in Human DNA Polymerase I: Molecular Dynamics Studies of Nucleotide Incorporation Opposite a Benzo[a]Pyrene-Derived Adenine Lesion. *Nucleic Acids Research* **2009,** *37* (21), 7095-7109.

54.	Koreeda, M.; Moore, P. D.; Wislocki, P. G.; Levin, W.; Conney, A. H.; Yagi, H.; Jerina, D. M., Binding of Benzo a Pyrene 7,8-Diol-9,10-Epoxides to DNA, Rna, and Protein of Mouse Skin Occurs with High Stereoselectivity. *Science* **1978,** *199* (4330), 778-780.

55.	Katz, A. K.; Carrell, H. L.; Glusker, J. P., Dibenzo[a,L]Pyrene (Dibenzo[Def,P]Chrysene): Fjord-Region Distortions. *Carcinogenesis* **1998,** *19* (9), 1641-1648.

56.	Chen, W. Y., Exogenous and Endogenous Hormones and Breast Cancer. *Best Practice & Research Clinical Endocrinology & Metabolism* **2008,** *22* (4), 573-585.

57.	Liehlr, J. G., Genotoxicity of the Steroidal Oestrogens Oestrone and Oestradiol: Possible Mechanism of Uterine and Mammary Cancer Development. *Human Reproduction Update* **2001,** *7* (3), 273-281.

58.	Bolton, J. L.; Trush, M. A.; Penning, T. M.; Dryhurst, G.; Monks, T. J., Role of Quinones in Toxicology. *Chemical Research in Toxicology* **2000,** *13* (3), 135-160.

59.	Okamoto, Y.; Chou, P. H.; Kim, S. Y.; Suzuki, N.; Laxmi, Y. R. S.; Okamoto, K.; Liu, X. P.; Matsuda, T.; Shibutani, S., Oxidative DNA Damage in Xpc-Knockout and Its Wild Mice Treated with Equine Estrogen. *Chemical Research in Toxicology* **2008,** *21* (5), 1120-1124.

60.	Ozcagli, E.; Sardas, S.; Biri, A., Assessment of DNA Damage in Postmenopausal Women under Hormone Replacement Therapy. *Maturitas* **2005,** *51* (3), 280-285.

61.	Malins, D. C.; Anderson, K. M.; Jaruga, P.; Ramsey, C. R.; Gilman, N. K.; Green, V. M.; Rostad, S. W.; Emerman, J. T.; Dizdaroglu, M., Oxidative Changes in the DNA of Stroma and Epithelium from the Female Breast - Potential Implications for Breast Cancer. *Cell Cycle* **2006,** *5* (15), 1629-1632.

62.	Benz, C. C.; Yau, C., Ageing, Oxidative Stress and Cancer: Paradigms in Parallax. *Nature Reviews Cancer* **2008,** *8* (11), 875-879.

63.	Wogan, G. N.; Hecht, S. S.; Felton, J. S.; Conney, A. H.; Loeb, L. A., Environmental and Chemical Carcinogenesis. *Seminars in Cancer Biology* **2004,** *14* (6), 473-486.

64.	Adetuniji, M. C.; Atanda, O. O.; Ezekiel, C. N.; Dipeolu, A. O.; Uzochukwu, S. V. A.; Oyedepo, J.; Chilaka, C. A., Distribution of Mycotoxins and Risk Assessment of

Maize Consumers in Five Agro-Ecological Zones of Nigeria. *European Food Research and Technology* **2014,** *239* (2), 287-296.

65.   Sharma, P.; Manderville, R. A.; Wetmore, S. D., Modeling the Conformational Preference of the Carbon-Bonded Covalent Adduct Formed Upon Exposure of 2′-Deoxyguanosine to Ochratoxin A. *Chemical Research in Toxicology* **2013,** *26* (5), 803-816.

66.   Raad, F.; Nasreddine, L.; Hilan, C.; Bartosik, M.; Parent-Massin, D., Dietary Exposure to Aflatoxins, Ochratoxin a and Deoxynivalenol from a Total Diet Study in an Adult Urban Lebanese Population. *Food and Chemical Toxicology* **2014,** *73* (0), 35-43.

67.   Mally, A.; Dekant, W., Mycotoxins and the Kidney: Modes of Action for Renal Tumor Formation by Ochratoxin a in Rodents. *Molecular Nutrition & Food Research* **2009,** *53* (4), 467-478.

68.   Huff, W. E.; Wyatt, R. D.; Hamilton, P. B., Nephrotoxicity of Dietary Ochratoxin-a in Broiler Chickens. *Applied Microbiology* **1975,** *30* (1), 48-51.

69.   Schwartz, G. G., Hypothesis: Does Ochratoxin a Cause Testicular Cancer? *Cancer Causes & Control* **2002,** *13* (1), 91-100.

70.   Mantle, P. G.; Faucet-Marquis, V.; Manderville, R. A.; Squillaci, B.; Pfohl-Leszkowicz, A., Structures of Covalent Adducts between DNA and Ochratoxin A: A New Factor in Debate About Genotoxicity and Human Risk Assessment. *Chemical Research in Toxicology* **2010,** *23* (1), 89-98.

71.   Obrecht-Pflumio, S.; Dirheimer, G., In Vitro DNA and Dgmp Adducts Formation Caused by Ochratoxin A. *Chemico-Biological Interactions* **2000,** *127* (1), 29-44.

72.   Obrecht-Pflumio, S.; Dirheimer, G., Horseradish Peroxidase Mediates DNA and Deoxyguanosine 3 '-Monophosphate Adduct Formation in the Presence of Ochratoxin A. *Archives of Toxicology* **2001,** *75* (10), 583-590.

73.   Dai, J.; Wright, M. W.; Manderville, R. A., Ochratoxin a Forms a Carbon-Bonded C8-Deoxyguanosine Nucleoside Adduct: Implications for C8 Reactivity by a Phenolic Radical. *Journal of the American Chemical Society* **2003,** *125* (13), 3716-3717.

74.   Faucet, V.; Pfohl-Leszkowicz, A.; Dai, J.; Castegnaro, M.; Manderville, R. A., Evidence for Covalent DNA Adduction by Ochratoxin a Following Chronic Exposure

to Rat and Subacute Exposure to Pig. *Chemical Research in Toxicology* **2004,** *17* (9), 1289-1296.

75.   Omumi, A.; Millen, A. L.; Wetmore, S. D.; Manderville, R. A., Fluorescent Properties and Conformational Preferences of C-Linked Phenolic-DNA Adducts. *Chemical Research in Toxicology* **2011,** *24* (10), 1694-709.

76.   Manderville, R. A., Structural and Biological Impact of Radical Addition Reactions with DNA Nucleobases. Richard, J. P., Ed. 2009; Vol. 43, pp 177-218.

77.   Poirier, M. C., Chemical-Induced DNA Damage and Human Cancer Risk. *Discovery Medicine* **2012,** *77*, 283-288.

78.   Donny-Clark, K.; Broyde, S., Influence of Local Sequence Context on Damaged Base Conformation in Human DNA Polymerase Iota: Molecular Dynamics Studies of Nucleotide Incorporation Opposite a Benzo a Pyrene-Derived Adenine Lesion. *Nucleic Acids Research* **2009,** *37* (21), 7095-7109.

79.   Jain, N.; Reshetnyak, Y. K.; Gao, L.; Chiarelli, M. P.; Cho, B. P., Fluorescence Probing of Aminofluorene-Induced Conformational Heterogeneity in DNA Duplexes. *Chemical Research in Toxicology* **2008,** *21* (2), 445-452.

80.   Liang, F. T.; Meneni, S.; Cho, B. P., Induced Circular Dichroism Characteristics as Conformational Probes for Carcinogenic Aminofluorene-DNA Adducts. *Chemical Research in Toxicology* **2006,** *19* (8), 1040-1043.

81.   Bichara, M.; Fuchs, R. P. P., DNA-Binding and Mutation Spectra of the Carcinogen N-2-Aminofluorene in Escherichia-Coli - a Correlation between the Conformation of the Premutagenic Lesion and the Mutation Specificity. *Journal of Molecular Biology* **1985,** *183* (3), 341-351.

82.   Brown, K.; Hingerty, B. E.; Guenther, E. A.; Krishnan, V. V.; Broyde, S.; Turteltaub, K. W.; Cosman, M., Solution Structure of the 2-Amino-1-Methyl-6-Phenylimidazo 4,5-B Pyridine C8-Deoxyguanosine Adduct in Duplex DNA. *Proceedings of the National Academy of Sciences of the United States of America* **2001,** *98* (15), 8507-8512.

83.   Humphreys, W. G.; Kadlubar, F. F.; Guengerich, F. P., Mechanism of C8 Alkylation of Guanine Residues by Activated Arylamines - Evidence for Initial Adduct Formation at the N7 Position. *Proceedings of the National Academy of Sciences of the United States of America* **1992,** *89* (17), 8278-8282.

84.   Kennedy, S. A.; Novak, M.; Kolb, B. A., Reactions of Ester Derivatives of Carcinogenic N-(4-Biphenylyl)Hydroxylamine and the Corresponding Hydroxamic Acid with Purine Nucleosides. *Journal of the American Chemical Society* **1997,** *119* (33), 7654-7664.

85.   Broyde, S.; Hingerty, B., Conformation of 2-Aminofluorene-Modified DNA. *Biopolymers* **1983,** *22* (11), 2423-2441.

86.   Bonala, R.; Torres, M. C.; Iden, C. R.; Johnson, F., Synthesis of the Phip Adduct of 2 '-Deoxyguanosine and Its Incorporation into Oligomeric DNA. *Chemical Research in Toxicology* **2006,** *19* (6), 734-738.

87.   Broyde, S.; Wang, L. H.; Zhang, L.; Rechkoblit, O.; Geacintov, N. E.; Patel, D. J., DNA Adduct Structure-Function Relationships: Comparing Solution with Polymerase Structures. *Chemical Research in Toxicology* **2008,** *21* (1), 45-52.

88.   Patel, D. J.; Mao, B.; Gu, Z. T.; Hingerty, B. E.; Gorin, A.; Basu, A. K.; Broyde, S., Nuclear Magnetic Resonance Solution Structures of Covalent Aromatic Amine-DNA Adducts and Their Mutagenic Relevance. *Chemical Research in Toxicology* **1998,** *11* (5), 391-407.

89.   Arlt, V. M.; Stiborova, M.; Schmeiser, H. H., Aristolochic Acid as a Probable Human Cancer Hazard in Herbal Remedies: A Review. *Mutagenesis* **2002,** *17* (4), 265-277.

90.   Debelle, F. D.; Nortier, J. L.; De Prez, E. G.; Garbar, C. H.; Vienne, A. R.; Salmon, I. J.; Deschodt-Lanckman, M. M.; Vanherweghem, J.-L., Aristolochic Acids Induce Chronic Renal Failure with Interstitial Fibrosis in Salt-Depleted Rats. *Journal of the American Society of Nephrology* **2002,** *13* (2), 431-436.

91.   Nitzsche, D.; Melzig, M. F.; Arlt, V. M., Evaluation of the Cytotoxicity and Genotoxicity of Aristolochic Acid I – a Component of Aristolochiaceae Plant Extracts Used in Homeopathy. *Environmental Toxicology and Pharmacology* **2013,** *35* (2), 325-334.

92.   Arlt, V. M.; Stiborova, M.; vom Brocke, J.; Simoes, M. L.; Lord, G. M.; Nortier, J. L.; Hollstein, M.; Phillips, D. H.; Schmeiser, H. H., Aristolochic Acid Mutagenesis: Molecular Clues to the Aetiology of Balkan Endemic Nephropathy-Associated Urothelial Cancer. *Carcinogenesis* **2007,** *28* (11), 2253-2261.

93.   Lukin, M.; Zaliznyak, T.; Johnson, F.; de los Santos, C., Structure and Stability of DNA Containing an Aristolactam Ii-Da Lesion: Implications for the Ner Recognition of Bulky Adducts. *Nucleic Acids Research* **2012,** *40* (6), 2759-2770.

94.   Mark, L.; Tanya, Z.; Francis, J.; Carlos de los, S., Structure and Stability of DNA Containing an Aristolactam Ii-Da Lesion: Implications for the Ner Recognition of Bulky Adducts. *Nucleic Acids Research* **2012,** *40* (6), 2759-2770.

95.   Millen, A. L.; McLaughlin, C. K.; Sun, K. M.; Manderville, R. A.; Wetmore, S. D., Computational and Experimental Evidence for the Structural Preference of Phenolic C-8 Purine Adducts. *Journal of Physical Chemistry A* **2008,** *112* (16), 3742-3753.

96.   Millen, A. L.; Manderville, R. A.; Wetmore, S. D., Conformational Flexibility of C8-Phenoxyl-2 '-Deoxyguanosine Nucleotide Adducts. *Journal of Physical Chemistry B* **2010,** *114* (12), 4373-4382.

97.   Millen, A. L.; Kamenz, B. L.; Leavens, F. M. V.; Manderville, R. A.; Wetmore, S. D., Conformational Flexibility of C8-Phenoxylguanine Adducts in Deoxydinucleoside Monophosphates. *Journal of Physical Chemistry B* **2011,** *115* (44), 12993-13002.

98.   Millen, A. L.; Churchill, C. D. M.; Manderville, R. A.; Wetmore, S. D., Effect of Watson-Crick and Hoogsteen Base Pairing on the Conformational Stability of C8-Phenoxyl-2 '-Deoxyguanosine Adducts. *Journal of Physical Chemistry B* **2010,** *114* (40), 12995-13004.

99.   Bolton, J. L., Quinoids, Quinoid Radicals, and Phenoxyl Radicals Formed from Estrogens and Antiestrogens. *Toxicology* **2002,** *177* (1), 55-65.

100.   Dix, T. A.; Aikens, J., Mechanisms and Biological Relevance of Lipid Peroxidation Initiation. *Chemical Research in Toxicology* **1993,** *6* (1), 2-18.

101.   Sturla, S. J., DNA Adduct Profiles: Chemical Approaches to Addressing the Biological Impact of DNA Damage from Small Molecules. *Current Opinion in Chemical Biology* **2007,** *11* (3), 293-299.

102.   Dai, J.; Sloat, A. L.; Wright, M. W.; Manderville, R. A., Role of Phenoxyl Radicals in DNA Adduction by Chlorophenol Xenobiotics Following Peroxidase Activation. *Chemical Research in Toxicology* **2005,** *18* (4), 771-779.

103.   Manderville, R. A., Ambient Reactivity of Phenoxyl Radicals in DNA Adduction. *Canadian Journal of Chemistry-Revue Canadienne De Chimie* **2005,** *83* (9), 1261-1267.

104.   Dai, J.; Wright, M. W.; Manderville, R. A., An Oxygen-Bonded C8-Deoxyguanosine Nucleoside Adduct of Pentachlorophenol by Peroxidase

Activation: Evidence for Ambient C8 Reactivity by Phenoxyl Radicals. *Chemical Research in Toxicology* **2003,** *16* (7), 817-821.

105.   Schlitt, K. M.; Sun, K. W.; Paugh, R. J.; Millen, A. L.; Navarro-Whyte, L.; Wetmore, S. D.; Manderville, R. A., Concerning the Hydrolytic Stability of 8-Aryl-2'-Deoxyguanosine Nucleoside Adducts: Implications for Abasic Site Formation at Physiological Ph. *The Journal of Organic Chemistry* **2009,** *74* (16), 5793-802.

106.   Omumi, A.; Millen, A. L.; Wetmore, S. D.; Manderville, R. A., Fluorescent Properties and Conformational Preferences of C-Linked Phenolic-DNA Adducts. *Chemical Research in Toxicology* **2011,** *24* (10), 1694-1709.

107.   Peterson, L. A., N-Nitrosobenzylmethylamine Is Activated to a DNA Benzylating Agent in Rats (Vol 10 Pg 19, 1997). *Chemical Research in Toxicology* **2000,** *13* (12), 1360-1360.

108.   Guarnieri, M.; Biser-Rohrbaugh, A.; Tyler, B. M.; Gabikian, P.; Bunton, T. E.; Wu, Q. Z.; Weingart, J.; Carson, B. S., Toxicity of Intracranial and Intraperitoneal O6-Benzyl Guanine in Combination with Bcnu Delivered Locally in a Mouse Model. *Cancer Chemotherapy and Pharmacology* **2002,** *50* (5), 392-396.

109.   Autrup, H.; Stoner, G. D., Metabolism of N-Nitrosamines by Cultured Human and Rat Esophagus. *Cancer Research* **1982,** *42* (4), 1307-1311.

110.   Wilson, K. A.; Wetmore, S. D., Complex Conformational Heterogeneity of the Highly Flexible O6-Benzyl-Guanine DNA Adduct. *Chemical Research in Toxicology* **2014,** *27* (7), 1310-1325.

111.   Prise, K. M.; Schettino, G.; Folkard, M.; Held, K. D., New Insights on Cell Death from Radiation Exposure. *Lancet Oncology* **2005,** *6* (7), 520-528.

112.   Wang, D.; Lippard, S. J., Cellular Processing of Platinum Anticancer Drugs. *Nature Reviews Drug Discovery* **2005,** *4* (4), 307-320.

113.   Geacintov, N. E.; Cosman, M.; Hingerty, B. E.; Amin, S.; Broyde, S.; Patel, D. J., Nmr Solution Structures of Stereoisomeric Covalent Polycyclic Aromatic Carcinogen−DNA Adducts: Principles, Patterns, and Diversity. *Chemical Research in Toxicology* **1997,** *10* (2), 111-146.

114.   Lukin, M.; de los Santos, C., Nmr Structures of Damaged DNA. *Chemical Reviews* **2006,** *106* (2), 607-686.

115.   Guengerich, F. P., Interactions of Carcinogen-Bound DNA with Individual DNA Polymerases. *Chemical Reviews* **2006,** *106* (2), 420-452.

116.   Wang, L. H.; Yu, X. Y.; Hu, P.; Broyde, S.; Zhang, Y. K., A Water-Mediated and Substrate-Assisted Catalytic Mechanism for Sulfolobus Solfataricus DNA Polymerase Iv. *Journal of the American Chemical Society* **2007,** *129* (15), 4731-4737.

Chapter 2. Insights into the Main Features of Damaged DNA O-linked Adducts: A DFT Study on the Conformational Behavior and Mutagenicity Associated with Unsubstituted Phenolic Carcinogens and Their Structural Difference with *ortho* and *para* C-linked Phenolic Adducts

2.1. Introduction:

Since DNA damage can alter genome integrity, understanding DNA damage mechanisms at the molecular level is of profound interest in chemistry, biology, toxicology, and medicine. The chemical bonds in DNA are exceptionally stable and their relative reactivity with electrophiles, nucleophiles, and light is low. The aqueous environment of the cell can alter the structure of DNA by formation of covalent bonds on nucleobases, dimers from the initial radicals, and other mutated complexes. The chemical modification of only a single nucleobase by various endogenous and exogenous factors can cause DNA damage. As explained in detail in chapter 1, damage to DNA can be classified in terms of base modifications, strand breaks, inter- or intrastrand crosslinks, and DNA–protein crosslinks.[1] In the three billion nucleotide human genome, DNA damage can have a catastrophic effect on the life of an individual. Due to loss of nucleobase functionality, damage to the nucleobases is a distinct and prevalent cause of DNA mutation. Among the four canonical DNA bases, guanine has the lowest oxidation potential (G (1.29V) < A (1.42V) < C (1.6V) < T (1.7V), and therefore guanine lesions are the most abundant.[2] Indeed, a diverse set of genotoxic compounds, such as those found in

cigarette smoke, manmade aromatic amines and aromatic hydrocarbons, commonly form addition products (adducts) by attacking guanine.[3-7]

Phenolic compounds have two opposed properties, they can show beneficial effects and they can be toxic, which refer to their damaging pro-oxidant properties that are associated with aging and disease.[8] Oxidative metabolism of phenols produces reactive phenoxyl radicals which rationalize their toxicity. Peroxidase activates enzymes with or redox-active transition metals catalyze oxidation of phenols by removal of one electron.[9] The formation, detection, and roles of these DNA C– and O–linked adducts have been reviewed,[10, 11] and they are believed to be critical lesions in phenolic-induced carcinogenesis. Ambient reactivity of substituted phenolic compounds at the C8 site of dG leads to the formation of *ortho* and *para* C8-phenoxyl-2′-deoxyguanosine (*o-PhOH*dG and *p-PhOH*dG), as well as O-linked adducts (*PhO*dG) (Scheme 2.1).
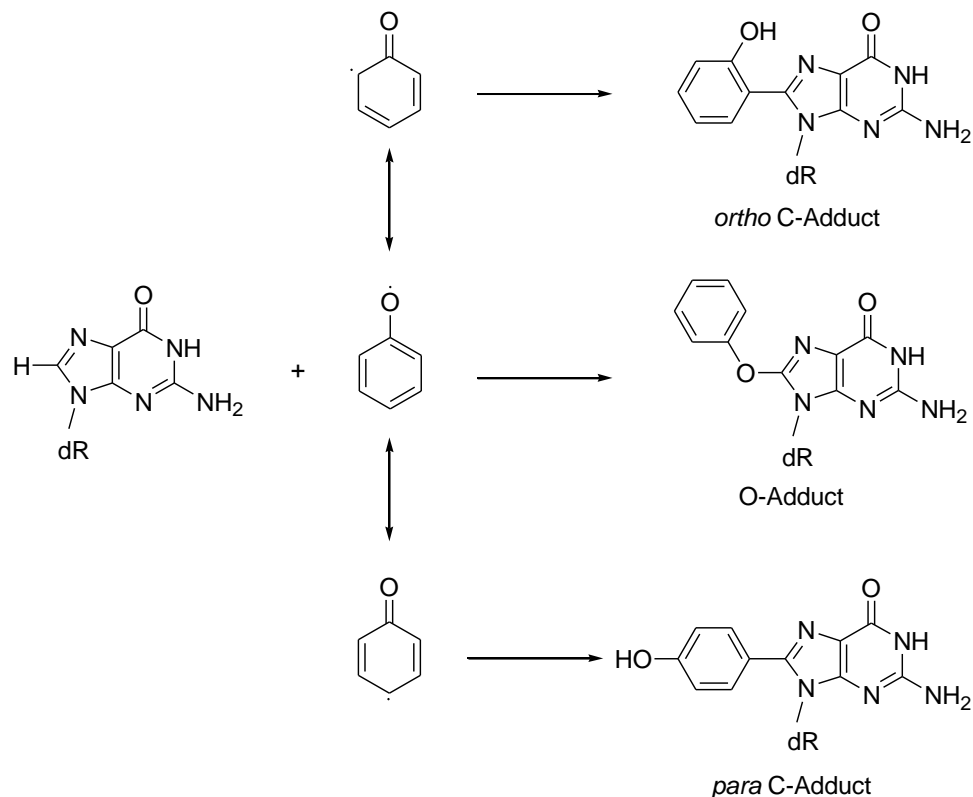
Scheme 2.1. Ambident reactivity of phenoxyl radical leading to the formation of O-linked and C-linked adducts.[10]

Additionally, it would be interesting to understand the possibility of the post-damage effects, since modification of nucleoside model by phenoxyl radical may increase the possibility of abasic site formation via deglycosylation.[12]

While several factors complicate the elucidation of their mutagenic profile, structural features and conformational preference can create a potential for mispairing in the DNA strand. The study of conformational heterogeneity allows us to assess the biological effects of induced damage, which derives from Watson-Crick and Hoogsteen hydrogen-bonding interaction between complementary nucleobases in DNA strands.[13-15] Nucleobases occupy two principal orientations about the sugar ring in nucleosides and nucleotides (Figure 2.1). The glycosidic

dihedral torsion angle (χ) acquires the *anti* conformation in most unmutated species where the H8 atom of the purines and the H6 of the pyrimidines is located above 2′-deoxyribose. Conversely, in the *syn* conformation, which is often the most stable conformer of damaged complexes, the N3 of purines and the O2 of pyrimidines are found in that position. There is a rapid interchange between *syn* and *anti* conformers in solution the speed of this interchange depends on the assembly and chemical conditions (Figure 2.1). One of the conformers can exist preferentially in a specific DNA structure. The furanose ring of (2′-deoxy) ribose at room temperature typically interconverts between the two conformations and ultimately assumes one of two: north (also known as C3′-endo) and south (or C2′-endo).[16] In B-DNA, the 2′-deoxyribose sugar acquires a C2′-endo conformation with the nucleobases exclusively in the *anti* position (Figure 2.1).[14, 17]

The sugar puckering C2′-endo and C3′-endo in Figure 2.1 is instructive just for disclosing two important different types of sugar puckering in the furanose ring while it is C2′-endo in both *anti* and *syn* conformers of natural guanosine in B-DNA.[18]
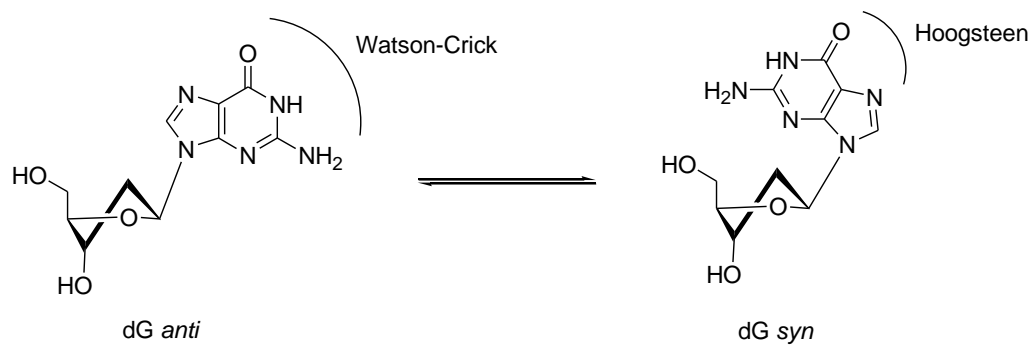
Figure 2.1. Equilibration between *anti* and *syn* conformations by rotating around glycosidic bond of 2′-deoxyguanosine and typical structures of sugar puckering for 2′-deoxyguanosine (C2′-endo, south).

The structural features and conformational properties of C-linked phenoxyl adducts have been studied experimentally and theoretically.[10, 19] The topological properties and stability of O-linked phenolic adducts at nucleobase, nucleoside, and nucleotide levels identified by performing DFT calculations. However, comparing and contrasting the structural properties of O-linked modified species with previously studied C-linked homologous[10, 20] provided a deeper insight in to conformational behaviour, stabilizing interactions and impact of oxygen-linkage in O-bonded adducts. Regarding the *ortho-* and *para*-C-linked structures, the experimental results demonstrated the most stable structure of the nucleobase adducts is planar.[10] Though, DFT calculations and particularly generation of the PESs by using scans command, and rotation about the glycosidic bond and the C–C linkage between the nucleobase and the phenoxyl moiety illustrate the twisted *syn* lowest energy structure conformation for both nucleoside adducts.[10]

To investigate the conformational and structural changes induced by formation of unsubstitued O-linked phenoxyl adduct (PhOdG) at the nucleobase,

nucleoside, and nucleotide level, quantum mechanical calculations were performed. Furthermore in order to evaluate the consequence of bridged oxygen, the structures of ortho- and para-phenolic C-linked (o-PhOHdG and p-PhOHdG) which previously has been studied by Dr. Wetmore and coworkers theoretically and experimentally, was take into account.[10, 13, 19, 21]

In this work we tried to gain deeper insight into the structural properties, conformational preference, and stability of ^{PhO}dG, o-^{PhOH}dG[10] and p-^{PhOH}dG[10] adducts.

2.2. Computational Details

2.2.1. Nucleobase Model.

To gain insight into the structural characteristics of the unsubstituted ^{PhO}G adduct, the B3LYP density functional method[22] in conjunction with the 6-31G(d) basis set was implemented to search the potential energy surface (PES). First, all local minima and transition states with respect to rotation about the dihedral angles θ (the angle controlling the relative orientation of the phenoxyl moiety at the C8-position and the nucleobase or $\angle$(N9–C8–O–C10)) and ϕ (the angle controlling the orientation within the O–substituted phenyl or $\angle$(C8–O–C10–C11)) were considered (Scheme 2.2). Specifically, θ and ϕ were systematically rotated and fixed in 20° increments from 0 to 360°. All minima and corresponding transition states were identified on the PES, and subsequently fully optimized and characterized using frequency calculations at the B3LYP/6-31G(d) level. B3LYP/6-311+G(2df,p) single-point energy calculations were performed and all corresponding relative energies include scaled (0.9806) zero-point

vibrational energy corrections. Calculations on the nucleobase model were carried out using Gaussian 09 (Revision A.02).[23]

2.2.2. Nucleoside Model

Minima for the [PhO]dG nucleoside adduct were initially identified through a conformational search using the internal coordinate Monte Carlo algorithm[24] in HyperChem 8.0.8. The AMBER molecular mechanics force field, with PM3[25, 26] charges, was implemented in the conformational search[12] and the following dihedral angles were scanned (Scheme 2.2): θ ($\angle$N9–C8–O–C) and ϕ ($\angle$C8–O–C–C4), χ($\angle$O4′–C1′–N9′–C4), as well as the endocyclic torsion angles of the sugar, namely υ_0 ($\angle$O4′–C1′–C2′–C3′), υ_1 ($\angle$C1′–C2′–C3′–C4′), υ_2 ($\angle$C2′–C3′–C4′–O4′), υ_3 ($\angle$C3′–C4′–O4′–C1′), υ_4 ($\angle$C4′–O4′–C1′–C2′), β ($\angle$C4′–C5′–O–H), ε ($\angle$C4′–C3′–O–H), and γ ($\angle$O3′–C4′–C5′–O). This procedure led to more than 200 conformers, with the 50 lowest energy conformers subsequently fully optimized with B3LYP/6-31G(d). Finally, B3LYP/6-311+G(2df,p) single-point energy calculations were carried out on the resulting ten lowest energy structures to identify the most stable conformer.
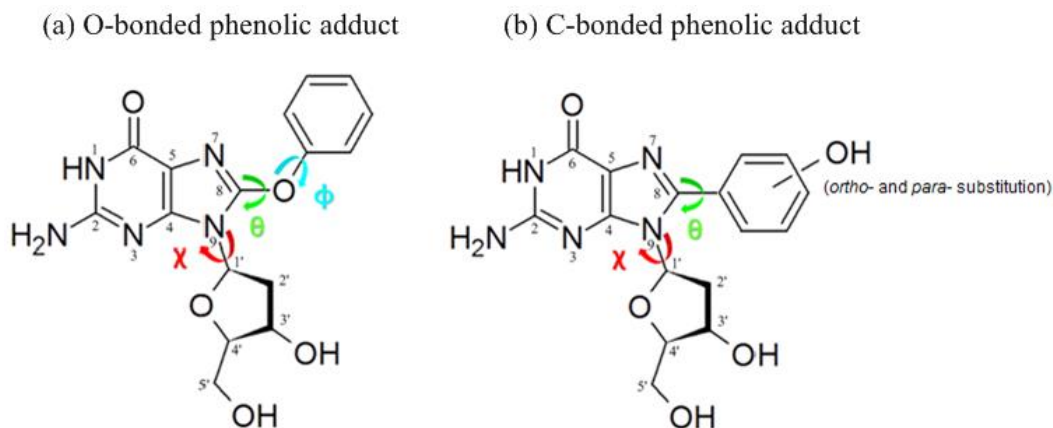
Scheme 2.2. Structures of (a) ^{PhO}dG, (b) $^{o\text{-}PhOH}$dG, and $^{p\text{-}PhOH}$dG and different θ (∠N9–C8–O–C), φ (∠C8–O–C–C4), and χ(∠O4′–C1′–N9′–C4) dihedrals.

Sstarting from the most stable conformer of the HyperChem conformational search, the PES of the nucleoside model was investigated by initially scanning θ and χ (the angle controlling the orientation of the nucleobase about the glycosidic bond or ∠(O4′–C1′–N9′–C4)), as done in previous work for the related $^{o\text{-}PhOH}$dG[10] and $^{p\text{-}PhOH}$dG[10] adducts.[27] Specifically, θ and χ were systematically altered and fixed in 10° increments from 0 to 360°, while the sugar puckering was constrained to C2′–endo which is the common sugar puckering in B-DNA. A conformational search also provided evidence for the preference of C2′–endo sugar puckering. The C3′–hydroxyl group was oriented such that ∠(H–C3′–O–H) approximately equaled –60°, and the C5′–hydroxyl group was directed toward the nucleobase (∠(C4′–C5′–O–H) approximately equals 50–90°). All the aforementioned constraints reserved similar to parameters in a previous research on *ortho*- and *para*- C-linked structures,[10] to be able to compare and contrast the structural properties of O-linked and C-linked modified species. A PES scan was also

conducted with respect to the χ and ɸ dihedral angles to determine the effect of the ɸ torsion. All stationary points on these PESs were fully optimized and characterized with B3LYP/6-31G(d). Scaled (0.9806) zero-point vibrational energies were added to the reported relative energies. All nucleoside calculations were performed using Gaussian 09 (Revisions A.02 or C.01).[23]

2.2.3. Nucleotide Model

Adding the 5′-monophosphate group to the β-constrained nucleoside models of the global (*syn*) and local (*anti*) minima identified from the PESs, generate the nucleotide model. Natural dG nucleotide, as well as C8-bonded phenoxyl dG adducts, has been investigated previously which provided a precise developed approach for including environmental effects and the representative model of the phosphate group that led us to the most biologically relevant nucleotide conformations.[28-32] Although other protocols use larger basis sets that include diffuse functions (6-31+G(d,p)), the mentioned computational model results are consistent with those obtained using larger basis sets that include diffuse functions (6-31+G(d,p)). There were minimal differences between the geometries obtained with the two basis sets (considering the diffuse functions or not).[20] Specifically, nucleotides can be accurately characterized by taking solvent effects (water) into account with PCM-B3LYP/6-31G(d) optimizations in water (ε = 78.4), and including an anionic phosphate model that is neutralized by a Na^+ counterion. For the nucleobase and nucleoside models, reported relative energies were obtained from gas-phase B3LYP/6-311+G(2df,p) single-point calculations, whereas the

corresponding calculations were carried out in water for the nucleotide models. All relative energies include scaled (0.9806) zero-point vibrational energy (ZPVE) corrections. All quantum chemical calculations were performed using Gaussian 09.

2.3. Results and discussions

2.3.1. Nucleobase Model

The PES was scanned as a function of θ versus ϕ at the nucleobase level for the unsubstituted O-linked adduct (^{PhO}G) with a goal to identify the preferred orientation of the phenoxyl moiety with respect to guanine in the absence of the sugar (Figure 2.2). This will not only illuminate the structural properties of the inserted phenoxyl group, but also shed light on the influence of the sugar segment on the preferred conformation and orientation of the mutated moiety with respect to the nucleobase. Following the PES scan, all minima and transition states were fully optimized by removing all restraints.
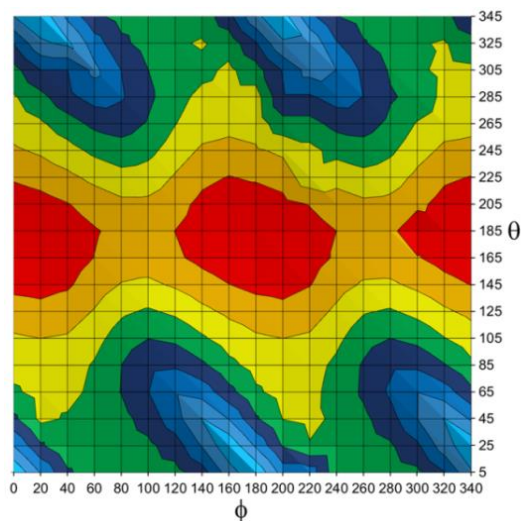
Figure 2.2. B3LYP/6-31G(d) θ versus ϕ potential energy surface for [PhO]G; The relative energy (kJ mol^{-1}) and color change is in 20 kJ mol^{-1} increments where the lowest energy regions are red.

On the PES for the nucleobase model, there is only one minimum ($\theta \sim 179.8°$, $\phi \sim 179.8°$) in which the phenoxyl moiety adopts a coplanar conformation, and is repeated after 180° rotation with respect to ϕ ($\theta \sim 180.0°$, $\phi \sim 359.8°$). A weak C−H•••N7 hydrogen bond stabilizes the planar minimum. Two transition states (($\theta \sim 180°$, $\phi \sim 91.8°$) and ($\theta \sim 180°$, $\phi \sim 267.6°$)) connect the symmetry-equivalent minima, which have a very small (3.0 kJ mol^{-1}) associated rotational barrier about ϕ (Figure 2.3).

$\theta=180.0°$;
$\Phi=91.8°$;
$\Delta E=3.0$

$\theta=180.0°$;
$\Phi=359.8°$;
$\Delta E=0.0$

$\theta=179.8°$;
$\Phi=179.9°$;
$\Delta E=0.0$
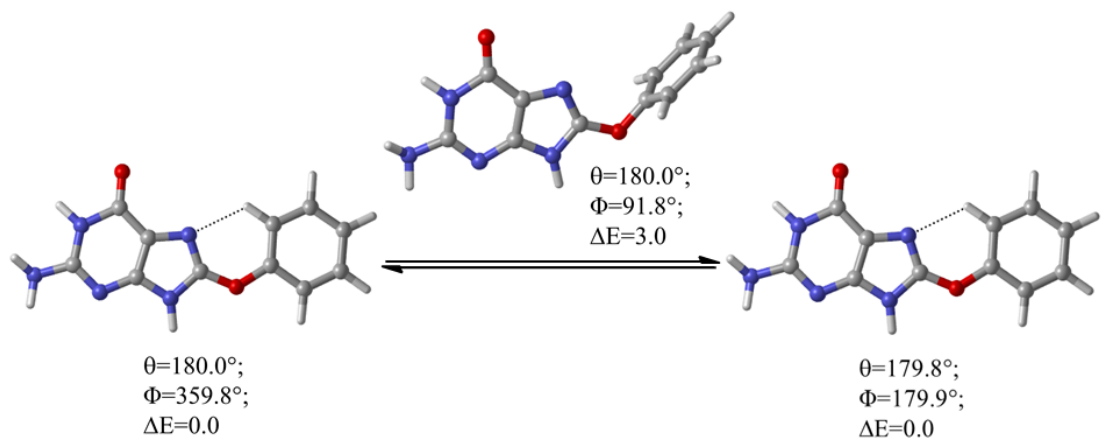
Figure 2.3. Comparison of relative energies (kJ mol^{-1}) between selected minima and transition states of ^{PhO}G at B3LYP/6-311+G(2df,p)//B3LYP/6-31G(d) as well as θ and ϕ angles.

Previous computational findings regarding the *ortho* C-bonded structure[10] suggested that the planar global minimum (θ =180.0°, *o*-^{PhOH}G) is stabilized by a slightly strong O–H•••N7 hydrogen bond while a lack of hydrogen bonding interactions between the hydroxyl on the aromatic ring and electronegative atoms of guanine in the *para*-nucleobase (*p*-^{PhOH}G) renders a less stable global minimum. The two local minima identified for the *o*-^{PhOH}G have the bulky moiety twisted approximately 25° from planarity with respect to the nucleobase (θ ~23.4° and ~336.3°), leading to destabilizing steric interactions between the N9 hydrogen of guanine and phenolic OH. Two skew transition states with an energy barrier of ~48 kJ mol^{-1} connect the global minimum and local minima. Efforts to clarify the structural features of the para C–linked[10] complex demonstrate that the perpendicular transition states connecting the local minima (θ ~359.5° and ~1.3 kJ mol^{-1}) have an associated rotational barrier of 20.8–20.9 kJ mol^{-1}.[10]

DFT calculations suggest that the bridged oxygen in the O-linked substituted adduct increases the flexibility of this species compared to $^{o-PhOH}$G and $^{p-PhOH}$G, which results in less steric and electronic interaction between the aromatic moiety with the bonded and nonbonded electrons on the nucleobase.

2.3.2. Nucleoside Model

There have been computational studies on the structural properties of C-linked adducts at the nucleoside level[10, 20] while unsubstituted O-linked mutated complexes have not been studied yet at the nucleoside level. In order to further explore the geometrical properties of ^{PhO}dG and provide a better understanding of oxygen as a bridging atom, the optimized structures of ^{PhO}dG was determined, and the structural properties are compared and contrasted with previously studied $^{o-PhOH}$dG[10, 32], and $^{p-PhOH}$dG[10, 33] adducts. Theoretical calculations were performed using the B3LYP method in conjunction with the 6-31G(d) and 6-311+G (2df, p) basis set and the most stable conformers were compared and contrasted. Understanding the properties of the these adducts at the nucleoside level in the presence of the sugar moiety is an important step in evaluating the effect of sugar moiety, and assessing the factors that may contribute to the conformation and binding preferences of these adducts in DNA helices.

The molecular geometry of mutated complexes of deoxyguanosine can be discussed in terms of the different structural units, the guanine ring and the furanose ring conformation. The furanose ring puckering can be defined via endo and exo, which refer to the displacement of an atom above or below the mean plane

of the ring (endo: on the same side as C5′ atom, exo: on the opposite side).[34]

Rotation around the C4′–C5′ bond leads to three possible conformers: "gauche-trans" (gt), "trans-gauche" (tg), and "gauche–gauche" (gg) (Scheme 2.3).[35] The intramolecular hydrogen bonds involving hydroxyl groups depend on the endo or exo character of the ribose and the nature of the possible interaction of the hydroxyl group and the guanine heterocyclic ring.
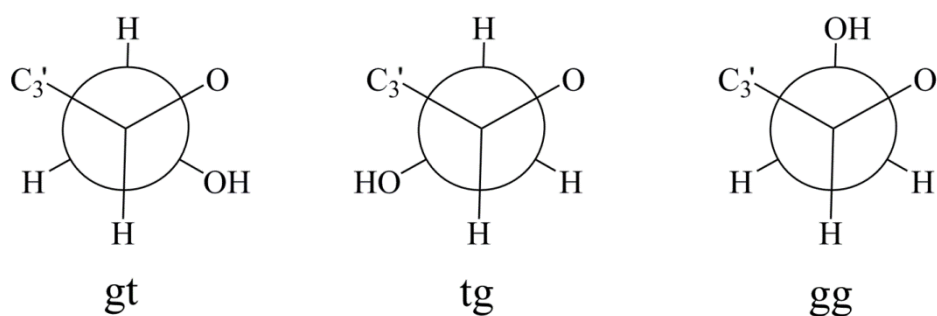
Scheme 2.3. Three possible conformations about C4′–C5′ bond.[34]

To probe the intrinsic properties of the ^{PhO}dG, θ versus χ, PESs were generated (Figure 2.4). Different conformers have different electronic structures and, in principle, may display significant variations in their physical properties, so the conformational preference of unsubstituted O–linked nucleobases have been examined by analyzing the relative stability of these structures in the gas phase. The same work has been performed on *ortho* and *para* C–bonded mutated complexes.[10]

Subsequently, PES for the fully optimized (B3LYP/6-311+G(2df,p)//B3LYP/6-31G(d)) structures of neutral unsubstituted O-linked deoxyguanosine (^{PhO}dG) identified from θ versus χ and ϕ versus χ appear in Figure 2.4 and Figure 2.7.

Global and local minima of *ortho* and *para* C-bonded structures (*o*-PhOHdG, and *p*-PhOHdG)[10] which were identified from the same procedure and level of theory are sketched in Figure 2.5 this model shows the important structural features of these types of damage to DNA.

The result of conformational search using HyperChem demonstrate that the most stable conformer acquire the C2′–endo sugar puckering which is the expected sugar puckering in B-DNA. Interestingly, the sugar puckering remained intact in unsubstituted O-linked adduct by implementing more accurate quantum mechanical method (DFT). This illuminates the determined sugar puckering in PhOdG modified complex, and it is consistent with the adopted sugar puckering of *o*-PhOHdG,[10] and *p*-PhOHdG.[10]

The orientation of the nucleobase about the glycosidic bond can be categorized as *anti* or *syn* according to χ, where *anti* is defined by χ values ranging from 90° to 270° and *syn* refers to χ values of 0–90° and 270–360°. The lowest energy conformer in PhOdG is C2′–endo/*syn* which namely (1.83 Å) N3•••HO5′ and N7•••H (2.23 Å), intramolecular hydrogen bonds. The *syn* orientation of the base unit with respect to the sugar is strongly stabilized by the formation of the N3•••HO5′ intramolecular hydrogen bond. Gauche–gauche (gg) is the stable conformer around the C4′–C5′ bond (Figure 2.5).
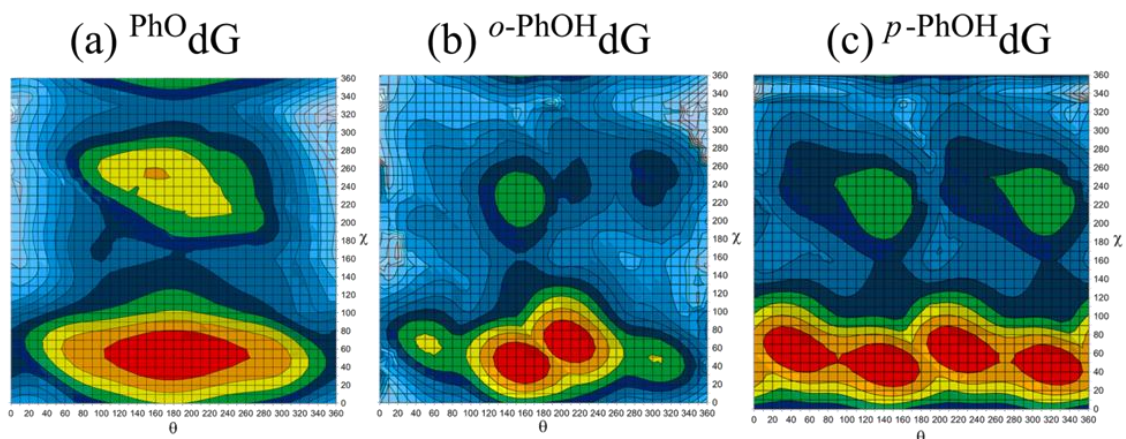
Figure 2.4. B3LYP/6-31G(d) θ versus χ potential energy surface for (a) PhOdG, (b) o-PhOHdG10, and (c) p-PhOHdG10; the relative energy (kJ mol−1) is represented by color, where the lowest energy regions are red, and each change in color represents a 10 kJ mol−1 increase in the relative energy.
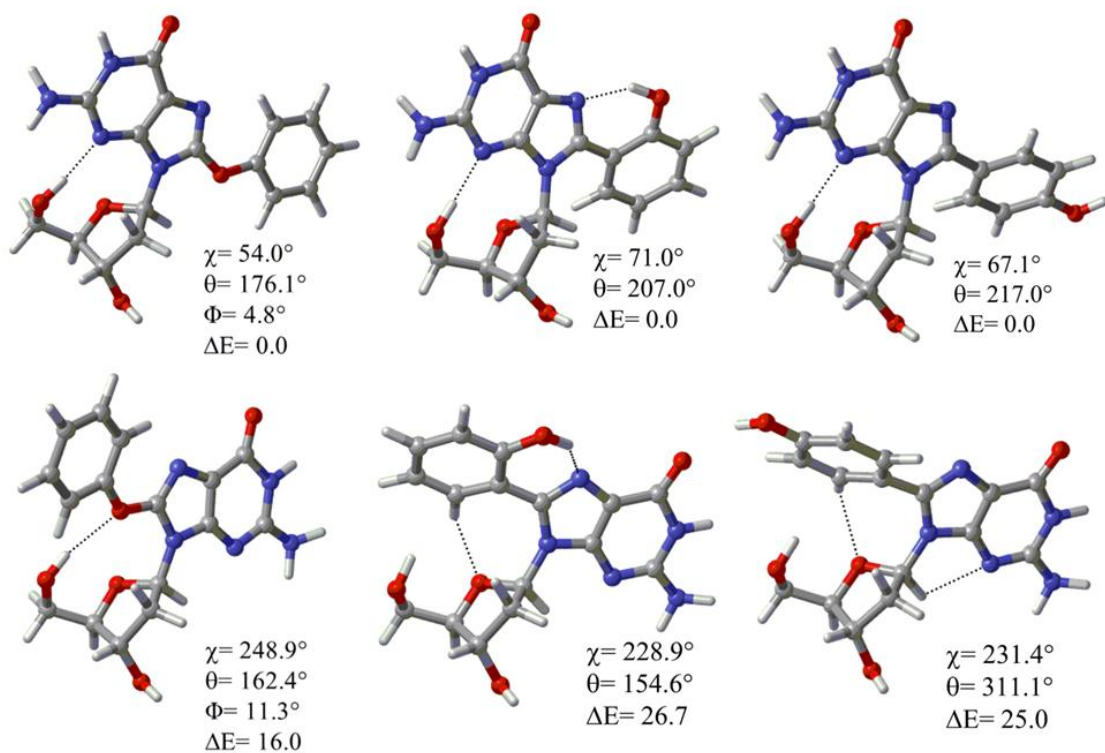
Figure 2.5. Fully optimized global and local minima of PhOdG, o-PhOHdG,10 and p-PhOHdG10 at B3LYP/6-31G(d) and relative energies of *anti*/*syn* conformers at B3LYP/6-311+G(2df,p) in kJ mol−1.

As seen in Figure 2.4 and Figure 2.5, the geometrical variations involving increasing energies on the PESs and formation of local minima involve structural changes in hydrogen bonding. The guanine should turn around the glycosidic linkage, leading to cleavage of the intramolecular N3•••H-O5′ hydrogen bond in the mutated complex and formation of a new hydrogen bond between O•••H-O5′ in an *anti* (gg) conformer. Previous study of *o*-PhOHdG,[10] and *p*-PhOHdG[10] analogues suggest *syn* and *anti* global and local minima in a gg conformer around the C4′–C5′ linkage (Scheme 2.3 and Figure 2.5).[10, 36]

Structural properties and conformational preference of these molecules PhOdG, *o*-PhOHdG,[10] and *p*-PhOHdG[10] at the nucleoside level can be interpreted by geometrical changes. Almost all adducts possess a certain degree of twist about the θ dihedral whereas it has been demonstrated that PhOdG ($\chi \sim 53.7°$, $\phi \sim 85.1°$, $\theta \sim 4.2°$) undergoes less deformation or conformational change. The phenoxyl moiety is less twisted with respect to the nucleobase, when compared to *o*-PhOHdG ($\chi = 71.0°$, $\theta = 207.0°$), and *p*-PhOHdG ($\chi = 67.1°$, $\theta = 217.0°$). This suggests that the sugar unit induces a further twist in the *ortho* and *para* C-linked[10] adducts and that its impact is not important due to the presence of the bridging oxygen which alleviates steric effects between the phenoxyl and guanosine units. In particular, the spatial orientation of phenoxyl and sugar moieties with respect to guanine result in less influential interactions compared to *ortho* and *para* C-linked structures. The alteration of dihedrals is not obvious in by gradually increasing the size of system from nucleobase (Figure 2.3) to nucleoside. Structural distortions of the *ortho* and

para C-linked structures become more obvious when changing the type of model due to closer weak (hydrogen bonding) and slightly strong (Van der Waals) interactions of the aryl moiety with the sugar segment and nucleobase.[10] In addition, the strength of hydrogen bonds should be taken into account, specifically the H-bond between the hydroxyl group of *o*-PhOHdG[10] and N7 as well as the hydrogen atom of the aryl moiety and the oxygen atom of the sugar in *p*-PhOHdG.[10] The latter is not as substantial leading to more twisted structures in *o*-PhOHdG[10] and *p*-PhOHdG[10] rather than PhOdG.
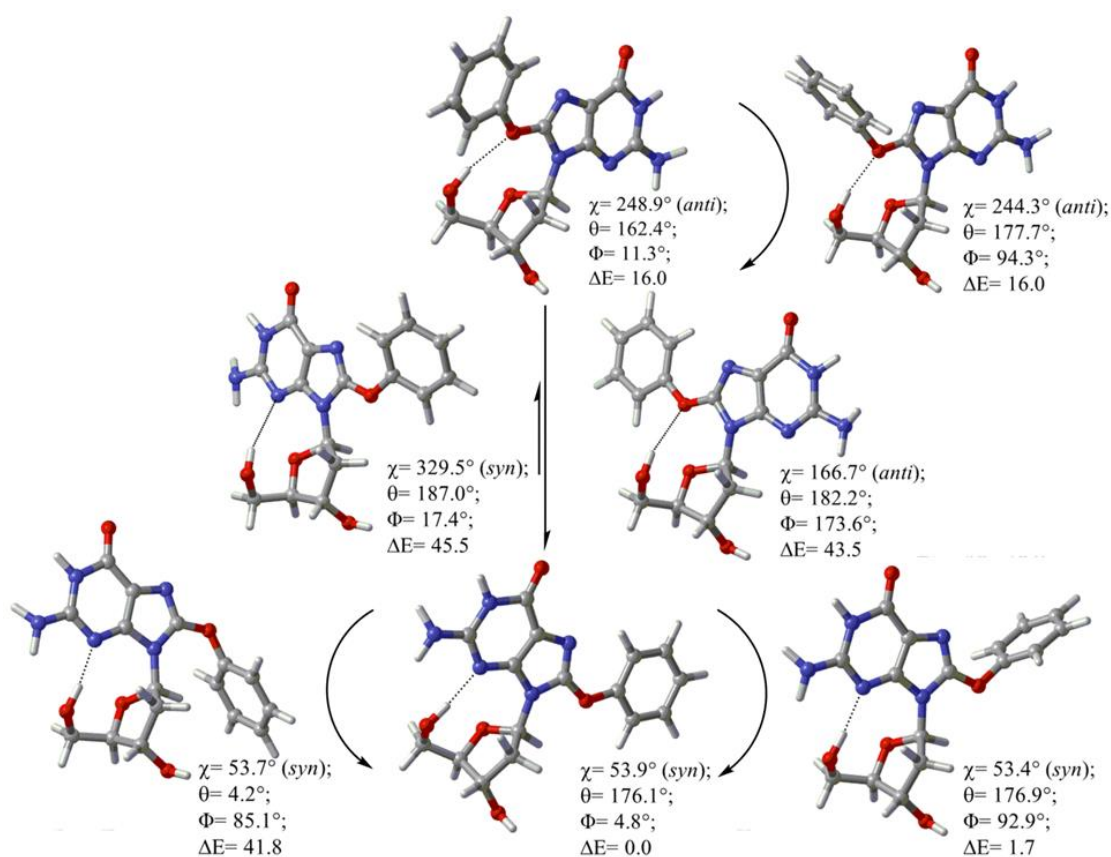
Figure 2.6. Fully optimized minima and transition states identified from PES of PhOdG, different dihedrals (θ, φ, and χ deg) and relative energies at B3LYP/6-311+G(2df,p) in kJ mol⁻¹.

Results of DFT calculations indicate that the phenoxyl unit is fairly planar with respect to guanosine in all global and local minima of the contour plots. The global minimum is connected to the local minima by three skew transition states on the blue regions of the counter plots: two *syn* ($\chi \sim 53.7°$, $\phi \sim 85.1°$, $\theta \sim 4.2°$), ($\chi \sim 329.5°$, $\phi \sim 17.4°$, $\theta \sim 187.0°$) and one *anti* conformer ($\chi \sim 166.7°$, $\phi \sim 173.6°$, $\theta \sim 182.2°$) with energy barriers of $\sim 41-45$ kJ mol^{-1}. There is a lower barrier to rotation (~ 16 kJ mol^{-1}) between the global and local minima for [PhO]dG in comparison with [o-PhOH]dG (~ 26.7 kJ mol^{-1})[10] and [p-PhOH]dG (~ 25.0 kJ mol^{-1})[10] as a result of less steric clashes and the higher flexibility of the phenoxy group (Figure 2.6). To provide better insight in to the effect of ϕ on the conformational flexibility and preference of [PhO]dG, quantitative analysis was conducted by generating a ϕ verses χ PES (Figure 2.7). The same global and local minima were identified from the new PES, and the greater flexibility of ϕ in comparison with θ observed in the nucleobase model ,was verified.
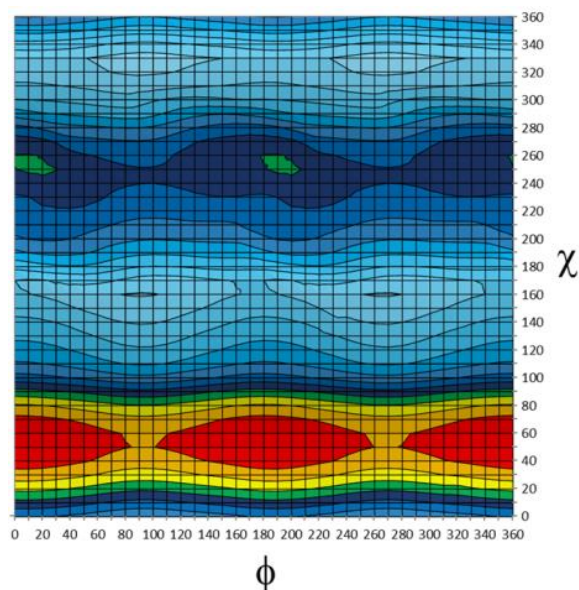
Figure 2.7. B3LYP/6-31G(d) ϕ verses χ potential energy surface for PhOdG; The relative energy (kJ mol^{-1}) is represented by color, where the lowest energyregions are red, and each change in color represents a 5 kJ mol^{-1} increase in the relative energy.

2.3.3. Nucleotide Model

The previous section suggests that fully optimized nucleoside models predict the *syn* orientation to be favored due to the presence of a non-native O5′–H•••N3 hydrogen bond. However, the nucleoside model cannot provide all essential information regarding the structure of damaged bases because the presence of the phosphate group can induce more steric hindrance which may lead to alteration in the conformational preference of modified bases at the nucleotide level. This increases the possibility of distorted DNA. To determine these effects, the nucleoside model is first exposed to a geometric constraint to yield the so-called β–constrained nucleoside model, which prevents the intramolecular hydrogen-bonding interaction between the C5′O–H and N3 site of mutated guanosine. This

71

renders the *syn* conformation less important than the initially predicted conformation at the nucleoside level.

To better predict the structure of the damaged bases in DNA helices, a nucleotide model is made by the including the 5′-monophosphate group. Previous computational methodology was developed for the accurate treatment of bulky adducts which could predict the correct and the biologically relevant *anti/syn* preference for the natural dG nucleotide as well as (damaged) *ortho* and *para* C8-phenoxyl-2′-deoxyguanosine 5′-monophosphate adducts.[20] Computational modeling reveals that the *anti* conformation is dominated for both types of damage at the nucleotide level. It seems that due to steric hindrance imposed by flanking and adjacent bases, there is an equilibrium between both the *anti* and *syn* conformations in the helix.[11] In the case of the O-linked unsubstituted structure, the nucleoside model preferentially adopts a *syn* orientation (by 16.0 kJ mol^{-1}) due to the presence of an O5′–H•••N3 interaction. Nevertheless, elimination of the mentioned hydrogen bond can better simulate the actual DNA environment, leading to a very small (<3.0 kJ mol^{-1}) *anti/syn* energy difference. Overall, inclusion of the 5′-monophosphate group leads to around 6.5 kJ mol^{-1} preference for the *syn* (nucleotide) conformation, due to a pseudo–skew structure of the *syn* conformer (χ=67.0°, θ=191.9°, ϕ=60.3° and β= 162.4°) stabilizing hydrogen bonding and less steric effects, while a pseudo–planar conformer with more steric effects was identified for the *anti* structure (χ=258.8°, θ=168.1°, ϕ=150.3° and β=191.9) at the nucleotide level (Figure 2.8).

The sugar puckering in the *anti* conformations of [o-PhOH]dG20 and [p-PhOH]dG20 at the nucleotide level changed to C1′−exo−O4′−endo, while no significant change has observed in the sugar puckering of the unsubstituted structure at the nucleotide level. C2′ *endo* sugar puckering is retained for both *anti* and *syn* conformers of O-linked mutated adducts after optimization. The steric effects, between the 5′−phosphate group and the phenoxyl moiety in the *anti* conformations destabilize this conformer in comparison with *syn*. Also, since the 5′−OH and the amino group of the nucleobase are far away, there is no distinctive stabilizing interaction. In contrast, the *syn* conformations of the [PhO]dG nucleotide adduct, which are stabilized by an O5′−H•••N3 intramolecular H−bond interactions, are preferred over the *anti* conformations by 6.3 kJ mol^{-1}. Nevertheless, a small energy difference has been identified between the *syn* and *anti* conformations. Since the energy difference in both *anti* and *syn* conformers of unsubstituted O-linked as well as *ortho* and *para* C-linked structures is low at the nucleotide level, both conformers may exist in more complicated systems such as the DNA model. Damage caused by C8-methylation of adenosine[37] and halogenation at the C8 position[38] results in stabilization of the *syn* conformation in both the gas phase and in solution in nucleotide models. On the other hand, C8-substitution is not the determining factor for the formation of the *syn* conformation at the nucleotide level. The presence of bulkier groups at the C8-site of guanosine results in the formation of the *anti* conformation in nucleotide models. Similar results has been identified for adenosine 5′-monophosphate modified at the C8-position by an n-butylamino group.[39] These observations can be interpreted as stabilizing effects, such as

hydrogen-bonding interactions between the C-8 substituent and the sugar, that combine to be greater than the destabizing impact of unfavorable steric effects. On the contrary, stacking interactions between flanking bases in the helix may affect the spatial orientation and interactions of the base–phosphate group. Ultimately, the impact of damage in the alteration of interactions among the nucleobase, sugar and modified moiety determines the overall structure of damaged DNA helices and the relative stability of each conformer. Particularly, the insignificant energy difference between the *syn* and *anti* conformers of unsubstituted O-linked, *ortho* and *para* C-linked adducts at the nucleotide level, emphasizes the importance of considering the larger DNA model and investigating the structures of larger DNA oligomers and their hydrogen-bonding ability. This will be discussed in detail in the next Chapter 3.
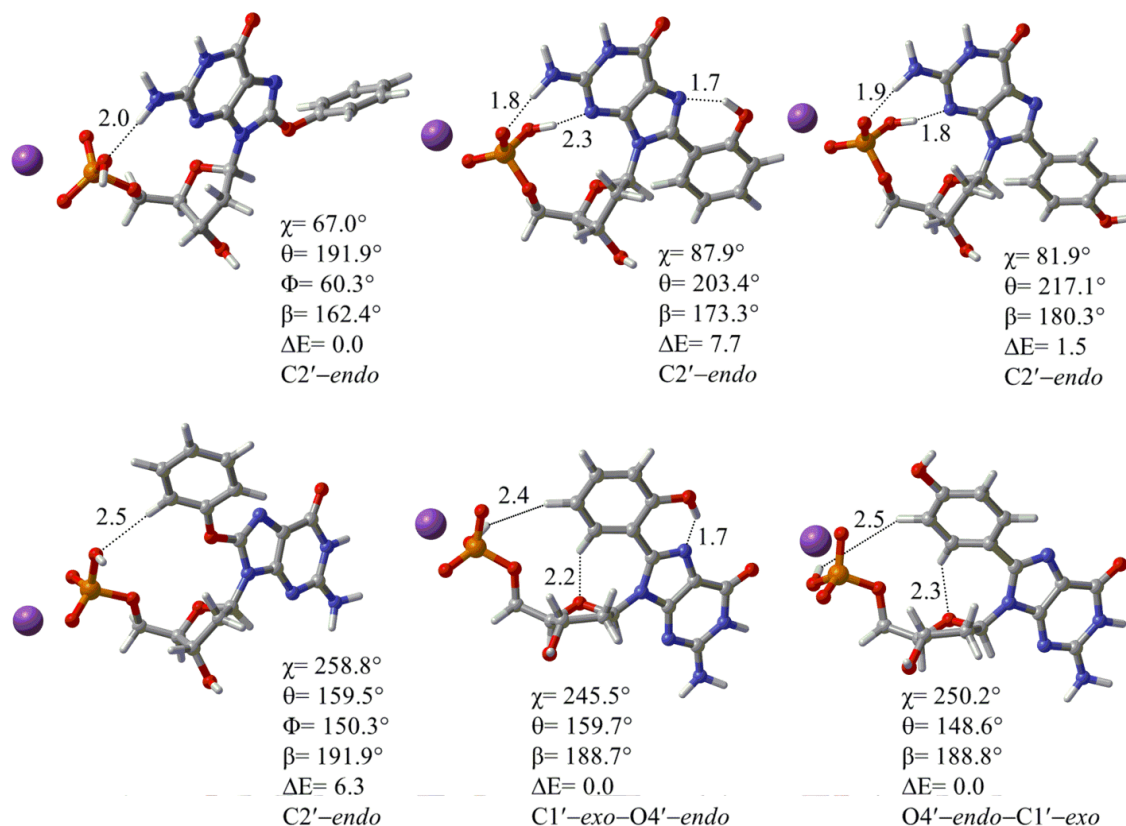
Figure 2.8. The conformations, *syn* (top) and *anti* (bottom), of [PhO]dG, [o-PhOH]dG[20] and [p-PhOH]dG[20] at the nucleotide level described by the counterion Na+ HPO4− model. Hydrogen-bonds (Å), dihedral angles (deg.), relative energies (kJ mol-1), and sugar puckering are determined by optimized structures at B3LYP/6-31G(d).

2.4. Conclusion

The unsubstituted adduct at the nucleobase level, adopts a planar minimum conformation and the transition state adopt a perpendicular arrangement of the rings. DFT calculations suggest that the barrier to rotation from the global minimum to transition state is greater in [o−PhOH]G and [p−PhOH]G (~20-50 kJ mol-1)[20] than [PhO]G (~3 kJ mol-1), which suggests that the bridged oxygen in the O-linked

adduct increases the flexibility of these species in comparison with $^{o-PhOH}$G and $^{p-PhOH}$G.

The nucleoside model shows that the gas-phase global minimum for ^{PhO}dG possesses a *syn* conformation with O5′–H•••N3 hydrogen bond and that the aryl ring is planer with respect the guanosine. The same is true for *ortho* and *para* C-linked structures (*syn* conformer as global minima). Due to the presence of a hydrogen bond between the hydroxyl group on the aromatic ring and N7 of the nucleobase in $^{o-PhOH}$dG[10] and more steric effects in $^{p-PhOH}$dG,[10] the aryl moiety is perpendicular with respect to guanosine giving a more twisted structure than the O-linked adduct. The impact of ϕ on the conformational flexibility and structural preference of ^{PhO}dG is identified by by generating a ϕ verses χ PES which resulted in the felexibility of ϕ dihedral and the same global and local minima.

At the nucleotide level, the *syn* conformer is the most stable structure for ^{PhO}dG mono phosphate, whereas both $^{o-PhOH}$dG mono phosphate[20] and $^{p-PhOH}$dG mono phosphate[20] adopt *anti* conformations. There is more variation in the sugar puckering in both conformers of C-linked[20] structures, while in the unsubstituted O-linked adduct the sugar puckering remained C2′-endo. The observed variation in the most stable conformer from small model perspective (nucleobae and nucleoside) to larger models (nucleotide) demonstrate that the sugar, phosphate, and the aryl substituent can affect the conformational flexibility. The situation can be more complicated due to more steric hindrance in the DNA double helix where the small models cannot provide a comprehensive perspective about the structural

properties of the adducts. In addition, the low energy difference at nucleotide level between *anti* and *syn* conformers of substituted O-linked adduct and *ortho* and *para* C-linked structures suggests that these modified species can acquire either or both of the aforementioned conformers in the double helix, so the details of conformational preferences will address in Chapter 3.

2.5. References

1.   Scharer, O. D., Chemistry and Biology of DNA Repair. *Angewandte Chemie International Edition* **2003,** *42* (26), 2946-74.

2.   Gurusamy, P.; Muthukumar, K.; Rajesh, S.; Muneeswaran, G.; Perumal, S.; Karunakaran, C., Theoretical Investigation of Quinone Metabolites of Dopamine Interaction with DNA – Insights into Toxicological Effects. *Journal of Structural Biology* **2012,** *180* (1), 125-131.

3.   Tretyakova, N.; Villalta, P. W.; Kotapati, S., Mass Spectrometry of Structurally Modified DNA. *Chemical Reviews* **2013,** *113* (4), 2395-2436.

4.   Doluca, O.; Withers, J. M.; Filichev, V. V., Molecular Engineering of Guanine-Rich Sequences: Z-DNA, DNA Triplexes, and G-Quadruplexes. *Chemical Reviews* **2013,** *113* (5), 3044-3083.

5.   Zhang, H.; Li, F.; Dever, B.; Li, X.-F.; Le, X. C., DNA-Mediated Homogeneous Binding Assays for Nucleic Acids and Proteins. *Chemical Reviews* **2012,** *113* (4), 2812-2841.

6.   Bloomfield, V. A.; Crothers, D. M.; Tinoco, I., Jr., *Nucleic Acids Structures, Properties, and Functions*. University Science Press: Mill Valley, CA, 2000.

7.   Geacintov, N. E.; Broyde, S., *The Chemical Biology of DNA Damage*. Wiley-VCH: Weinheim, 2010.

8.   Sturla, S. J., DNA Adduct Profiles: Chemical Approaches to Addressing the Biological Impact of DNA Damage from Small Molecules. *Current Opinion in Chemical Biology* **2007,** *11* (3), 293-299.

9.   Manderville, R. A., Ambient Reactivity of Phenoxyl Radicals in DNA Adduction. *Canadian Journal of Chemistry-Revue Canadienne De Chimie* **2005,** *83* (9), 1261-1267.

10.   Millen, A. L.; McLaughlin, C. K.; Sun, K. M.; Manderville, R. A.; Wetmore, S. D., Computational and Experimental Evidence for the Structural Preference of Phenolic C-8 Purine Adducts. *Journal of Physical Chemistry A* **2008,** *112* (16), 3742-3753.

11.   Sharma, P.; Manderville, R. A.; Wetmore, S. D., Modeling the Conformational Preference of the Carbon-Bonded Covalent Adduct Formed Upon Exposure of 2′-

Deoxyguanosine to Ochratoxin A. *Chemical Research in Toxicology* **2013,** *26* (5), 803-816.

12.   Kuska, M. S.; Majdi Yazdi, M.; Witham, A. A.; Dahlmann, H. A.; Sturla, S. J.; Wetmore, S. D.; Manderville, R. A., Influence of Chlorine Substitution on the Hydrolytic Stability of Biaryl Ether Nucleoside Adducts Produced by Phenolic Toxins. *The Journal of Organic Chemistry* **2013,** *78* (14), 7176-7185.

13.   Millen, A. L.; Churchill, C. D. M.; Manderville, R. A.; Wetmore, S. D., Effect of Watson-Crick and Hoogsteen Base Pairing on the Conformational Stability of C8-Phenoxyl-2 '-Deoxyguanosine Adducts. *Journal of Physical Chemistry B* **2010,** *114* (40), 12995-13004.

14.   Millen, A. L.; Sharma, P.; Wetmore, S. D., C8-Linked Bulky Guanosine DNA Adducts: Experimental and Computational Insights into Adduct Conformational Preferences and Resulting Mutagenicity. *Future Medicinal Chemistry* **2012,** *4* (15), 1981-2007.

15.   Blackburn, G. M.; Gait, M., J, *Nucleic Acids in Chemistry and Biology*. 2nd ed.; Oxford University Press: New York, 1996.

16.   Pogozelski, W. K.; Tullius, T. D., Oxidative Strand Scission of Nucleic Acids: Routes Initiated by Hydrogen Abstraction from the Sugar Moiety. *Chemical Reviews* **1998,** *98* (3), 1089-1107.

17.   Kuska, M. S.; Witham, A. A.; Sproviero, M.; Manderville, R. A.; Majdi Yazdi, M.; Sharma, P.; Wetmore, S. D., Structural Influence of C8-Phenoxy-Guanine in the Nari Recognition DNA Sequence. *Chemical Research in Toxicology* **2013,** *26* (9), 1397-1408.

18.   Imanishi, T.; Obika, S., Bnas: Novel Nucleic Acid Analogs with a Bridged Sugar Moiety. *Chemical Communications* **2002,** (16), 1653-1659.

19.   Schlitt, K. M.; Sun, K. W.; Paugh, R. J.; Millen, A. L.; Navarro-Whyte, L.; Wetmore, S. D.; Manderville, R. A., Concerning the Hydrolytic Stability of 8-Aryl-2'-Deoxyguanosine Nucleoside Adducts: Implications for Abasic Site Formation at Physiological Ph. *The Journal of Organic Chemistry* **2009,** *74* (16), 5793-802.

20.   Millen, A. L.; Manderville, R. A.; Wetmore, S. D., Conformational Flexibility of C8-Phenoxyl-2 '-Deoxyguanosine Nucleotide Adducts. *Journal of Physical Chemistry B* **2010,** *114* (12), 4373-4382.

21.   Omumi, A.; Millen, A. L.; Wetmore, S. D.; Manderville, R. A., Fluorescent Properties and Conformational Preferences of C-Linked Phenolic-DNA Adducts. *Chemical Research in Toxicology* **2011,** *24* (10), 1694-709.

22.   Becke, A. D., Density-Functional Thermochemistry .3. The Role of Exact Exchange. *Journal of Chemical Physics* **1993,** *98* (7), 5648–5652.

23.   Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Jr., J. A. M.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.02; Gaussian, Inc.: Wallingford CT, 2009.

24.   Chang, G.; Guida, W. C.; Still, W. C., An Internal-Coordinate Monte Carlo Method for Searching Conformational Space. *Journal of the American Chemical Society* **1989,** *111* (12), 4379-4386.

25.   Stewart, J. J. P., Optimization of Parameters for Semiempirical Methods V: Modification of Nddo Approximations and Application to 70 Elements. *Journal of Molecular Modeling* **2007,** *13* (12), 1173–1213.

26.   Rezac, J.; Fanfrlik, J.; Salahub, D.; Hobza, P., Semiempirical Quantum Chemical Pm6 Method Augmented by Dispersion and H-Bonding Correction Terms Reliably Describes Various Types of Noncovalent Complexes. *Journal of Chemical Theory and Computation* **2009,** *5* (7), 1749–1760.

27.   Millen, A. Properties of C-Linked C8-Phenoxyl Guanine DNA Adducts. University of Lethbridge, Lethbridge, Alberta, Canada, 2006.

28.   Millen, A. L.; Manderville, R. A.; Wetmore, S. D., Conformational Flexibility of C8-Phenoxyl-2'-Deoxyguanosine Nucleotide Adducts. *The Journal of Physical Chemistry B* **2010,** *114* (12), 4373-82.

29.   Foloppe, N.; Hartmann, B.; Nilsson, L.; MacKerell, A. D., Intrinsic Conformational Energetics Associated with the Glycosyl Torsion in DNA: A Quantum Mechanical Study. *Biophysical Journal* **2002,** *82* (3), 1554-1569.

30.   Churchill, C. D. M.; Wetmore, S. D., Developing a Computational Model That Accurately Reproduces the Structural Features of a Dinucleoside Monophosphate Unit within B-DNA. *Physical Chemistry Chemical Physics* **2011,** *13* (36), 16373-16383.

31.   Shishkin, O. V.; Gorb, L.; Zhikol, O. A.; Leszczynski, J., Conformational Analysis of Canonical 2-Deoxyribonucleotides. 1. Pyrimidine Nucleotides. *Journal of Biomolecular Structure & Dynamics* **2004,** *21* (4), 537-553.

32.   Wiechelman, K.; Taylor, E. R., Anti-Syn Conformational Range of Pyrimidines with Deoxyribofuranose. *Journal of Biomolecular Structure & Dynamics* **1998,** *15* (6), 1181-1194.

33.   Schlitt, K. M.; Sun, K. W. M.; Paugh, R. J.; Millen, A. L.; Navarro-Whyte, L.; Wetmore, S. D.; Manderville, R. A., Concerning the Hydrolytic Stability of 8-Aryl-2'-Deoxyguanosine Nucleoside Adducts: Implications for Abasic Site Formation at Physiological Ph. *The Journal of Organic Chemistry* **2009,** *74* (16), 5793-5802.

34.   Aliakbar Tehrani, Z.; Fattahi, A.; Pourjavadi, A., Influence of Metal Complexation on Acidity of Cytosine Nucleosides: Part I, Li+, Na+ and K+ Cation. *Scientia Iranica* **2012,** *19* (3), 535-545.

35.   Javan, M. J.; Tehrani, Z. A.; Fattahi, A., Structural Behavior of Sugar Radicals Formed by Proton Transfer Reaction of Deoxycytidine Cation Radical: Detailed View from Nbo Analysis. *Structural Chemistry* **2012,** *23* (4), 1185-1192.

36.   Millen, A. L.; Kamenz, B. L.; Leavens, F. M. V.; Manderville, R. A.; Wetmore, S. D., Conformational Flexibility of C8-Phenoxylguanine Adducts in Deoxydinucleoside Monophosphates. *Journal of Physical Chemistry B* **2011,** *115* (44), 12993-13002.

37.   Yasuniwa, M.; Tokuoka, R.; Ogawa, K.; Yamagata, Y.; Fujii, S.; Tomita, K.-I.; Limn, W.; Ikehara, M., The Crystal and Molecular Structrue of 8-Methyladenosine 3′-Monophosphate Dihydrate. *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis* **1979,** *561* (1), 240-247.

38.   Tavale, S. S.; Sobell, H. M., Crystal and Molecular Structure of 8-Bromoguanosine and 8-Bromoadenosine, Two Purine Nucleosides in the Syn Conformation. *Journal of Molecular Biology* **1970,** *48* (1), 109-123.

39.   Leng, M.; Ptak, M.; Rio, P., Conformation of Acetylaminofluorene and Aminofluorene Modified Guanosine and Guanosine Derivatives. *Biochemical and Biophysical Research Communications* **1980,** *96* (3), 1095-1102.

Chapter 3. Molecular Dynamics Simulations of DNA Containing Unsubstituted O-linked, *Ortho* and *Para* C-Linked Adducts Paired with Cytosine and Guanine Mismatch

3.1. Introduction

It is known that modification of dG at the C8 position not only can affect the barrier between the *syn* and *anti* conformation for the free nucleoside, but also that these lesions have an apparent impact when they are incorporated within the DNA duplex.[1, 2]

Studies have shown that modification of the C8 site of 2′-deoxyguanosine results in the formation of the *syn* conformation. Calculations show that for a class of C8-aryl-dG adducts there is an ~25 kJ mol^{-1} preference for the *syn* conformer of all nucleoside adducts.[3, 4] The most important factors that stabilize the *syn* conformer are the steric bulk of the adducted moiety, as well as an intermolecular hydrogen bond between O5'–H•••N3.[1, 5]

Within the duplex, the *syn*/*anti* conformation equilibration is more complicated due to the complexity of the helix environment. The steric and electronic effects result in three major conformations, which are in equilibrium.[1, 6, 7] There are various factors affecting the conformations including the preference for the formation of the *Watson-Crick* H-bonding interaction vs. possible Hoogsteen H-bonding, as well as bulkiness, spatial orientation and planarity of the attached moiety to the nucleobase, which determine the π-stacking ability of the adducted moiety.[6]

Conformational heterogeneity for the *N*-linked C8-dG adducts, specifically for the carcinogens 2-(acetylamino)fluorene (AAF-dG), 2-aminofluorene (AF-dG) and 1-aminopyrene (AP-dG) have been investigated meticulously.[8, 9] Three main conformations identified for C8-arylamine dG adducts within the duplex have been determined by using techniques such as ^{1}H and ^{19}F NMR spectroscopy, crystallographic analysis, and circular dichroism. The "B-type" (B) conformer emerge from the *anti* conformation of the modified guanine residue. While, the Watson-Crick hydrogen-bonding interaction with the complementary base in the opposing strand is maintained. The adducted moiety is located in the solvent exposed major groove. "B-type" is the least perturbed conformation. A base-displaced "stacked" (S) conformer has been termed as a second conformer in which the modified guanine residue acquires the *syn* orientation, as the adducted moiety is flipped into the interior of the helix. The opposing base is flipped out of the helix, so the Watson-Crick hydrogen-bonding interaction is ruptured. On the other hand, there is often a modification in stacking interactions between adducts and the flanking bases. Flipping of the adducted moiety into the minor groove leads to the adoption of the third conformation, known as the "wedge" (W) (Scheme 3.1). In this case, the modified guanine is mispaired opposite another purine base and acquires a *syn* conformation. However the guanine portion remains stacked within the helix.[10] Therefore, a variety of mutated nucleotides are capable of forming more than one conformer.[11-16] Also, it is well known that there might be conversion between conformations within the duplex because both *syn* and *anti* conformers can be considered as stable helices. Different factors are

involved in the proportions of each conformer. For instance, the influence of ring size on the conformational equilibrium regarding the *N*-linked aromatic adducts has been studied by Shapiro et al.[6] *N*-linked aniline (AN-dG), aminofluorene (AF-dG) and aminopyrene (AP-dG) are examples of specific C8 modified N-linked structures. ^{1}H NMR data supplemented with computational analyses was performed for different mutated sequences. The results prove that the AN-dG modified duplex embraces the B conformation, while the AP-dG adopts mainly the S conformer in the duplex, with a minority of the W conformer. Of the three B, W, and S, conformations identified for the modified duplex with AF-dG, the major (70%) conformer is determined to be the S.[6]
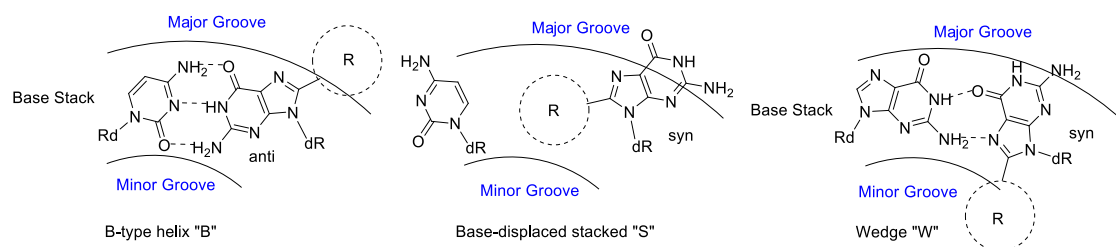
Scheme 3.1. Different three major conformations formed after insertion of mutated nucleobase in DNA double helix.[17]

In general, larger planar ring systems have preference for a stacked, intercalated conformation. Zhou and coworkers (using ^{19}F NMR) performed a conformational analysis for fluorinated substrates.[18] They investigated the conformational heterogeneity of C8-fluoroaminophenyl modified dG (FAF-dG) and C8-fluoroaminobiphenyl modified dG (FABP-dG) (FAF without the methylene linkage to keep the phenyl rings planar relative to each other). Freely rotatable

FABP-dG, has only a B conformer, while the S:B ratio of 55:45 was identified for FAF-dG. This consequence is in accordance with previous ^{1}H NMR structural studies of a specific sequence using ^{19}F NMR spectroscopy.[19] Base sequence is another key feature in the conformational preferences for this class of adducts. For instance, analysis of the conformers for FAAF-dG based on ^{19}F NMR resonances by Cho et al., uncovered a ratio of 5:65:30 for the conformational preference of B:S:W with two cytosines flanking the modified base, and a 5:30:65 proportion with a thymine and adenosine flanking the mutated base.[20]

Conformation and structure of *N*-linked adducts and *C*-linked adducts has had the attention of many scientists.[1, 3, 21, 22] Molecular dynamics simulations implemented by Wetmore *et al*. illustrated that the energy level of *syn* and *anti* conformers of each *ortho* and the *para* C-linked phenolic adducts were similar.[11, 23] Hence, more than one conformation within the duplex was expected. Experimental results provided by circular dichroism experiments involving C8-aryl adducted species revealed regular B-form DNA for both *ortho* and *para* phenol mutated complexes. In the case of the 8-benzothiophenyl-dG adduct, the acquired *syn* conformer and following S-structure in the double helix can be verified by the induced CD band.[24] Further investigations on the effect of the opposing base on the conformations of phenolic-, furanyl-, and benzothiophenyl-dG modified duplexes revealed that C8-aryl adducts incorporated opposite the correctly paired C were found to prefer *anti* B-type structures, while base pairing to a mismatched G lead to W structures.[22, 25]

It is worth considering the toxicological properties of C8 Adducts. Genes are expressed by codons which are made up of three nucleobases, the insertion or deletion of a nucleotide can change the reading frame, known as a frameshift mutation. In general, a frameshift mutation, causes a disorder in the reading of the codons. Frameshift mutations are apparent in some severe genetic diseases such as Tay-Sachs disease and Cystic Fibrosis.[26, 27] They also increase susceptibility to certain cancers and classes of familial hypercholesterolaemia.[28] A frameshift mutation is also involved in the resistance to infection by the HIV retrovirus.[29] Principally, frameshift mutations often result from the bulge stabilization by base displaced S structures.

The unpredictability of C8 adduct conformation, flanking sequences and opposing bases are distinct features emerging from the flexibility of the DNA double helix. For instance, different conformers which are structurally similar in the case of aromatic amines, cause a wide range of toxicity.[30] It has been observed that there is a direct dependence between mutagenicity and conformational preference. The promutagenic conformations of C8 aryl-amine adducts are determined to be the *syn* conformations which trigger S and W conformers in the DNA double helix while the *anti* conformer does not show any toxic effect. Similarly, in modified duplexes with AAF adducts, exclusively the *syn* conformer of the adducted nucleobase can form frameshift mutations. Heterogeneous, *syn/anti* AF mutated duplexes generally cause to G↔T transversions rather than frameshift mutations.[8] AF is known to stabilize G:A mismatches through the exclusive

adoption of the W conformer which evolves when it is mispaired against A.[31] Also, stabilization of the G:G mismatch by AF-dG rationalizes G↔C transversions.[8]

Frameshift mutations occur with especially high frequencies in regions of repeated base sequences, the hexanucleotide 5′-$G^1G^2CG^3CC$-3′ which forms the recognition sequence of the *Nar*I is the most vulnerable hot spot for this type of mutation.[32] The name *Nar*I is derived for *Nocardia aregentinesis* which is the name of a bacterium in E-coli carrying this sequence. Modification at the C8 position of the G^3 guanine by bulky arylamines is known to induce -2 frameshift mutations. Frameshift mutation occurs when nucleobases are deleted or inserted in a DNA sequence and the number of nucleotides added or removed are not a multiple of three, leading to a shift in the way that the sequence is read).[33-35]

Rizzo's findings suggest that modification of the *Nar*I sequence at G^3 by the bulky arylamine IQ, removes two bases from the mutated sequence in the 5′ direction during replication, which leads to a frameshift mutation.[36] Misalignment of base pairing in regions where base repeats can be stabilized by bulky adducts like IQ, increases the probability of this type of lesion. Formation of the S conformation by IQ stabilizes the bulged out sequence, due to loss of two base pairs from the replicated strand. Likewise, acquiring the S conformation in the case of *N*-linked arylamine adducts accounts for the propensity of this lesion to form -2 frameshift mutations.[37]

Moreover, MD and DFT calculations indicated that the C-linked phenoxyl adducts against guanine mismatch is stable. The decanucleotide with flanking purine bases are more affected by the stabilization phenomenon of guanine

mismatch. This is interpreted mainly by stronger base-base interactions.[11, 23] These results are further confirmed by experimental melting temperatures. Hence there are several factors in determining the mutagenic properties of modified C-linked species among them the specific preferred conformation (*syn*), and the variety of complementary base (adduct against guanine mismatch) while taking the specific type of the sequence in to account which can delineate the impact of flanking bases. These results are comparable with aforementioned corresponding N-linked C8-dG adducts derived from arylamine carcinogens.[8] Specifically, since the phenoxyl adduct lesions are better able to stabilize G:G mismatches when flanked by purine bases, purine-rich sites would be expected to show a greater propensity for misincorporation of G opposite the adduct.

This allows for comparison with the N-linked aromatic amine adducts, which have been well-studied within this sequence context. However, the work in this chapter indicates that the conformational and base-pairing preferences of the adducts in *NarI* oligonucleotides can be identified by MD simulations.

In summary, it is known that C8-dG adducts when are incorporated in a double helix against the normal pyrimidine partner C, will typically exist in either the B-form or S-form conformation, or a combination of both in fast exchange. Studies has elucidated that preference of an S-conformation leading to stabilization of the 2-base bulge damage. In the other word, a 12mer NaI sequence in complement with *NarI*′(-2) 10mer sequence due to this damage. For the modified duplex with [4F-PhO]dG studied in *NarI*, there was no stabilization of the bulge. This implies unfavourablity *para*- flouro substituted phenolic moiety of modified guanosine to

flip out of the helix and acquiring a base displaced stacked S conformation.[17] The experimental data showed that the conformation of the adduct [4F-PhO]dG within the duplex hybridized to the complementary strand containing a normally base-pairing C opposite was a B-type conformation. To further support this argument, incorporation of the modified residues in the sequence with abasic strand implied the disfavouring the S conformation is typical of adducts containing small ring systems, or ring systems with considerable flexibility.[11, 17, 23, 38] Conformational studies on the structurally similar aniline adduct (AN-dG) indicates the exclusive preference of B conformer opposite C.[12, 13, 23]

In the B conformation, the adduct exists in an *anti* glycosidic orientation. Experimental studies (measuring melting point of strand, T_m) on AN-dG modified with complementary cytosine base pair confirms the computational results,[13, 15] whereas the [4F-PhO]dG adduct induced a greater destabilization in the studied sequence. This similar impact on the duplex stability can be ascribed to the uptake of the same B conformation. An extra energy is required for solvation of the hydrophobic aromatic ring. This is due to the adduct in major groove which is exposed to solvent. Moreover, the steric interactions of the modified nucleobase with complementary and flanking bases contributes to in stabilization of the double helix. [39]

To investigate the conformational and structural changes induced by the formation of unsubstitued O-linked phenoxyl adduct ([PhO]dG) at the nucleobase and nucleoside level,[40] quantum mechanical calculations were performed. To evaluate the consequence of bridged oxygen, the structure of *ortho*- and *para*- phenolic C-

linked mutagenic products ([o-PhOH]dG and [p-PhOH]dG) previously studied experimentally and theoretically, were taken into account.[3, 22, 41, 42] Furthermore, molecular dynamics simulation (MD) are presented in this chapter to evaluate important physical features of mutated adducts in the *Nar*I sequence duplex at molecular time scales. The high probability of -2 deletion mutation, caused by modification of guanine at G^3 site of *Nar*I sequence, is due to the nature of nucleobase and position of the G^3 site.

Mutation of the guanine in the G^3 position induces an especially high frequency of -2 deletion mutations. Tendency to mutate of guanine is modulated by the nature of the nucleotide in the G^3 position. Free energies were assessed relative to natural *Nar*I sequence for each mutation to provide the most stable conformer of adducts after incorporation in double helix.

In this chapter we try provide valuable information for evaluating interaction between nucleotides in a duplex by MD simulations on unsubstituted O-linked, as well as *ortho*- and *para*- C-linked adducted DNA strand.

3.1.1. Molecular simulation approaches

Molecular dynamics is proven versatile tool for investigating normal and damaged DNA oligonucleotides. Molecular models are capable of providing insight into chemical and biological function by giving information about conformation, hydration, and even thermodynamic quantities.

Perhaps the most noted success of this approach was discovering the structure of DNA by Watson and Crick using a wire model.[40] The concept of extending

molecular modeling to dynamic systems was triggered by the advent of computers. The first attempt at applying a molecular dynamics approach had poor accuracy and necessarily involved many simplifications, such as *in vacuo* simulation, as well as removing explicit hydrogen atoms due to the inadequate computational resources. Fundamentally, MD modeling of a system is based on using classical Newtonian mechanics. Indeed, it is an eminent methodology in identifying the dynamics, and thermodynamics of macromolecules. Following the advent of modern computers, simulations on large systems with over 10^6 atoms[43] and also timescales of around one microsecond[44] have become possible. The improvement of hardware and software have produced promising results using this method.

In molecular mechanics and dynamics, the total energy of a system consists of the kinetic energy plus the potential energy. The energy of atoms while moving determines the kinetic energy. On the other hand, the potential energy is defined as the sum of the electrostatic and van der Waals energies. Several simplifications have been taken into account in molecular mechanics and dynamics calculations to achieve more tractable computations than quantum mechanical methods. Atoms are considered as soft or hard spheres, whereas in quantum mechanical methods, atoms are collections of electrons and nuclei. Newton's laws of motion (classical potential energy equations) are used instead of the quantum Schrödinger equation. The kinetic energy of each atom is defined as half of the mass (m) of the object times the velocity (v) squared

$$K = \frac{1}{2}mv^2$$

A computer program sums various forces acting upon each atom, such as electrostatic interactions and van der Waals forces, which have an effect on the bonds that constrain atoms, and finally calculates the positions that the atoms will occupy in a specific time interval. When a system is evaluated within multiple time steps, a trajectory is produced that shows the dynamic behavior of the system over time (position as function of time). In fact, molecular dynamic trajectories contain snapshots of the simulated system.

3.1.1.1. Force fields

MM and MD use potential energy equations that are defined to calculate the energy. They describe the atoms and their connections parameters (bond lengths, torsions etc.) that are used in the potential energy equations. In other words, energetic properties can be described by force fields which illustrate the forces acting on each atom of the molecule. There are many force fields designed for different purposes based on the type of system that is investigated (protein, DNA, inorganic complexes, and small drug molecules) such as OPLS, MMFF, SIBFA, GROMACS, CHARMM, and AMBER.

Popular modern force fields include those associated with the *Assisted Model Building with Energy Refinement* (AMBER)[45] and *Chemistry at HARvard Molecular Mechanics* (CHARMM)[46] software suites. A modern force field attempts to mimic the behavior of molecules as precisely as possible while taking into account the cost of computation.

3.1.1.2. The AMBER potential energy function

In this study we have employed the AMBER force fields and associated simulation software. [45] The total energy in the Amber force field consists of the following function to describe the potential energy of the system:

$$V(r) = \sum_{Bonds} K_b (b - b_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 +$$

$$\sum_{tortions} (\frac{V_n}{2}) \, (1 + \cos(n\phi - \delta)) + \sum_{nonb_{ij}} \left(\frac{A_{ij}}{r_{ij}{}^{12}}\right) - \left(\frac{B_{ij}}{r_{ij}{}^{6}}\right) + (\frac{q_i q_j}{r_{ij}})$$

There are three terms which include the energy of the bonds between atoms in the system, plus a nonbonded term that accounts for van der Waals and electrostatic interactions. With the intention of making the calculations tractable a cutoff is used for the van der Waals term. This cutoff is usually between 7 and 10 Å, and in the 2008 version of AMBER the cutoff is employed at 8 Å by default.[47] This cutoff is also applied in the Ewald summation for the electrostatic calculations. This function results in the potential energy of a molecule for a given physical conformation, so the flexibility and dynamics of the molecule leads to a change in the initial potential energy.

The first and second terms describe the energy of the bond, bond length and bond angle. The bond length constant, K_b(kcal/mol·Å²), and the angle constant, K_θ(kcal/mol·rad²), come from experimental data. The balanced bond length is denoted by b_0(Å) and the measured equilibrium bond angle is denoted by θ_0(rad). In molecular dynamic simulations, b(Å) and θ(rad) are the temporal bond length and angle. As a bond can be modeled by a spring with its own equilibrium length,

so can the energy required for stretching or compressing approximated by Hooke's law for an ideal spring. The quantity V_n (kcal/mol) is derived to reproduce the torsional profile where n is the period between the maxima and minima. The phase angle Φ (rad) gives the location of the minima and maxima, and the torsion angle is denoted by δ (rad). The nonbonded term consists of two important components: an electrostatic term using Coloumb's law, and the van der Waals potential, which is given by the Lennard-Jones function. It is well known that one of the most commonly used functions for the van der Waals potential is the Lennard-Jones approximation, representing the attractive and repulsive van der Waals forces between atoms with no net electrostatic charge. These forces vary for different atom types. r_{ij} (Å) is the distance between two atom centers, A_{ij}(kcal·Å^{12}/mol) *and* B_{ij} (kcal·Å^6/mol) are the roots of the product of the A and B van der Waals parameters of the two interacting atoms.

$$A_{ij} = \sqrt{A_i A_j}$$

$$B_{ij} = \sqrt{B_i B_j}$$

The constant parameters A and B stem from experimental data, whereas the van der Waals radius (r (Å)) and potential energy (ε(kcal/mol)) are determined from the following relationships:

$$A = 4\varepsilon r^{12} \qquad\qquad B = 4\varepsilon r^6$$

q_0 is known as the atomic point charge, and the sign of the combination of the charges determines whether the electrostatic force is attractive or repulsive

$$F = \frac{q_1 q_2}{4\pi\varepsilon r^2}$$

3.1.1.3. Parameter development

An important process which is involved in molecular dynamic calculations is the development of appropriate atom and bond parameters. As explained in the previous section there are constant values in the AMBER force fields [48-50] and they are taken from empirical parameters. The other parameters are calculated theoretically using quantum mechanical methods. The empirical constants for the bond length K_b and the bond angle K_θ are identified using experimental techniques, such as high-resolution X-ray crystal structures, NMR studies, Raman spectroscopy and infrared spectroscopy. There is a direct correlation between the force constant and the force needed to deviate from the equilibrium value, namely, the more rigid the bond or angle the larger the force constants. This is related to the rigidity of an ideal spring based on Hooke's law.

Since, in this study we modeled DNA lesions containing bonds or atom types that are not found in the force field parameter set, we developed new parameters. Bond parameters and atom types are usually assigned by comparison with chemically similar bonds and atoms which are defined by default in the force field. One of the most difficult steps is to assign correct partial charges to new molecules. We calculated the electrostatic potential using the HF method with the 6-31G* basis set[51] using Gaussian software.[52] Subsequently, during MD simulations the restrained electrostatic potential (RESP) algorithm[48, 53] was then used to fit the charges to atomic centers. From the dependence of the partial charges on the conformation of the molecule, we can determine the partial charges from multiple conformations of the flexible mutated complexes by rotating the glycosidic bond

to adopt both *syn* and *anti* conformers. Since there is no change in the partial charges during the MD simulation, we can use the geometrical parameters of conformations which can be calculated through the quantum mechanical method.

3.1.1.4. Electrostatics

Because of complexity, stochastic relations and other variables, not all real world problems can be represented in a model. Attempting to use analytical models for such systems usually require too many assumptions so that the resulting solutions are likely to be inferior and inadequate for implementation. In order to render more approachable force fields in this study, some simplifications are exerted on the electrostatic component which contributes to the reduction of the computational complexity. Firstly, the charges on atoms are assigned as a single point charge at the center of the atoms. Secondly, the surrounding electrostatic environment does not affect the fixed point charges during the phase of the simulation and finally, the particle mesh Ewald approximation (a method for computing long-range interactions, e.g. Coulombic interactions) is employed during calculating the electrostatic interactions. The cutoff which has been applied to the electrostatic interaction is quite similar to the cutoff used for the Lennard-Jones simplification. An overview of typical molecular modeling and dynamics procedures

To accurately model a large system such as the damaged DNA described in this work, it is necessary to first obtain reasonable starting coordinates for the atoms in the system. The topological properties which describe the covalent connectivity

of the molecules are the starting coordinates for the atoms in the system. In addition, structural properties of the starting conformation of the molecule (taken from physical measurements such as X-ray structures, NMR data or a theoretical model) should be considered as well.

Some modifications are often applied to these coordinates for various purposes such as altering DNA sequences, substituting covalent DNA adducts, addition of ions to neutralize the system, and solvation with explicit water molecules. Subsequently the system is subjected to energy minimization, which leads to relaxation of the solute and solvent. During subsequent equilibration, the system is slowly heated to the desired temperature, and decreasing restraints are used to allow the solute to slowly move toward a lower-energy state. Once the desired temperature has been reached, all molecules are unrestrained and the actual production molecular dynamics can begin. The simulation is run for a length of time, often on the nanosecond timescale, and then the trajectory is analyzed for structural features of interest.

3.1.1.5. Energy minimization

These factors raise the potential energy of mutated complexes, which prohibit generation of appropriate atomic velocities, bond lengths, and other quantities when a force field is applied. It is essential to adjust the conformation of the damaged bases prior to performing the production step in molecular dynamics. During the minimization step, the geometrical and topological parameters (bond lengths, angles, and dihedrals) are improved. This reduces the number of atomic

collisions, which assists in forming hydrogen bonds, and ultimately decreases the potential energy of the system to a reasonable level. During this process, the potential energy of a system decreases. Another significant highlight in this step is that there is no restriction placed on the total energy while atoms are moved and bonds are changed. Also, time is not taken in to account in minimization. An algorithm needs to be defined for minimization. Steepest descent (SD) and conjugate gradient descent (CG)[54, 55] are among the most common minimization algorithms employed in molecular modeling and dynamics. Reaching a local energetic minimum is possible via a combination of both of these algorithms. The SD algorithm is more efficient in the beginning of a process, but as we get closer to the local minimum the CG algorithm becomes more efficient. In other words, SD is implemented for the preliminary minimizations of a structure, because the initial energy of a molecule is expected to be very high. At this time, moving towards a local minimum will result in a significant drop in the level of total potential energy. However the CG algorithm for minimization can handle shallow gradients more appropriately than SD minimization, so SD is usually followed by CG descent minimization.

3.1.1.6. Molecular dynamics simulations

Due to the complex nature of bio molecular systems which usually consist of a large number of particles, it is analytically impossible to calculate the properties of these systems. A possible technique to circumvent and examine a complex system at the atomic level, and model the changes of that system within a specific time, is

Molecular Dynamics (MD). During the process of a simulation, a force field is applied to a molecular system, so the interaction of atoms in a limited period of time under the common laws of physics are considered. MD simulation is a virtual laboratory and connects experiments with theory precisely, which can help scientists develop new drugs and medical technologies, and prompt experimental research by creating new hypotheses.

Molecular dynamics simulations calculate the motion of the atoms in a molecular assembly. A MD trajectory is generated using Newton's second law of motion to determine the atoms' positions, velocities, net force and acceleration in different times.

$$\frac{d^2 x_i}{dt^2} = \frac{F_{x_i}}{m_i}$$

The above equation shows that the acceleration of an atom in the system in a direction equals the force in that direction divided by the atomic mass. In the initial step of MD, the velocities of the atoms in the starting structure are assigned velocities drawn from Boltzman distribution reflecting atomic velocities at the temperature of the simulation. This is a Boltzman distribution generated around the average velocity, v_0, at temperature T, given by the following equation:

$$v_0 = (\frac{K_B T}{m})^{\frac{1}{2}}$$

K_B is the Boltzman constant, T is the temperature (K), and m is the atomic mass. The force F acting on an atom at position x, can be determined using the potential energy V(x), generated by the force field:

$$F = \frac{-\partial V\left(x\right)}{\partial x}$$

If the force acting on an atom is known, the acceleration a, can be determined using Newton's Second Law:

$$a = \frac{F}{m}$$

Once the initial velocity and acceleration of an atom are known, the new velocity after one time step can be determined for each atom. The potential energy function for the new conformation leads to evaluating a new acceleration. By repeating this process over the duration of the simulation, the velocities and acceleration can be obtained. The fastest motions of the system are limiting factors of the time step. Vibration of bonds with hydrogen atoms are among the fastest motions. In general, the SHAKE algorithm[56] is used to constrain the fast vibrations, which allow a longer time step, and result in the generation of long trajectories more efficiently. However, to make the calculations more feasible for modeling the solvent, the angle between the hydrogens of water is kept fixed. Insertion of constrained on hydrogen atoms of water and mobile hydrogen atoms of the DNA are essential because they decrease inaccuracies in the calculations. These restriction allow to increase time steps without facing instability in the simulation.

3.1.1.7. Temperature and pressure regulation

An important facet of MD simulations is treating the system under actual temperature, pressure, and energy conditions leading to statistical mechanical ensembles. In ensembles, several properties of the system are constant for the

duration of the simulation. A microscopic state of the system is given by a point space of the system, which is defined by the positions and the momenta of the *N* atoms of the system. The macroscopic state is defined by the number of particles, volume, energy, pressure, temperature, etc. A thermodynamical ensemble is a collection of microscopic states that all feel an identical macroscopic state.

In this research we sought to represent the conditions inside a cell, and therefore "isothermal-isobaric" or "NPT" ensembles were employed which clarifies that the number of molecules in the system N, the temperature T, and the pressure P, are held constant. Normally, the pressure and temperature are set at 1 atm and 310 K, respectively.

Several methods[57] are used to mimic a heat bath, and a thermostat attempts to keep and stabilize the temperature. The Berendsen thermostat[57] used for the production step in MD, and also the Langevin thermostat[58] for equilibration in some simulations are the most common. The latter is employed in this survey. Absorbing the energy/heat or contributing the kinetic energy of the system is the key feature of a heat bath which helps to maintain the temperature. Thereby, a system which has been properly equilibrated during simulation has an even temperature, with very small fluctuations (up to ~1°C around the desired temperature in each time step). A barostat keeps the pressure constant by increasing or decreasing the volume of the system.[59] The fluctuation of pressure is approximately 200 atms which, on average, is at or very close to the target pressure. A user-supplied coupling constant adjusts the temperature of the heat bath and the pressure at the determined interval. It is possible to apply more

computationally efficient larger constants, but deviations from the target temperature and pressure should be taken into account.

3.1.1.8. Modeling of the solvent

Explicit modeling of water molecules is employed in this project. Several common explicit water models have been specified which vary in the defined parameters for bond angles, charge, and van der Waals radius. In the MD simulations of DNA, the TIP3P water model is often applied.[60] In the TIP3P water model, a rigid triangle is defined in which three point charges are centered on the oxygen atom and the two hydrogens. For computing Lennard-Jones interactions, only the van der Waals radius of the oxygen is considered while the hydrogens are excluded.

1.1.1.1.1. Periodic boundary methodology

Since explicit modeling of a realistic volume of water for biochemical reactions is difficult computationally, periodic boundary conditions are used in MD simulations. This technique allows the modeling of very large systems, but introduces a level of periodicity that is not present in nature. The periodic boundary approach in simulations represents a box that contains the solute and surrounding solvent. Subsequently, replication of this box in an infinite lattice of images of itself means that any given solvent atom has neighbors on all sides, with no solvent/vacuum boundaries in the system.[61] Identical boxes are an important condition of this model, which allows those atoms or particles that are leaving the

central box to reenter from an image box that is directly opposite their exit point. Interactions between atoms are calculated using only the closest replicas, so interactions are not double counted. Adjusting the size of the box is an important condition when using a periodic boundary as it prevents the atoms of the solute from interacting with image atoms. This gives rise to the concept of so-called boundary effects. Ultimately a box in which walls are at least 10 Å (in this project it set to 8 Å as mentioned in computational details section) the from the nearest solute molecule is depicted in MD simulations.

3.1.1.9. Particle mesh Ewald evaluation of electrostatic interactions

Imposition of the periodic boundary condition has the advantage of facilitating the use of the particle mesh Ewald (PME) method, which considers particles in a finite distance and applies some modification to the Coulomb's law to evaluate electrostatic interactions.[62, 63] The basis of PME is that the sum of the short-range potential and the long-range potential of an atom is equal to the total electrostatic potential of the given atom. As there is a direct ratio between of the number of atoms and an increase in electrostatic potential, considering both components during calculation would not be tractable for models that may exceed 50,000 atoms. In order to decrease the computational complexity, only short range interactions, atoms with a cutoff distance similar to the Lennard-Jones calculations, are considered and calculated directly. PME requires the generation of a charge distribution mesh for the simulation volume, with the local charge distribution represented as a point in the mesh within periodic boundary

conditions. A distinct disadvantage of PME is the introduction of some error into MD simulations which can be rationalized by the described approximations.[45]

3.1.1.10. Limitations of molecular dynamics simulations

MD simulations are a very beneficial and important tool for modeling a wide variety of large systems of proteins, nucleic acids and their complexes. It can give us information about small and large scale conformational changes, dynamic processes such as ion transport in biological systems and can determine and construct 3D structures. However, there are some limitations to this method that mostly stem from restricted computational resources. On the other hand, simplifications applied in the molecular mechanical force fields cause error. Furthermore, the basis of this approach, which treats the system using Newton's laws of motion, may be inaccurate due to the inevitable presence of features (e.g. Bond formation and dissociation) that would require quantum mechanical calculations to get precise results. Limitations in timescale, especially in many interesting cell-level biological events, are another source of error. The study of systems of very large size, such as the DNA and some (cell-level protein/substrate complexes), is precluded as a result of computational limitations.

Finally, traditional MM/MD simulations facilitate treating larger systems for longer timescales. Increasing the power and decreasing the cost of computational resources, providing greater detail and longer timescales in all kinds of simulations is the main purpose and future of MD simulation.

3.2. Computational details

3.2.1. DNA Model

Classical molecular dynamics (MD) simulations were performed on the canonical 5′-CTCG^{1}G^{2}CG^{3}CCATC 12-mer oligonucleotide from the *Nar*I recognition sequence, while sequences with the unsubstituted O-linked (^{PhO}dG), *ortho* ($^{o\text{-}PhOH}$dG) or *para* C–bonded ($^{p\text{-}PhOH}$dG) adducts at the G^3 site were also considered. The G^3 site in the *Nar*I sequence was specifically chosen since this is a hot spot for frameshift mutations induced by N–linked C8-dG adducts.[12, 16] These have similar properties to the phenoxyl dG adducts of interest in the present work. The initial coordinates for the θ and χ dihedral angles in the *syn* and *anti*-conformations of ^{PhO}dG were set as close to the values in the DFT nucleoside minima as possible based on steric constraints imposed by the duplex environment (the steric hindrances from flanking and complementary bases in DNA do not allow to set the angles based on the calculated results from nucleoside or nucleotide models). Partial atomic charges for the O– and C–linked dG phenoxyl adducts were determined using the RED.v.III.4[64] program. MD simulations were conducted for 20 ns (as the result of 40 ns and 20 ns simulations were pretty similar based on RMSD plots, to save the time and computational cost we performed 20 ns simulations) using the PMEMD module of the AMBER 11 or 12 software.[65] The parmbsc0 modification to the parm99[66] force field was implemented for natural DNA, while generalized AMBER force field (GAFF)[67] parameters were used to describe the adducts and assigned using ANTECHAMBER 1.4.[68] Initial structures

were prepared using the NAB[65] program and GaussView to modify G[3] to the adducts. The LEaP module of AMBER 11 was used to prepare the systems for MD. The adducted *Nar*I sequences were solvated with modified TIP3P[69] water in an octahedral box extending up to 8.0 Å from each solute species. The SHAKE[70] method was used to constrain all bonds to hydrogen. To avoid edge effects, the periodic boundary condition was applied in all calculations. The addition of 22 Na^+ ions neutralized the system. The nonbonded cutoff was set at 10 Å for the Lennard-Jones interactions, while long-range electrostatic interactions were treated with the particle-mesh Ewald (PME) method. A force of 500.0 kcal mol^{-1} was used to restrain the positions of the adduct atoms. Within these constraints, initial minimization of the system was performed for a total of 1000 steps (500 steps under the steepest descent algorithm followed by another 500 steps under the conjugate gradient minimization algorithm with the DNA helix held fixed via a force constant of 500 kcal mol^{-1} Å^{-2}). In addition, 1000 steps of steepest descent minimization followed by 1500 steps of conjugate gradient minimization were performed on the entire system and the system was minimized without restraints. The system was then gradually heated from 0 to 300 K. The temperature-bath coupling was achieved by the Langevin temperature equilibration, and a 10 kcal mol^{-1} Å^{-2} force constant of was applied to restrain the solute over the first 20 ps. However, heating followed by maintaining the temperature and pressure constant at 300 K and 1 atm, during a total of 20 ns unrestrained MD simulation. The DNA helix in the solvated systems was restrained to the initial coordinates using a weak force constant of 10 kcal mol^{-1} Å^{-2} while warming. The isobaric condition was

enforced for each system in order to relax the positions of the solvent molecules. Finally, MD simulations were performed without position restraints for 20 ns with 2 fs time steps.

The stability of the MD trajectories was assessed using the root mean square deviations (RMSD). The results were analyzed using the ptraj module of Amber 11 or 12. The PyMOL and VMD software were used to visualize the trajectories and depict structural representations. Clustering with respect to the atoms forming the θ, φ and χ dihedral angles of ^{PhO}dG was performed using the ptraj module of AMBER to obtain the representative structure of the adduct. H–bonding and stacking interactions were evaluated for MD structures. H–bonding interactions and stacking binding strengths were determined using DFT calculations.

3.2.2. Free energy calculations

Free energy calculations were performed using the molecular mechanics/Poisson–Boltzmann surface area method (MM-PBSA method).[71] Specifically, snapshots were taken every 400 and 40 ps, and water was removed to calculate the entropy and other free energy terms, respectively. Furthermore, the total free energy (G_{Tot}) was estimated as the sum of the molecular mechanics energy (E_{MM}), the solvation free energy (E_{sol}) and the entropy as:

$$G_{Tot} = E_{MM} + E_{sol} - TS$$

where T is the temperature and S is the sum of the rotational, translational and vibrational entropies for DNA as estimated using the MMPBSA method. The E_{MM}

term was calculated as the sum of the internal energy (E_{int}), the van der Waals interaction (E_{vdW}) and the electrostatic energy (E_{elec}):

$$E_{MM} = E_{int} + E_{vdW} + E_{elec.}$$

E_{int} is the sum of the energy terms stemming from deviations of the bond lengths (E_{bonds}), bond angles (E_{angles}) and dihedral angles ($E_{dihedrals}$) in the DNA helix from their equilibrium values:

$$E_{int} = E_{bonds} + E_{angles} + E_{dihedrals}$$

E_{sol} was estimated from the electrostatic solvation energy (E_{PB}) and the nonpolar contribution to the solvation free energy (E_{NP}):

$$E_{sol} = E_{PB} + E_{NP}$$

The E_{PB} term was calculated using the Poisson-Boltzmann method as implemented in the MMPBSA program and distributed as a part of AMBERTOOLS.

3.3. Results and discussion

Carefully designed computational models are of great importance. Since nucleobase and nucleoside models cannot provide a realistic picture about conformational flexibility of bulky adducts meticulously, their geometrical properties identified by incorporating into the duplex.

Theoretical and experimental results show that, the Watson-Crick hydrogen bonding, which forms normally with *anti* conformers is generally stronger than the Hoogsteen bonding, which stabilizes the *syn* conformer in DNA. However, the modified guanine with *ortho* substituted aryl moiety against cytosine prefers

Hoogsteen hydrogen bonding regardless of adopted *anti/syn.* Consequently, substitution of the *ortho* adduct in the (ODN1 = 5′-CCATXCTACC-3′ and ODN2 = 5′-GGTAGXATGG-3′) DNA sequences, irrespective of the type of the studied DNA sequence, may not lead to base-substitution mutations in DNA.[22, 23] In contrast, in the case of the *para* adduct there is not any noticeable difference in the stability level of Watson–Crick and Hoogsteen pairs with cytosine. Interestingly, the probability of forming Hoogsteen pairs of the *para* adducts with guanine and cytosine is similar. It is worth noting that in this sequence *para* adduct prefers *syn* conformation which may lead to stability of base mismatch in DNA.[22, 23]

There is a direct relation between the conformational flexibility of some bulky adducts such as N-linked adducts, and the type of sequence.[72-76] Steric clashes are among the most important determining factors in analyzing the conformational preference of bulky adducts in DNA.[77] MD simulations on the *anti* or *syn* con-formation of [o-PhOH]dG or [p-PhOH]dG in two types of DNA sequences (ODN1 = 5′-CCATXCTACC-3′ and ODN2 = 5′-GGTAGXATGG-3′) differ in the type of flanking bases (pyrimidines or purines) with the damaged base; and complementary bases do not have any effect on the conformational preference of modified bases.[22, 78] These results indicate that the phenoxyl moiety resides inside the DNA helix, which is different from other bulky groups. In addition, MD simulation provided complementary data that explains the observed experimental differences in the strand such as melting temperatures. Experimental evidence shows that in the case of C8-dG adducts the complementary base has a significant effect on stability of the double helix, which leads to less decrease in duplex stability when the adduct

is mismatched against quinine compared to cytosine. A *syn* preference of the adducts is determined theoretically in both duplexes in agreement with experimental results. DFT (B3LYP and M06–2X) calculations on MD structures were implemented to evaluate the specific stacking interactions between the adduct and the flanking and surrounding nucleobases. *De facto*, the conformation adopted by the adducts in DNA could not be determined based only on the measured melting temperatures, and these experimental results should be used along with the DFT stabilities to confirm a *syn* preference regardless of the sequence or opposing base.

MD and complementary DFT calculations also suggested the stabilization of the G:G mismatch by the C-linked phenoxyl adducts. These results clarify the mutagenic properties of C–linked phenoxyl adducts which are comparable with the formerly studied corresponding N–linked C8–dG adducts derived from arylamine carcinogens. Purine-rich sites may induce G:G mismatch stabilization, because the phenoxyl adduct flanked by purine bases shows more stability. MD simulations will be crucial to provide greater insight into the conformational and base-pairing preferences of the adducts in biologically relevant systems which will aid in explaining and interpreting the experimental results.

Molecular dynamics (MD) simulations were conducted with the ^{PhO}dG, $^{o\text{-}PhOH}$dG, and $^{p\text{-}PhOH}$dG adducts placed in the *syn* and *anti* conformations at the G^3 position in the *Nar*I (5′–CTCG^{1}G^{2}CG^{3}CCATC–3′) sequence and paired against either the complementary C or G mismatch. Two possible orientations of the phenoxyl moiety with respect to the nucleobase, corresponding to θ values of ~0° and ~180°

111

(Figure 3.1 definition of θ), were considered for the *syn* and *anti* conformations of the adduct. Because of structural symmetry in $p\text{-PhOH}$dG, rotation of the θ dihedral to 0° and 180° lead to the same structure, so θ values were not considered for the *para* adduct. The two orientations of θ deviate from planarity (~0° and ~180°) around 10−55° to eliminate the steric clashes with neighboring residues in the initial structures for the minimization step (the first step in MD simulation).[17]
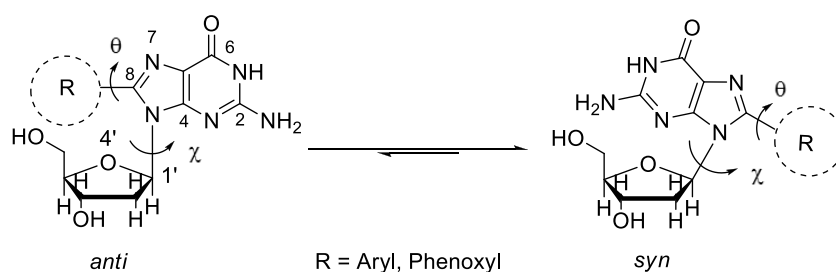
Figure 3.1. *anti−syn* equilibrium for C8-dG adducts $^{\text{PhO}}$dG (R = phenoxy) and $^{o\text{-PhOH}}$dG and $^{p\text{-PhOH}}$dG (R = aryl). Dihedral angle χ [∠(O4′−C1′−N9−C4)] defines the glycosidic bond orientation to be *anti* (χ = 180 ± 90°) or *syn* (χ = 0 ± 90°), and θ defines the degree of twist between the nucleobase and the C8 substituent R.

3.3.1. Adducts against cytosine

Figure 3.2 summarizes the distribution of χ, θ, and ϕ, which was achieved by performing ptraj calculations on the results gathered from the production step of the MD simulation.

The distribution of θ, χ, and ϕ of $^{\text{PhO}}$dG:C show a unimodal distribution of χ, and bimodal distribution of θ and ϕ for the *anti (θ~0°)* structure. In the case of the *anti (θ~180°)* structure, a bimodal distribution of χ and ϕ is identified, while the distribution of θ is unimodal. As explained, the distribution of χ, θ, and ϕ is similar in *syn (θ~0°)*, *syn (θ~180°)* and *anti (θ~0°)*. The ϕ dihedral is not defined in C−linked

adducts (o-PhOHdG, and p-PhOHdG), so a bimodal distribution of χ and θ is determined for *anti (θ~0°)* and *anti* mutated species of *ortho* and *para* C–linked adducts respectively. In contrast, the distribution of χ and θ is unimodal for *anti (θ~180°)*, *syn (θ~0°)* and *syn (θ~180°)* of o-PhOHdG. In the case of p-PhOHdG, unimodal and bimodal distribution of χ and θ is observed, respectively. Overall, each mutated species with a specific theta value can be separated into various types based on the distributions of different dihedrals.
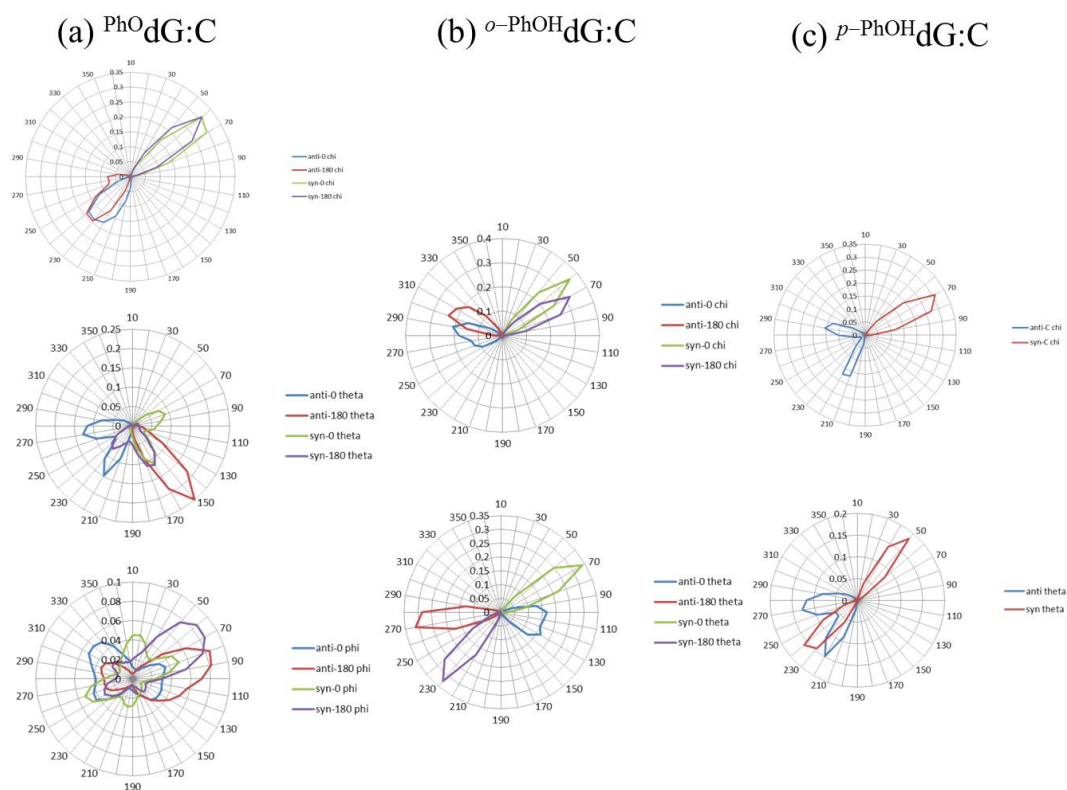
Figure 3.2. Percent distribution of the χ (degrees) θ (degrees) and ϕ (degrees) dihedral angles throughout the 20 ns trajectories for the (a) ^{PhO}dG, (b) o-^{PhOH}dG and (c) p-^{PhOH}dG adducts paired against cytosine, *anti*-0 (blue), *anti*-180 (red), *syn*-0 (green), *syn*-180 (purple).

113

3.3.1.1. The free energy calculations

To investigate the affinity for binding of all mutated compounds when paired against cytosine in the intercalation binding site, the free energies of these compounds were calculated by the MM-PBSA method. The calculated free energies suggest that the *anti* conformation of the adduct is more stable than the *syn* conformation as is known for natural dG (Table 3.1). Based on the distribution of dihedral angles, the lowest energy conformer calculated for [PhO]dG is *anti* ($\theta\sim180°$)-2 with $\theta\sim151.4°$, $\chi\sim225.69°$ and $\phi\sim268.18°$ and the most stable

conformer for [o-PhOH]dG is *anti ($\theta\sim180°$)* with $\theta\sim270.79°$, $\chi\sim304.35°$. Finally, the *anti*-3 conformer identified for [p-PhOH]dG with $\theta\sim231.06°$ and $\chi\sim283.81°$, comparable to the average χ ($\sim249°$) adopted by dG at the G^3 position in the same sequence.

It has been shown from NMR studies with C8-dG adducts of N-linked adducts that it is possible to have more than one conformation in the same DNA sequence.[79, 80] Based on this, the structure of the lowest energy (<5 kcal/mol) has the same characteristics.[81] The O-linked and C-linked adducts are predicted to show behavior similar to N-linked adducts; thus, in addition to the conformer selected as the global conformer, there are also some other stable conformers. Based on the Gibbs free energy values obtained (see Table 3.1 and Figure 3.4), there are five coexistent stable conformers for [PhO]dG (i.e. four conformers plus the global conformers), three coexistent stable conformers for [p-PhOH]dG (i.e. two

conformers plus the global conformers). Moreover, based on the same table, there is no stable conformer for o-PhOHdG.

On the other hand, the *syn (θ~180°)-1* (θ~237.38°, χ~59.33° and ϕ~292.43°, ΔE=84.43 kJ mol⁻¹), *anti (θ~0°)-4* (θ~111.48° and χ~249.59°, ΔE=90.27 kJ mol⁻¹), and *syn*-1 (θ~234.82° and χ~60.92°, ΔE=44.31 kJ mol⁻¹) are the least stable structures for PhOdG, o-PhOHdG, and p-PhOHdG respectively (Table 3.1). The stability of the most stable structures can be explained by less steric hindrance between the modified nucleoside and its complementary base (cytosine), and less distortion of the dihedral angles mutated nucleotide. In the case of O-linked adducts, *anti (θ~0°)-2* (θ~148.34°, χ~215.17° and ϕ~287.73°, ΔE=12.55 kJ mol⁻¹). So the mutated residue could assume a number of orientations (*syn/anti*), in the global minimum and other mentioned local minima.

Table 3.1. MM-PBSA total free energy (kJ mol^{-1}) for the *Nar*I sequence with the mutated adduct incorporated at the G^3 position opposing cytosine from 20 ns MD simulations.

Adduct	Adduct conformation	E	E_{int}	E_{elec}	E_{vdW}	E_{MM}	E_{NP}	E_{PB}	E_{sol}	$-TS$	G	ΔE
^{PhO}dG:C	*anti* ($\theta\sim0°$)-1	-21247.94	4761.81	15879.66	-1804.78	18932.43	107.49	-26073.56	-25966.07	-2607.47	-23855.41	18.03
	anti ($\theta\sim0°$)-2	-21252.80	4747.46	15900.25	-1801.69	18941.60	107.32	-26100.71	-25993.39	-2608.10	-23860.89	12.55
	anti ($\theta\sim0°$)-3	-21248.40	4752.90	15880.92	-1810.04	18919.80	107.19	-26079.63	-25972.43	-2610.10	-23858.51	14.93
	anti ($\theta\sim0°$)-4	-21245.72	4758.46	15866.02	-1798.02	18921.85	107.32	-26060.80	-25953.48	-2607.30	-23853.03	20.41
	anti ($\theta\sim180°$)-1	-21225.52	4787.96	15961.00	-1819.41	19026.07	106.48	-9051.92	-8945.43	-2599.85	-23825.33	48.11
	anti ($\theta\sim180°$)-2	-21261.00	4740.76	15908.99	-1805.05	18940.47	107.61	-26081.72	-25974.10	-2612.45	-23873.44	0.00
	anti ($\theta\sim180°$)-3	-21251.04	4765.32	16006.02	-1821.40	19046.61	106.69	-26214.56	-26107.87	-2598.93	-23849.93	23.51
	anti ($\theta\sim180°$)-4	-21249.32	4750.47	15955.22	-1808.50	18993.18	107.32	-26141.46	-26034.15	-2608.39	-23857.71	15.73
	syn ($\theta\sim0°$)-1	-21233.38	4737.71	16042.67	-1813.62	19062.97	106.11	-26180.38	-26074.27	-2600.27	-23833.65	39.79
	syn ($\theta\sim0°$)-2	-21228.07	4743.74	16077.86	-1814.68	19103.18	106.15	-26226.86	-26120.71	-2600.36	-23828.42	45.02
	syn ($\theta\sim0°$)-3	-21227.02	4737.33	16090.41	-1808.67	19115.02	106.40	-26223.60	-26117.20	-2603.83	-23830.85	42.59
	syn ($\theta\sim0°$)-4	-21225.35	4748.88	16110.62	-1812.60	19143.06	106.02	-26267.65	-26161.63	-2595.84	-23821.19	52.25
	syn ($\theta\sim180°$)-1	-21188.86	4757.04	16198.65	-1806.90	19244.64	105.77	-26337.82	-26232.05	-2600.15	-23789.01	84.43
	syn ($\theta\sim180°$)-2	-21224.93	4740.18	16121.87	-1811.72	19146.44	105.86	-26255.48	-26149.62	-2599.94	-23824.87	48.57
	syn ($\theta\sim180°$)-3	-21220.79	4732.48	16371.62	-1819.32	19381.29	105.27	-26511.79	-26406.52	-2594.29	-23815.08	58.36
	syn ($\theta\sim180°$)-4	-21219.32	4746.20	16156.01	-1819.32	19179.41	105.90	-26290.04	-26184.14	-2599.52	-23818.84	54.60
$^{o\text{-}PhOH}$dG:C	*anti* ($\theta\sim0°$)-1	-21195.93	4755.87	15726.48	-1802.44	18775.53	106.69	-26029.71	-25922.98	-2603.87	-23799.81	85.22
	anti ($\theta\sim0°$)-2	-21204.51	4743.23	15746.40	-1790.82	18793.77	107.07	-26030.04	-25922.93	-2602.36	-23806.88	78.13
	anti ($\theta\sim0°$)-3	-21204.68	4762.19	15850.04	-1816.67	18891.93	106.36	-26152.05	-26045.73	-2599.02	-23803.74	81.27
	anti ($\theta\sim0°$)-4	-21188.74	4756.66	15756.36	-1802.71	18805.95	106.94	-26018.66	-25911.72	-2606.00	-23794.74	90.27
	anti ($\theta\sim180°$)	-21283.51	4772.02	15968.40	-1723.33	19017.21	106.70	-26320.15	-26213.50	-2589.11	-23885.01	0.00
	syn ($\theta\sim0°$)	-21238.60	4749.45	16117.14	-1708.13	19158.30	107.11	-26395.75	-26288.63	-2588.72	-23839.62	45.39
	syn ($\theta\sim180°$)	-21226.82	4762.84	15563.57	-1677.90	18648.45	108.40	-25873.31	-25764.91	-2610.03	-23849.35	35.66
$^{p\text{-}PhOH}$dG:C	*anti*-1	-21196.40	4752.31	15798.41	-1807.13	18839.42	106.98	-26229.79	-26122.80	-2602.41	-23798.76	25.61
	anti-2	-21193.47	5107.99	15518.25	-1790.25	18930.97	107.36	-25931.01	-25823.65	-2602.36	-23795.83	28.54
	anti-3	-21223.55	4726.16	15834.10	-1816.27	18840.34	106.65	-26251.55	-26144.85	-2600.86	-23824.37	0.00
	anti-4	-21202.67	4742.27	15428.17	-1803.28	18462.82	106.90	-25841.43	-25734.53	-2603.24	-23805.91	18.46
	syn-1	-21171.00	4746.62	15470.34	-1781.28	18530.18	107.24	-25881.72	-25774.49	-2609.06	-23780.06	44.31
	syn-2	-21199.66	4724.15	15915.85	-1810.35	18925.70	106.11	-26298.87	-26192.76	-2600.69	-23800.35	24.02

3.3.1.2. Structure analysis

3.3.1.2.1. Clustering

Clustering is used to get the representative structures, which provide information about geometrical properties, hydrogen bonding, adduct location and ultimately structural stability. Figure 3.3 and Figure 3.5 depict the representative structures for the most stable and the least stable structures of PhOdG, *o*-PhOHdG, and *p*-PhOHdG when paired against cytosine. In all the preferred structures (O–linked and C–linked), the adduct adopts a B conformer and consequently the phenoxyl/phenyl moiety is located in the solvent-exposed major groove and does not stack with the neighboring bases. It is nonplanar with respect to the nucleobase. Conversely, the least stable conformers (based on the free energy calculations) in the case of PhOdG and *p*-PhOHdG is *syn,* in which the adduct adopts a W-type wedge conformation. Nevertheless, the orientation of the phenoxyl moiety with respect to the nucleobase is nonplanar while for the *o*-PhOHdG conformer, the aryl moiety in the *anti* conformer is located in a major groove and is a B type conformer (Table 3.2 and Table 3.4). It is important to note that in the case of the *syn* (θ~180°) (θ~346.01° and χ~64.22°, ΔE=35.66 kJ mol^{-1}) and *syn*-2 (θ~41.55° and χ~70.26°, ΔE=24.02 kJ mol^{-1}) the guanine flipped out of the helix to provide the vacancy for residing the phenoxyl moiety inside the double helix which leading to an intercalated conformer.

(a) ^{PhO}dG:C   (b) ^{o-PhOH}dG:C   (c) ^{p-PhOH}dG:C

$\chi = 225.7°$;
$\theta = 151.5°$;
$\Phi = 268.2°$;

$\chi = 304.4°$;
$\theta = 270.8°$;

$\chi = 283.8°$;
$\theta = 231.1°$;
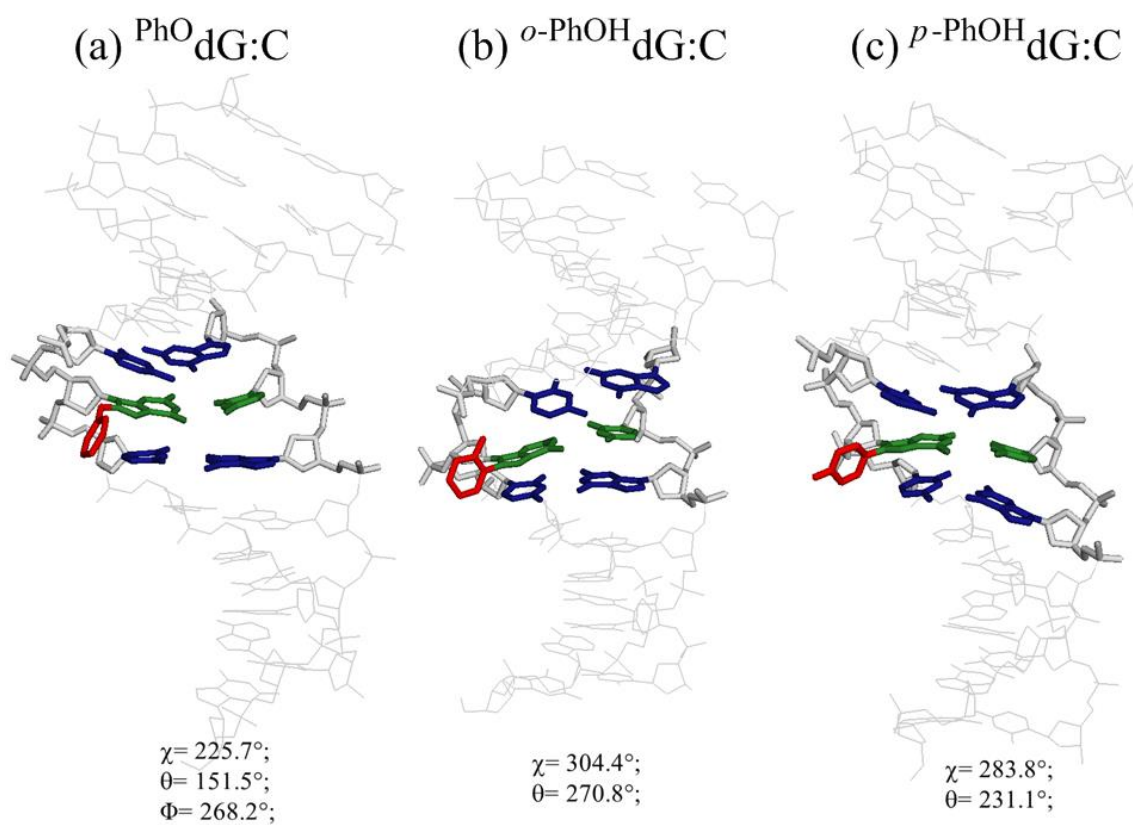
Figure 3.3. Representative structures of the most stable conformers of the (a) ^{PhO}dG, (b) ^{o-PhOH}dG and (c) ^{p-PhOH}dG adduct in the *Nar*I helix paired against the complementary cytosine identified from clustering calculations.

^{PhO}dG:C $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $^{p\text{-}PhOH}$dG:C

χ= 227.8°;
θ= 153.1°;
Φ= 286.4°;
ΔE= 18.0

χ= 215.2°;
θ= 148.3°;
Φ= 287.5°;
ΔE= 12.6

χ= 212.5°;
θ= 137.1°;
Φ= 103.3°;
ΔE= 14.9

χ= 233.7°;
θ= 141.5°;
Φ= 92.0°;
ΔE= 15.7

χ= 210.3°;
θ= 217.7°;
ΔE= 18.5

χ= 70.3°;
θ= 41.6°;
ΔE= 24.0
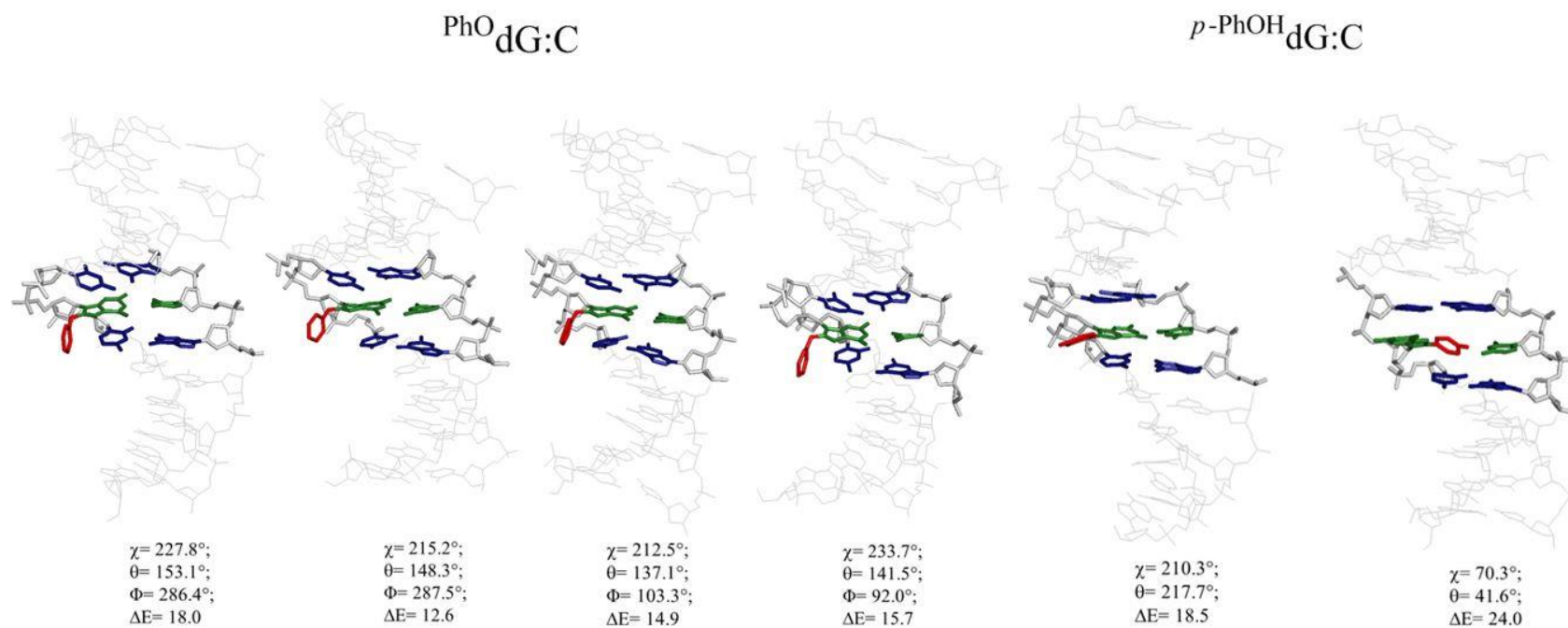
Figure 3.4. Representative structures of the other stable conformers of the ^{PhO}dG and $^{p\text{-}PhOH}$dG adduct in the *Nar*I helix paired against the complementary cytosine identified from clustering calculations.

(a) ^{PhO}dG:C (b) $^{o\text{-}PhOH}$dG:C (c) $^{p\text{-}PhOH}$dG:C

χ= 59.3°;
θ= 237.4°;
Φ= 292.4°;

χ= 249.6°;
θ= 111.5°;

χ= 60.9°;
θ= 234.8°;
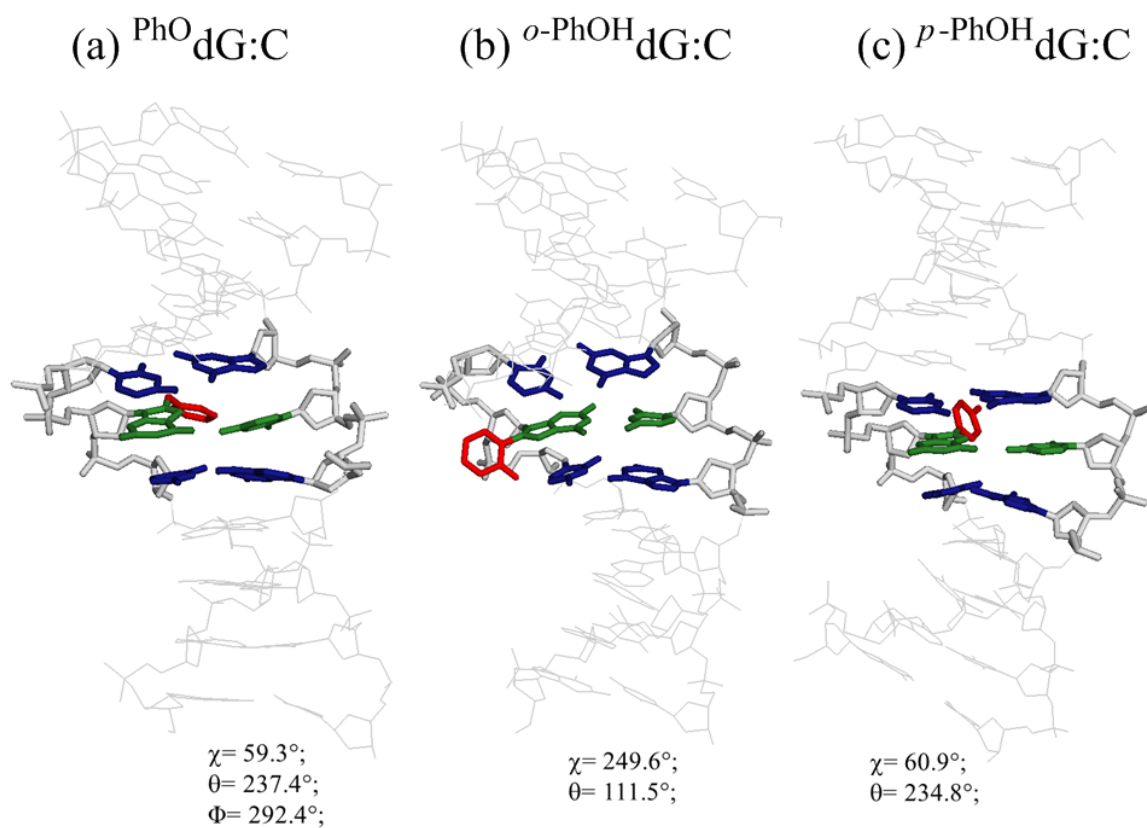
Figure 3.5. Representative structures of the least stable conformers of the (a) ^{PhO}dG, (b) $^{o\text{-}PhOH}$dG and (c) $^{p\text{-}PhOH}$dG adduct in the *Nar*I helix paired against the complementary cytosine identified from clustering calculations.

(b) $^{o\text{-PhOH}}$dG:C

χ= 70.3°;
θ= 41.6°;
ΔE= 24.0

(c) $^{p\text{-PhOH}}$dG:C

χ= 64.2°;
θ= 346.0°;
ΔE= 35.7
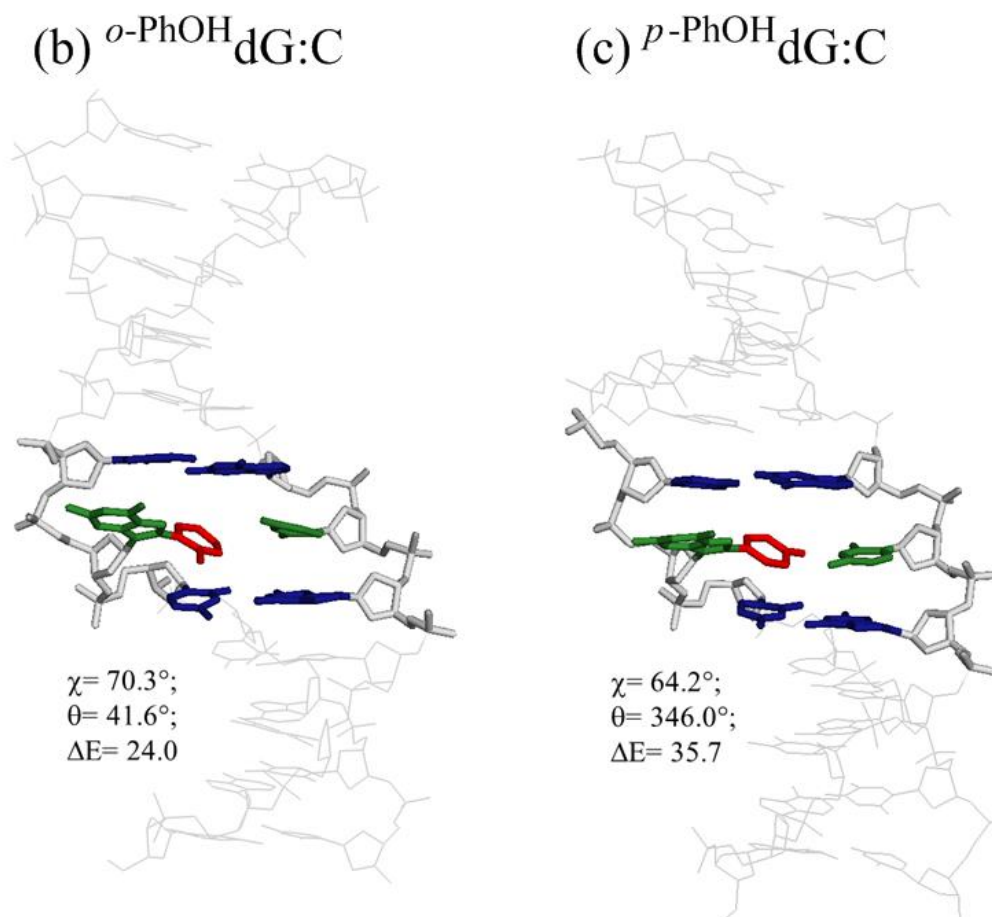
Figure 3.6. Representative structures of the pseudo base displaced stacked conformers of the (a) $^{o\text{-PhOH}}$dG and (b) $^{p\text{-PhOH}}$dG adduct in the *Nar*I helix paired against the complementary cytosine identified from clustering calculations.

For both $^{o\text{-PhOH}}$dG:C (*syn*-180) and $^{p\text{-PhOH}}$dG:C (*syn*-2), the pseudo-base-displaced stacked conformation is identified in the *Nar*I sequence (Figure 3.6). The representative structures illustrate that intercalated conformer in the duplex which pushes the complementary nucleobase away from the initial position (flipping out a little) may increase the probability for the formation of the 2-base

bulge mutation and therefore result in greater mutagenicity for C-linked adducts in comparison to O-linked structures (Figure 3.6 and Table 3.4).

In the case of ^{PhO}dG in the natural (paired with cytosine) *Nar*I sequence, the oxygen bridge allows more conformational flexibility and results in fewer steric collisions. In contrast, the *ortho* and *para* C–linked structures lack this flexibility due to direct attachment of aryl moiety to the C8 site of guanosine.

The average C1′–C1′ distance is slightly wider in the $^{p\text{-}PhOH}$dG:C pair (11.5±0.4 Å) than in the canonical G:C variant (10.9±0.4 Å). It also slightly wider than the distance in many adducts against cytosine and the structures adopted by nucleobase pairs in natural helices (10.6 ±0.2 Å, see Table 3.2).[38]

3.3.1.2.2. RMSD and sugar puckering

B-DNA is the most flexible conformer compared to the other common DNA structures.[82] Flexibility within the duplex structure can take place at different sites. Defined sites of rotation for a nucleotide within the duplex are given in Scheme 3.2, which illustrates that the torsional flexibility within the duplex mainly occurs within the backbone. However these sites are not freely rotatable, so each backbone angle is restricted to a discrete range of rotation.[83] The largest degree of rotational freedom is commonly identified for the δ and χ torsional angles. The conformation of the furanose (C2′-*endo* vs. C3′-*endo*) is determined by δ angle. Furthermore, this conformation is highly dependent on the glycosidic angle, χ.[84] Four atoms, O4′–C1′–N9–C4 in purines are involved in the glycosidic angle χ.
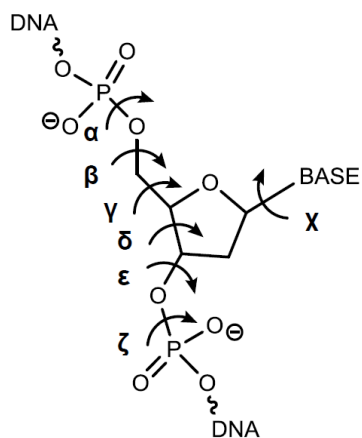
Scheme 3.2. Different sites of rotation for a nucleotide within the duplex.

It has been suggested that guanine exhibits the highest energy barrier for *anti/syn* interconversion.[85]   This can be explained by electrostatic contacts between exocyclic N2 amino group of guanosine and the 5′ phosphate.[85] Table 3.2 demonstrates the probability of finding C2′-*endo* conformers during the 20ns simulation. The percentage of finding this sugar puckering is the highest percentage (37.33%) in ^{PhO}dG *anti ($\theta\sim0°$)-3*. The percentages of finding C2′-*endo* sugar puckering are 2.96, 18.47, and 32.18 in ^{PhO}dG *anti* ($\theta\sim180°$)-2, $^{o\text{-}PhOH}$dG *anti* ($\theta\sim180°$), and $^{p\text{-}PhOH}$dG *anti-3* correspondingly. The RMSDs as a function of the simulation time for the species bound to the G^3 base in *Nar*I sequence were analyzed. The average RMSD was in the range of 1.0–3.5 A° during the simulation time as shown in Table 3.2. These results show the equilibrium state of the simulation for the mutated complexes. All simulations showed reasonable stability over 20ns; however, the deviation of RMSD in some mutated complexes from the normal range implies that these are more flexible.   The fluctuations in RMSDs are mainly assigned to the mobility of 3′ and 5′ end of the duplex.

Table 3.2. Parameters from 20 ns MD simulations of the ^{PhO}dG, $^{o\text{-}PhOH}$dG, and $^{p\text{-}PhOH}$dG adducts against cytosine, incorporated into the G[3] position of the *Nar*I DNA sequence.

Adduct	Adduct conformation	Cluster[a]	RMSD (Å)	*C2'-endo* (%)	Sugar Pucker	Sugar Pucker (°)	Adduct-θ (°)	Adduct-χ (°)	Adduct-φ (°)	C1'-C1' (A)
	anti (θ∼0°)-1	Bimodal (θ, φ)	4.38±0.82	34.75	$_1$E (*C1'-exo*)	123.72±44.49	153.11	227.84	286.37	10.78±0.17
	anti (θ∼0°)-2	Bimodal (θ, φ)	2.04±0.37	32.90	$_1$E (*C1'-exo*)	120.38±45.43	148.34	215.17	287.73	10.78±0.17
	anti (θ∼0°)-3	Bimodal (θ, φ)	4.15±0.67	37.33	$_1$E (*C1'-exo*)	127.17±41.09	137.11	212.50	103.31	10.81±0.26
	anti (θ∼0°)-4	Bimodal (θ, φ)	3.09±0.58	25.66	$_1$E (*C1'-exo*)	115.82±41.59	136.69	224.76	111.72	10.77±0.17
	anti (θ∼180°)-1	Bimodal (χ, φ)	1.97±0.44	32.13	$_1$E (*C1'-exo*)	138.72±11.90	182.02	278.28	315.02	10.67±0.19
	anti (θ∼180°)-2	Bimodal (χ, φ)	4.18±0.68	2.96	^{0}E (*C4'-endo*)	102.96±21.31	151.46	225.69	268.18	10.73±0.17
	anti (θ∼180°)-3	Bimodal (χ, φ)	2.17±0.52	35.29	$_1$E (*C1'-exo*)	139.80±11.89	150.67	269.24	126.65	10.69±0.18
^{PhO}dG:C	*anti* (θ∼180°)-4	Bimodal (χ, φ)	3.67±0.76	4.30	^{0}E (*C4'-endo*)	105.05±22.03	141.53	233.69	92.05	10.73±0.17
	syn (θ∼0°)-1	Bimodal (θ, φ)	1.98±0.46	13.02	$_1$E (*C1'-exo*)	124.13±21.77	171.77	54.28	249.76	11.71±0.27
	syn (θ∼0°)-2	Bimodal (θ, φ)	3.89±0.86	27.49	$_1$E (*C1'-exo*)	134.11±18.76	112.50	56.05	266.78	11.79±0.37
	syn (θ∼0°)-3	Bimodal (θ, φ)	2.10±0.45	21.01	$_1$E (*C1'-exo*)	128.82±22.09	166.97	54.89	91.52	11.73±0.27
	syn (θ∼0°)-4	Bimodal (θ, φ)	3.93±0.86	30.34	$_1$E (*C1'-exo*)	136.12±17.35	84.96	57.82	40.30	11.81±0.38
	syn (θ∼180°)-1	Bimodal (θ, φ)	4.36±1.11	3.21	$_1$E (*C1'-exo*)	109.55±20.87	237.38	59.33	292.43	11.02±0.87
	syn (θ∼180°)-2	Bimodal (θ, φ)	4.09±0.82	7.65	$_1$E (*C1'-exo*)	124.74±14.44	158.44	64.62	265.05	11.63±0.41
	syn (θ∼180°)-3	Bimodal (θ, φ)	4.39±0.97	1.14	^{0}E (*C4'-endo*)	100.83±19.47	218.71	39.67	71.36	10.27±0.96
	syn (θ∼180°)-4	Bimodal (θ, φ)	4.02±0.82	7.78	$_1$E (*C1'-exo*)	125.16±14.22	147.56	57.31	80.82	11.65±0.38
	anti (θ∼0°)-1	Bimodal (θ, χ)	2.71±0.54	15.06	$_1$E (*C1'-exo*)	130.16±16.47	129.51	271.73	–	10.64±0.20
	anti (θ∼0°)-2	Bimodal (θ, χ)	2.88±0.75	3.12	^{0}E (*C4'-endo*)	102.74±20.75	128.26	252.67	–	10.70±0.18
$^{o\text{-}PhOH}$dG:C	*anti* (θ∼0°)-3	Bimodal (θ, χ)	3.50±0.74	16.03	$_1$E (*C1'-exo*)	133.72±10.65	99.10	285.91	–	10.62±0.20
	anti (θ∼0°)-4	Bimodal (θ, χ)	4.50±0.88	6.75	$_1$E (*C1'-exo*)	109.91±22.09	111.48	249.59	–	10.72±0.18
	anti (θ∼180°)	Unimodal	4.26±0.81	18.47	$_1$E (*C1'-exo*)	133.49±14.20	270.79	304.35	–	10.65±0.22
	syn (θ∼0°)	Unimodal	4.59±0.65	3.67	$_1$E (*C1'-exo*)	113.69±18.55	59.92	56.75	–	11.57±0.31
	syn (θ∼180°)	Unimodal	4.08±0.68	13.18	$_1$E (*C1'-exo*)	129.49±13.90	346.01	64.22	–	10.58±0.89
	anti-1	Bimodal (θ, χ)	4.16±0.90	15.74	$_1$E (*C1'-exo*)	132.03±13.00	268.40	281.98	–	10.63±0.20
	anti-2	Bimodal (θ, χ)	4.52±0.66	0.05	^{0}E (*C4'-endo*)	89.58±16.00	280.73	234.87	–	10.71±0.17
$^{p\text{-}PhOH}$dG:C	*anti*-3	Bimodal (θ, χ)	3.25±0.80	32.18	$_1$E (*C1'-exo*)	138.95±11.46	231.06	283.81	–	10.64±0.21
	anti-4	Bimodal (θ, χ)	2.12±0.48	0.00	^{0}E (*C4'-endo*)	77.60±10.41	217.70	210.33	–	10.67±0.16
	syn-1	Bimodal (θ)	2.64±0.64	11.87	$_1$E (*C1'-exo*)	125.58±18.01	234.82	60.92	–	10.75±0.40
	syn-2	Bimodal (θ)	4.09±0.79	20.81	$_1$E (*C1'-exo*)	129.92±20.75	41.55	70.26	–	11.49±0.43

3.3.1.2.3. Hydrogen bond occupancy

To study the influence of mutations on direct hydrogen bonds between adducts and complementary nucleobases, and for detailed information regarding hydrogen-bonding properties, the percentage of hydrogen-bond occupancy was calculated between donor and acceptor atoms (Table 3.3). In the hydrogen-bond analysis, a certain criterion for a hydrogen bond has defined by donor−acceptor distance (d_{DA}) ≤3.4 Å and the donor−hydrogen−acceptor angle (α_{DHA}) ≥120°. The stability of hydrogen bonds was measured by the percentage of their presence during the simulation, where interactions populated over 15 % in the 20 ns-long trajectories were considered. In the natural helix, Watson−Crick hydrogen bonds have ~100% occupancy. Upon mutation, the overall pattern of hydrogen bond interactions and Watson−Crick hydrogen bonds between the opposing C and each of ^{PhO}dG, $^{o\text{-}PhOH}$dG, and $^{p\text{-}PhOH}$dG is maintained intact throughout the simulation (~100% occupancy). The important hydrogen bond interactions in native and mutated complexes are portrayed in representative structures shown in Figure 3.7. The number and types of hydrogen bonds formed between nucleobases in adducted structures and those in natural DNA are nearly the same whereas there is slightly more bonding in the native duplex (Table 3.3, Table 3.4 and Figure 3.7). Overall, the results illustrate strong H-bonding interactions between modified guanosine at G3 position and the complementary base which decreases from *anti* to *syn* conformers. Indeed, intra5′ and intra3′ interactions are stronger than inter5′ and inter3′ contacts which expectedly reduces from *anti* to *syn* conformers.
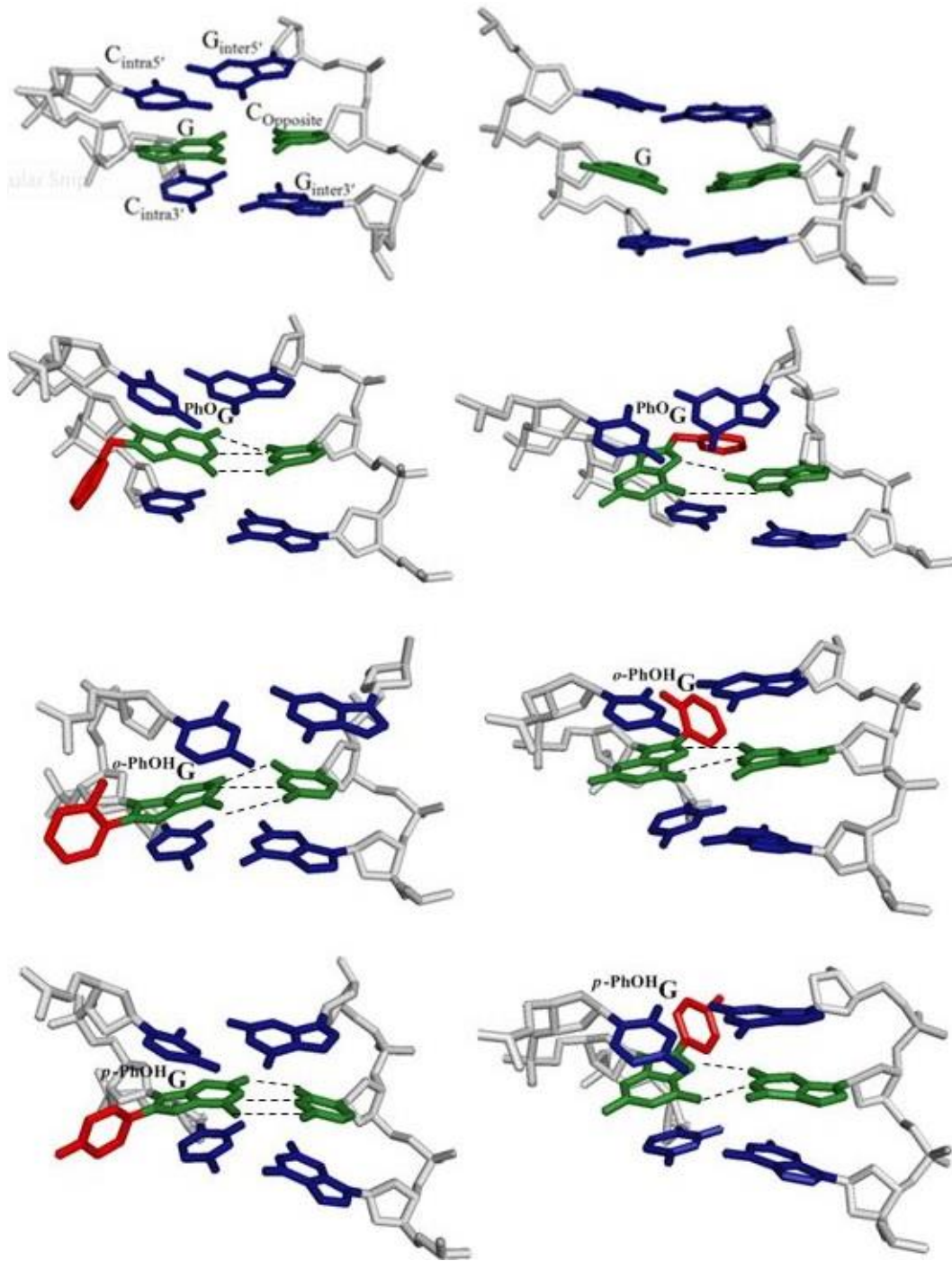
Figure 3.7. Portion of the representative structure from MD simulations which depicts the hydrogen bonds and stacking interactions of intrastrand base at the 5′-side, intrastrand base at the 3′-side, interstrand base at the 5′-side, and interstrand base at the 3′-side consciously for all PhOdG, o-PhOHdG and p-PhOHdG adducts with *anti*-G paired against cytosine in the B-DNA conformation, *syn*-G Hoogsteen hydrogen bonded with the guanine mismatch,

Table 3.3. AMBER interaction energies (kJ mol^{-1}) between ^{PhO}dG, $^{o\text{-}PhOH}$dG, and $^{p\text{-}PhOH}$dG adducts against cytosine in the G^3 position of the NarI sequence and the surrounding nucleobases from 20 ns MD simulations.

Adduct	Adduct conformation	ΔE_{Hbond}[a]	$\Delta E_{intra5'}$[b]	$\Delta E_{intra3'}$[c]	$\Delta E_{inter5'}$[d]	$\Delta E_{inter3'}$[e]
^{PhO}dG:C	*anti* ($\theta\sim0°$)-1	-110.92±6.86	-28.53±7.49	-59.32912±6.99	-21.17±5.86	20.88±5.10
	anti ($\theta\sim0°$)-2	-111.04±6.86	-28.41±7.53	-59.16±6.74	-22.47±6.44	-21.13±5.19
	anti ($\theta\sim0°$)-3	-109.95±9.37	-27.78±6.94	-59.50±6.99	-22.59±6.40	-21.59±5.56
	anti ($\theta\sim0°$)-4	-110.62±7.07	-28.79±7.03	-59.83±6.90	-22.17±6.32	20.84±5.06
	anti ($\theta\sim180°$)-1	-110.08±7.32	-39.58±8.37	-54.18±7.11	-14.02±7.20	20.29±6.32
	anti ($\theta\sim180°$)-2	-110.21±7.15	-28.12±6.78	-60.29±6.74	-24.39±6.36	20.67±5.06
	anti ($\theta\sim180°$)-3	-110.75±7.15	-34.35±6.86	-53.14±7.07	-17.07±6.82	18.24±6.07
	anti ($\theta\sim180°$)-4	-110.67±6.95	-27.45±6.95	-60.84±7.07	-24.10±6.32	20.42±5.48
	syn ($\theta\sim0°$)-1	-45.10±20.46	-4.14±5.27	-8.83±7.57	-13.81±6.53	-8.91±11.84
	syn ($\theta\sim0°$)-2	-33.26±15.98	-7.32±6.19	-2.176±6.99	-11.38±8.66	-12.34±12.34
	syn ($\theta\sim0°$)-3	-36.86±29.45	-2.89±5.65	-7.61±7.70	-14.64±6.69	-12.01±12.09
	syn ($\theta\sim0°$)-4	-30.42±17.74	-7.49±5.98	-1.088±6.69	-10.25±8.16	-12.51±12.72
	syn ($\theta\sim180°$)-1	-46.02±11.46	-7.49±6.78	-8.41±6.95	-9.92±5.44	-9.08±13.97
	syn ($\theta\sim180°$)-2	-48.91±7.61	-5.36±5.06	-7.11±6.69	-13.60±7.07	-9.71±10.84
	syn ($\theta\sim180°$)-3	-52.13±9.25	-12.30±7.87	-9.29±6.53	-8.62±4.94	0.50±15.31
	syn ($\theta\sim180°$)-4	-50.08±7.78	-4.77±5.19	-8.45±6.69	-14.69±7.28	-8.66±11.17
$^{o\text{-}PhOH}$dG:C	*anti* ($\theta\sim0°$)-1	-112.34±7.32	-40.96±6.57	-55.81±7.45	-10.71±7.66	19.62±5.86
	anti ($\theta\sim0°$)-2	-112.55±7.11	-37.07±7.32	-61.04±7.24	-15.31±7.61	20.75±5.15
	anti ($\theta\sim0°$)-3	-111.13±7.53	-36.90±6.95	-59.16±8.37	-10.04±7.78	19.08±6.53
	anti ($\theta\sim0°$)-4	-112.63±7.20	-34.22±7.78	-60.04±7.20	-16.19±8.58	20.29±5.48
	anti ($\theta\sim180°$)	-108.83±7.53	-47.45±12.89	-50.88±8.24	-5.69±8.70	-11.46±6.28
	syn ($\theta\sim0°$)	-41.09±7.82	3.56±4.48	-6.15±6.36	-17.32±6.99	-5.82±11.51
	syn ($\theta\sim180°$)	-18.99±11.92	-35.69±8.19	-33.22±9.50	-2.55±5.56	-0.25±6.99
$^{p\text{-}PhOH}$dG:C	*anti*-1	-107.86±7.78	-35.94±6.95	-46.82±8.12	-11.38±7.28	19.62±6.19
	anti-2	-108.78±6.95	-24.85±9.33	55.35±6.90	-19.79±6.40	21.13±5.19
	anti-3	-107.90±7.36	-34.4332±6.57	-45.61±8.12	-11.97±6.53	18.07±5.82
	anti-4	-109.54±6.57	-25.65±6.02	-56.94±5.94	-20.96±5.27	21.84±5.19
	syn-1	-39.96±12.22	-8.58±6.53	-5.10±7.41	-7.03±6.90	-11.09±12.76
	syn-2	-46.19±8.03	-37.36±9.41	-28.41±8.33	-11.51±5.69	-2.55±5.98

[a]ΔE_{Hbond} is the hydrogen bond strength in the dimer consisting of the nucleobase at the G^3 position and the opposing base. [b]$\Delta E_{intra5'}$ is the stacking interaction energy between the nucleobase at the G^3 position and the intrastrand base at the 5′-side of the adduct. [c]$\Delta E_{intra3'}$ is the stacking interaction energy between the nucleobase at the G^3 position and the intrastrand base at the 3′-side of the adduct. [d]$\Delta E_{inter5'}$ is the stacking interaction energy between the nucleobase at the G^3 position and the base in the opposing strand at the 5′-side of the adduct. [e]$\Delta E_{inter3'}$ is the interaction energy between the nucleobase at the G^3 position and the base on the opposite strand stacked at the 3′-side of the adduct.

Duplex stability can be rationalized precisely by determining the distinctive

base-pairing energies which are intrastrand and interstrand stacking interactions

between the adduct and the flanking bases. The data obtained from the hydrogen bond analysis indicate that the there is a slight decrease in the H-bonding interaction energy for the most stable conformer of ^{PhO}dG, $^{o-PhOH}$dG, and $^{p-PhOH}$dG with ΔE_{Hbond} –110.2, –108.8, –107.9 kJ mol^{-1}, respectively (Table 3.1 and Table 3.4). An opposite trend is observed with $\Delta E_{intra5'}$ and $\Delta E_{inter5'}$ with the lowest –28.12 kJ mol^{-1} and –24.39 kJ mol^{-1} for intra5′ and inter5′ interactions of ^{PhO}dG *anti (θ~180°)-2*, respectively. Due to the helical twist, intrastrand stacking (Table 3.3) at the 3′-side of the adduct (−60.29 kJ mol^{-1}) is greater than at the 5′-side (−28.12 kJ mol^{-1}). Significant interstrand stacking is observed at the 5′-side of the adduct (−24.39 kJ mol^{-1}), but the interstrand contact at the 3′–side (inter3′) is more repulsive (by ∼ 10 kJ mol^{-1}) than in the unmodified duplex. The same trend is observed for $^{p-PhOH}$dG *anti*-3. Conversely, this contact is favorable and somewhat strong in $^{o-PhOH}$dG (-11.46 kJ mol^{-1}). This can be justified by stabilizing hydrogen bonding between the hydroxyl group and the oxygen of phosphate backbone. In the case of the *ortho* adduct, the Hoogsteen face forms a plausibly strong pairing with cytosine which results in preferential binding of the *ortho* adduct to cytosine regardless of the type of acquired conformation (*anti/syn*). Consequently, substitution of the *ortho* adduct in (ODN1 = 5′-CCATXCTACC-3′ and ODN2 = 5′-GGTAGXATGG-3′) sequence, irrespective of the type of the studied DNA sequence, may not lead to base-substitution mutations in DNA.[15, 58] In contrast, the *anti* conformer, with strong Watson-Crick hydrogen-bond, is stable for the *ortho* adduct in the *Nar*I sequence. Interestingly, the probability of forming Hoogsteen pairs of *para* adducts with guanine and cytosine is similar. It is worth noting that

the *syn* conformation of the *para* adduct is preferred in DNA double helix which may cause to base mismatches.[1558] While, again for *para* adduct in *Nar*I sequence the *anti* conformer is stable against cytosine. This clarifies the importance of considering different sequences as a result of encountering to different steric and electrostatic interactions and definitely new conformational behaviors.

Table 3.4. Hydrogen bonding occupancies (%) for the hydrogen bonds between the adduct and the opposing base (cytosine) over the duration of the MD Simulations also the type of hydrogen bonds, position of adduct in a specific groove and conformation of DNA provided.

Adduct	Adduct conformation	N2-H(G³)···O2	N1-H(G³)···N3	N4-H···O6(G³)	N3-H(C)···N7(G³)	H-bond	Position adduct	DNA conformation
PhOdG:C	*anti* (θ~0°)-1	99.67	99.99	98.74	–	Watson-Crick	major groove	B
	anti (θ~0°)-2	99.72	99.92	98.75	–	Watson-Crick	major groove	B
	anti (θ~0°)-3	98.60	98.80	98.38	–	Watson-Crick	major groove	B
	anti (θ~0°)-4	99.65	99.95	98.36	–	Watson-Crick	major groove	B
	anti (θ~180°)-1	99.67	99.85	98.74	–	Watson-Crick	major groove	B
	anti (θ~180°)-2	99.94	99.94	98.11	–	Watson-Crick	major groove	B
	anti (θ~180°)-3	99.88	99.95	98.44	–	Watson-Crick	major groove	B
	anti (θ~180°)-4	99.85	99.96	98.62	–	Watson-Crick	major groove	B
	syn (θ~0°)-1	–	–	85.84	45.68	Hoogsteen	minor groove	W
	syn (θ~0°)-2	–	–	76.58	47.83	Hoogsteen	minor groove	W
	syn (θ~0°)-3	–	–	74.84	39.99	Hoogsteen	minor groove	W
	syn (θ~0°)-4	–	–	73.36	43.59	Hoogsteen	minor groove	W
	syn (θ~180°)-1	–	–	73.19	28.81	Hoogsteen	minor groove	W
	syn (θ~180°)-2	–	–	92.36	43.15	Hoogsteen	minor groove	W
	syn (θ~180°)-3	–	–	59.44	37.17	Hoogsteen	minor groove	W
	syn (θ~180°)-4	–	–	93.12	46.20	Hoogsteen	minor groove	W
o-PhOHdG:C	*anti* (θ~0°)-1	99.66	99.96	98.34	–	Watson-Crick	major groove	B
	anti (θ~0°)-2	99.61	99.97	98.69	–	Watson-Crick	major groove	B
	anti (θ~0°)-3	99.67	99.95	98.55	–	Watson-Crick	major groove	B
	anti (θ~0°)-4	99.71	99.97	98.42	–	Watson-Crick	major groove	B
	anti (θ~180°)	99.50	99.87	98.56	–	Watson-Crick	major groove	B
	syn (θ~0°)	–	–	92.14	51.55	Hoogsteen	minor groove	W
	syn (θ~180°)	–	–	46.93	23.78	Hoogsteen	inside helix	S
p-PhOHdG:C	*anti*-1	99.58	99.88	98.06	–	Watson-Crick	major groove	B
	anti-2	99.76	100.00	98.12	–	Watson-Crick	major groove	B
	anti-3	100.00	100.00	98.09	–	Watson-Crick	major groove	B
	anti-4	99.92	99.92	99.05	–	Watson-Crick	major groove	B
	syn-1	–	–	86.35	45.84	Hoogsteen	minor groove	W
	syn-2	–	–	84.84	31.69	Hoogsteen	inside helix	S

The implemented H-bond distance cutoff was a 3.40 Å heavy atom separation and a 120° X–H–X angle. Only H-bonds with occupancy of >15% are reported.

130

3.3.2. Adducts Against Guanine Mismatch

According to Figure 3.8, the distributions of θ, χ, and ϕ of ^{PhO}dG:G show a unimodal distribution of θ, and bimodal distribution of χ and ϕ for the *anti ($\theta{\sim}0°$)* structure. In the case of *anti ($\theta{\sim}180°$)* only a bimodal distribution of ϕ is seen, while the distribution of θ and χ are unimodal. A similar distribution of χ, θ, and ϕ is identified for *syn ($\theta{\sim}0°$)* and *syn ($\theta{\sim}180°$)* where the distribution of χ is unimodal and the distribution of θ and ϕ are bimodal. Since the ϕ dihedral is not defined in C–linked adducts $^{o\text{-PhOH}}$dG and $^{p\text{-PhOH}}$dG, unimodal distribution of χ and θ is observed for *anti ($\theta{\sim}0°$)*, *anti ($\theta{\sim}180°$)*, *syn ($\theta{\sim}0°$)*, and *syn ($\theta{\sim}180°$)* conformers of $^{o\text{-PhOH}}$dG. On the other hand, in the case of $^{p\text{-PhOH}}$dG bimodal distribution of χ and θ is observed for the *anti* conformer while a bimodal distribution of χ and θ is observed for *syn* conformer. Based on the distributions of different dihedrals, each mutated species with specific theta value can be separated into various types.
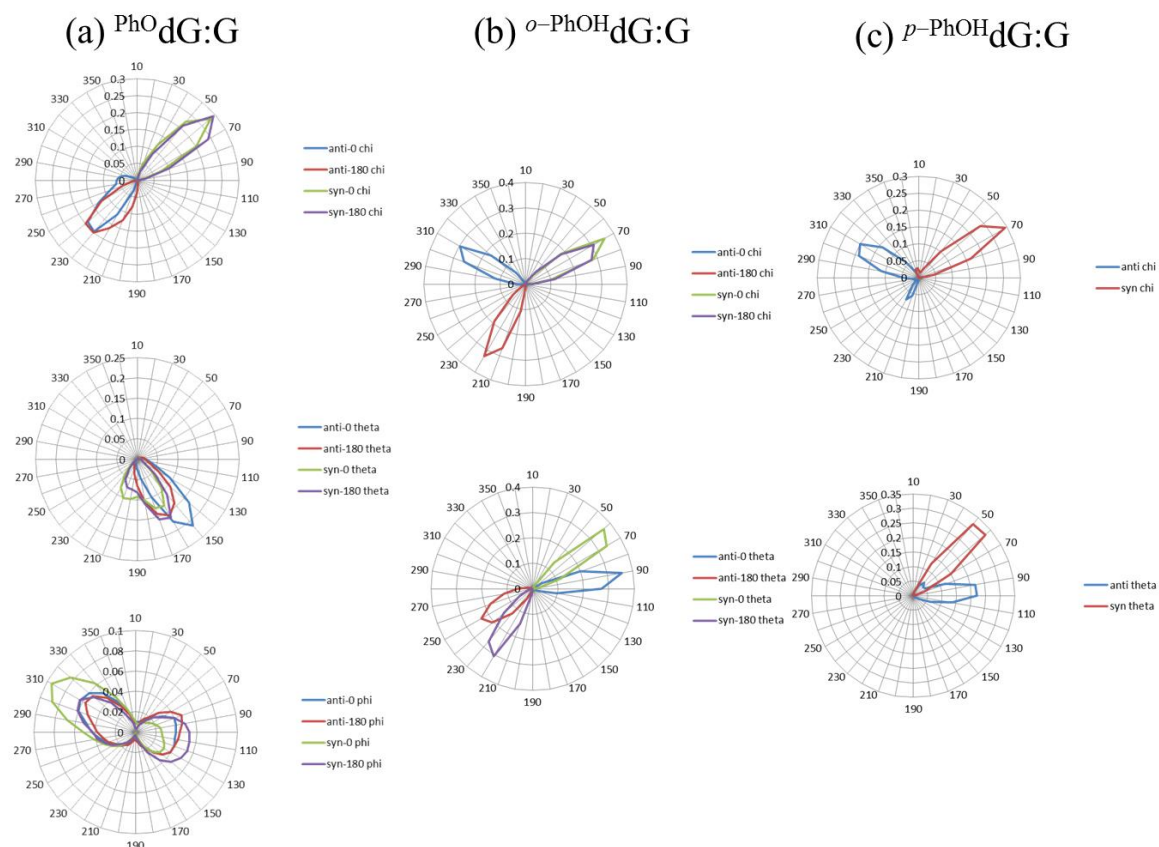
Figure 3.8. Percent distribution of the χ (degrees) θ (degrees) and ϕ (degrees) dihedral angles throughout the 20 ns trajectories for the (a) PhOdG, (b) o-PhOHdG and (c) p-PhOHdG adducts against guanine mismatch, *anti*-0 (blue), *anti*-180 (red), *syn*-0 (green), *syn*-180 (purple).

As explained above, the free energies of all adducts against a guanine mismatch were calculated by the MM-PBSA method. Although adducts in complement with cytosine in *Nar*I sequence adopt an *anti* conformation, the lowest energy structure of the adduct is a *syn* conformation when it is mismatched with guanine (Table 3.5). Based on the distribution of dihedrals, the lowest energy calculated free energy of PhOdG is *syn* (θ~180°)-3 with θ~201.57°, χ~50.63° and ϕ~114.06°. The most stable conformer for o-PhOHdG is *syn* (θ~0°) with θ~58.15°, χ~67.68°. Finally, the *syn-1* conformer is identified as the structure with lower energy for p-PhOHdG compound,

with $\theta\sim53.45°$, $\chi\sim59.44°$. The average χ ($\sim53.4°$) for natural dG is consistent with the χ values of O–linked and *ortho* and *para* C–linked adducts. Based on the obtained Gibbs free energy values (see Table 3.5 and Figure 3.10), there are eight coexistent stable conformers for [Ph]dG (i.e. seven conformers plus the global conformers), two coexistent stable conformers for *p*-[PhOH]dG (i.e. one conformer plus the global conformers), and two coexistent stable conformers for *o*-[PhOH]dG (i.e. one conformer plus the global conformers).

The *anti* ($\theta\sim0°$)-3 ($\theta\sim168.32°$, $\chi\sim280.97°$ and $\phi\sim117.47°$, $\Delta E=74.06$ kJ mol^{-1}), *anti* ($\theta\sim180°$) ($\theta\sim242.18°$ and $\chi\sim211.88°$, $\Delta E=47.40$ kJ mol^{-1}), and *anti*-2 ($\theta\sim80.52°$ and $\chi\sim223.72°$, $\Delta E=61.23$ kJ mol^{-1}) are the least stable structures for [Ph]dG, *o*-[PhOH]dG, and *p*-[PhOH]dG respectively (Table 3.5).

Table 3.5. MM-PBSA total free energy (kJ mol^{-1}) for the *Nar*I sequence with the mutated adducts incorporated at the G^3 position opposing guanine from 20 ns MD simulations.

Adduct	Adduct conformation	E	E_{int}	E_{elec}	E_{vdW}	E_{MM}	E_{NP}	E_{PB}	E_{sol}	$-TS$	G	ΔE
	anti (θ∼0°)-1	-21079.47	4760.40	16000.46	-1738.34	19022.51	106.94	-26277.29	-26170.33	-2613.61	-23693.07	42.89
	anti (θ∼0°)-2	-21077.64	4751.55	15955.61	-1726.83	18980.34	107.79	-26201.03	-26093.26	-2613.80	-23691.44	44.52
	anti (θ∼0°)-3	-21046.53	4775.76	16021.50	-1732.08	19065.18	107.57	-26290.63	-26183.05	-2615.36	-23661.90	74.06
	anti (θ∼0°)-4	-21072.18	4762.33	15972.69	-1730.47	19004.55	107.73	-26206.23	-26098.49	-2614.68	-23686.84	49.12
	anti (θ∼180°)-1	-21064.60	4763.19	15990.70	-1734.08	19019.81	107.65	-26244.22	-26136.57	-2620.74	-23685.33	50.63
	anti (θ∼180°)-2	-21067.47	4758.38	15952.51	-1723.51	18987.38	107.77	-26209.23	-26101.47	-2616.49	-23683.95	52.01
	syn (θ∼0°)-1	-21098.83	4750.46	15891.43	-1737.52	18904.36	106.68	-26139.83	-26033.14	-2618.55	-23717.38	18.58
^{PhO}dG:G	*syn* (θ∼0°)-2	-21106.97	4752.10	15830.13	-1743.60	18838.63	106.27	-26092.77	-25986.49	-2614.12	-23721.10	14.85
	syn (θ∼0°)-3	-21111.84	4740.31	15879.95	-1742.46	18877.79	106.65	-26126.51	-26019.88	-2619.07	-23730.89	5.06
	syn (θ∼0°)-4	-21099.43	4757.04	15919.66	-1745.24	18931.46	106.27	-26166.90	-26060.63	-2615.68	-23715.12	20.84
	syn (θ∼180°)-1	-21109.29	4746.04	15880.86	-1739.34	18887.56	106.29	-26144.26	-26037.95	-2613.21	-23722.49	13.47
	syn (θ∼180°)-2	-21105.34	4748.58	15884.52	-1748.56	18884.54	106.16	-26131.29	-26025.15	-2614.93	-23720.27	15.69
	syn (θ∼180°)-3	-21123.35	4735.31	15871.19	-1739.26	18867.24	106.27	-26139.88	-26033.60	-2612.59	-23735.96	0.00
	syn (θ∼180°)-4	-21113.58	4746.99	15811.89	-1741.79	18817.08	106.35	-26069.86	-25963.52	-2611.40	-23725.00	10.96
	anti (θ∼0°)	-21036.92	4744.30	15901.20	-1728.09	18917.41	107.25	-26258.36	-26151.09	-2617.05	-23653.99	32.43
	anti (θ∼180°)	-21024.64	4757.59	15763.39	-1727.30	18793.67	107.25	-26112.18	-26004.94	-2614.38	-23639.01	47.40
$_{o\text{-PhOH}}$dG:G	*syn* (θ∼0°)	-21075.63	4732.08	15814.73	-1738.19	18808.63	106.78	-26177.77	-26070.96	-2610.79	-23686.42	0.00
	syn (θ∼180°)	-21056.82	4734.26	15863.00	-1732.49	18864.76	107.00	-26201.65	-26094.65	-2612.17	-23669.01	17.41
	anti-1	-21017.59	4760.22	15679.07	-1729.40	18709.90	107.09	-26183.42	-26076.32	-2613.21	-23630.81	43.45
	anti-2	-21006.11	4765.21	15683.14	-1724.64	18723.71	107.24	-26133.22	-26025.99	-2606.90	-23612.99	61.23
	anti-3	-21021.61	4736.90	15658.08	-1710.77	18684.21	107.37	-26186.50	-26079.12	-2606.83	-23628.43	45.80
	anti-4	-21030.21	4749.50	15540.23	-1727.41	18562.32	107.01	-26007.22	-25900.22	-2609.80	-23640.02	34.23
$_{p\text{-PhOH}}$dG:G	*syn*-1	-21066.28	4744.26	15708.82	-1741.89	18711.19	106.28	-26213.13	-26106.86	-2607.96	-23674.24	0.00
	syn-2	-21044.08	4766.60	15624.97	-1740.91	18650.66	105.83	-26142.81	-26036.99	-2608.69	-23652.78	21.47

Clustering and representative structures show that in mutated complexes with the *anti* conformer, the DNA adopts a B-type conformer and the phenoxyl or aryl moiety is located in solvent exposed major groove. In adducts that adopt the *syn* conformer, the phenoxyl/aryl moiety is in the minor groove and the DNA adopts the Wedge conformation.

Steric interactions between adduct (modified bases) and the opposing G induce a slight non-planarity to the phenoxyl moiety (Figure 3.9, Figure 3.11 and Table 3.6).

(a) ^{PhO}dG:G   (b) $^{o\text{-PhOH}}$dG:G   (c) $^{p\text{-PhOH}}$dG:G

$\chi = 50.6°$;
$\theta = 201.6°$;
$\Phi = 114.1°$;

$\chi = 67.7°$;
$\theta = 58.2°$;

$\chi = 59.4°$;
$\theta = 53.5°$;
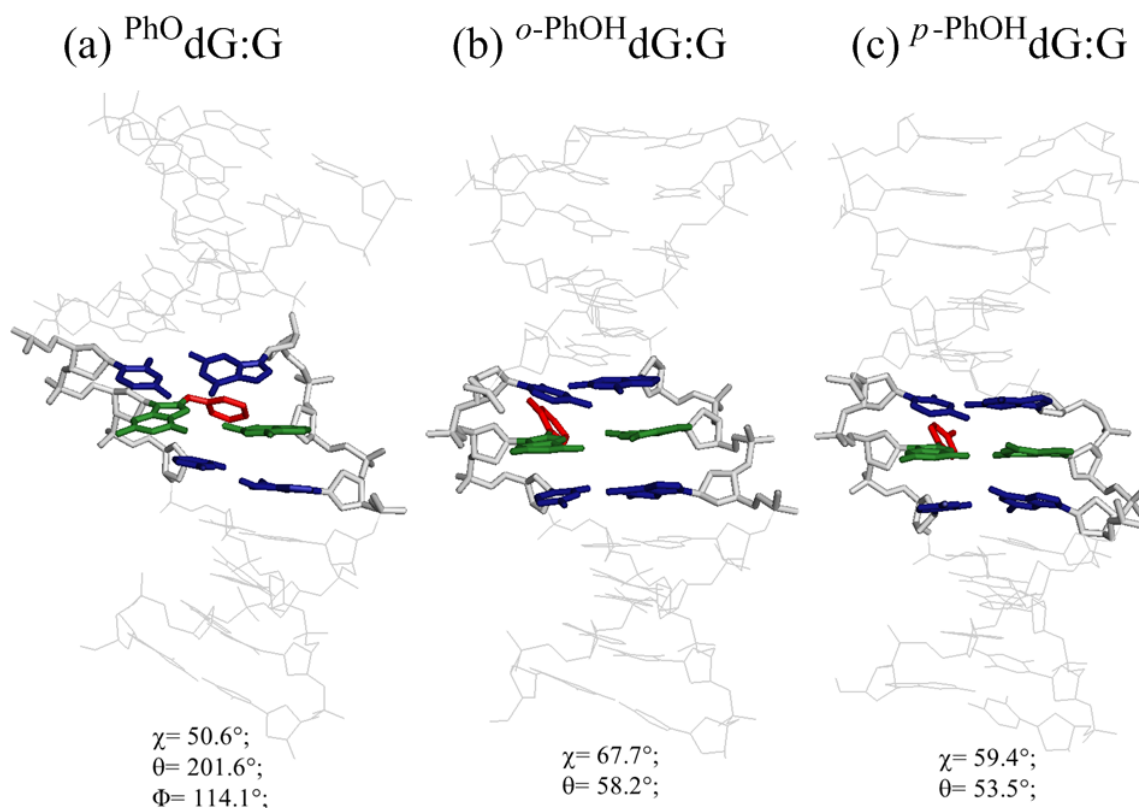
Figure 3.9. Representative structures of the most stable conformers of the (a) ^{PhO}dG, (b) $^{o\text{-PhOH}}$dG and (c) $^{p\text{-PhOH}}$dG adduct in the *Nar*I helix paired against a guanine mismatch identified from clustering calculations.

PhO dG:G

χ= 48.6°;
θ= 208.8°;
Φ= 314.8°;
ΔE= 18.6

χ= 57.8°;
θ= 157.0°;
Φ= 297.4°;
ΔE= 14.8

χ= 51.3°;
θ= 202.5°;
Φ= 113.2°;
ΔE= 5.1

χ= 48.5°;
θ= 156.0°;
Φ= 96.0°;
ΔE= 20.8

o-PhOH dG:G

χ= 68.3°;
θ= 230.5°;
ΔE= 17.4

χ= 38.9°;
θ= 213.0°;
Φ= 296.0°;
ΔE= 13.5

χ= 62.3°;
θ= 163.7°;
Φ= 280.7°;
ΔE= 15.7

χ= 62.5°;
θ= 109.3°;
Φ= 151.8°;
ΔE= 11.0

p-PhOH dG:G

χ= 13.2°;
θ= 62.8°;
ΔE= 21.5
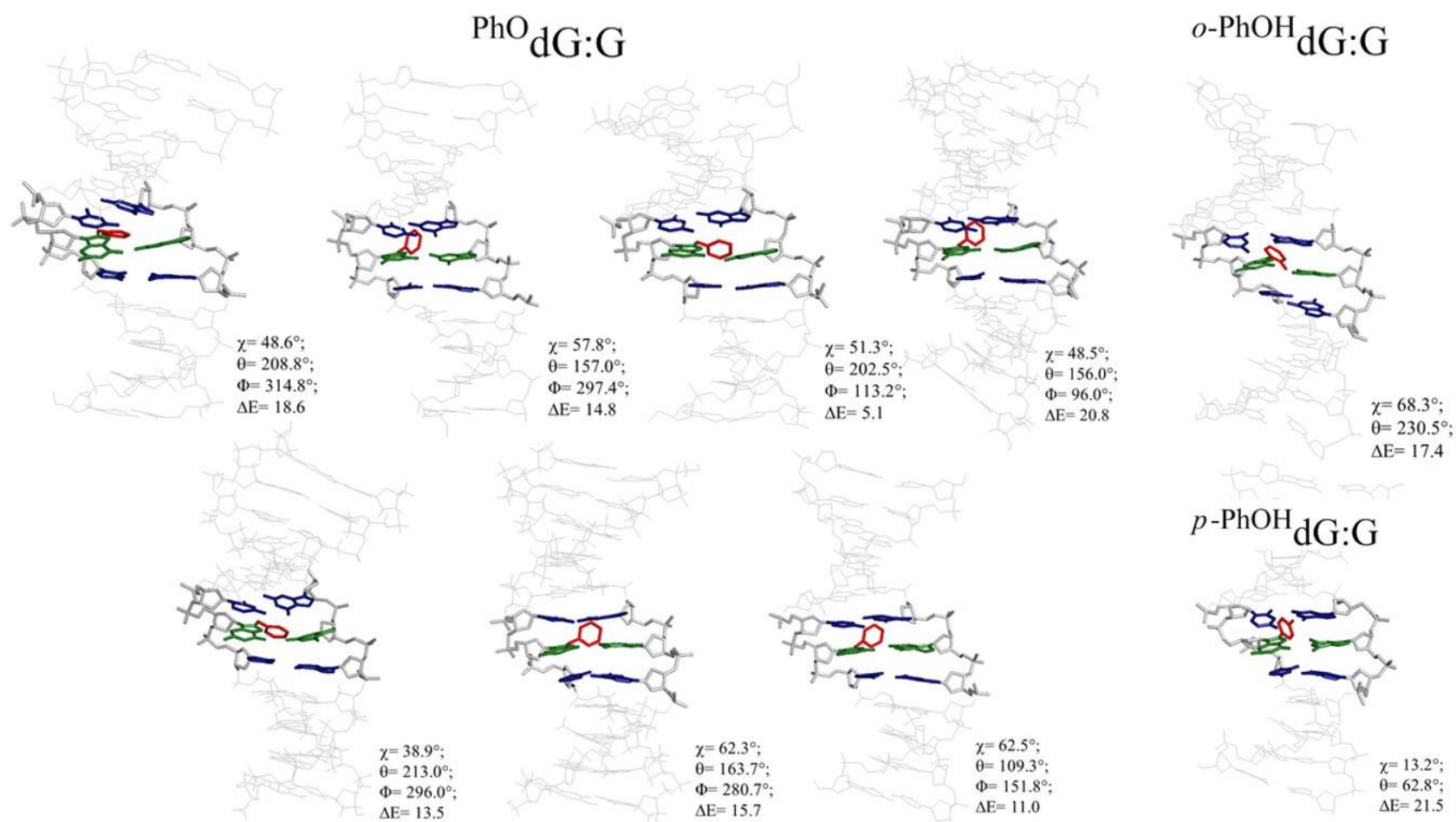
Figure 3.10. Representative structures of the other stable conformers of the PhO dG, o-PhOH dG and p-PhOH dG adduct in the *Nar*I helix paired against the complementary guanine identified from clustering calculations.

136
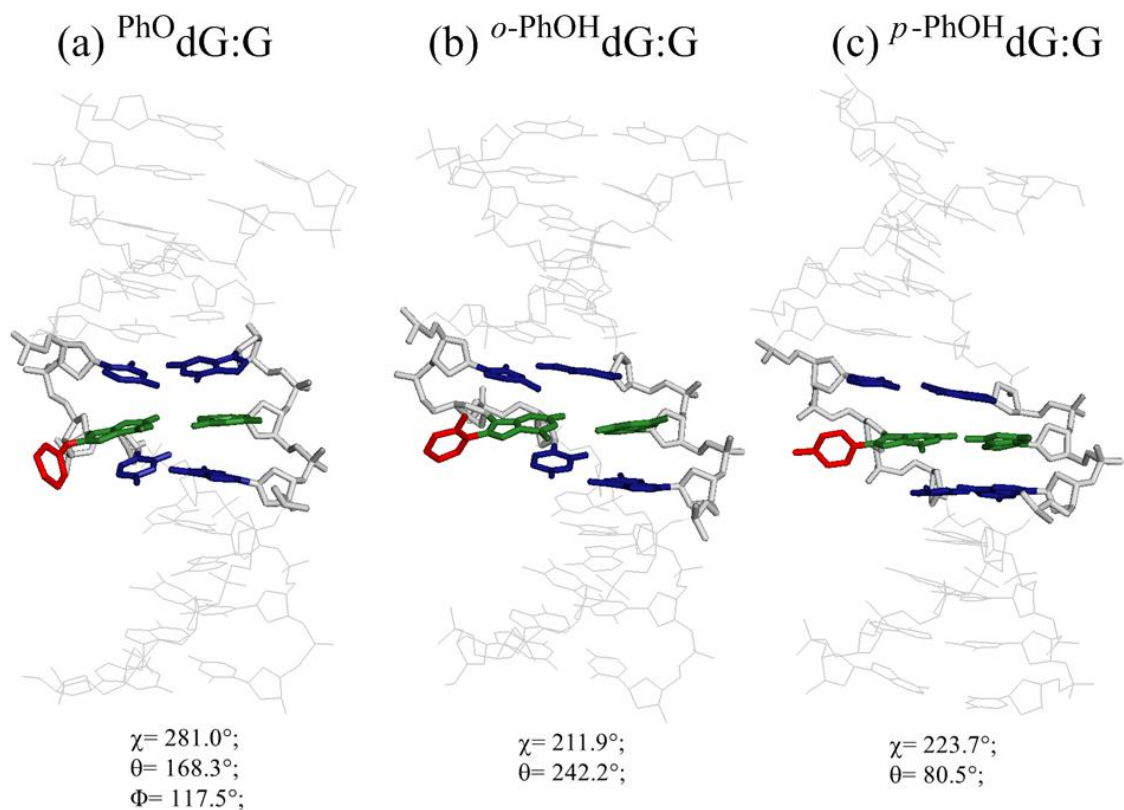
Figure 3.11. Representative structures of the least stable conformers of the (a) [PhO]dG, (b) [o-PhOH]dG and (c) [p-PhOH]dG adduct in the *Nar*I helix paired against guanine mismatch, identified from clustering calculations.

The $C1'$-$C1'$ distance for all adducts against a guanine mismatch (~11.5 Å) is greater than the corresponding distance in adducts paired with cytosine (~10.9 Å), which can be explained by the presence of guanine instead of cytosine, as the complement base (Table 3.6).

Table 3.6. Parameters from 20 ns MD simulations of the ^{PhO}dG, $^{o\text{-}PhOH}$dG, and $^{p\text{-}PhOH}$dG adducts against guanine mismatch, incorporated into the G^3 position of the *Nar*I DNA sequence.

Adduct	Adduct conformation	Cluster	RMSD (Å)	*C2'-endo* (%)	Sugar Pucker	Sugar Pucker (°)	Adduct-θ (°)	Adduct-χ (°)	Adduct-φ (°)	C1'-C1' (Å)
	anti (θ~0°)-1	Bimodal (χ, φ)	3.17±0.79	53.64	^{2}E (*C2'-endo*)	145.57±11.67	168.72	281.07	116.40	11.20±0.37
	anti (θ~0°)-2	Bimodal (χ, φ)	4.58±0.86	4.47	^{0}E (*C4'-endo*)	101.27±25.34	151.15	233.47	94.98	11.13±0.40
	anti (θ~0°)-3	Bimodal (χ, φ)	2.59±0.53	49.31	^{2}E (*C2'-endo*)	142.43±24.13	168.32	280.97	117.47	11.16±0.40
	anti (θ~0°)-4	Bimodal (χ, φ)	4.80±0.76	4.04	^{0}E (*C4'-endo*)	98.69±25.19	132.99	234.96	96.80	11.10±0.42
	anti (θ~180°)-1	Bimodal (φ)	3.89±0.87	30.35	^{0}E (*C4'-endo*)	94.42±89.48	179.06	214.01	102.66	11.45±0.71
	anti (θ~180°)-2	Bimodal(φ)	3.79±0.85	21.36	$_1$E (*C1'-exo*)	114.81±40.65	124.22	227.32	102.60	11.78±0.87
^{PhO}dG:G	*syn* (θ~0°)-1	Bimodal(θ, φ)	1.86±0.34	3.13	^{0}E (*C4'-endo*)	96.80±28.44	208.76	48.63	314.77	10.71±0.49
	syn (θ~0°)-2	Bimodal (θ, φ)	4.57±0.80	30.67	$_1$E (*C1'-exo*)	132.46±23.87	156.99	57.76	297.43	11.60±0.46
	syn (θ~0°)-3	Bimodal (θ, φ)	1.98±0.36	4.33	^{0}E (*C4'-endo*)	97.64±31.37	202.48	51.30	113.22	11.79±0.54
	syn (θ~0°)-4	Bimodal (θ, φ)	4.59±0.87	22.07	^{0}E (*C4'-endo*)	127.41±23.02	156.02	48.47	95.99	11.55±0.49
	syn (θ~180°)-1	Bimodal(θ, φ)	4.31±0.70	2.35	^{0}E (*C4'-endo*)	103.08±20.44	213.03	38.91	295.95	10.75±0.55
	syn (θ~180°)-2	Bimodal (θ, φ)	3.72±0.58	10.99	$_1$E (*C1'-exo*)	125.55±16.57	163.70	62.31	280.67	11.66±0.44
	syn (θ~180°)-3	Bimodal (θ, φ)	2.70±0.55	1.03	^{0}E (*C4'-endo*)	103.31±21.08	201.57	50.63	114.06	10.91±0.54
	syn (θ~180°)-4	Bimodal (θ, φ)	2.30±0.44	9.34	$_1$E (*C1'-exo*)	125.53±15.39	109.30	62.52	151.83	11.70±0.37
	anti (θ~0°)	Unimodal	3.97±0.78	50.62	$_1$E (*C1'-exo*)	141.55±27.91	87.76	300.7	–	11.69±0.60
	anti (θ~180°)	Unimodal	4.11±0.77	2.67	^{0}E (*C4'-endo*)	78.16±24.07	242.18	211.88	–	11.33±0.49
$^{o\text{-}PhOH}$dG:G	*syn* (θ~0°)	Unimodal	4.53±0.74	16.95	$_1$E (*C1'-exo*)	129.01±16.93	58.15	67.68	–	11.62±0.34
	syn (θ~180°)	Unimodal	4.47±0.81	17.83	$_1$E (*C1'-exo*)	126.98±21.52	230.49	68.35	–	11.53±0.41
	anti-1	Bimodal (θ, χ)	3.78±0.70	34.65	$_1$E (*C1'-exo*)	137.71±21.44	92.81	301.28	–	11.52±0.59
	anti-2	Bimodal (θ, χ)	4.08±1.03	0.00	^{0}E (*C4'-endo*)	77.42±13.41	80.52	223.72	–	11.20±0.33
$^{p\text{-}PhOH}$dG:G	*anti*-3	Bimodal (θ, χ)	3.41±0.91	45.63	$_1$E (*C1'-exo*)	138.99±36.71	59.04	320.62	–	11.67±0.52
	anti-4	Bimodal (θ, χ)	4.14±0.78	0.03	^{0}E (*C4'-endo*)	74.44±10.16	44.77	213.49	–	11.08±0.31
	syn-1	Bimodal (χ)	4.03±0.73	24.34	$_1$E (*C1'-exo*)	126.09±30.96	53.45	59.44	–	11.68±0.27
	syn-2	Bimodal (χ)	2.21±0.44	3.89	^{0}E (*C4'-endo*)	95.55±27.91	62.83	13.12	–	11.66±0.20

[a]See Figure 1 for definitions of important dihedral angles. [b]Percentage of the simulation that mutated adduct adopts the C2'–*endo* pucker and the most common sugar pucker adopted by the deoxyribose . [c]Distance between C1' of mutated complex and C1' of the pairing nucleotide.

In the case of the lowest energy structure of ^{PhO}dG:G, *syn* ($\theta{\sim}180°$)-3, two strong Hoogsteen H-bonds [N1-H•••O6 (71.40% occupancy) and N2-H•••N7 (35.75% occupancy)] are formed between the adduct and guanine. A third H-bond (N2-H•••O6) that is present for a smaller portion of the simulation (43.49% occupancy) is also identified. Although the N2-H•••O6 H-bond also appears over the course of the simulation of the natural sequence (19.8%), the occupancy is greater in adducted DNA because steric interactions between the phenoxyl moiety and the opposing G are slightly alleviated when the G shifts toward the major groove and forms the N2-H•••O6 H-bond with the adduct. The same rationalization can be applied to $^{o\text{-}PhOH}$dG:G *syn*-0 with [N1-H•••O6 (96.91% occupancy), N2-H•••N7 (81.75% occupancy), N2-H•••O6 (41.68)] and $^{p\text{-}PhOH}$dG:G *syn*-1 with [N1-H•••O6 (98.81% occupancy), N2-H•••N7 (88.54% occupancy), N2-H⋯O6 (30.17)]. In comparison, the hydrogen-bond occupancies of dG against a guanine mismatch are N1-H•••O6 (86.20% occupancy), N2-H•••N7 (90.70% occupancy) and N2-H•••O6 (19. 80) (Table 3.7). The N1-H•••O6 occupancy is higher in $^{p\text{-}PhOH}$dG:G because of less steric effects and symmetry in the structure while increasing the steric interactions of the hydroxyl group in $^{o\text{-}PhOH}$dG:G reduces the occupancies. The N1-H•••O6 occupancy reduces to 71.40% in ^{PhO}dG:G which is even less than dG:G. This suggests that the distortion of the double helix structure results from a reduction of this H-bonding interaction with the phenoxyl moiety present during the course of the simulation.

Table 3.7 Hydrogen bonding occupancies for the hydrogen bonds between the adduct and the opposing base (guanine) over the duration of the MD Simulations also the type of hydrogen bonds, position of adduct in a specific groove and conformation of DNA provided.

Adduct	Adduct conformation	N1-H(G³)···O6	N1-H(G³)···N7	N2-H(G³)···N7	N1-H···O6(G³)	N2-H···N7(G³)	N2-H···O6(G³)	H-bond	Position adduct	DNA conformation
PhOdG:G	*anti* (θ~0°)-1	67.66	34.28	94.62	–	–	–	Hoogsteen	major groove	B
	anti (θ~0°)-2	54.03	87.36	95.89	–	–	–	Hoogsteen	major groove	B
	anti (θ~0°)-3	65.21	33.80	96.41	–	–	–	Hoogsteen	major groove	B
	anti (θ~0°)-4	52.08	51.04	95.66	–	–	–	Hoogsteen	major groove	B
	anti (θ~180°)-1	56.93	45.94	79.14	–	–	–	Hoogsteen	major groove	B
	anti (θ~180°)-2	70.78	30.55	62.17	–	–	–	Hoogsteen	major groove	B
	syn (θ~0°)-1	–	–	–	81.13	18.93	24.95	Hoogsteen	minor groove	W
	syn (θ~0°)-2	–	–	–	88.23	80.49	36.92	Hoogsteen	minor groove	W
	syn (θ~0°)-3	–	–	–	72.23	27.81	33.09	Hoogsteen	minor groove	W
	syn (θ~0°)-4	–	–	–	84.49	66.00	35.43	Hoogsteen	minor groove	W
	syn (θ~180°)-1	–	–	–	78.61	23.59	30.40	Hoogsteen	minor groove	W
	syn (θ~180°)-2	–	–	–	91.48	83.47	33.90	Hoogsteen	minor groove	W
	syn (θ~180°)-3	–	–	–	71.40	35.75	43.49	Hoogsteen	minor groove	W
	syn (θ~180°)-4	–	–	–	93.64	84.54	36.07	Hoogsteen	minor groove	W
o-PhOHdG:G	*anti* (θ~0°)	38.84	70.07	89.56	–	–	–	Hoogsteen	major groove	B
	anti (θ~180°)	66.96	34.39	93.04	–	–	–	Hoogsteen	major groove	B
	syn (θ~0°)	–	–	–	96.91	81.75	41.68	Hoogsteen	minor groove	W
	syn (θ~180°)	–	–	–	95.70	74.44	53.19	Hoogsteen	minor groove	W
p-PhOHdG:G	*anti*-1	26.95	80.74	80.62	–	–	–	Hoogsteen	major groove	B
	anti-2	63.21	39.22	96.42	–	–	–	Hoogsteen	major groove	B
	anti-3	39.30	74.46	88.07	–	–	–	Hoogsteen	major groove	B
	anti-4	53.70	54.61	96.46	–	–	–	Hoogsteen	major groove	B
	syn-1	–	–	–	98.81	88.54	30.17	Hoogsteen	minor groove	W
	syn-2	–	–	–	99.06	95.31	13.87	Hoogsteen	minor groove	W

The implemented H-bond distance cutoff was a 3.40 Å heavy atom separation and a 120° X–H–X angle. Only H-bonds with occupancy of >15% are reported.

The strength of Hoogsteen H-bonding is enhanced by ~23 kJ mol^{-1} in ^{PhO}dG:G, *syn* (θ~180°)-3 compared to that of the natural strand, which is likely due to the greater planarity of paired bases in the adducted helix in the representative structure. The stacking interactions are slightly enhanced for intra 3′ (Figure 3.7) by ~15 kJ mol^{-1} in the adducted *syn* (θ~180°)-3 helix compared to those in the natural helix. Changes in the relative arrangement of the flanking bases help accommodate the additional phenoxyl moiety in the duplex. Conversely, the reduction in stacking interactions in intra 5′ of ~7 kJ mol^{-1} is attributed to the new orientation of adducted guanosine in the double helix. Subsequently, the inter 3′ interactions are more repulsive. The same trend is observed for the lowest energy structure of the *ortho* and *para* C–linked structures with the presence of less attractive inter 3′ interactions in the case of *syn*-0 $^{o\text{-}PhOH}$dG:G adduct in comparison to other counterparts against a guanine mismatch (Table 3.8). Ultimately, the same trend identified for adducts in the duplex against a guanine mismatch compares with their counterparts in complement with cytosine. Strong hydrogen-bonding interaction energies, which are lower than the identified values due to Hoogsteen hydrogen-bonding and guanine-guanine interactions. Regarding the stable conformers, stronger intra 3′ and 5′ and inter 3′ and 5′ interaction energies is determined which are consistent with the stability of the conformers.

Table 3.8. AMBER interaction energies (kJ mol^{-1}) between ^{PhO}dG, $^{o\text{-}PhOH}$dG, and $^{p\text{-}PhOH}$dG adducts against guanine mismatch in the G^3 position of the *Nar*I sequence and the surrounding nucleobases from 20 ns MD simulations.

Adduct	Adduct conformation	ΔE_{Hbond}[a]	$\Delta E_{intra5'}$[b]	$\Delta E_{intra3'}$[c]	$\Delta E_{inter5'}$[d]	$\Delta E_{inter3'}$[e]
	anti (θ~0°)-1	-42.47±16.23	-28.12±3.39	-24.81±3.76	-13.18±4.35	-8.49±2.47
	anti (θ~0°)-2	-39.12±16.28	-30.46±6.20	-58.20±7.49	-29.162±6.32	-13.60±7.57
	anti (θ~0°)-3	-43.64±15.52	-39.75±7.20	-53.30±7.74	-20.17±5.77	12.59±7.20
	anti (θ~0°)-4	-38.49±16.53	-30.54±6.53	-59.79±6.86	-28.37±6.15	-14.56±6.57
	anti (θ~180°)-1	-42.68±19.33	-33.97±9.08	-59.16±7.15	-22.17±10.12	17.32±6.15
	anti (θ~180°)-2	-49.62±19.83	-36.41±10.08	-58.99±7.20	-19.75±11.88	17.03±6.23
^{PhO}dG:G	*syn* (θ~0°)-1	-59.96±7.49	-9.79±5.27	-19.71±8.74	-11.30±4.73	3.01±5.10
	syn (θ~0°)-2	-65.35±7.87	-5.27±5.69	-10.29±7.45	-11.42±6.69	-4.56±7.74
	syn (θ~0°)-3	-59.83±7.95	-9.08±5.23	-19.21±8.91	-10.54±4.85	3.64±6.11
	syn (θ~0°)-4	-63.43±8.49	-5.61±5.15	-12.10±7.66	-11.42±6.69	-2.34±8.53
	syn (θ~180°)-1	-60.33±7.36	-10.29±5.94	-18.33±8.53	-12.05±5.10	2.13±5.73
	syn (θ~180°)-2	-64.94±8.37	-5.73±5.15	-9.79±7.071	-11.13±6.32	-5.69±8.16
	syn (θ~180°)-3	-61.17±7.78	-8.91±5.15	-17.49±8.33	-10.88±4.64	2.89±5.56
	syn (θ~180°)-4	-64.39±8.58	-5.27±4.81	-10.88±6.74	-11.46±6.78	-4.64±7.32
	anti (θ~0°)	-55.65±13.60	-42.38±8.37	-59.25±8.62	-16.53±7.49	12.93±7.53
	anti (θ~180°)	-47.82±16.15	-37.36±7.03	-56.90±6.690	-28.24±7.24	15.27±6.82
$^{o\text{-}PhOH}$dG:G	*syn* (θ~0°)	-57.82±8.16	-17.07±10.67	-7.82±5.77	-7.03±5.94	-3.18±7.95
	syn (θ~180°)	-70.71±9.25	-2.51±6.82	-9.29±10.04	-8.95±7.95	-4.10±9.41
	anti-1	-50.84±15.56	-41.46±8.83	-48.12±7.49	-13.05±8.49	11.42±7.82
	anti-2	-40.46±15.65	-25.86±5.90	-54.77±6.44	-29.20±5.23	14.69±8.12
$^{p\text{-}PhOH}$dG:G	*anti*-3	-55.02±14.23	-46.19±8.99	-44.85±8.03	-6.02±9.46	-8.45±7.61
	anti-4	-33.47±15.56	-30.50±5.98	-55.23±6.40	-27.28±5.48	15.82±5.94
	syn-1	-59.71±8.41	-7.03±6.69	-13.51±6.53	-8.07±6.32	2.26±9.08
	syn-2	-57.45±8.58	-12.092±5.98	-16.23±6.11	-7.82±5.86	10.21±7.49

[a]ΔE_{Hbond} is the hydrogen bond strength in the dimer consisting of the nucleobase at the G^3 position and the opposing base. [b]$\Delta E_{intra5'}$ is the stacking interaction energy between the nucleobase at the G^3 position and the intrastrand base at the 5′-side of the adduct. [c]$\Delta E_{intra3'}$ is the stacking interaction energy between the nucleobase at the G^3 position and the intrastrand base at the 3′-side of the adduct. [d]$\Delta E_{inter5'}$ is the stacking interaction energy between the nucleobase at the G^3 position and the base in the opposing strand at the 5′-side of the adduct. [e]$\Delta E_{inter3'}$ is the interaction energy between the nucleobase at the G^3 position and the base on the opposite strand stacked at the 3′-side of the adduct.

3.4. Conclusion

Since molecules are dynamic, experimental structures alone may not give the entire picture. Furthermore, the limitations in synthetic methods, facilities and difficulties in interpreting spectra, to provide clear idea for structural, conformational, and chemical properties of molecules, result in an increased demand for theoretical calculations. So an interdisciplinary approach is required. Molecular simulations are a necessary complement to experimental studies.

The results of our studies suggest that the single-ring oxygen linked C8-phenoxy-dG adduct (^{PhO}dG) adopts an *anti* conformation opposite C within the G^3 position of the 12–mer *Nar*I recognition sequence (5′–CTCG^{1}G^{2}CG^{3}CCATC–3′). In this conformation, the phenoxyl moiety resides in the major groove. No evidence of additional stacking interactions with ^{PhO}dG in a *syn* conformation opposite C was obtained. In contrast, both $^{o\text{-}PhOH}$dG:C (*syn*-180) and $^{p\text{-}PhOH}$dG:C (*syn*-2) showed an interclated conformation in the *Nar*I sequence, which may result in the formation of the 2-base bulge mutation. As a result, greater mutagenicity may expect for C-linked adducts in comparison to O-linked structures.

Within the double helix where the ^{PhO}dG:G is guanine mismatched, the ^{PhO}dG adopts a *syn* conformation and the phenoxyl ring is twisted from planarity.[17] Furthermore, the phenoxy moiety is residing in the minor groove and ^{PhO}dG adopts a "W-type" wedge conformation.. The same result is observed for the $^{o\text{-}PhOH}$dG and $^{p\text{-}PhOH}$dG adducts against a guanine mismatch. Induction of frameshift mutation is more probable in *ortho* and *para* C–linked structures. This is thought to be the effect of the oxygen bridge that allows more conformational flexibility and less

steric collisions in ^{PhO}dG in natural (paired with cytosine) *Nar*I sequence (pairing with cytosine) and against a guanine mismatch.

Overall, the *syn* conformer of the O-linked adduct induces more steric clashes when it is mismatched against a guanine complementary base compared to cytosine. However, the phenoxyl moiety does not cause a significant distortion of the double helix due to the small size of the aromatic link while the deformation of the double helix is more apparent in *ortho* and *para* C-linked structures.

3.5. References

1.   Millen, A. L.; Manderville, R. A.; Wetmore, S. D., Conformational Flexibility of C8-Phenoxyl-2 '-Deoxyguanosine Nucleotide Adducts. *Journal of Physical Chemistry B* **2010,** *114* (12), 4373-4382.

2.   Uesugi, S.; Ikehara, M., C-13 Magnetic-Resonance Spectra of 8-Substituted Purine Nucleosides - Characteristics Shifts for Syn Conformation. *Journal of the American Chemical Society* **1977,** *99* (10), 3250-3253.

3.   Millen, A. L.; McLaughlin, C. K.; Sun, K. M.; Manderville, R. A.; Wetmore, S. D., Computational and Experimental Evidence for the Structural Preference of Phenolic C-8 Purine Adducts. *Journal of Physical Chemistry A* **2008,** *112* (16), 3742-3753.

4.   Sessler, J. L.; Sathiosatham, M.; Doerr, K.; Lynch, V.; Abboud, K. A., A G-Quartet Formed in the Absence of a Templating Metal Cation: A New 8-(N,N-Dimethylaniline)Guanosine Derivative. *Angewandte Chemie-International Edition* **2000,** *39* (7), 1300-+.

5.   Gannett, P. M.; Sura, T. P., Base-Pairing of 8-Oxoguanosine and 8-Oxo-2'-Deoxyguanosine with 2'-Deoxyadenosine, 2'-Deoxycytosine, 2'-Deoxyguanosine, and Thymidine. *Chemical Research in Toxicology* **1993,** *6* (5), 690-700.

6.   Shapiro, R.; Ellis, S.; Hingerty, B. E.; Broyde, S., Effect of Ring Size on Conformations of Aromatic Amine-DNA Adducts: The Aniline-C8 Guanine Adduct Resides in the B-DNA Major Groove. *Chemical Research in Toxicology* **1998,** *11* (4), 335-341.

7.   Liang, F. T.; Meneni, S.; Cho, B. P., Induced Circular Dichroism Characteristics as Conformational Probes for Carcinogenic Aminofluorene-DNA Adducts. *Chemical Research in Toxicology* **2006,** *19* (8), 1040-1043.

8.   Patel, D. J.; Mao, B.; Gu, Z. T.; Hingerty, B. E.; Gorin, A.; Basu, A. K.; Broyde, S., Nuclear Magnetic Resonance Solution Structures of Covalent Aromatic Amine-DNA Adducts and Their Mutagenic Relevance. *Chemical Research in Toxicology* **1998,** *11* (5), 391-407.

9.   Jain, N.; Reshetnyak, Y. K.; Gao, L.; Chiarelli, M. P.; Cho, B. P., Fluorescence Probing of Aminofluorene-Induced Conformational Heterogeneity in DNA Duplexes. *Chemical Research in Toxicology* **2008,** *21* (2), 445-452.

10.   Abuaf, P.; Hingerty, B. E.; Broyde, S.; Grunberger, D., Solution Conformation of the N-(Deoxyguanosin-8-Yl)Aminofluorene Adduct Opposite Deoxyinosine and Deoxyguanosine in DNA by Nmr and Computational Characterization. *Chemical Research in Toxicology* **1995,** *8* (3), 369-378.

11.   Omumi, A.; Millen, A. L.; Wetmore, S. D.; Manderville, R. A., Fluorescent Properties and Conformational Preferences of C-Linked Phenolic-DNA Adducts. *Chemical Research in Toxicology* **2011,** *24* (10), 1694-1709.

12.   Broyde, S.; Hingerty, B., Conformation of 2-Aminofluorene-Modified DNA. *Biopolymers* **1983,** *22* (11), 2423-2441.

13.   Broyde, S.; Hingerty, B. E.; Srinivasan, A. R., Influence of the Carcinogen 4-Aminobiphenyl on DNA Conformation. *Carcinogenesis* **1985,** *6* (5), 719-725.

14.   Broyde, S.; Wang, L. H.; Zhang, L.; Rechkoblit, O.; Geacintov, N. E.; Patel, D. J., DNA Adduct Structure-Function Relationships: Comparing Solution with Polymerase Structures. *Chemical Research in Toxicology* **2008,** *21* (1), 45-52.

15.   Brown, K.; Hingerty, B. E.; Guenther, E. A.; Krishnan, V. V.; Broyde, S.; Turteltaub, K. W.; Cosman, M., Solution Structure of the 2-Amino-1-Methyl-6-Phenylimidazo 4,5-B Pyridine C8-Deoxyguanosine Adduct in Duplex DNA. *Proceedings of the National Academy of Sciences of the United States of America* **2001,** *98* (15), 8507-8512.

16.   Broyde, N. E. G. a. S., *The Chemical Biology of DNA Damage.* © 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim: New York University.

17.   Kuska, M. S.; Witham, A. A.; Sproviero, M.; Manderville, R. A.; Majdi Yazdi, M.; Sharma, P.; Wetmore, S. D., Structural Influence of C8-Phenoxy-Guanine in the Nari Recognition DNA Sequence. *Chemical Research in Toxicology* **2013,** *26* (9), 1397-1408.

18.   Zhou, L.; Rajabzadeh, M.; Traficante, D. D.; Cho, B. P., Conformational Heterogeneity of Arylamine-Modified DNA: 19f Nmr Evidence. *Journal of the American Chemical Society* **1997,** *119* (23), 5384-5389.

19.   van de Poll, M. L. M.; Lafleur, M. V. M.; Van Gog, F.; Vrieling, H.; Meerman, J. H. N., N-Acetylated and Deacetylated 4'-Fluoro-4-Aminobiphenyl and 4-Aminobiphenyl Adducts Differ in Their Ability to Inhibit DNA Replication of Single-Stranded M13 in Vitro and of Single-Stranded Φx174 in Escherichia Coli. *Carcinogenesis* **1992,** *13* (5), 751-758.

20.   Patnaik, S.; Cho, B. P., Structures of 2-Acetylaminofluorene Modified DNA Revisited: Insight into Conformational Heterogeneity. *Chemical Research in Toxicology* **2010,** *23* (11), 1650-1652.

21.   Liang, F. T.; Cho, B. P., Conformational and Thermodynamic Impact of Bulky Aminofluorene Adduction on Simulated Trans Lesion DNA Synthesis. *Chemical Research in Toxicology* **2011,** *24* (4), 597-605.

22.   Omumi, A.; Millen, A. L.; Wetmore, S. D.; Manderville, R. A., Fluorescent Properties and Conformational Preferences of C-Linked Phenolic-DNA Adducts. *Chem Res Toxicol* **2011,** *24* (10), 1694-709.

23.   Millen, A. L.; Sharma, P.; Wetmore, S. D., C8-Linked Bulky Guanosine DNA Adducts: Experimental and Computational Insights into Adduct Conformational Preferences and Resulting Mutagenicity. *Future Medicinal Chemistry* **2012,** *4* (15), 1981-2007.

24.   Omumi, A.; Beach, D. G.; Baker, M.; Gabryelski, W.; Manderville, R. A., Postsynthetic Guanine Arylation of DNA by Suzuki-Miyaura Cross-Coupling. *Journal of the American Chemical Society* **2011,** *133* (1), 42-50.

25.   Rankin, K. M.; Sproviero, M.; Rankin, K.; Sharma, P.; Wetmore, S. D.; Manderville, R. A., C-8-Heteroaryl-2 '-Deoxyguanosine Adducts as Conformational Fluorescent Probes in the Nari Recognition Sequence. *Journal of Organic Chemistry* **2012,** *77* (23), 10498-10508.

26.   Farrell, P. M.; Rosenstein, B. J.; White, T. B.; Accurso, F. J.; Castellani, C.; Cutting, G. R.; Durie, P. R.; LeGrys, V. A.; Massie, J.; Parad, R. B.; Rock, M. J.; Campbell, P. W., Guidelines for Diagnosis of Cystic Fibrosis in Newborns through Older Adults: Cystic Fibrosis Foundation Consensus Report. *Journal of Pediatrics* **2008,** *153* (2), S4-S14.

27.   Myerowitz, R., Tay-Sachs Disease-Causing Mutations and Neutral Polymorphisms in the Hex a Gene. *Human Mutation* **1997,** *9* (3), 195-208.

28.   Rader, D. J.; Cohen, J.; Hobbs, H. H., Monogenic Hypercholesterolemia: New Insights in Pathogenesis and Treatment. *Journal of Clinical Investigation* **2003,** *111* (12), 1795-1803.

29.   Manderville, R. A., Structural and Biological Impact of Radical Addition Reactions with DNA Nucleobases. In *Advances in Physical Organic Chemistry, Vol 43*, Richard, J. P., Ed. 2009; Vol. 43, pp 177-218.

30.   Sugimura, T., Overview of Carcinogenic Heterocyclic Amines. *Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis* **1997,** *376* (1-2), 211-219.

31.   Meneni, S.; Liang, F.; Cho, B. P., Examination of the Long-Range Effects of Aminofluorene-Induced Conformational Heterogeneity and Its Relevance to the Mechanism of Translesional DNA Synthesis. *Journal of Molecular Biology* **2007,** *366* (5), 1387-1400.

32.   Jain, N.; Li, Y.; Zhang, L.; Meneni, S. R.; Cho, B. P., Probing the Sequence Effects on Nari-Induced −2 Frameshift Mutagenesis by Dynamic 19f Nmr, Uv, and Cd Spectroscopy†. *Biochemistry* **2007,** *46* (46), 13310-13321.

33.   Fuchs, R. P. P.; Schwartz, N.; Daune, M. P., Hot Spots of Frameshift Mutations Induced by the Ultimate Carcinogen N-Acetoxy-N-2-Acetylaminofluorene. *Nature* **1981,** *294* (5842), 657-659.

34.   Hoffmann, G. R.; Fuchs, R. P. P., Mechanisms of Frameshift Mutations: Insight from Aromatic Amines. *Chemical Research in Toxicology* **1997,** *10* (4), 347-359.

35.   Elmquist, C. E.; Wang, F.; Stover, J. S.; Stone, M. P.; Rizzo, C. J., Conformational Differences of the C8-Deoxyguanosine Adduct of 2-Amino-3-Methylimidazo 4,5-F Quinoline (Iq) within the Nari Recognition Sequence. *Chemical Research in Toxicology* **2007,** *20* (3), 445-454.

36.   Wang, F.; DeMuro, N. E.; Elmquist, C. E.; Stover, J. S.; Rizzo, C. J.; Stone, M. P., Base-Displaced Intercalated Structure of the Food Mutagen 2-Amino-3-Methylimidazo[4,5-F]Quinoline in the Recognition Sequence of the Nari Restriction Enzyme, a Hotspot for −2 Bp Deletions. *Journal of the American Chemical Society* **2006,** *128* (31), 10085-10095.

37.   Dutta, S.; Li, Y.; Johnson, D.; Dzantiev, L.; Richardson, C. C.; Romano, L. J.; Ellenberger, T., Crystal Structures of 2-Acetylaminofluorene and 2-Aminofluorene in Complex with T7 DNA Polymerase Reveal Mechanisms of Mutagenesis. *Proceedings of the National Academy of Sciences of the United States of America* **2004,** *101* (46), 16186-16191.

38.   Wilson, K. A.; Wetmore, S. D., Complex Conformational Heterogeneity of the Highly Flexible O6-Benzyl-Guanine DNA Adduct. *Chemical Research in Toxicology* **2014,** *27* (7), 1310-1325.

39.   Kuska, M. S.; Majdi Yazdi, M.; Witham, A. A.; Dahlmann, H. A.; Sturla, S. J.; Wetmore, S. D.; Manderville, R. A., Influence of Chlorine Substitution on the Hydrolytic Stability of Biaryl Ether Nucleoside Adducts Produced by Phenolic Toxins. *The Journal of Organic Chemistry* **2013,** *78* (14), 7176-7185.

40.   Watson, J. D.; Crick, F. H. C., Molecular Structure of Nucleic Acids - a Structure for Deoxyribose Nucleic Acid. *Nature* **1953,** *171* (4356), 737-738.

41.   Schlitt, K. M.; Sun, K. W.; Paugh, R. J.; Millen, A. L.; Navarro-Whyte, L.; Wetmore, S. D.; Manderville, R. A., Concerning the Hydrolytic Stability of 8-Aryl-2'-Deoxyguanosine Nucleoside Adducts: Implications for Abasic Site Formation at Physiological Ph. *J Org Chem* **2009,** *74* (16), 5793-802.

42.   Millen, A. L.; Churchill, C. D. M.; Manderville, R. A.; Wetmore, S. D., Effect of Watson-Crick and Hoogsteen Base Pairing on the Conformational Stability of C8-Phenoxyl-2 '-Deoxyguanosine Adducts. *Journal of Physical Chemistry B* **2010,** *114* (40), 12995-13004.

43.   Freddolino, P. L.; Arkhipov, A. S.; Larson, S. B.; McPherson, A.; Schulten, K., Molecular Dynamics Simulations of the Complete Satellite Tobacco Mosaic Virus. *Structure* **2006,** *14* (3), 437-449.

44.   Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M., Refinenement of the Amber Force Field for Nucleic Acids: Improving the Description of Alpha/Gamma Conformers. *Biophysical Journal* **2007,** *92* (11), 3817-3829.

45.   Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber Biomolecular Simulation Programs. *Journal of Computational Chemistry* **2005,** *26* (16), 1668-1688.

46.   Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M., Charmm: The Biomolecular Simulation Program. *Journal of Computational Chemistry* **2009,** *30* (10), 1545-1614.

47.   Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.;

Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *Amber Tools*, Version 1.0; University of California: San Francisco, 2008.

48.   Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society* **1995,** *117* (19), 5179-5197.

49.   Wang, J.; Cieplak, P.; Kollman, P. A., How Well Does a Restrained Electrostatic Potential (Resp) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *Journal of Computational Chemistry* **2000,** *21* (12), 1049-1074.

50.   Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and Testing of a General Amber Force Field. *Journal of Computational Chemistry* **2004,** *25* (9), 1157-1174.

51.   Hehre, W. J.; Ditchfield, R.; Pople, J. A., Self—Consistent Molecular Orbital Methods. Xii. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *The Journal of Chemical Physics* **1972,** *56* (5), 2257-2261.

52.   Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J., Gaussian 09, Revision B.01. Wallingford CT, 2009.

53.   Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A., A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The Resp Model. *The Journal of Physical Chemistry* **1993,** *97* (40), 10269-10280.

54.   Hestenes, M. R.; Stiefel, E., Methods of Conjugate Gradients for Solving Linear Systems. *Journal of Research of the National Bureau of Standards* **1952,** *49* (6), 409-436.

55.   Snyman, J. A., Practical Mathematical Optimization. *Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms* **2005,** *97*, 1-257.

56.   Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C., Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *Journal of Computational Physics* **1977,** *23* (3), 327-341.

57.   Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R., Molecular Dynamics with Coupling to an External Bath. *The Journal of Chemical Physics* **1984,** *81* (8), 3684-3690.

58.   Izaguirre, J. A.; Catarello, D. P.; Wozniak, J. M.; Skeel, R. D., Langevin Stabilization of Molecular Dynamics. *Journal of Chemical Physics* **2001,** *114* (5), 2090-2098.

59.   Berendsen, H. J. C.; Postma, J. P. M.; Vangunsteren, W. F.; Dinola, A.; Haak, J. R., Molecular-Dynamics with Coupling to an External Bath. *Journal of Chemical Physics* **1984,** *81* (8), 3684-3690.

60.   Jorgensen, W. L.; Madura, J. D., Quantum and Statistical Studies of Liquids .25. Solvation and Conformation of Methanol in Water. *Journal of the American Chemical Society* **1983,** *105* (6), 1407-1413.

61.   Smith, B. B., Dynamics of Proteins and Nucleic Acids: By Ja Mccammon and Sc Harvey. Pp 234. Cambridge University Press, Ny. 1987. $19.95 (Pbk). *Biochemical Education* **1989,** *17* (4), 220-220.

62.   York, D. M.; Darden, T. A.; Pedersen, L. G., The Effect of Long-Range Electrostatic Interactions in Simulations of Macromolecular Crystals - a Comparison of the Ewald and Truncated List Methods. *Journal of Chemical Physics* **1993,** *99* (10), 8345-8348.

63.   Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A Smooth Particle Mesh Ewald Method. *Journal of Chemical Physics* **1995,** *103* (19), 8577-8593.

64.   Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P., The R.E.D. Tools: Advances in Resp and Esp

Charge Derivation and Force Field Library Building. *Physical Chemistry Chemical Physics* **2010,** *12* (28), 7821–7839.

65.   Case, D.; Darden, T. A.; Cheatham, T. E.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Crowley, M.; Walker, R.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P., *Amber 11*.

66.   Cheatham, T. E.; Cieplak, P.; Kollman, P. A., A Modified Version of the Cornell Et Al. Force Field with Improved Sugar Pucker Phases and Helical Repeat. *Journal of Biomolecular Structure & Dynamics* **1999,** *16* (4), 845–862.

67.   Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and Testing of a General Amber Force Field. *Journal of Computational Chemistry* **2004,** *25* (9), 1157–1174.

68.   Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *Journal of Molecular Graphics and Modelling* **2006,** *25*, 247–260.

69.   Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of Simple Potential Functions for Simulating Liquid Water. *Journal of Chemical Physics* **1983,** *79* (2), 926-935.

70.   Gonnet, P., P-Shake: A Quadratically Convergent Shake in O(N(2)). *Journal of Computational Physics* **2007,** *220* (2), 740-750.

71.   Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E., Mmpbsa.Py: An Efficient Program for End-State Free Energy Calculations. *Journal of Chemical Theory and Computation* **2012,** *8* (9), 3314-3321.

72.   Hang, B.; Chenna, A.; Guliaev, A. B.; Singer, B., Miscoding Properties of 1,N-6-Ethanoadenine, a DNA Adduct Derived from Reaction with the Antitumor Agent 1,3-Bis(2-Chloroethyl)-1-Nitrosourea. *Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis* **2003,** *531* (1-2), 191-203.

73.   Seo, K.-Y.; Jelinsky, S. A.; Loechler, E. L., Factors That Influence the Mutagenic Patterns of DNA Adducts from Chemical Carcinogens. *Mutation Research/Reviews in Mutation Research* **2000,** *463* (3), 215-246.

74.   Mantle, P. G.; Faucet-Marquis, V.; Manderville, R. A.; Squillaci, B.; Pfohl-Leszkowicz, A., Structures of Covalent Adducts between DNA and Ochratoxin A: A

New Factor in Debate About Genotoxicity and Human Risk Assessment. *Chemical Research in Toxicology* **2009,** *23* (1), 89-98.

75.   Pfohl-Leszkowicz, A.; Manderville, R. A., Ochratoxin A: An Overview on Toxicity and Carcinogenicity in Animals and Humans. *Molecular Nutrition & Food Research* **2007,** *51* (1), 61-99.

76.   Gates, K. S.; Nooner, T.; Dutta, S., Biologically Relevant Chemical Reactions of N7-Alkylguanine Residues in DNA. *Chemical Research in Toxicology* **2004,** *17* (7), 839-856.

77.   Cho, B. P., Dynamic Conformational Heterogeneities of Carcinogen-DNA Adducts and Their Mutagenic Relevance. *Journal of Environmental Science and Health, Part C* **2004,** *22* (2), 57-90.

78.   Pfohl-Leszkowicz, A.; Manderville, R. A., Ochratoxin A: An Overview on Toxicity and Carcinogenicity in Animals and Humans (Vol 51, Pg 61, 2007). *Molecular Nutrition & Food Research* **2007,** *51* (9), 1192-1192.

79.   Cho, B. P.; Beland, F. A.; Marques, M. M., Nmr Structural Studies of a 15-Mer DNA Duplex from a Ras Protooncogene Modified with the Carcinogen 2-Aminofluorene: Conformational Heterogeneity. *Biochemistry* **1994,** *33* (6), 1373-1384.

80.   Ohandley, S. F.; Sanford, D. G.; Xu, R.; Lester, C. C.; Hingerty, B. E.; Broyde, S.; Krugh, T. R., Structural Characterization of an N-Acetyl-2-Aminofluorene (Aaf) Modified DNA Oligomer by Nmr, Energy Minimization, and Molecular-Dynamics. *Biochemistry* **1993,** *32* (10), 2481-2497.

81.   Nolan, S. J.; McNulty, J. M.; Krishnasamy, R.; McGregor, W. G.; Basu, A. K., C8-Guanine Adduct-Induced Stabilization of a-1 Frame Shift Intermediate in a Nonrepetitive DNA Sequence. *Biochemistry* **1999,** *38* (42), 14056-14062.

82.   Leslie, A. G. W.; Arnott, S.; Chandrasekaran, R.; Ratliff, R. L., Polymorphism of DNA Double Helices. *Journal of Molecular Biology* **1980,** *143* (1), 49-72.

83.   Schneider, B.; Neidle, S.; Berman, H. M., Conformations of the Sugar-Phosphate Backbone in Helical DNA Crystal Structures. *Biopolymers* **1997,** *42* (1), 113-124.

84.   Packer, M. J.; Hunter, C. A., Sequence-Dependent DNA Structure: The Role of the Sugar-Phosphate Backbone. *Journal of Molecular Biology* **1998,** *280* (3), 407-420.

85.	Neidle, S., Principles of Nucleic Acid Structure. 1st ed.; Oxford University Press: Oxford, 2008.

Chapter 4. DFT Calculations on the Stability of Chloro Substituted O–linked Adducts, at Nucleobase, Nucleoside, and Nucleotide Levels. The Influence of Chlorine Substitution and Protonation on the Hydrolytic Stability of Biaryl Ether Nucleoside Adducts Produced from Phenolic Toxins

4.1. Introduction

The basic hypothesis of chemical carcinogenesis is that formation of a covalent bond between a chemical and DNA to form a DNA adduct represents the first essential step in the tumor initiation process. Nucleobases, especially 2′-deoxyguanosine (dG), are preferential targets of electrophilic attachment.[1] A single electrophile can react with different sites on the guanine nucleobase of DNA. Various carbon and nitrogen atoms on dG can be modified by several electrophiles derived from chemical carcinogens. Exocyclic and endocyclic N and the O6 site of dG are reactive towards electrophilic addition by alkylating agents which can be derived from aflatoxins, nitroamines and mustards. Endocyclic N1 and N3 can react with α, β–unsaturated carbonyl groups. PAHs undergo bioactivation to electrophiles which tend to react with the exocyclic N2 site of dG. The C8 site of dG is also a target site for modification formation of lesions (Scheme 4.1).[2] This is an interesting aspect of C8 reactivity because this site is not thought to be highly nucleophilic and does not react with most electrophilic species. However, it is the preferred site of certain radical addition reactions. This preference stems, in part, from the fact that dG has the lowest oxidation potential (1.3 V vs. NHE, pH 7) of all four nucleobases making it most susceptible to oxidation. Hydroxyl radicals,

nitrogen dioxide, alkyl radicals, arylamines, nitroaromatics, and phenoxyl radicals are all known to form covalent adducts at C8 of dG.[3-5]
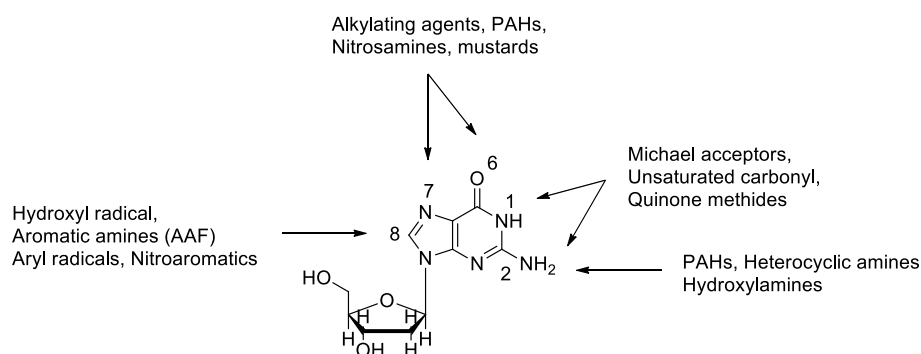
Scheme 4.1. Preferential sites of reactivity with 2'-deoxyguanosine by various electrophiles.[2]

In addition to carcinogenesis, the relationships between cellular DNA damage caused by endogenous and environmental genotoxic agents, the cellular response, and the development and prevention of human diseases and aging have recently attracted great interest in the medical, biological, and chemical research communities.[6, 7]

Phenols, as common organic compounds, show beneficial biological activities which are observed in vitamin E and other phenolic counterparts that possess antioxidant properties.[8] On the other hand the carcinogenicity of phenol is evident in the chlorophenols, which are among the external causes of DNA damage. Chlorophenols can be found in pesticides, herbicides and wood preservatives.[1] Breakdown of some natural products can produce some agents that can display deleterious pro-oxidant properties. These agents and enzymes with peroxidase activity can initiate oxidation of phenols by elimination of one electron and formation of phenoxyl radicals as intermediates. Examples include

156

pentachlorophenol (PCP), 2,4,6-trichlorophenol, 2,4-dichlorophenol, other isomers and the mycotoxin ochratoxin A. Experimental studies on modification of DNA by chlorophenols clarify the tendency of adduction at C8 site of guanosine.[9, 10] It has been identified that phenols undergo peroxidase-mediated oxidation into phenoxyl radicals.[11, 12] Chlorophenol carcinogenesis could also be linked with metabolic activation to phenoxyl radicals and quinones that covalently bind to DNA. Likewise, quinone metabolites of simple chlorophenols, react with DNA and multiple adducts can be formed.[10, 11, 13] In vitro reactivity profiles for chlorophenol metabolites indicate the potential for direct DNA adduction by oxygen-based radicals, carbon-based radicals and quinones (Scheme 4.2).
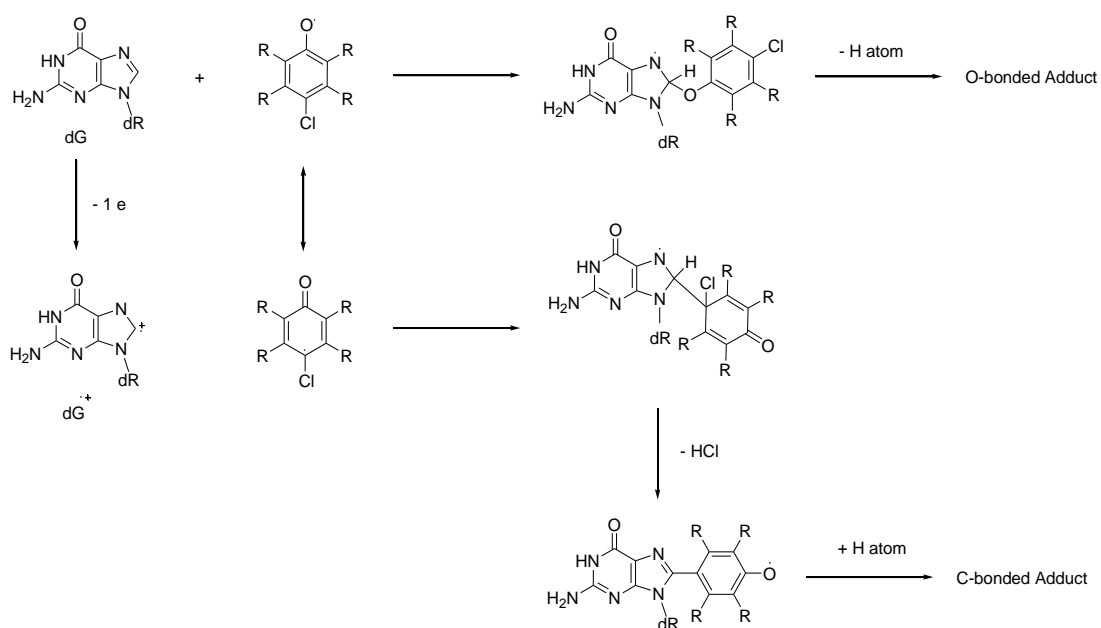
Scheme 4.2. Summary of HRP/H_2O_2-Mediated Reaction of a CP with dG.[11]

Depurination is known to play a major role in cancer initiation.[14, 15] Hydrolytic cleavage of β-N-glycosidic bond (1′-N9) of purines which lead to releasing a nucleic base is known as depurination. The second fragment formed in the depurination

157

of deoxyribonucleosides and ribonucleosides is the sugar, e.g. 2′–deoxyribose and ribose (Figure 4.1). Larger molecules which are more complex, such as nucleoside residues, nucleotides and nucleic acids, also undergo depurination. Deoxyribonucleosides and their derivatives are substantially more prone to depurination than their corresponding ribonucleoside counterparts. [15]
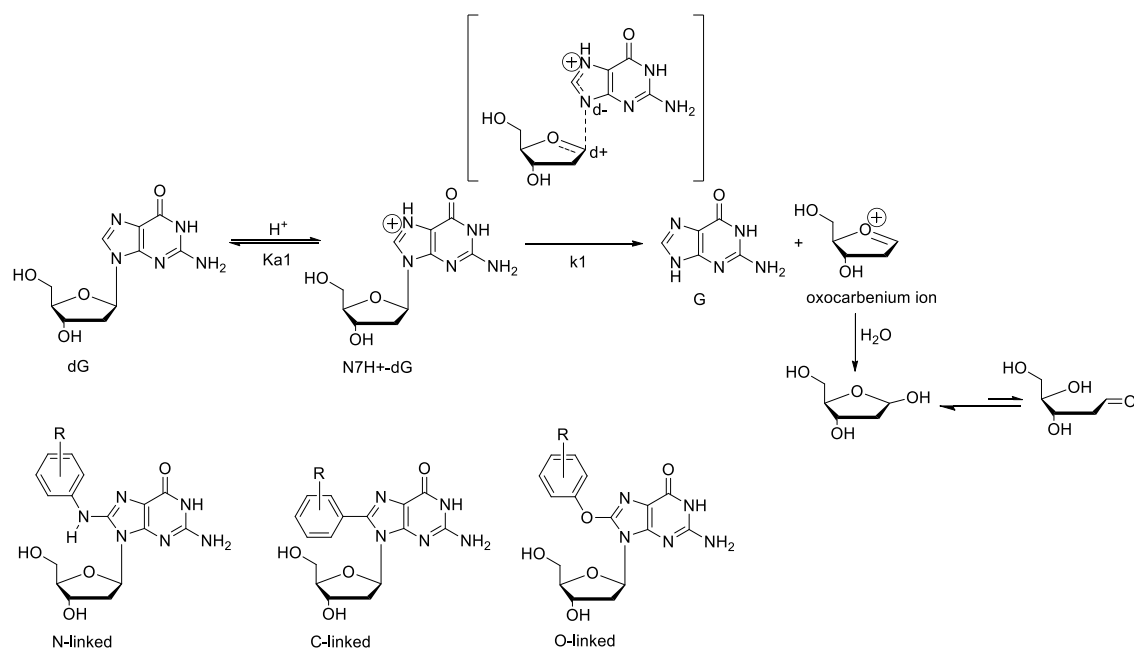
Figure 4.1. Acid-catalyzed hydrolysis of dG; the same mechanism for O-linked, C-linked and N-linked adducts.[16]

Purine is a good leaving group via the N9-nitrogen, hence depurination takes place easily and is not an uncommon reaction. Studies estimate that around 5,000 purines are lost in this way each day in a typical human cell.[3, 16] In cells, one of the main causes of depurination takes place by reaction with endogenous metabolites. The anomeric carbon is especially reactive towards nucleophilic substitution: the carbon-oxygen bond is shorter, stronger and more polar, while the carbon-purine

bond is longer and weaker. This makes the bond especially susceptible to hydrolysis. In the chemical synthesis of oligonucleotides, depurination is one of the major factors limiting the length of synthetic oligonucleotides.[17]

Proton transfer reactions are of great importance in chemistry and in the biomolecular processes of living organisms. The latter include most enzymatically catalyzed reactions. The protonation state of chemical groups, e.g. the side chains of amino acids, is fundamentally related to their biomolecular function. It has been identified that acid can accelerate the rate of the depurination in unmodified 2′-deoxyguanosine (dG).

Protonation increases the rate of depurination for unmodified 2′-deoxyguanosine. A stepwise mechanism is proposed for the acid-catalyzed hydrolysis of dG (Figure 4.1). The protonation occurs in the first step and is defined by the acid dissociation constant K_{a1}. The second step, which is the rate determining step, is the cleavage of the glycosidic bond and proceeds in a unimolecular fashion and is defined by k_1. Protonation makes the guanine a good leaving group. An oxocarbenium ion converts to the 1′-hydroxylated sugar after a hydration step. Finally, this sugar is tautomerized to the aldehyde isomer which may lead to the formation of an interstrand DNA cross-link lesion in duplex DNA.[18, 16] Covalent modification of dG, by affecting the accumulation of electrons at N9 site of guanosine, increases the rate of depurination and forms abasic site. It has been demonstrated that N7 position of guanosine is the most nucleophilic site and this is due to resonance.[19] Therefore, the protonation of guanosine at the N7 site by formation of a positive charge on guanosine forms a better leaving group; also,

alteration of the C8 site of guanosine with other electrophiles can accelerate the depurination.

Interestingly, C8 of dG is also susceptible to covalent modification by numerous electrophiles. The type of C8-substituent is a determining factor in inducing the properties that make the 8-guanine adduct an excellent leaving group for depurination. In general, electron-withdrawing substituents at the 8-position of dG can stabilize the developing negative charge at N9 during rate-limiting cleavage of the glycosidic bond.[20] Bulky aryl ring systems can be attached to C8 of dG to form nitrogen-, carbon-, or oxygen-linked adducts. N-linked adducts are generated from reaction of the DNA nucleobases with nitrenium ion metabolites derived from arylamine carcinogens.[21, 22] N-linked adducts can easily be depurinated under mildly acidic to neutral conditions, where dG has little reactivity.[21] The corresponding C-linked adducts having direct attachment of the aryl ring at C8 of dG has a small impact on pK_{a1}.[23] Nevertheless, protonation of adducts and increasing acidity of solution enhance the reactivity of adducts in comparison with dG. As a matter of fact, functionalization of the phenyl ring with electron-withdrawing substituents causes an increase in the rate of hydrolysis related to natural guanosine. Although there are several factors accelerating depurination, the relief of steric strain during the process of depurination by removal of sugar moiety is the most important factor.[23] Finally, substitution of the aromatic ring of phenoxyl by chlorine atoms, as electron withdrawing groups, enhances the rate of hydrolysis.

O−linked C8-dG adducts bearing electron-deficient polychlorinated phenoxyl ring systems were expected to have enhanced susceptibility to hydrolysis. In this chapter, the structural features of chloro substituted O−linked adducts are studied computationally. In addition, a theoretical approach to determine proton affinities and deglycosylation barriers and draw comparisons between the hydrolysis rates of all chloro O−linked complexes and protonated structures is presented. This will shed light on the impact of modified mutated complexes, functionalization of the aryl group on the rates of depurination and formation of abasic sites.

4.2. Computational Details

4.2.1. Nucleobase Model

To characterize the conformations of the functionalized O−linked phenol adducts, the same procedure used for unsubstituted adducts was applied to the chloro substituted O−linked structures. DFT (density functional theory)[24] was used in the form of the B3LYP functional. The use of B3LYP in this project was justified based on past work done on the related C-bonded adducts in the Wetmore laboratory.[25] Specifically, the B3LYP/6-311+G(2df,p)//B3LYP/6-31G(d) level of theory was found to provide accurate geometries. For all chloro-substituted mutated complexes, B3LYP/6-31G(d) potential energy surfaces were initially searched for selected adducts. First, all local minima and transition states for systematic rotation about the C8−O (θ) and O−C bonds (ϕ) for the O−linked (unsubstituted) G adduct were considered. Specifically, the dihedral angles θ (theta) and ϕ (phi) were rotated and fixed in 20° increments from 0° to 360°. In

161

the case of the [4-Cl-PhO]dG monochlorosubstituted adduct, one chlorine atom was added to the phenoxyl ring of the fully optimized minima and transition states of the unsubstituted O−linked G adduct. The geometries of the [PCP-O]G (pentachloro) adduct were found by addition of two chlorine atoms to the minima and transition state structures of the [TCP-O]G (trichloro) adduct. B3LYP/6-31G(d) full optimizations followed by B3LYP/6-31G(d) frequency and B3LYP/6-311+G(2df,p) single- point calculations were then run on the new fully-optimized structures in each case.

4.2.2. Nucleoside Model

The lowest energy conformation of the nucleobase adducts were used to build the nucleoside model. Minima for the [4-Cl-PhO]dG, [DCP-O]dG, [TCP-O]dG, [PCP-O]dG nucleoside adducts were initially identified through a conformational search using the internal coordinate Monte Carlo algorithm in HyperChem 8.0.8. The AMBER molecular mechanics force field, with PM3 charges, was implemented in the conformational search and the θ ($\angle$N9−C8−O−C) and ϕ ($\angle$C8−O−C−C4), χ($\angle$O4′−C1′−N9′−C4), as well as the endocyclic torsion angles of the sugar, namely υ_0 ($\angle$O4′−C1′−C2′−C3′), υ_1 ($\angle$C1′−C2′−C3′−C4′), υ_2 ($\angle$C2′−C3′−C4′−O4′), υ_3 ($\angle$C3′−C4′−O4′−C1′), υ_4 ($\angle$C4′−O4′−C1′−C2′), β ($\angle$C4′−C5′−O−H), ε ($\angle$C4′−C3′−O−H), and γ ($\angle$O3′−C4′−C5′−O) dihedral angles were scanned.[26, 27] This procedure led to more than 200 conformers, out of which the 50 lowest energy conformers subsequently fully optimized with B3LYP/6-31G(d). Finally, B3LYP/6-311+G(2df,p) single-point energy calculations were carried out on the resulting ten lowest energy structures to identify the most stable conformer. In order to

identify the global minima and local minima and transition states, systematic gas-phase B3LYP/6-31G(d) PES scans were performed with the θ (∠N9–C8–O–C) and χ(∠O4′–C1′–N9′–C4) dihedral angles constrained in 10° increments from 0° to 360°. Upon addition of the 2′-deoxyribose sugar the sugar puckering must be considered. The starting models for these PES scans were optimized with C2′-endo sugar puckering, which is present in the B-form of DNA and is the preferred pucker in the related C8-phenoxyl-2′-deoxyguanosine adducts. Moreover, the 5′–OH was allowed to adopt the lowest energy orientation, which involves interactions with the nucleobase (denoted as β-unconstrained), and is relevant when the nucleoside is at the 5′-terminal position in a duplex. As a result of this process, the PESs for all the chloro-substituted O-linked nucleoside models were generated.

4.2.3. Nucleotide Model

Previous studies on the natural dG nucleotide, as well as C8-bonded phenoxyl dG adducts indicate that nucleotide conformations[28] can be accurately characterized when solvation (water, $\varepsilon = 78.4$) is taken into account during PCM-B3LYP/6-31G(d) optimizations (further explanation about PCM is provided in *4.2.4. Proton affinity and deglycosilation barriers*), and an anionic phosphate model is neutralized with a Na⁺ counterion. Geometries obtained from this approach are compatible with the structures that were generated by using larger basis sets and specifically including diffuse functions (6-31+G(d,p)) to take the charge of the phosphate into account.[28, 29] This developed protocol was used for the nucleotide models. The nucleotide model was generated by adding a 5′–monophosphate

group to the lowest energy *anti* and *syn* conformations identified using the β–constrained nucleoside model. For the nucleobase and nucleoside models, reported relative energies were obtained from gas-phase B3LYP/6-311+G(2df,p) single-point calculations, whereas the corresponding calculations were carried out in water for the nucleotide models. All relative energies include scaled (0.9806) zero-point vibrational energy (ZPVE) corrections.

4.2.4. Proton affinity and deglycosilation barriers

The resulting gas-phase global minima from PESs for $^{4\text{-Cl-PhO}}$dG, $^{DCP\text{-O}}$dG, $^{TCP\text{-O}}$dG, $^{PCP\text{-O}}$dG were optimized by IEF-PCM B3LYP/6-31G(d) in water to obtain the corresponding structures (further explanation about IEF is provided in *4.2.4. Proton affinity and deglycosilation barriers*). The lowest energy gas and solvent phase structures were then used to determine the proton affinities (PA) and the deglycosylation barrier of O-linked structures. The B3LYP/6-311+G(2df,p) PA (including B3LYP/6-31G(d) ZPVE corrections) at the N3 and N7 sites of the O-linked 8-dG adducts were determined in the gas and solvent (water) phases as the negative of the enthalpy change for protonation. To scrutinize the effects of the phenoxyl moiety, Cl-substitution and protonation on sugar loss, the deglycosylation reaction at the nucleoside level was investigated by altering and fixing the glycosidic bond length (1′–N9) in 0.1 Å increments from 1.4 Å to 3.5 Å in both the gas phase and water for all chloro substituted adducts and N3 and N7 protonated species. All quantum chemical B3LYP calculations were performed using Gaussian 09 revisions A.02 and C.01.[30]

The type of model which was used for determining the solvent effect is distinctive. Indeed, in the Self-Consistent Isodensity Polarized Continuum Model (SCI-PCM), the cavity is an isodensity surface of the molecule. The isodensity PCM model (SCRF=IPCM) are calculations that are used for determining the PA in water. It employs a converged SCF numerically, and iterates until the cavity shape no longer changes. The Self-Consistent IPCM model (SCRF=SCIPM) is implemented based on the different embedded cavity in the SCF calculation and includes the effect of solvation.

The Polarizable Continuum Model (PCM) using the integral equation formalism variant (IEFPCM) is the default SCRF method.

Explicit solvent (all-atom description) methods, which reflects the realistic physical picture of the system, was not used since it is computationally expensive due to consideration of the full details of molecular structure. Processing many atoms requires long runs in order to equilibrate solvent and solute which is time consuming and requires lots of memory. Additionally, this model cannot specify the polarizability of solvent and solute. In addition, large fluctuations are provoked due to the small system size. Therefore, we used the implicit solvent (continuum description) approach. The implicit solvent method has also some disadvantages compared to QM, the highest possible level accredited for treating the solute. Second, an artificial boundary between the solute and solvent in the implicit solvent model reduces the efficiency of this method in treating short range effects. To address these issues we modeled the solvent as a polarizable continuum and not as individual molecules which also makes *ab initio* computations feasible.

4.3. Results and Discussions

4.3.1. Nucleobase Model

4.3.1.1. Structure of ^{PhO}G

The structural and conformational properties of unsubstituted O-linked guanine adducts has been discussed in Chapter 2. According to the B3LYP potential energy surface (Figure 4.2) the lowest energy structure (global minimum) was found to have a theta (θ) value of ~180 ° and a phi (Φ) value of ~359.8°. Another planar minimum (local minimum) resulting from molecular symmetry appears at ~179.8° (θ) and ~179.9° (ϕ) which is quite similar to the global minimum. In the transition state the nucleobase and the phenol ring are approximately perpendicular, with a θ ~180° and Φ ~91.8°, were also characterized (Figure 2.3). Thus, in summary, the O-linked unsubstituted adduct has one planar global minimum that is connected to a mirror image of itself by one perpendicular transition state, which is only 3 kJ mol^{-1} higher in energy than the global minimum. Thus, the barrier to rotation about θ is very small. The planar minima and perpendicular transition states identified for the *ortho* and *para* C- linked nucleobase adducts support the conclusions in the present work that the O-linked adduct also adopts such conformations.
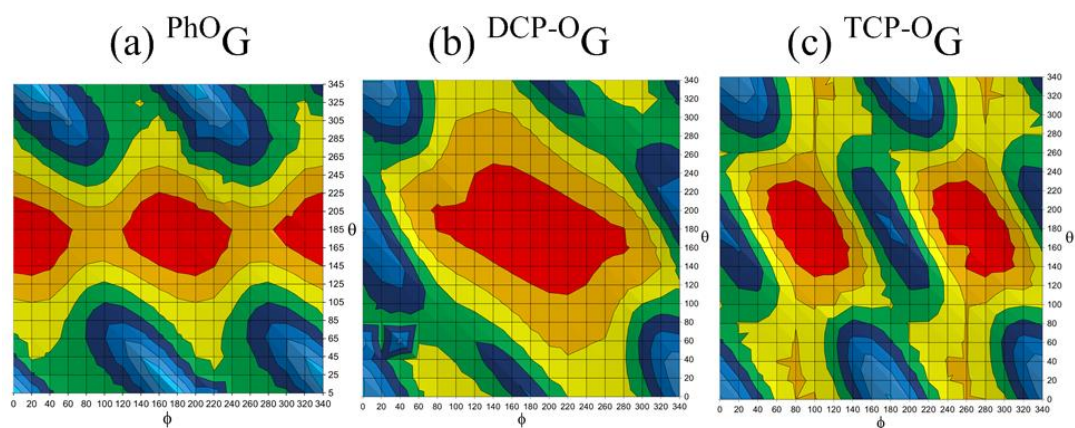
Figure 4.2. Potential energy surfaces of ^{PhO}G, $^{DCP-O}$G and $^{TCP-O}$G.

4.3.1.2. Structure of $^{4-Cl-PhO}$G

According to B3LYP/6-31G(d) full optimizations, the monochloro adduct adopts a planar geometry. Indeed, with all constraints released, the monochloro substituted adduct acquire the following geometrical parameters $\theta = 180°$ and $\phi = 180°$ (and 360°). This global minimum is connected by two identical perpendicular transition states at $\theta = 180°$ and $\phi = 95°$ (or 265°) due to molecular symmetry. Thus, the monochloro substituted adduct has one planar global minimum connected by a transition state with perpendicular ring arrangements. The barrier for rotation about θ is again very low ($\sim$4.5 kJ mol^{-1}).

The minimum identified for the monochloro adduct is comparable to the unsubstituted O-linked adduct due to the lack of intramolecular attraction between the nucleobase and the chlorine atom on the phenoxyl ring. The planar orientation and small barrier for rotation is maintained upon chlorination. Ultimately, the similar structural and conformational properties for the monochloro substituted adduct and unsubstituted adduct leads us to conclude that

the exchange of one chlorine atom at the *para* position of the phenoxyl ring does

not have a significant impact on the molecular structure (Figure 4.3).

$\theta = 180.0°;$
$\Phi = 94.8°;$
$\Delta E = 4.6$

$\theta = 180.0°;$
$\Phi = 360.0°;$
$\Delta E = 0.0$

$\theta = 180.0°;$
$\Phi = 264.7°;$
$\Delta E = 4.5$

$\theta = 180.0°;$
$\Phi = 180.0°;$
$\Delta E = 0.0$

Figure 4.3. Comparison of relative energies (kJ mol^{-1}) between selected minima and transition states of [4-Cl-PhO]G at B3LYP/6-311+G(2df,p)//B3LYP/6-31G(d) as well as θ and ϕ angles.

4.3.1.3. Structure of [DCP-O]G

The adduct [DCP-O]G has two chlorine atoms on the phenoxyl ring of the

unsubstituted mutated complex at the *ortho* and *para* positions. The B3LYP

potential energy surface was searched (Figure 4.2). A planar global minimum with

$\theta = 180°$ and $\Phi = 180°$ was identified. Geometrical properties of the fully optimized

distinguished minima, without any constraints, are consistent with this prediction

($\theta = 180°$ and $\Phi = 180°$). This shows that the global minimum for this adduct is

again planar, which directly correlates with the global minima found for both the

unsubstituted and monochloro substituted adducts. It also shows the negligible

effect of functional groups at a single *ortho* position of the aromatic ring. This is a

result of the flexibility of the phenoxyl moiety as a result of the bridging oxygen. Therefore, the dichloro mutated nucleobase adopts a conformation in which the *ortho* chlorine atom has the least interactions with the atoms of nucleobase. The transition state geometries identified from the potential energy surface has θ = ~180° and Φ = ~0°. Conversely, when the constraints on χ dihedral and sugar moiety dihedrals were released on this structure, a geometry with θ = 80.1° and Φ = 89.6° was determined. The topological properties and orientations are not consistent with the point identified on the potential energy surface. A B3LYP frequency calculation was used to get more information about the speculated transition state and revealed a single imaginary frequency at ~ −25 cm^{-1}. This frequency is too small to be assigned to a transition state. To gain more information about the nature of the optimized geometry, an IRC (Intrinsic Reaction Coordinate) calculation was run. In the input, the default step size was decreased to half the original value, and the step size was increased to 30. However, after only 5 optimization steps, the structure converged and the structures in both directions looked very similar. Additional calculations will be required to confirm the nature of this point on the potential energy surface and its connection to other minima (Figure 4.2 and Figure 4.4).

Minimum

θ= 179.8°;
Φ= 179.8°;
ΔE= 0.0

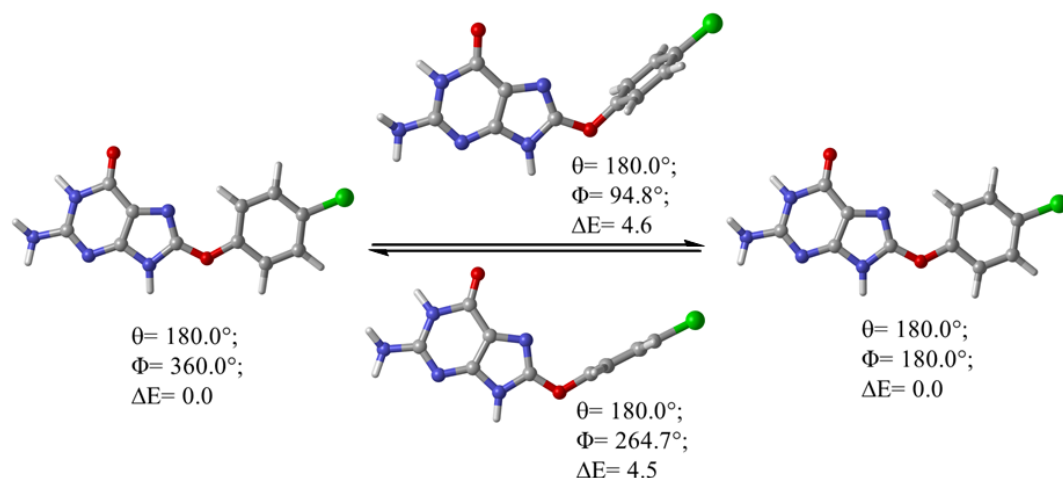Transition state

θ= 256.0°;
Φ= 89.6°;
ΔE= 30.4

Figure 4.4. Comparison of relative energies (kJ mol^{-1}) between a selected minimum and transition state of $^{DCP-O}$G at B3LYP/6-311+G(2df,p)//B3LYP/6-31G(d) as well as θ and φ angles.

4.3.1.4. Structure of $^{TCP-O}$G

Chlorination of the phenoxyl ring by less than three chlorine atoms slightly alters the geometries and energies of the unsubstituted and monochloro adducts. This is not true for the addition of the third chlorine atom at the second *ortho* position since it leads to conspicuous interactions between the halogen atom and the nucleobase. Specifically, the lone pairs of the *ortho* chlorine atom interact with the N7 site of guanine. From the B3LYP potential energy surface (Figure 4.2), two identical minima were found at θ = ~180° and Φ = ~100° (or 260°). Additionally, two identical transition states were seen at θ = ~180° and Φ = 0° (or 180°). Once all constraints on χ dihedral and sugar moiety dihedrals were released, a slight change in the geometrical properties of minima and transition states was identified. The minima fell to θ = 180° and Φ = 267.7° (or 92.3°) (**Error! Reference ource not found.**) while the transition states adopt θ = 265° and Φ = 189.3° or θ = 85.5° Φ = 9.3° (**Error! Reference source not found.**). Therefore, there are two identical minima connected by two identical transition states that are

170

approximately ~31 kJ mol^{-1} higher in energy. Thus, the θ rotational barrier found

for the trichloro adduct is substantially higher than that for both the monochloro

and unsubstituted counterparts. This increase in relative energy is likely due to the

interaction between the chlorine atoms in the *ortho* position and the nucleobase.

Since both *ortho* sites are functionalized with two chlorine atoms, the flexibility of

the phenoxyl moiety does not have a significant effect on arbitrating the steric and

electronic interactions (Figure 4.2 and Figure 4.5).

θ= 85.5°;
Φ= 9.3°;
ΔE= 30.9

θ= 180.0°;
Φ= 267.7°;
ΔE= 0.0

θ= 265.5°;
Φ= 189.3°;
ΔE= 30.9

θ= 180.0°;
Φ= 92.3°;
ΔE= 0.0

Figure 4.5. Comparison of relative energies (kJ mol^{-1}) between selected minima and transition States of $^{TCP-O}$G at B3LYP/6-311+G(2df,p)//B3LYP/6-31G(d) as well as θ and φ angles.

4.3.1.5. Structure of $^{PCP-O}$G

Functionalization of the phenoxyl moiety with two other chlorine atoms at the

meta positions forms the pentachloro adduct. To model this, we began with the

global minima (θ = 180°, Φ = 92°) and transition states (θ = 85.5°, Φ = 9.3°) that

were identified from the trichloro adduct and added two chlorine atoms. Since the

addition of these two extra chlorine atoms did not lead to additional interactions

with the nucleobase, no distinctive change was discerned in the geometrical properties of the [PCP-O]dG adduct. After full optimization, the global minimum for the pentachloro adduct was found at $\theta = 180°$ and $\Phi = 92.2°$, which is identical to the minimum found for the trichloro example. Therefore, these results indicate that the addition of the chlorine atoms in the meta position had no effect on the geometry of the minima. Similarly, with the additional chlorination, the transition states of the pentachloro adduct are not drastically different in geometry or energy from that found in the trichloro adduct. Releasing all constraints in the pentachloro transition state (Figure 4.6) gave structures with $\theta = 86.1°$ and $\Phi = 8.9°$. The rotational barrier about θ in the pentachloro structure is ~33 kJ mol^{-1}, which is only ~2–3 kJ mol^{-1} more than the barrier in the trichloro structure. It can be inferred that the pentachloro adduct has a similar geometry and relative energies as the trichloro adduct due to molecular symmetry, and negligible interaction between chloro atoms at the meta position and the nucleobase.

Minimum

$\theta = 180.0°$;
$\Phi = 92.2°$;
$\Delta E = 0.0$

Transition state

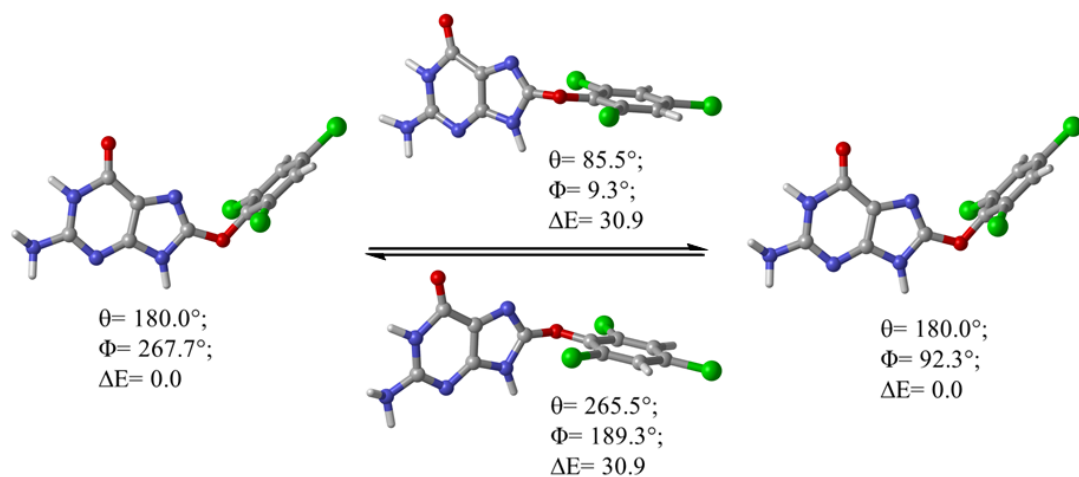$\theta = 86.1°$;
$\Phi = 8.9°$;
$\Delta E = 33.1$
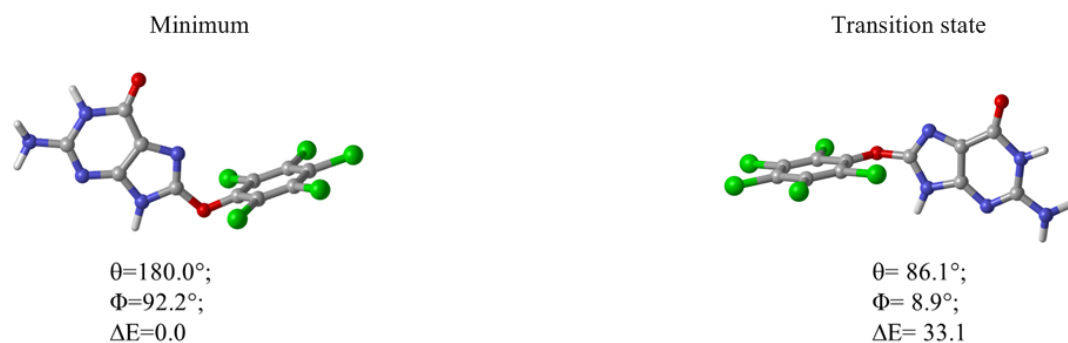
Figure 4.6. Comparison of relative energies (kJ mol^{-1}) between selected minimum and transition states of [PCP-O]G at B3LYP/6-311+G(2df,p)//B3LYP/6-31G(d) as well as θ and ϕ angles.

Consequently, analyzing both the unsubstituted and chlorosubstituted O-linked G adducts, it is apparent that the addition of chlorine atoms to the aromatic ring of the phenoxyl moiety of the unsubstituted template may generate new interactions with the nucleobase. The type and implication of this interaction depend on whether the halogen atoms are added to the *ortho*, *meta* or *para* position. *Ortho* substitution has a greater impact since there are more interactions with the atoms of the nucleobase. With increasing the number of chlorine atoms on the *ortho* positions, the barrier to rotation (change in θ) increases. The conformations adopted by the unsubstituted adduct show that, without a discernible nucleobase–Cl interaction, both θ and Φ fall to a planar minimum, while the transition states correlating to this planar minimum adopt a perpendicular orientation of phenoxyl relative to θ. A planar minimum and perpendicular transition state are also seen in the *para*-monochloro adduct due to a lack of interaction between the nucleobase and the halogen atom at the para site. Rotation of the phenoxyl about ϕ results in a specific orientation of phenoxyl in which the *ortho* chlorine of $^{DCP-O}G$ has the least interactions with the nucleobase atoms whereas the second *ortho* chlorine atom. This causes a change in the planar and perpendicular geometries in the trichloro adduct. Additional chlorine atoms at the *meta* position of the phenoxyl ring introduces no additional outstanding interactions with the nucleobase, therefore geometries and relative energies are quite similar for $^{TCP-O}G$ and $^{PCP-O}G$.

4.3.2. Nucleoside Model

Examination at the nucleoside level can provide insight into the mechanism of toxicity for adduct-forming toxins by investigating the hydrolytic stability of chloro substituted phenolic O-linked adducts. As clarified in Chapter 2, modification of dG at the C8 position is known to affect the barrier between the *syn* and *anti* conformation for the free nucleoside. Studies have shown that modification at the C8 site of 2'-deoxyguanosine often results in a favourable *syn* conformation.[25,29] Computational analyses of a class of C8-aryl-dG adducts demonstrated a ~25 kJ mol^{-1} preference for the *syn* conformer for all nucleoside adducts. This preference was rationalized due to the steric bulk of the adducted moiety as well as an intramolecular hydrogen bond between the 5′O–H and N3 site of guanosine (Figure 2.4, 2.5 and 2.6).[25]

After examining the structures of both the unsubstituted and chlorosubstituted O–linked guanine adducts, the next step should be adding the deoxyribose sugar moiety as previously employed for the ortho and para C–bonded 2′–deoxyguanosine (dG) adducts. It is assumed that the deoxyribose sugar moiety may impose a twist by further interactions that alter the geometry of the adducts and greater steric clashes. In this section a comparison is made between chlorosubstituted O–linked adducts and each of $^{o\text{-PhOH}}$dG and $^{p\text{-PhOH}}$dG and $^{\text{PhO}}$dG.

Chapter 2 clarified the structural properties of unsubstituted adduct as well as a conformational assessment of C-linked adducts at the nucleoside level.[25, 28] There

is no theoretical evidence about the lowest energy conformers of chloro substituted O-linked mutated complexes at the nucleoside level.[29] In the present study, we extend our models to represent the structural properties and conformational flexibility of some Cl-substituted O-linked mutated species ([4-Cl-PhO]dG, [DCP-O]dG, [TCP-O]dG, and [PCP-O]dG). To better demonstrate the effect of the oxygen linkage between the functionalized aromatic ring and guanosine, the impact of ϕ is considered specifically. DFT calculations using B3LYP/6-31G(d) were performed to generate PESs for each structure to provide insight into the stability, flexibility, preferred conformations and the effects of the sugar moiety on the conformational distortion of the desired nucleobases.

To understand the role of steric strain, the discussion will focus on the more stable regions of the surface, since it is anticipated that the corresponding high-energy structures do not play a significant role. The contour plot (Figure 4.7) for [4-Cl-PhO]dG is quite similar to [PhO]dG which reveals that the global minimum (red region) adopts the *syn* conformation (χ=54.12°, θ= 175.62°, ϕ=5.02°). Indeed, there is another minimum with (χ= 248.37°, θ=164.00°, ϕ=10.50°) known as local minimum that is higher in energy than the *syn* minimum. Three transition states are identified in the range of 42.66–46.30 kJ mol^{-1} with χ= 54.11°, θ=4.05° ΔE=43.90 kJ mol^{-1}, χ= 166.55°, θ=181.65° ΔE=42.66 kJ mol^{-1}, and χ= 330.40°, θ=185.28° ΔE=46.30 kJ mol^{-1}. Rotational barriers between minima can also be estimated from the graph to be 15.89 kJ mol^{-1} for conversion between *syn (anti)* minima (Figure 4.7 and Figure 4.8).

(a) ^{PhO}dG (b) $^{4-Cl-PhO}dG$ (c) $^{DCP-O}dG$

(d) $^{TCP-O}dG$ (e) $^{PCP-O}dG$
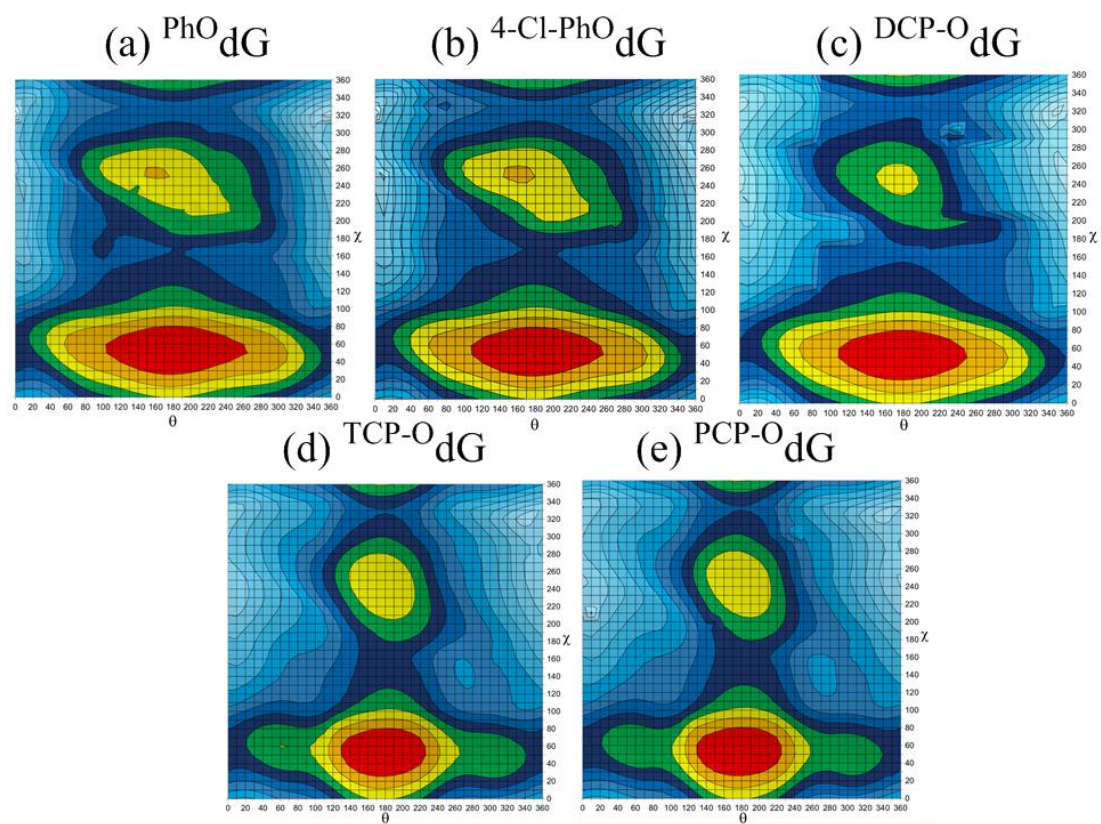
Figure 4.7. Potential energy surfaces of unsubstituted and chloro substituted O-linked adducts.

Figure 4.8. Fully optimized minima and transition states identified from PES of $^{4\text{-Cl-PhO}}$dG, different dihedrals (θ, χ, and φ deg) and relative energies at B3LYP/6-311+G(2df,p) in kJ mol^{-1}.

There is no significant difference in the stability of the global minimum in the case of $^{4\text{-Cl-PhO}}$dG in comparison with $^{\text{PhO}}$dG because addition of a chlorine atom at the farthest site will incur the least interactions with sugar segment and the other atoms on the nucleobase. Important geometrical parameters, as well as the relative energies, for fully optimized minima and transition states are provided in Figure 4.8 when full optimizations (i.e., all constraints released) are performed on important regions of contour plot (i.e. global minima, local minima and transition states), the geometries change very little. χ and θ deviate by approximately 5° in the minima and less than 10° in the transition states. The PES of $^{\text{DCP-O}}$dG illustrates

(Figure 4.7 and Figure 4.9) that one *ortho* chlorine cannot distort the structure of the new mutated complex in comparison with ^{PhO}dG and $^{4-Cl-PhO}$dG. This again identifies the significance of the conformational flexibility of O-linked adducts and the effect of oxygen as the bridge, which decrease the spatial interactions of the sugar moiety and the phenoxyl moiety, so that the global minimum is almost planer (χ= 53.08°, θ=176.07°, and ϕ=1.47°). Perpendicular transition states are found (χ= 54.50°, θ=3.50°, and ϕ=90.87°), (χ= 165.56°, θ=178.16°, and ϕ=171.20°), and (χ= 329.32°, θ=168.30°, and ϕ=158.88°) with energy barriers of ΔE = 44.48, ΔE = 45.73 and ΔE = 46.91 kJ mol^{-1}, respectively. The barrier to rotation for *syn* to *anti* conversion is ~18.14 kJ mol^{-1} which has increased as the result of the *ortho* chlorine. The *anti* and *syn* minima are connected by a transition state with χ= 244.28°, θ=173.06°, and ϕ=115.86°.

Figure 4.9. Fully optimized minima and transition states identified from PES of $^{DCP-O}$dG, different dihedrals (θ, χ, and ϕ deg) and relative energies at B3LYP/6-311+G(2df,p) in kJ mol^{-1}.

More significant deviations in structural properties from unsubstituted and mono chloro substituted adducts occur in some instances, which will be discussed in more detail below. The sugar moiety and functionalization of the aromatic ring have more impact on the geometrical properties of $^{TCP-O}$dG, and particularly $^{PCP-O}$dG. In the case of $^{TCP-O}$dG, the second chlorine atom at the *ortho* site of aromatic ring, as well as steric crowding which the phenoxyl ring passes over the sugar

179

moiety by adopting a twisted conformation, induces an obvious distortion. The *syn* conformer is the lowest energy minimum for $^{TCP-O}dG$ with geometrical parameters of χ= 53.26°, θ=178.92°, and ϕ=93.09°. The *anti* local minimum has interactions between the C5′–OH and the oxygen atom of phenoxyl (χ= 247.19°, θ=175.06°, and ϕ=96.73°). These interactions lead to steric crowding of the *ortho* chlorine atom which is partially alleviated by inducing a greater degree of twist in the molecule and less stability (14.25 kJ mol^{-1}) in comparison with the global minimum. Figure 4.7 and Figure 4.10 show the energy minima (global and local minima) and the transition states which connect these structures. Two transition states with energy barriers of 40.09 kJ mol^{-1} and 43.68 kJ mol^{-1} connect the global and local minima. Another transition state with the energy barrier of 43.62 kJ mol^{-1} connects the global minimum to itself.
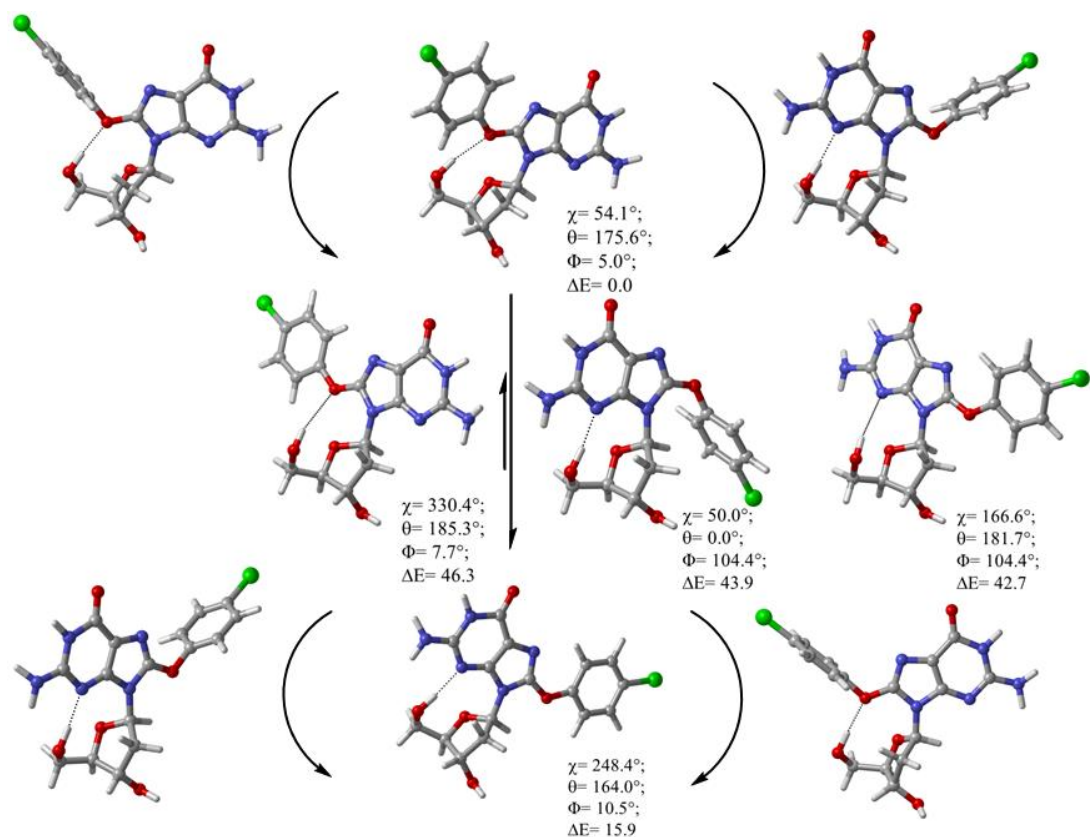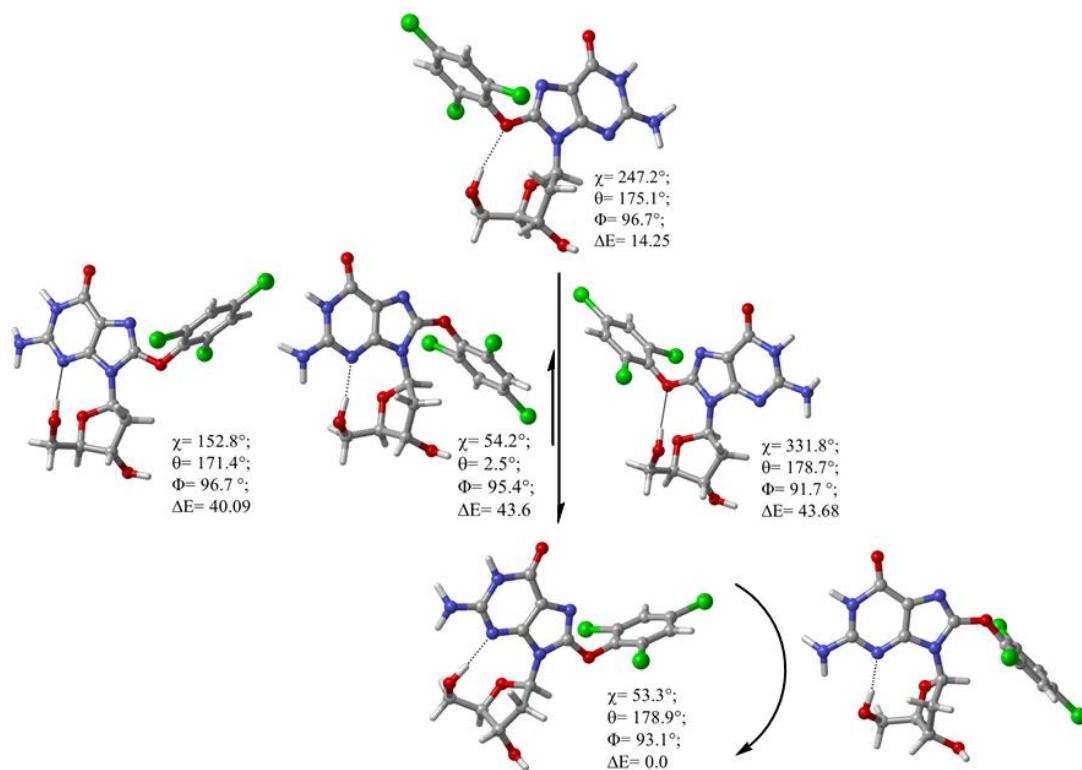
Figure 4.10. Fully optimized minima and transition states identified from PES of $^{TCP\text{-}O}$dG, different dihedrals (θ, χ, and φ deg) and relative energies at B3LYP/6-311+G(2df,p) in kJ mol^{-1}.

Functionalization of the aromatic ring at the *meta* position did not lead to any conspicuous significant interaction, so they did not affect the lowest energy conformations significantly in the global minimum (χ= 53.50°, θ=177.39°, and φ=89.54°), and local minimum (χ= 246.51°, θ=175.89°, and φ=90.69° ΔE=14.45 kJ mol^{-1}). The phenoxyl substituent is twisted with respect to the nucleobase by approximately 90.69° and this conformer exhibits a perpendicular orientation of the aromatic ring (Figure 4.11). Despite this twisting, both global and local minima of the mutated nucleoside contain an O5′−H•••N7 hydrogen bond (1.83 Å), and O5′-H•••O (oxygen atom of phenoxyl moiety) hydrogen bond (2.21 Å). Barriers to rotation from the global minimum to the local minima were predicted to be around

14.45 kJ mol^{-1}, indicating that conversion of the global minimum to local minimum might be experimentally difficult. The sugar puckering generally remains C2'-*endo* in all the chloro substituted O–linked adducts; this result was obtained from HyperChem conformational search without applying any constraint on sugar moiety of nucleoside model.
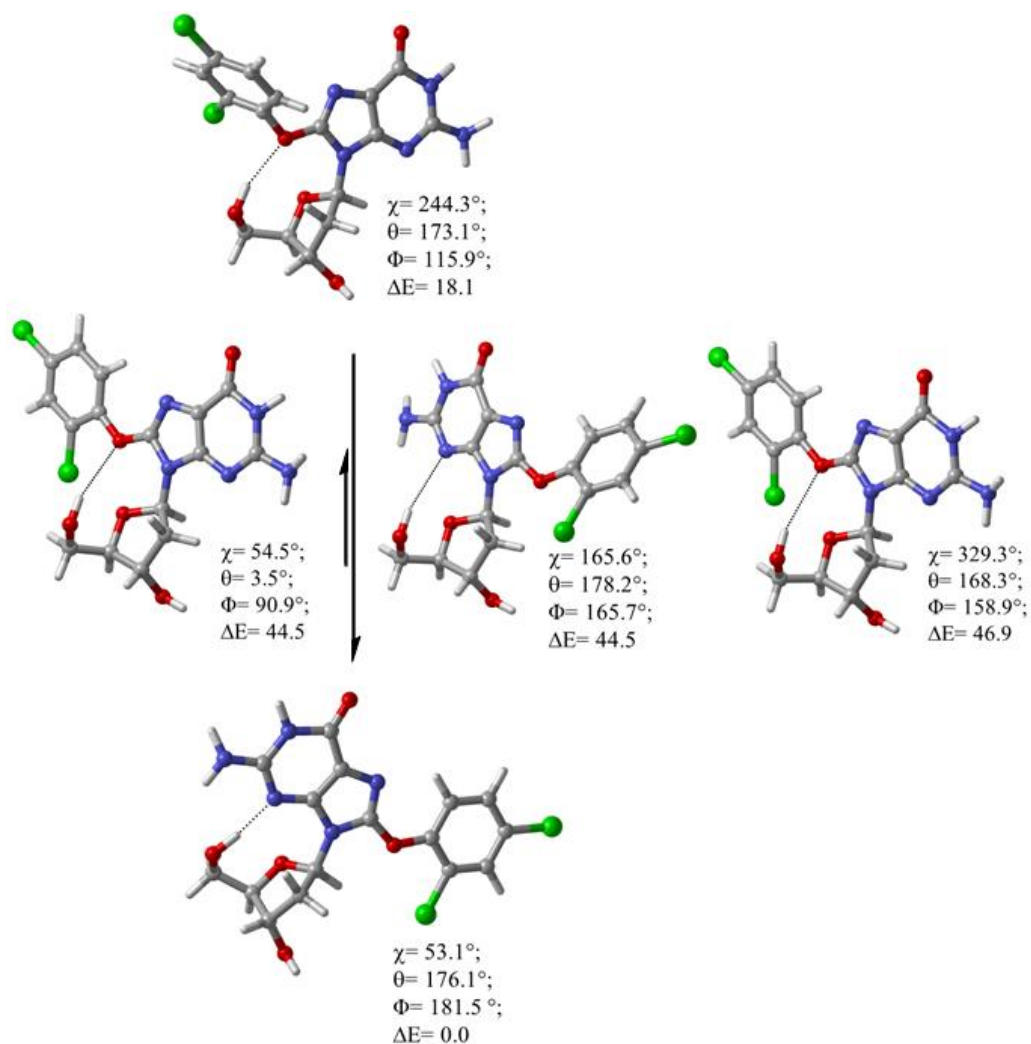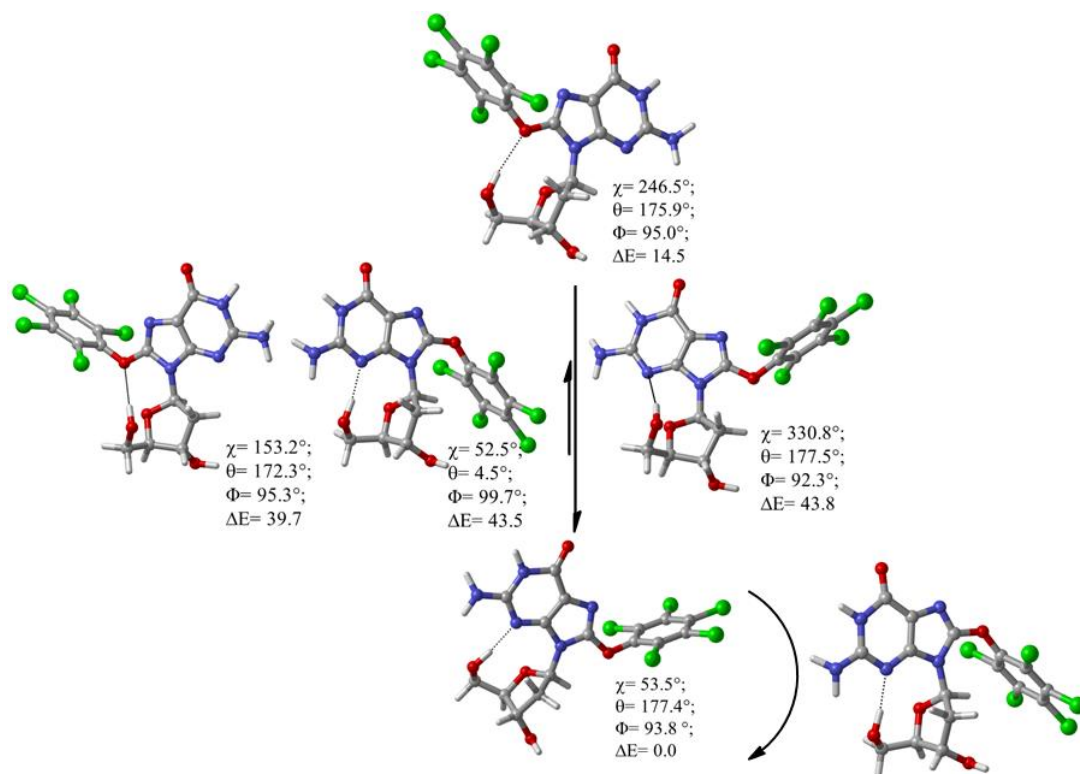

Figure 4.11. Fully optimized minima and transition states identified from PES of $^{PCP-O}$dG, different dihedrals (θ, χ, and ϕ deg) and relative energies at B3LYP/6-311+G(2df,p) in kJ mol^{-1}.

The barrier to rotation between *anti* and *syn* conformers is quite similar for unsubstituted and monochloro mutated complexes whereas there is a sharp decrease in stability for the local minima of the dichloro species since the global minimum is planar and local minimum is perpendicular. This barrier includes the required energy for deviation from the planar orientation of the phenoxyl moiety

182

to the perpendicular alignment. There is a reduction in this barrier for $^{TCP-O}$dG and $^{PCP-O}$dG in comparison with $^{Ph-O}$dG and $^{4-Cl-Ph-O}$dG. This can be rationalized by an initial perpendicular orientation of the phenoxyl moiety in the global minima of both tri- and pentachloro adducts that remains unchanged in the local minima (*anti* conformer). The stability of the local minimum is comparable with the global minimum in $^{TCP-O}$dG and $^{PCP-O}$dG.

Although the PESs have been generated for all Cl substituted O-linked adducts, the creation of PESs for the monochloro and pentachloro adducts was deemed unnecessary since all previous structures adopted a planar (θ) 180° orientation which indicated that the extra chlorine atoms would do not have a conspicuous interaction with the nucleobase or sugar moiety.

After determining the optimal orientations about the glycosidic and C8–O bonds, other PES scans were performed with respect to the χ and ϕ bonds for all unsubstituted and chloro substituted adducts in an effort to determine the possible dependence of these dihedral angles on each other. Indeed, there was no significant change in the geometrical properties of mutated complexes. Comparison of the global minima for the nucleobase and nucleoside models allowed us to determine the effect of the deoxyribose moiety on the structure of the adducts. In all nucleosides, the *syn* conformer of the base is stabilized by an intramolecular O5′–H•••N3 hydrogen bond and the *anti* conformer of the base is stabilized by an intramolecular O5′–H•••O phenoxyl hydrogen bond (Figure 4.12).
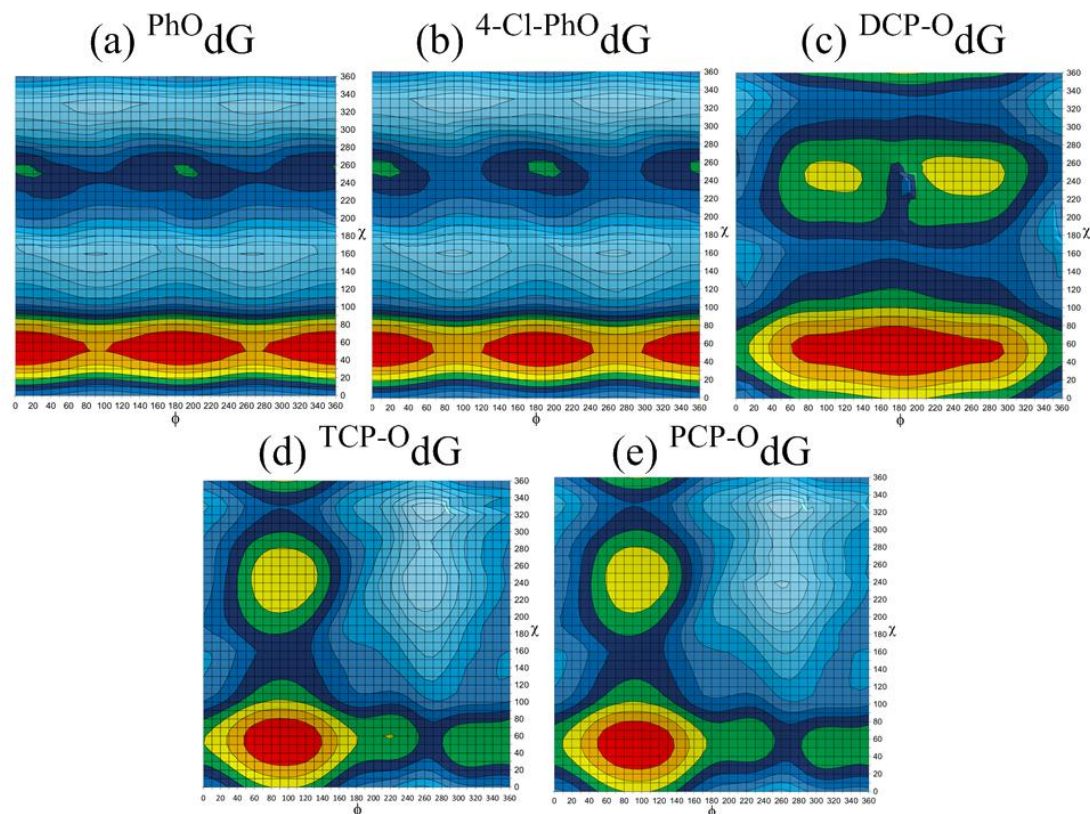
Figure 4.12. B3LYP/6-31G(d) ϕ versus χ potential energy surface for unsubstituted and chloro substituted; The relative energy (kJ mol^{-1}) is represented by color, where the lowest energy regions are red, and each change in color represents a 5 kJ mol^{-1} increase in the relative energy.

4.3.3. Proton Affinity (PA)

Protonation reactions, i.e., A + H$^+$ → AH$^+$, are among the most important reactions in chemistry and biology. Many fundamental chemical rearrangements in most enzymatic reactions are initiated with protonation or deprotonation. The ability of a molecule to accept a proton in the gas phase can be determined by two quantities. The first is the gas-phase basicity, which is the negative of the free energy change associated with the protonation reaction. The other and more

frequently used index is the proton affinity defined as the negative of the enthalpy change under standard conditions. Experimental determination of these parameters is not easy, so with the recent phenomenal growth in computer power, a growing attention has been given to the possibility of calculating these parameters by quantum methods. *Ab initio* approaches are very successful in providing reliable values of proton affinities and gas-phase basicities for small molecules even at lower levels of theory. However, the computational expense of this method limits the application of *ab initio* methods to estimate the proton affinities and it is still impractical for larger molecules. On the other hand, semiempirical methods such as AM1, MNDO and PM3, are not consistently reliable in calculations of proton affnities. DFT methods are more prevalent since, in principle, they include electron correlation energy.[27, 31-37]

It is well-known that the formation of positive charge on N7 of natural dG by protonation or methylation, which lead to formation of four covalent bonds on nitrogen atom, facilitates the rate of hydrolysis. It has been identified that the acid-catalyzed hydrolysis of dG involves equilibrium protonation; and therefore, monomolecular cleavage of the C-N9 bond is the rate limiting step. [21, 22, 38]

DFT calculations were used to determine structures of the protonated species (Figure 4.13) in the gas phase and solvent phase (water). Proton affinities (PA, kcal·mol^{-1}) for both N7H$^+$ and N3H$^+$ adducts are shown in Table 4.1. The *anti* conformation is the most stable structure for both the neutral and N7H$^+$ species of dG, however the *syn* conformation is favored for the N3H$^+$ species due to an H-bonding interaction between 5′-O of the sugar moiety and hydrogen at the N3 site

of the nucleobase (Table 4.1). The same conformer is identified for all the chloro substituted N7 protonated species (*syn*-skew) with the difference that there is a stabilizing hydrogen bond between O5′-OH and N3. As previously explained, neutral O-linked 8-dG adducts prefer the *syn* conformer (Figure 4.13, Table 4.1). In all protonated and neutral structures, adopting the *syn* conformation reduces steric interactions between the phenolic ring and sugar moiety. Hydrogen bonding between 5′-OH and N3 is a stabilizing interaction (Figure 4.13).

The conformation around the diaryl ether bond is governed by both the substituent electronic effect on the π-electronic system and steric effects. An electron-withdrawing substituent at the *para* position of a benzene ring leads to the formation of a stable skew conformation ($\theta 1 = 90°$ and $\theta 1 = 0°$) while derivatives with an electron-donating substituent prefer the twist or skew form in which the C-O bond to ring B is planar (Scheme 4.3). The preferable skew form of diphenyl ethers with an electron-withdrawing substituent is stabilized by interaction between the bridging oxygen π-lone pair and the substituent through ring B. A competition between the substituent and the ether oxygen for an intramolecular charge-transfer interaction allowed the twist or skew ($\theta 1 = 0°$, $\theta 2 = 90°$) conformation. In the case of alkyl substituents at *para* position of aromatic ring, their electronic effect is insufficient to affect the conformation of diaryl ethers.[39]
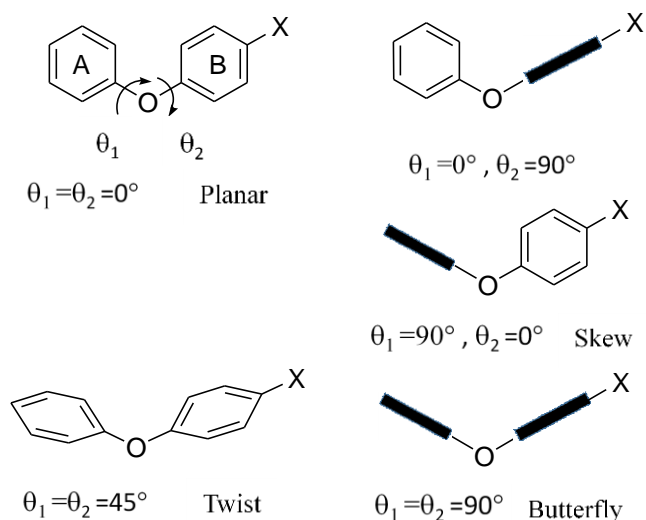
Scheme 4.3. Typical conformations for 4-Substituted diphenyl ethers.[39]

It has been mentioned that neutral [Ph-O]dG, [4-Cl-Ph-O]dG and [DCP-O]dG were present in a planar conformation, while neutral [TCP-O]dG and [PCP-O]dG (Figure 4.4, Figure 4.5, Figure 4.6, and Figure 4.7) (Table 4.1) adopted a skew conformation to diminish interactions between the dG moiety and the TCP and PCP ring system. To explore another aspect of the damaged structures, each adduct was protonated at the N3 and N7 site. Addition of a proton to the minimum structure would allow us to identify whether the adducts are stable in acidic environments or whether deglycosylation occurs. To evaluate the effect of solvent on the stability of charged nucleosides, PCM computations were carried out in water. As anticipated, protonation at N7 and N3 of all O-linked 8-dG adducts leads to a preferred skew conformation attributable to steric and electronic interactions between the added proton and substituted phenoxyl (Table 4.1 and **Error! Reference source not ound.**). In the gas-phase the calculated N7 PA for dG (non-adducted) is 232.1 kcal mol^{-1} (Table 4.1) consistent with the experimental PA (234.4 kcal mol^{-1}).[40] This is

~15.5 kcal mol^{-1} above the gas-phase N3 PA which can be explained by alteration of the *anti* conformer of dG to *syn* conformer by protonation at N3 site. These results illustrate that dG is the only base that undergoes a conformational change (*anti → syn*) upon N3-protonation.

Protonated unsubstituted adduct (N3H$^+$) has a gas-phase PA of 219.8 kcal mol^{-1} which increases to 257.8 kcal mol^{-1} in water. The same trend is seen for the N7H$^+$ adduct with PA values of 227.8 kcal mol^{-1} and 259.6 kcal·mol^{-1} in the gas and solvent phase, respectively. Addition of one chlorine changed the N3H$^+$ PA to 217.9 kcal mol^{-1} in gas phase and 257.4 kcal mol^{-1} in water. The N7 PA is 225.2 and 258.8 kcal mol^{-1} in gas phase and solvent, respectively. According to Table 4.1 chlorine substitution of the phenoxyl moiety reduces the electron density on N3 and N7 of the nucleobase in the case of di-, tri- and pentachloro substituted adducts. The calculated proton affinities for N3 and N7 of $^{DCP-O}$dG (which include an *ortho* chlorine atom) are 219.1 and 223.9 in the gas phase and 259.1 and 258.0 kcal mol^{-1} in solvent phase, respectively (Table 4.1).

In comparison to the dichloro mutated complex, the second *ortho* chlorine of $^{TCP-O}$dG did not have any significant effect on the N3H$^+$ PA in the gas phase. PAs of N3 (water) and N7 (water) and N7 (in the gas phase) adducts reduced to 257.4, 223.2 and 254.8 kcal·mol^{-1}, respectively. However, the computational values show that in $^{PCP-O}$dG, there is a reduction in N7 PA in the gas phase and solvent with values equal to 257.4 and 218.0 kcal mol^{-1}, respectively, while the PA of the N3 adduct in water did not change. The decrease in the N7 PA in the gas and water

phase upon increasing the number of Cl atoms on the phenolic ring shows the expected inductive effect of chlorine (Table 4.1).

The DFT calculations show that the most favorable site for protonation is N7 in water which is attributed to higher PA values in comparison with the gas phase as well as N3 protonated counterparts in solvent and gas phase. There is only a ~5 kcal·mol⁻¹ energy difference between the PA at N7 and N3 sites in water for natural guanosine (Table 4.1). Attachment of the phenoxyl substituent to C8 of dG to form ^PhO dG reduces the gas-phase N7 PA by 4.3 kcal mol⁻¹, and enhances N3 PA by 3.2 kcal mol⁻¹. The N7 proton affinity in the gas and water phase indicates that protonation at N7 is favored. The PA for N7 of ^PhO dG is 8 kcal·mol⁻¹ higher than N3 of the same analogues. In water, the preference of protonation at the N7 site is only 1.8 kcal·mol⁻¹. Calculations indicate attachment of Cl-substituents to the phenyl ring further diminishes the N7 PA, giving a PA of 221.3 kcal mol⁻¹ for ^PCP-O dG in the gas-phase which is 6.5 kcal mol⁻¹ less than the N7 PA for ^PhO dG in the same phase. In contrast, functionalization of the aromatic moiety with chlorine did not affect the PA at N3 site of guanosine in the gas phase, which is just around 1.4−3.2 kcal·mol⁻¹ more than the N3 PA calculated for dG.

The effect of solvent has been considered by performing calculations in water which demystifies that the PA at N3 is greater than N7 site especially for ^DCP-O dG, ^TCP-O dG and ^PCP-O dG. Thus, it can be concluded that the protonation is preferred at the N3 site rather than at N7. This can affect abasic site formation. There were no changes in conformations, all of them being *syn*-skew structures because of the steric effects of the proton at N7. It is assumed that the substitution of the other

chlorine atoms at *ortho* position on the aromatic ring of phenoxyl moiety have caused this big drop in PA from ^{DCP-O}dG to ^{TCP-O}dG. Chlorine attached to phenyl exerts an inductive effect but the resonance effect is small. The inductive effects of *ortho* Cl will be greater than *meta* and *para* due to proximity. So the drop in PA was not notable in ^{PCP-O}dG.

For the neutral species ^{PhO}dG, ^{4-Cl-PhO}dG, ^{DCP-O}dG, ^{TCP-O}dG, all adopt *syn*-planar conformers, while ^{TCP-O}dG and ^{PCP-O}dG acquired *syn*-skew conformers because of the steric effects of the second *ortho* chlorine.

All N3 protonated structures are more folded than the N7 protonated structures as a result of the formation of new hydrogen bonds between O5′ hydroxyl and the inserted proton at N3 site of adducted guanosine, as well as *steric interactions between proton at N7 site and chlorine atoms on aromatic ring,* which leads to more steric interactions. All of the protonated adducts adopt *syn*-skew conformers except protonated dG. In summary, DFT calculations highlight that when water is the solvent, modification of the C8 site of the guanosine with highly chlorinated phenoxyl moiety significantly affects the basicity of N7 and results in favorability of protonation at N3 site. The computational values are in accordance with experimental findings.

There are different effects that account for the identified trends for unsubstituted, and chloro substituted species; the effect solvent as well as site of protonation (N3 and N7 protonation). Furthermore, contribution of

190

intramolecular hydrogen bonds in *syn* and *anti* conformers (steric effects) can induce distortion in mutated complexes.

Calculations demonstrated that the PA strongly depends on the site of protonation on the nucleoside as well as functional groups on the aromatic ring of phenoxyl. The protonated structures are stabilized through resonance, hydrogen bonding and electronic interactions with chlorine(s) on the phenoxyl.
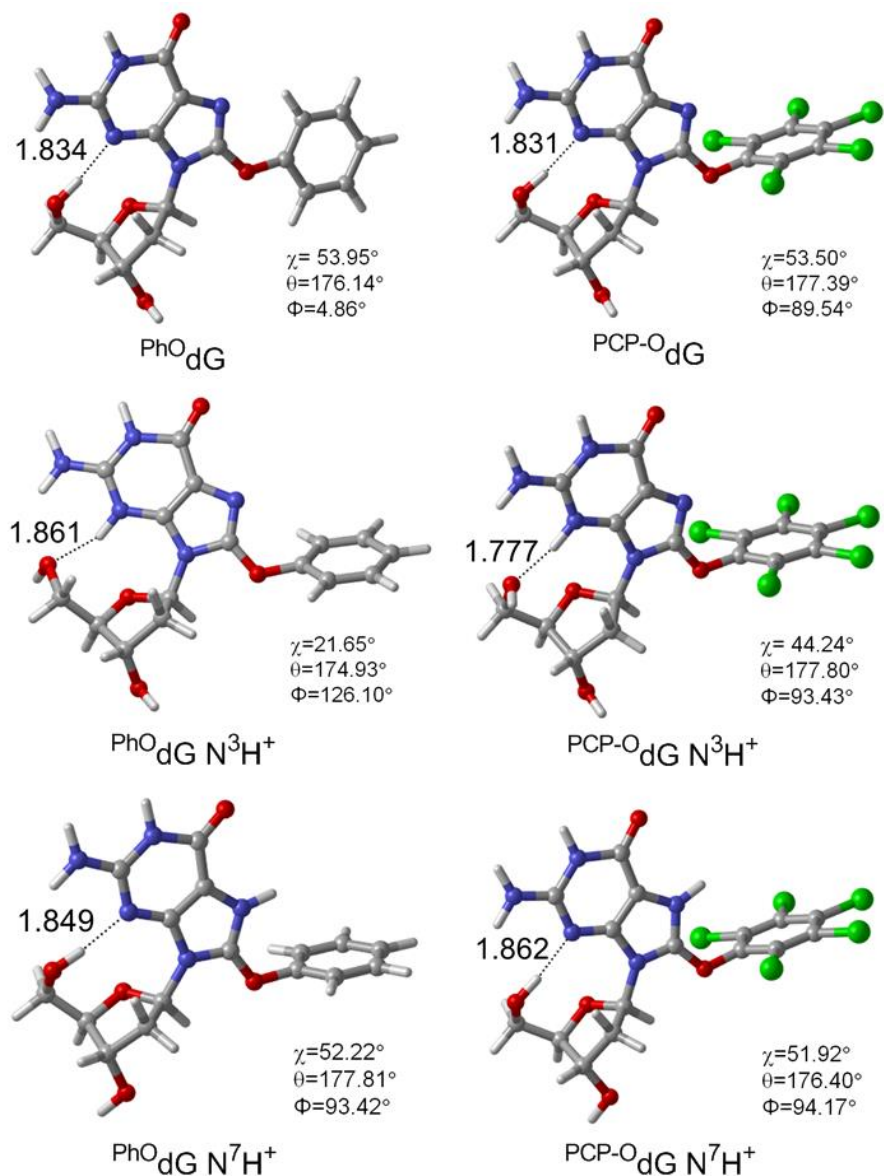
Figure 4.13. The most stable B3LYP/6-311+G(2df,p)//B3LYP/6-31G(d) conformers for the [PhO]dG and [PCP-O]dG adducts, as well as their N3-(N3H+) and N7-(N7H+) protonated analogues (Select hydrogen bond lengths (Å) are provided).

Table 4.1. Low-Energy Conformations[a] and Proton Affinities (PA, kcal mol⁻¹) at the N3 and N7 Site of O-linked C8-dG adducts.[b]

Adduct	Neutral	N⁷H⁺	N³H⁺	N³ PA$_{(gas)}$	N³ PA$_{(Water)}$	N⁷ PA$_{(gas)}$	N⁷ PA$_{(Water)}$
PhO-dG	*syn*-planar	*syn*-skew	*syn*-skew	219.8	257.8	227.8	259.6
4Cl-PhO-dG	*syn*-planar	*syn*-skew	*syn*-skew	217.9	257.4	225.2	258.8
DCP-O-dG	*syn*-planar	*syn*-skew	*syn*-skew	219.1	259.1	223.9	258.0
TCP-O-dG	*syn*-skew	*syn*-skew	*syn*-skew	219.1	257.4	223.2	254.8
PCP-O-dG	*syn*- skew	*syn*- skew	*syn*- skew	218.0	257.4	221.3	253.8
dG	*anti*	*anti*	*syn*	216.6	255.7	232.1	260.8

[a]The most stable B3LYP/6-311+G(2df,p)//B3LYP/6-31G(d) conformers for the neutral, N⁷-(N⁷H⁺) and N³-(N³H⁺) protonated analogs. [b]N³ and N⁷ proton affinity (PA) was calculated as the negative of the enthalpy change for protonation (kcal mol⁻¹).

4.3.4. Deglycosylation and the influence of the protonation site on deglycosylation

The hydrolytic stability of O-linked phenolic C8-dG adducts is of great interest which we address in this section. We focused on adducts that carry electron-deficient polychlorinated phenoxyl ring systems, because they are expected to have high susceptibility to hydrolysis. This high susceptibility is due to stabilization of the negative charge at N9 site of guanosine, which is caused by presence of the electron deficient groups of chlorine. We studied the properties of O-linked C8-dG adducts as well as N3 and N7 protonated structures computationally and compared them with previous results on hydrolysis rates of dG and the corresponding C– and N–linked 8–dG adducts.[18, 41, 42]

Electron-withdrawing substituents at the para position of the attached phenyl ring caused the greatest increase the rate of hydrolysis relative to dG. As mentioned before, relief of steric strain upon removal of the deoxyribose sugar moiety is one of the main reasons of increasing the rate of hydrolysis. Another prevailing factor is stabilization of the developing negative charge at N9 by electron-withdrawing para-substituents.[23]

4.4. Acidic Hydrolysis

The N7 nitrogen is often targeted by electrophiles as it has the most electron density of all atoms in deoxyguanosine.[2] Bond formation at this site consigns a formal positive charge on the nitrogen. The formation of this formal positive charge greatly accelerates the rate of glycosidic bond cleavage during the depurination process. It is well known that hydrolysis of dG under acidic conditions occurs through a two-step process which is kinetically pseudo-first order (Figure 4.1).[21, 43, 44] This mechanism can be expanded for depurination, which is the most common type of damage to the nucleobase (in single or double stranded DNA). Step 1 is a preequilibrium which involves the protonation at N7 (N7H$^+$dG denoted by K_{a1}). Step 2 is the unimolecular rate-limiting S_N1-type dissociation of the base from the sugar moiety, represented by k_1, followed by the cleavage of the glycosidic bond.[21]

Overall, acidic conditions enhance the favorability of depurination since the heterolytic breakage of the glycosidic bond leads to the neutralization of the charge at N7. In the other step, the short lived (10^{-12} to 10^{-11} s)[21] charged oxocarbenium

ion is created and readily undergoes reaction with water. It is well known that the presence of the positive charge on the nucleobase makes it a better leaving group than the neutral base.[45]

Modification of dG at the C8 site can also affect the hydrolytic stability. Increases in the rate of hydrolysis by electron withdrawing groups at the C8 position has been observed.[16] Chlorine atoms can be used as electron withdrawing group; thus, functionalization of the phenol ring with chlorine enhances favorability of the O-linked adduct formation. This enhancement of O-linked adduct formation preference occurs by increasing electrophilicity of the phenolic radical intermediate and reducing the rate of bimolecular phenolic radical coupling.[10] It is known that attachment of electron withdrawing groups to the 8-position of dG can increase rates of hydrolysis to afford abasic sites.

DFT calculations predict a greater N3 PA than N7 PA for [DCP-O]dG, [TCP-O]dG and [PCP-O]dG in water (Table 4.1). **Error! Reference source not found.** illustrates the alculated deglycosylation profile for neutral [Ph]OdG (purple trace), [4-Cl-Ph-O]dG (blue trace), [DCP-O]dG (green trace), [TCP-O]dG (orange trace) and [PCP-O]dG (red trace), as well as the corresponding N3- and N7-protonated species (kJ mol^{-1}). The theoretical results show that in both water and the gas phase, N7-protonation has a more conspicuous effect on the barrier for deglycosylation than does N3-protonation. This can be attributed to the direct effect of the protonation and positive charge at N7 site on the N9–C1′ glycosidic bond.[16] In water (Figure 4.14), the deglycosylation barrier for the N7-protonated O-linked adducts is in the range of ~60−70 kJ mol^{-1},

while the barrier for deglycosylation in N3- protonated adducts is ~100 kJ mol^{-1}. There is a further increase to ~120 kJ mol^{-1} in the case of the neutral adducts in water. So we have three groups: neutral, N3H$^+$ and N7H$^+$ species, and each group specifically consist of five types of adducts (^{PhO}dG, $^{4\text{-}Cl\text{-}PhO}$dG, $^{DCP\text{-}O}$dG, $^{TCP\text{-}O}$dG, $^{PCP\text{-}O}$dG). In all the groups it has been found computationally that $^{PCP\text{-}O}$dG has the lowest barrier, followed by $^{TCP\text{-}O}$dG, $^{DCP\text{-}O}$dG, $^{4\text{-}Cl\text{-}PhO}$dG and finally ^{PhO}dG. This trend shows the impact of the electron-withdrawing chlorine as a substituent on phenoxyl. It is worth to mention that as a general rule in organic chemistry, the electron-withdrawing substituents to the C8-position of dG increases the rate of depurination through stabilization of the developing negative charge at N9 during rate-limiting step which is the cleavage of the glycosydic bond. For neutral dG there is negative charge development in the ionization transition state that is stabilized by electron withdrawing groups. For the cations, the ionization step involves a loss of positive charge. This is facilitated by electron-withdrawing groups because they destabilize the cation starting material more than they destabilize the TS (which has less positive charge).

To provide a better understanding of the phenomena which occur during deglycosylation, the process was compared in solvent (water) and the gas phase. In the gas phase (Figure 4.14**Error! Reference source not found.**), the eglycosylation barrier is slightly diminished for the N7-protonated adducts versus the corresponding barrier in water, while the barrier for the N3-protonated adducts is increased sharply in the gas phase and becomes comparable to the neutral adducts. It has been found that there was no change in the trend for chloro

substituted adducts in solvent and the gas phase. In essence, the quantum chemical data imply that the chlorophenoxyl substituents can play a determining role on the site of protonation in that N3H$^+$ adducts are relatively more stable to hydrolysis compared to their N7H$^+$ analogues.

We probed the deglycosylation barriers in the gas phase by DFT calculations which suggested $^{PCP-O}$dG to be the most reactive. In the DFT calculations, barriers were determined for the monoprotonated species (Figure 4.14) showing that the monoprotonated adduct was the least reactive.

Regardless of the exact mechanism, removal of the sugar moiety from the phenoxy C8-substituted O-linked adducts at 1'-N9 can proceed. Despite the fact that it is unlikely that steric factors resulting from chlorine substitution affect the deglycosylation rate, it has been identified that electronic factors play a main role. A distinct justification for the resistance to hydrolysis emerges from the impact of the phenoxyl substituent to diminish N7 proton affinity in comparison with natural guanosine. However, DFT calculations imply that the barrier to deglycosylation becomes progressively smaller for neutral, N3H$^+$, and N7H$^+$ adducts bearing increased numbers of Cl-substituents. This observation suggests that the trend for hydrolysis of 8-dG adducts stems from the electron withdrawing properties of the C8-substituents that accelerate the rate of depurination through stabilization of the developing negative charge at N9 during rate-limiting cleavage of the glycosyl bond. DFT calculations suggest that N3-protonation may compete effectively with N7-protonation for both $^{TCP-O}$dG and $^{PCP-O}$dG. Also, the monoprotonated N3H$^+$

species possess much greater barriers to deglycosylation than their N7 protonated counterparts. This is because of the direct effect of the positive charge at N3 stabilizing the negative charge at N9 atom after removing the sugar moiety.
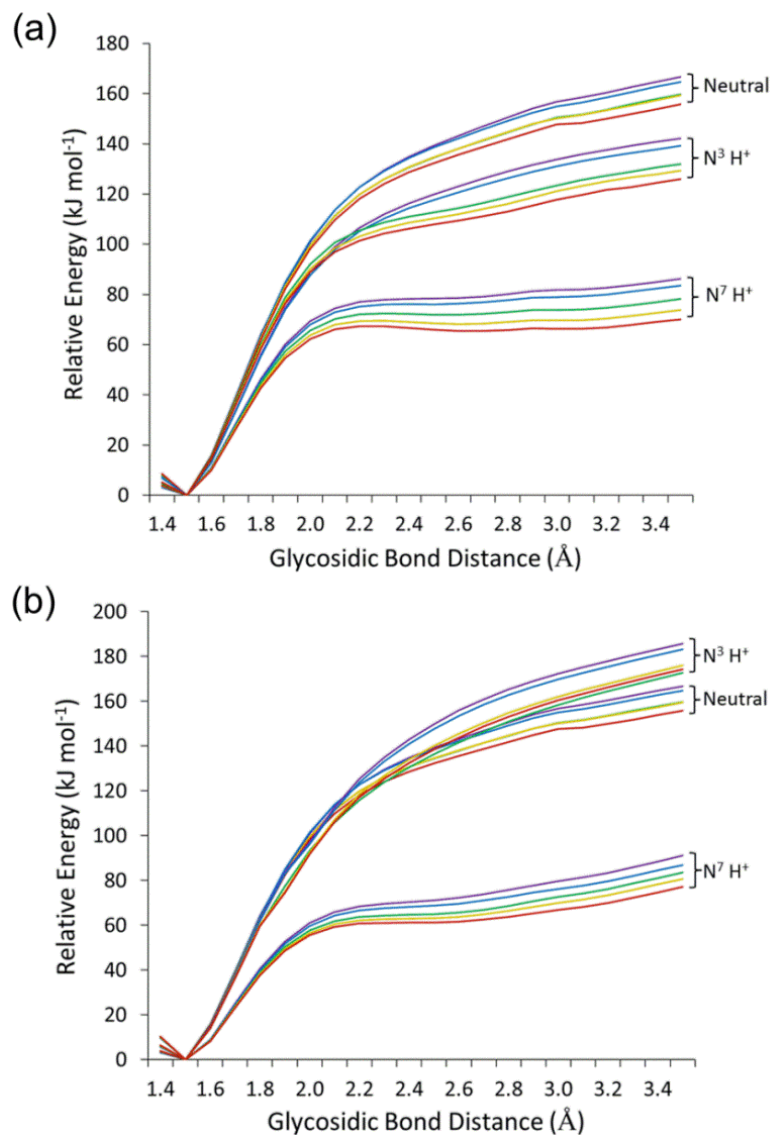
Figure 4.14. Constrained IEF-PCM-B3LYP/6-31G(d) deglycosylation barriers calculated in (a) water and (b) Fthe gas phase for PhOdG (purple), 4Cl-PhOdG (blue), DCP-OdG (green), TCP-OdG (orange) and PCP-OdG (red), as well as the corresponding N3 (N3 H+) and N7 (N7 H+) protonated species (kJ mol-1).

4.4.1. Nucleotide

Studies of the nucleoside and nucleotide adducts will aid in establishing a protocol for using small models to predict the preferred conformation of bulky damaged bases in physiological environments. To better predict the structure of the damaged bases in DNA helices, a nucleotide model that includes the 5′-monophosphate group was studied.[28, 29] However, as discussed for the ^{PhO}dG, $^{o\text{-}PhOH}$dG, and $^{p\text{-}PhOH}$dG nucleotide adducts, the 5'-monophosphate group was added to the lowest energy *syn* and *anti* conformations, determined by the (β-constrained) nucleoside model, to generate the nucleotide model.

The importance of including diffuse functions is clear due to the presence of negative charge in these structures. Ultimately, minimal differences between the geometries for a specific adduct are obtained regardless of the basis set implemented so the basis set without diffuse functions was used for simplicity.

The *anti/syn* conformational preference of the $^{o\text{-}PhOH}$dG, $^{p\text{-}PhOH}$dG and ^{PhO}dG nucleotide adducts was investigated in Chapter 2 of this thesis. Nucleoside models predict the *syn* conformation to be favored due to the presence of a C5'-H•••N3 hydrogen bond, which cannot occur in DNA. When an additional geometric constraint is imposed on the C5′-hydroxyl group that prevents this interaction, the *syn* conformation becomes less important than initially predicted. Furthermore, when this constraint is released in subsequent optimizations, the resulting *anti* structures are nearly thermoneutral with the original *syn* global minimum. Hence, even though the nucleoside model has certain advantages and can be implemented

to get some valuable information about the structure of damaged bases, it is not giving a comprehensive picture the *anti/syn* conformational preference of modified nucleobases in DNA. To overcome this deficiency nucleotide model that includes the 5′-monophosphate group was used.

This methodology was subsequently applied to the damaged nucleotides to determine whether the [o-PhOH]dG and [p-PhOH]dG mono phosphate adducts share the natural preference for *anti* or whether they adopt a *syn* conformation, which would suggest a greater potential for mutagenicity. Although various types of damage show a weak preference for the *anti* conformation, the [o-PhOH]dG monophosphate adduct has a stronger preference for the *anti* conformation than the natural nucleotide. The smallest *anti/syn* energy difference in the adducts is 2.5 kJmol^{-1}, observed for [DCP-O]dG mono phosphate which preferred the *syn* conformer. This suggests that the preferred structure of bulky nucleotide adducts in DNA may depend on small structural differences (i.e., the location of the chlorine or other functional group on the phenyl group) (Figure 4.15).

The same approach was used to explore the conformers at the nucleotide level of all chloro substituted O-linked structures ([PhO]dG, [4-Cl-PhO]dG, [DCP-O]dG, [TCP-O]dG and [PCP-O]dG). It has been identified that all chloro O-linked adducts prefer *syn* conformers relative to *anti* conformers at the nucleotide level, except [TCP-O]dG and [PCP-O]dG which adopt *anti* conformations at the nucleotide level with energy differences of 33.0 and 0.7 kJ mol^{-1} relative to the *syn* conformer, respectively.
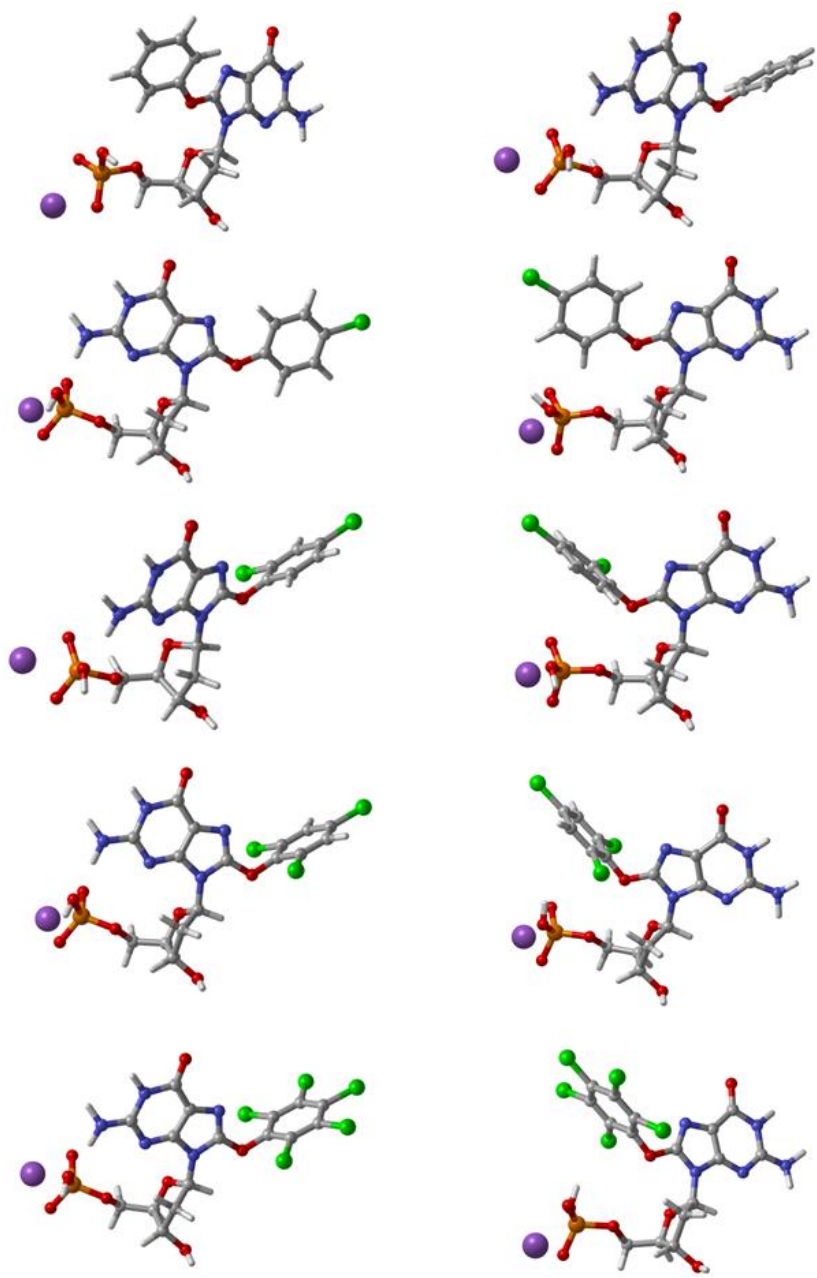
Figure 4.15. The biologically relevant *anti* (right) and *syn* (left) structures optimized in water for natural 2′- deoxyguanosine 5′-monophosphate (described by the counterion Na+ HPO4-) of unsubstituted and all chlorosubstituted structures.

Table 4.2. Dihedral angles (χ and ϕ deg.), and B3LYP/6-311+G(2df,p) relative energies (kJ mol^{-1}) for the biologically relevant unsubstituted and all chlorosubstituted structures.

Adducts	Conformation	χ	ϕ	E_{rel}
PhOdG	*syn*	66.9	60.3	0.0
	anti	258.8	150.3	6.3
4-Cl-PhOdG	*syn*	67.3	186.5	0.0
	anti	250.8	78.8	7.7
DCP-OdG	*syn*	68.1	230.7	0.0
	anti	247.1	88.4	2.5
TCP-OdG	*syn*	68.5	93.3	33.0
	anti	243.9	88.7	0.0
PCP-OdG	*syn*	67.3	186.5	0.7
	anti	250.8	178.8	0.0

As discussed for the other nucleotide models, the biologically-relevant *anti* conformation isolated using the counterion model for 4-Cl-PhOdG has an *anti* conformation that is 7.7 kJ mol^{-1} higher in energy than the *syn* nucleotide. Nevertheless, the preference for the *syn* conformer is actually less for PhOdG (6.3 kJ mol^{-1}) and DCP-OdG (2.5 kJ mol^{-1}) adducts than the *anti* structure. This prefrence is changed to *anti* conformer for tri- and penta- chloro mutated complexes. The energy difference between the preferred *syn* and *anti* conformer is quite similar for natural dG mono phosphate and the unsubstituted O-linked structure. In the case of 4-Cl-PhOdG, the chlorine atom at the *para* position of the aromatic ring is too far away to alter the topological properties of the monochloro adduct perceptibly. Thus, no more steric and electronic interactions are created which increase the stability of *syn* conformer more than PhOdG.

On the other hand, the *ortho* chlorine in $^{DCP-O}$dG directly influences the structural properties of this mutated complex. In particular, we note the steric effects of the lone pairs of *ortho* chlorine with the guanosine as well as attractions and repulsions between atoms, and other intermolecular forces. Figure 4.15shows that the interactions between *anti* and *syn* are much more comparable in dichloro substituted adducts. This leads to the lowest difference between *anti* and *syn* energies (2.5 kJ mol^{-1}) while, more steric and electronic effects are seen for the *syn* conformer in unsubstituted and monochloro mutated complexes.

Manifest attractive or repulsive electrostatic interactions as well as London dispersion forces in the $^{TCP-O}$dG and $^{PCP-O}$dG lead to a contrary trend in their conformational stability compared to unsubstituted, mono, and di chloro counterparts. This changes the structural properties of the nucleotide mono phosphate model such that the *anti* conformer is more stable than *syn*. However the impact of two chlorine atoms at *ortho* site is distinctly different in that a high barrier of 33.0 kJ mol^{-1} exists between *syn* and *anti* adducts. The interactive forces between the produce a very big difference between the stability of *anti* and *syn* structures, even though the *ortho* chlorine atoms still play the main role in stabilizing the *anti* conformer with an energy difference of 0.7 kJ mol^{-1} relative to *syn*.

This hypothesis is supported by the geometrical data in the Table 4.2. The χ and ϕ dihedral angles, which can reveal the geometrical and structural properties of mutated adducts, confirm the thermodynamic results. The C2′−endo sugar

puckering remained unchanged in all chloro substituted structures. The *syn* conformer of all Cl substituted O-linked adducts at nucleotide level shows an increase in χ values from 66.9° in $^{\text{PhO}}$dG to 68.5° in $^{\text{TCP-O}}$dG. In the case of $^{\text{PCP-O}}$dG, χ= 67.3° is fairly similar to the χ value for $^{\text{4-Cl-PhO}}$dG. Conversely, there is a decrease in χ values for *anti* chloro substituted nucleotide monophosphate models which changes to χ=250.8° in $^{\text{PCP-O}}$dG, the same value found in $^{\text{4-Cl-PhO}}$dG.

Interactions that are present between the chlorine of the aromatic ring of the bulky phenoxyl group and the phosphate emphasizes the importance of considering nucleotide models.

4.5. Conclusion

The computational approach outlined in this chapter has provided greater knowledge of the structure of the chloro substituted adducts which are formed due to exposure to phenolic carcinogens, as well as a better understanding of the likely stability, structural and conformational properties of these adducts at the nucleobase, nucleoside, and nucleotide level. However, the implications of these findings within the context of DNA have not yet been fully explored. Since all $^{\text{4-Cl-PhO}}$dG, $^{\text{DCP-O}}$dG, $^{\text{TCP-O}}$dG, and $^{\text{PCP-O}}$dG adducts adopt the *syn* conformer at the nucleoside level, there could be a potential for mispairing in the DNA strand. Furthermore, investigating the nucleotide model implies that the *syn* conformer is preferred for $^{\text{4-Cl-PhO}}$dG and $^{\text{DCP-O}}$dG species, whereas the *anti* conformer is dominant for $^{\text{TCP-O}}$dG, and $^{\text{PCP-O}}$dG mutated O-linked complexes. Therefore, it will be indispensable to investigate the hydrogen-bonding base pair preferences of the

adducts in both the *anti* and *syn* conformations to determine the potential for mismatch stabilization in the DNA double helix.

Experimental and computational data show that overall, the adducts were found to hydrolyze ~4 to 30 times faster than native dG under similar conditions. Although all chloro substituted O-linked adducts have more affinity to protonation at N7 rather than N3 on guanosine, an increase in the electron withdrawing character of the adducted moiety causes a decrease in the experimental rate of acidic hydrolysis. This is inconsistent with the theoretical results and the proposed reaction mechanism and the rate-determining step cleavage of the glycosidic bond.[16]

Although protonation at N7 can decrease the deglycosylation barriers, the hydrolytic stability of the O-linked adducts are considerable enough to attribute the toxicity of phenols to the frameshift mutation and generation of abasic sites through the acid catalyzed mechanism at physiological pH. In fact, a comparison of unsubstituted and chloro substituted adducts indicates that [PCP-O]dG is the most toxic adduct, and should be considered more in carcinogenesis studies.[46]

Other sources of stabilization (such as intrastrand interactions or steric clashes) may play an important role in determining adduct structure in damaged DNA helices, where a complex conformational heterogeneity may exist depending on the sequence context of the adduct. Therefore, it will be imperative to consider intrastrand interactions with flanking bases in a larger model.

4.6. References

1.   Sturla, S. J., DNA Adduct Profiles: Chemical Approaches to Addressing the Biological Impact of DNA Damage from Small Molecules. *Current Opinion in Chemical Biology* **2007,** *11* (3), 293-299.

2.   Manderville, R. A., Structural and Biological Impact of Radical Addition Reactions with DNA Nucleobases. Richard, J. P., Ed. 2009; Vol. 43, pp 177-218.

3.   Scharer, O. D., Chemistry and Biology of DNA Repair. *Angewandte Chemie International Edition* **2003,** *42* (26), 2946-74.

4.   Millen, A. L.; Sharma, P.; Wetmore, S. D., C8-Linked Bulky Guanosine DNA Adducts: Experimental and Computational Insights into Adduct Conformational Preferences and Resulting Mutagenicity. *Future Medicinal Chemistry* **2012,** *4* (15), 1981-2007.

5.   Broyde, N. E. G. a. S., *The Chemical Biology of DNA Damage.* © 2010 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim: New York University.

6.   Benigni, R.; Bossa, C., Mechanisms of Chemical Carcinogenicity and Mutagenicity: A Review with Implications for Predictive Toxicology. *Chemical Reviews* **2011,** *111* (4), 2507-2536.

7.   Lukin, M.; de los Santos, C., Nmr Structures of Damaged DNA. *Chemical Reviews* **2006,** *106* (2), 607-686.

8.   Burton, G. W.; Ingold, K. U., Vitamin-E - Application of the Principles of Physical Organic-Chemistry to the Exploration of Its Structure and Function. *Accounts of Chemical Research* **1986,** *19* (7), 194-201.

9.   Dai, J.; Park, G.; Perry, J. L.; Il'Ichev, Y. V.; Bow, D. A. J.; Pritchard, J. B.; Faucet, V.; Pfohl-Leszkowicz, A.; Manderville, R. A.; Simon, J. D., Molecular Aspects of the Transport and Toxicity of Ochratoxin A. *Accounts of Chemical Research* **2004,** *37* (11), 874-881.

10.   Dai, J.; Wright, M. W.; Manderville, R. A., An Oxygen-Bonded C8-Deoxyguanosine Nucleoside Adduct of Pentachlorophenol by Peroxidase Activation: Evidence for Ambident C8 Reactivity by Phenoxyl Radicals. *Chemical Research in Toxicology* **2003,** *16* (7), 817-821.

11.   Dai, J.; Sloat, A. L.; Wright, M. W.; Manderville, R. A., Role of Phenoxyl Radicals in DNA Adduction by Chlorophenol Xenobiotics Following Peroxidase Activation. *Chemical Research in Toxicology* **2005,** *18* (4), 771-779.

12.   Kornyushyna, O.; Stemmler, A. J.; Graybosch, D. M.; Bergenthal, I.; Burrows, C. J., Synthesis of a Metallopeptide - Pna Conjugate and Its Oxidative Cross-Linking to a DNA Target. *Bioconjugate Chemistry* **2005,** *16* (1), 178-183.

13.   Advances in Molecular Toxicology. Fishbein, J. C., Ed. Elsevier Science Amsterdam, 2006; Vol. 1, pp 1-188.

14.   Lindahl, T., Instability and Decay of the Primary Structure of DNA. *Nature* **1993,** *362* (6422), 709-715.

15.   Cavalieri, E.; Saeed, M.; Zahid, M.; Cassada, D.; Snow, D.; Miljkovic, M.; Rogan, E., Mechanism of DNA Depurination by Carcinogens in Relation to Cancer Initiation. *Iubmb Life* **2012,** *64* (2), 169-179.

16.   Kuska, M. S.; Majdi Yazdi, M.; Witham, A. A.; Dahlmann, H. A.; Sturla, S. J.; Wetmore, S. D.; Manderville, R. A., Influence of Chlorine Substitution on the Hydrolytic Stability of Biaryl Ether Nucleoside Adducts Produced by Phenolic Toxins. *The Journal of Organic Chemistry* **2013,** *78* (14), 7176-7185.

17.   Roger, M.; Hotchkiss, R. D., Selective Heat Inactivation of Pneumococcal Transforming Deoxyribonucleate. *Proceedings of the National Academy of Sciences of the United States of America* **1961,** *47* (5), 653-&.

18.   Zoltewicz, J. A.; Clark, D. F.; Sharpless, T. W.; Grahe, G., Kinetics and Mechanism of the Acid-Catalyzed Hydrolysis of Some Purine Nucleosides. *Journal of the American Chemical Society* **1970,** *92* (6), 1741-1750.

19.   Gates, K. S.; Nooner, T.; Dutta, S., Biologically Relevant Chemical Reactions of N7-Alkylguanine Residues in DNA. *Chemical Research in Toxicology* **2004,** *17* (7), 839-856.

20.   Hovinen, J.; Glemarec, C.; Sandstrom, A.; Sund, C.; Chattopadhyaya, J., Spectroscopic, Kinetic and Semiempirical Molecular-Orbital Studies on 8-Amino-Adenosines, 8-Methylamino-Adenosines and 8-Dimethylamino-Adenosines. *Tetrahedron* **1991,** *47* (26), 4693-4708.

21.   Novak, M.; Ruenz, M.; Kazerani, S.; Toth, K.; Nguyen, T. M.; Heinrich, B., Kinetics of Hydrolysis of 8-(Arylamino)-2 '-Deoxyguanosines. *Journal of Organic Chemistry* **2002,** *67* (7), 2303-2308.

22.   Patel, D. J.; Mao, B.; Gu, Z. T.; Hingerty, B. E.; Gorin, A.; Basu, A. K.; Broyde, S., Nuclear Magnetic Resonance Solution Structures of Covalent Aromatic Amine-DNA Adducts and Their Mutagenic Relevance. *Chemical Research in Toxicology* **1998,** *11* (5), 391-407.

23.   Schlitt, K. M.; Sun, K. W.; Paugh, R. J.; Millen, A. L.; Navarro-Whyte, L.; Wetmore, S. D.; Manderville, R. A., Concerning the Hydrolytic Stability of 8-Aryl-2'-Deoxyguanosine Nucleoside Adducts: Implications for Abasic Site Formation at Physiological Ph. *The Journal of Organic Chemistry* **2009,** *74* (16), 5793-802.

24.   Becke, A. D., Density-Functional Thermochemistry .4. A New Dynamical Correlation Functional and Implications for Exact-Exchange Mixing. *Journal of Chemical Physics* **1996,** *104* (3), 1040–1046.

25.   Millen, A. L.; McLaughlin, C. K.; Sun, K. M.; Manderville, R. A.; Wetmore, S. D., Computational and Experimental Evidence for the Structural Preference of Phenolic C-8 Purine Adducts. *Journal of Physical Chemistry A* **2008,** *112* (16), 3742-3753.

26.   Stewart, J. J. P., Optimization of Parameters for Semiemperical Methods. 1. Method. *Journal of Computational Chemistry* **1989,** *10* (2), 209–220.

27.   Rezac, J.; Fanfrlik, J.; Salahub, D.; Hobza, P., Semiempirical Quantum Chemical Pm6 Method Augmented by Dispersion and H-Bonding Correction Terms Reliably Describes Various Types of Noncovalent Complexes. *Journal of Chemical Theory and Computation* **2009,** *5* (7), 1749–1760.

28.   Millen, A. L.; Manderville, R. A.; Wetmore, S. D., Conformational Flexibility of C8-Phenoxyl-2 '-Deoxyguanosine Nucleotide Adducts. *Journal of Physical Chemistry B* **2010,** *114* (12), 4373-4382.

29.   Sharma, P.; Manderville, R. A.; Wetmore, S. D., Modeling the Conformational Preference of the Carbon-Bonded Covalent Adduct Formed Upon Exposure of 2′-Deoxyguanosine to Ochratoxin A. *Chemical Research in Toxicology* **2013,** *26* (5), 803-816.

30.   Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Jr., J. A. M.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.;

Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.02; Gaussian, Inc.: Wallingford CT, 2009.

31.   Yang, S. Y.; Hristov, I.; Fleurat-Lessard, P.; Ziegler, T., Optimizing the Structures of Minimum and Transition State on the Free Energy Surface. *Journal of Physical Chemistry A* **2005,** *109* (1), 197–204.

32.   Dal Peraro, M.; Ruggerone, P.; Raugei, S.; Gervasio, F. L.; Carloni, P., Investigating Biological Systems Using First Principles Car-Parrinello Molecular Dynamics Simulations. *Current Opinion in Structural Biology* **2007,** *17* (2), 149–156.

33.   Becke, A. D., Density-Functional Exchange-Energy Approximation with Correct Asymptotic-Behavior. *Physical Review A* **1988,** *38* (6), 3098–3100.

34.   Becke, A. D., Density-Functional Thermochemistry .3. The Role of Exact Exchange. *Journal of Chemical Physics* **1993,** *98* (7), 5648–5652.

35.   Pople, J. A.; Schlegel, H. B.; Krishnan, R.; Defrees, D. J.; Binkley, J. S.; Frisch, M. J.; Whiteside, R. A.; Hout, R. F.; Hehre, W. J., Molecular-Orbital Studies of Vibrational Frequencies. *International Journal of Quantum Chemistry* **1981**, 269–278.

36.   Amos, T.; Snyder, L. C., Unrestricted Hartree-Fock Calculations .I. Improved Method of Computing Spin Properties. *Journal of Chemical Physics* **1964,** *41* (6), 1773–1774.

37.   Stewart, J. J. P., Optimization of Parameters for Semiempirical Methods V: Modification of Nddo Approximations and Application to 70 Elements. *Journal of Molecular Modeling* **2007,** *13* (12), 1173–1213.

38.   Dipple, A., DNA-Adducts of Chemical Carcinogens. *Carcinogenesis* **1995,** *16* (3), 437-441.

39.   Uno, B.; Iwamoto, T.; Okumura, N., Importance of Substituent Intramolecular Charge-Transfer Effect on the Molecular Conformation of Diphenyl Ethers. *The Journal of Organic Chemistry* **1998,** *63* (26), 9794-9800.

40.   Greco, F.; Liguori, A.; Sindona, G.; Uccella, N., Gas-Phase Proton Affinity of Deoxyribonucleosides and Related Nucleobases by Fast-Atom-Bombardment Tandem Mass-Spectrometry. *Journal of the American Chemical Society* **1990,** *112* (25), 9092-9096.

41.   Hevesi, L.; Wolfson-Davidson, E.; Nagy, J. B.; Nagy, O. B.; Bruylants, A., Contribution to the Mechanism of the Acid-Catalyzed Hydrolysis of Purine Nucleosides. *Journal of the American Chemical Society* **1972,** *94* (13), 4715-4720.

42.   Zoltewicz, J. A.; Clark, D. F., Kinetics and Mechanism of the Hydrolysis of Guanosine and 7-Methylquanosine Nucleosides in Perchloric Acid. *The Journal of Organic Chemistry* **1972,** *37* (8), 1193-1197.

43.   Hevesi, L.; Wolfsond.E; Bruylant.A; Nagy, J. B.; Nagy, O. B., Contribution to Mechanism of Acid-Catalyzed Hydrolysis of Purine Nucleosides. *Journal of the American Chemical Society* **1972,** *94* (13), 4715-+.

44.   Romero, R.; Stein, R.; Bull, H. G.; Cordes, E. H., Secondary Deuterium-Isotope Effects for Acid-Catalyzed Hydrolysis of Inosine and Adenosine. *Journal of the American Chemical Society* **1978,** *100* (24), 7620-7624.

45.   Schlitt, K. M.; Sun, K. W. M.; Paugh, R. J.; Millen, A. L.; Navarro-Whyte, L.; Wetmore, S. D.; Manderville, R. A., Concerning the Hydrolytic Stability of 8-Aryl-2 '-Deoxyguanosine Nucleoside Adducts: Implications for Abasic Site Formation at Physiological Ph. *The Journal of Organic Chemistry* **2009,** *74* (16), 5793-5802.

46.   Michalowicz, J.; Duda, W., Phenols - Sources and Toxicity. *Polish Journal of Environmental Studies* **2007,** *16* (3), 347-362.

Chapter 5. Concluding Remarks and Future Perspectives

5.1. Concluding Remarks

The aim of this thesis was to develop a computational model for the study of O–linked DNA adducts, which predicts their conformational flexibility and base-pairing preferences by following a bottom-up approach (small nucleobase model to large DNA duplex perspective). Investigating the structural properties to unsubstituted and chloro substituted O–linked species and comparing the results with the geometrical properties of *ortho*– and *para*– C–linked counterparts provides a better understanding of steric and electronic effects involving the sugar segment and the effects on the backbone of DNA. This provides information that is complementary to experimental results that are available for some of these models. Ultimately, the stability and conformational pattern in the context of DNA duplexes containing phenoxyl damaged lesions were investigated by employing molecular dynamic simulation.

Theoretical results involving the small (nucleobase, nucleoside) models implied an expected twisted *syn* conformation of the adducts at the nucleoside level and also showed the presence of an O5′–H•••N3 hydrogen bond as the predicted favoured conformer. To understand the effect of protonation on the destabilization of the glycosidic bond, proton affinity at N^3 and N^7 of guanosine were determined. Computational models in both solvent and gas phase illustrate the structural changes in the orientation of the phenoxyl moiety with respect to

the nucleobase from planar to skew. The PA decreases upon functionalization of the aromatic ring with increasing numbers of chlorine. The probability of abasic site formation was explored by determining the deglycosylation barrier which becomes progressively smaller for N^7H^+ and N^3H^+ adducts bearing increased chlorine substitution. Nevertheless, accumulation of electron density at N^7 enhances basicity. Since formation of a positive charge on the nucleobase makes it a better leaving group, this enhances the favorability of depurination. In fact, the monochloro substituted N^3H^+ species possesses a much greater barrier to deglycosylation than the N^7H^+ counterparts. In addition, chlorine substitution of the phenoxyl ring will inductively stabilize the negative charge on N9 nitrogen during the process of heterolytic cleavage of the glycosidic bond.

Due to conformational alteration which can arise from steric constraints imposed by phosphate group, the nucleoside model cannot necessarily provide all the essential structural information. We expanded the model to include the phosphate component of the backbone. Calculations at nucleotide level identified a *syn* conformer for unsubstituted–, mono– and di– chloro O–linked adducts as the most stable. Conversely, tri– and pentachloro- substituted O–linked species as well as *ortho*– and *para*– C–linked homologs preferred *anti* conformers. However, steric and electronic effects of the phenoxyl moiety are more obvious in C-linked mutated complexes via alteration in sugar puckering. Taken together, this model predicted a greater stability for the *syn* conformation of unsubstituted, monochloro, and dichloro species, whereas tri- and pentachloro species prefer *anti* conformers for *ortho* and *para* adducts. Therefore, small calculated energy

differences suggest the possibility of a complex conformational heterogeneity in DNA helices.[1-2]

Structural properties and the mutagenicity of the adducts can be extrapolated to biological systems by studying their stability in the DNA duplex (*Nar*I sequence) which was addressed in Chapter 3. Conformational flexibility of the adducts can affect the base-pairing preferences of the adducts which may increase the possibility of mispairing in the helix. While the *anti* conformation exhibited a preference for base-pairing with cytosine, the *syn* conformation formed stable base pairs with a guanine mismatch. This illuminates the importance of steric effects particularly from the backbone of DNA; complementary and flanking bases cause to reverse the trends in the relative energies of the *anti* and *syn* conformations compared to the nucleoside and nucleotide model. The *Nar*I sequence represents a major area for frameshift mutation especially by N–linked structures. As a result of the availability of extended studies on the other adducts in this sequence, it has also attracted attention as the specific sequence in this project.[3-5] In addition, the structural dependence of the adducts sequence can be proved by comparing its structural properties with previous studies in other sequences of DNA which are new arrangements of nucleotides and definitely varying the type of flanking and complementary bases. Indeed, the *anti* conformation against cytosine is preferred with this model for all adducts. This suggests that the neighboring sugar and backbone play an important role in destabilizing the predicted *syn* conformation of phenoxyl adducts at the nucleoside level and in the DNA model. This suggests that the *anti* conformation may be more relevant to the *Nar*I sequence of the DNA

double helix. On the other hand, the *syn* conformer for unsubstituted O-linked and C—linked structures against a guanine mismatch is the lowest energy structure. This endorses the significance of studying mutated complexes at the DNA level and that insufficient structural information is obtained from small nucleoside and nucleotide models. For both $^{o\text{-PhOH}}$dG:C (syn-2: $\theta{\sim}346.0°$ and $\chi{\sim}64.2°$) and $^{p\text{-PhOH}}$dG:C (syn-180:$\theta{\sim}41.6°$ and $\chi{\sim}70.3°$), the base-displaced stack conformation is identified in the *Nar*I sequence which causes the frameshift mutation. These results suggest greater mutagenicity for C-linked adducts in comparison with O-linked structures. Furthermore, in the case of *ortho* and *para* C-linked adducts in two decanucleotides (oligonucleotide (ODN1) and (ODN2)), the simulations predicted that the adducts preferentially adopt a *syn* conformation in DNA regardless of sequence or the identity of the base in the complementary strand. [6]

Although the small models could not precisely predict the conformational preference of adducted nucleobases in the DNA double helix, nucleotide and nucleoside models were able to provide insight into other questions, such as the preferred orientation of the bulky group with respect to dG, sugar puckering, and preferred values of χ for both the *anti* and *syn* conformations. Therefore, small nucleoside, nucleotide monophosphate models can provide different structural information leading to a better understanding about base-pairing preferences and hydrogen bonding.

Due to adopting mainly B or W type conformations against both cytosine and guanine, little distortion of the backbone was observed in the duration of simulation for all phenoxyl adduct lesions, which decreases the possibility of

214

repairing by enzymes. Hoogsteen bonding to guanine resulted in mismatch stabilization, however hydrogen bonding occupancies and base pairing energies clarified that the guanine mismatch cannot stabilize the double helix as effectively as full complementary duplexes. Alternatively, *ortho* and *para* C-linked adducts in two decanucleotides ODN1 and ODN2 sequences, show evidence for mismatch stabilization leading to preferential insertion of guanine opposite of the C-linked phenoxyl adducts, or G↔C mutations upon replication. The same outcome has been monitored for structurally-related ([Ph]dG) adducts.[7] Computational studies in this thesis tried to shed light on the importance of geometrical and conformational properties of mutated complexes in stabilizing/destabilizing interstrand and intrastrand interactions which influence the duplex stability. It would be striking to consider other imperative factors in stabilizing helices more comprehensively For instance, by examination of the impact of flanking and complementary bases on the duplex stability in different bulky adducted DNA sequences in the future.

5.2. Future Work

In the very beginning, modeling consisted of putting balls and sticks together by hand to visualize molecules. With today's amazing advances in computer technology, it is possible to not only visualize static molecules but also to follow conformational changes, vibrations of atoms, and to examine different segments of molecules over time. A disadvantage of standard MD simulations is that it cannot properly explore conformational properties due to sampling for an inadequate length of time. Many different enhanced sampling methods have been developed to address this problem. As computational power increases over time and more

efficient algorithms are developed, more accurate MD simulations may identify new conformational and geometrical changes other than those observed experimentally.

Overall, previous computational studies on bulky C8-dG adducts[8-10] show that MD simulation is an efficient tool in verifying the structural properties and the presence or lack of discrete interactions within adducted duplexes. This can provide deeper and complementary rationalizations for experimental observations. As a result of the complexity of calculation and computational expense, QM and DFT calculations provide more accurate results for small models which are limited in the specific strand flanking and complementary bases, so they are not able to reproduce a thorough picture of all interactions in larger models. In the absence of NMR data or crystal structures, structural data obtained from (MD) calculations in combination with experimental melting temperatures can reveal the preferred conformation of mutagenic DNA lesions.

The identification of related DNA adducts that can serve as biomarkers for measuring carcinogenic exposure is also an important future goal. Indeed, such molecules could permit the exploration of other biological implications of bulky adducts.

For every question answered, more questions appear that are equally interesting and important. For example, when different chloro substituted adducts are incorporated in the double helix does the conformation remain the same as the geometrical properties obtained from small nucleobase, nucleoside or nucleotide models? Since chlorine atoms can cause significant steric and electronic effects,

chloro-substituted and especially highly chloro substituted species may cause the formation of S conformer in DNA and induce formation of abasic sites which are among the most serious distortions in the duplex.

To what extent are the MD simulations of Cl-substituted adducts consistent with results obtained from unsubstituted and *ortho* and *para* C-linked structures? How do the steric and electronic effects of chlorine substitution on the aromatic ring distort the double helix? In MD simulation, what will be the effect of the second *ortho* chlorine atom in comparison to the monochloro adduct and unsubstituted species adduct? Since MD simulation of unsbstituted O-linked adduct demonstrated B and W conformers for the aforementioned adduct against cytosine and guanine, respectively we may expect other types of conformers or more conspicuous distortion of double helix by presence of heavy Chlorine atom at *para*-and *ortho*- position of aromatic ring of phenoxyl moiety. Can we observe similar results about the effect of the chlorine atom as identified in small models? The same questions apply to the effect of *meta* chlorine atoms of the pentachloro substituted species in comparison with trichloro mutated complexes. Since protonation of modified guanosine could increase the probability of formation of abasic sites, it would be worthwhile to consider the effect of N^7 and N^3 protonation of chloro substituted adducts against an abasic site. Subsequently, effects of the bulky lesions should be taken into account through studies on large models (MD simulation of DNA) while investigating other sequences of DNA except *Nar*I, that will determine the significance of flanking bases in determining adduct conformation and base-pairing properties. Finally duplex stability should be

217

investigated in other types of sequences. Understanding the function of Nucleotide Excision Repair (NER) or Base Excision Repair (BER) enzymes would be interesting, so the *in vivo* biological effects of these adducts can be addressed in the future. Although there is a direct relationship between DNA adducts and human cancer, very few studies on understanding adduct structure and formation have been undertaken. Expanding the studies in this field may lead to the development of innovative medicinal strategies for earlier detection and prevention of cancer.

Although a conformational search using Hyperchem established that the most stable conformers acquire the C2' sugar puckering, it would be interesting to generate new PESs for unsubstituted and all chloro substituted O-linked adducts without considering any constraints during scan calculations. This may result in more realistic results for the conformational preferences of modified guanosine with respect to sugar puckering.

These future research avenues may be helpful for researchers interested in repairing the damage induced by exposure to certain molecules, so that they are no longer detrimental to the human population. Molecular dynamic simulations will hopefully be more widely applied in different fields of research such as chemistry and the pharmaceutical sciences, where this technique may lead to new possibilities in drug design.

5.3. References

1.   Kuska, M. S.; Majdi Yazdi, M.; Witham, A. A.; Dahlmann, H. A.; Sturla, S. J.; Wetmore, S. D.; Manderville, R. A., Influence of Chlorine Substitution on the Hydrolytic Stability of Biaryl Ether Nucleoside Adducts Produced by Phenolic Toxins. *The Journal of Organic Chemistry* **2013,** *78* (14), 7176-7185.

2.   Sturla, S. J., DNA Adduct Profiles: Chemical Approaches to Addressing the Biological Impact of DNA Damage from Small Molecules. *Current Opinion in Chemical Biology* **2007,** *11* (3), 293-299.

3.   Millen, A. L.; Sharma, P.; Wetmore, S. D., C8-Linked Bulky Guanosine DNA Adducts: Experimental and Computational Insights into Adduct Conformational Preferences and Resulting Mutagenicity. *Future Medicinal Chemistry* **2012,** *4* (15), 1981-2007.

4.   Patel, D. J.; Mao, B.; Gu, Z. T.; Hingerty, B. E.; Gorin, A.; Basu, A. K.; Broyde, S., Nuclear Magnetic Resonance Solution Structures of Covalent Aromatic Amine-DNA Adducts and Their Mutagenic Relevance. *Chemical Research in Toxicology* **1998,** *11* (5), 391-407.

5.   Kuska, M. S.; Witham, A. A.; Sproviero, M.; Manderville, R. A.; Majdi Yazdi, M.; Sharma, P.; Wetmore, S. D., Structural Influence of C8-Phenoxy-Guanine in the NarI Recognition DNA Sequence. *Chemical Research in Toxicology* **2013,** *26* (9), 1397-1408.

6.   Omumi, A.; Millen, A. L.; Wetmore, S. D.; Manderville, R. A., Fluorescent Properties and Conformational Preferences of C-Linked Phenolic-DNA Adducts. *Chemical Research in Toxicology* **2011,** *24* (10), 1694-1709.

7.   Kohda, K.; Tsunomoto, H.; Kasamatsu, T.; Sawamura, F.; Terashima, I.; Shibutani, S., Synthesis and Miscoding Specificity of Oligodeoxynucleotide Containing 8-Phenyl-2′-Deoxyguanosine. *Chemical Research in Toxicology* **1997,** *10* (12), 1351-1358.

8.   Omumi, A.; Millen, A. L.; Wetmore, S. D.; Manderville, R. A., Fluorescent Properties and Conformational Preferences of C-Linked Phenolic-DNA Adducts. *Chem Res Toxicol* **2011,** *24* (10), 1694-709.

9.   Sharma, P.; Manderville, R. A.; Wetmore, S. D., Modeling the Conformational Preference of the Carbon-Bonded Covalent Adduct Formed Upon Exposure of 2′-Deoxyguanosine to Ochratoxin A. *Chemical Research in Toxicology* **2013,** *26* (5), 803-816.

10.   Wilson, K. A.; Wetmore, S. D., Complex Conformational Heterogeneity of the Highly Flexible O6-Benzyl-Guanine DNA Adduct. *Chemical Research in Toxicology* **2014,** *27* (7), 1310-1325.