STRUCTURAL STUDIES OF BUTADIENE-DERIVED DNA ADDUCTS

By

WEN XU

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Chemistry

December, 2008

Nashville, Tennessee

Approved:

Professor Michael P. Stone

Professor Martin Egli

Professor Brandt F. Eichman

Professor Jens Meiler

Professor Carmelo J. Rizzo

ACKNOWLEDGEMENTS

I am grateful to all of those who made this dissertation possible. First and foremost, I would like to thank my advisor, Dr. Michael P. Stone, for his guidance and help during my whole Ph.D. period. Appreciation also goes to my Ph.D. committee members: Dr. Martin Egli, Dr. Carmelo J. Rizzo, Dr. Brandt F. Eichman, and Dr. Jens Meiler for their time and effort.

I would like to thank Dr. Feng Wang for her nice and patient help in my early stages of this project. I would also like to thank Dr. Tandace A. Scholdberg and Dr. W. Keither Merritt for their initial work on the butadiene projects. I appreciate the help from Dr. Ivan D. Kozekov, Ms. Albena Kozekov, and Dr. Richard Hodge for the DNA sample synthesis and purification; Dr. Stephen Lloyd and all involved members of their laboratories for their biochemical study in this project; and Dr. Markus Voehler, Dr. Jason Jacob and Dr. Donald Stec for their assistance on NMR experiments. I would like to thank Ms. Karen C. Angel and Dr. Robert L. Eoff for their help on protein expression and purification; thank Ms. Ms. Lioudmila V. Loukachevitch for her help on crystallization; and thank Dr. Adriana Irimia, Dr. Feng Li, Dr. Pradeep S. Pallan, and Dr. Joel Harp for their help on X-ray crystallography. I would like to thank Ms. Sarrah Musser for proofreading my dissertation. I would like to thank people in the Center for Structural Biology, who helped me on the computational work of this research project. And also, I would like to thank people in the Center for Molecular Toxicology who helped me on the biological experiments performed in my research.

I have been very fortunate to work in Stone Group and I am grateful for the time

with all lab friends: Dr. Kyle Brown, Dr. Young Jin Cho, Dr. Yazhen Wang, Dr. Hai Huang, Dr. Surajit Banerjee, Dr. Ganesh Shanmugam, Ms. Ewa Kowal, Ms. Heather Day, and Mr. Travis Nielson.

I would like to thank the National Institutes of Health and the Center for Molecular Toxicology for financial support.

Finally, I would like to thank my entire family for their support throughout my graduate study.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

LIST OF SCHEMES

LIST OF ABBREVIATIONS

1-D	one-dimensional
2-D	two-dimensional
8-oxoG	7,8-dihydro-8-oxoguanine
AAF-dG	*N*-(deoxyguanosin-8-yl)-2-acetylaminofluorene
BD	butadiene
BPDE	benzo[*a*]pyrene diol epoxide
CCD	charged coupled device
CORMA	complete relaxation matrix analysis
COSY	correlation spectroscopy
δ	chemical shift
dA	deoxyadenosine
dC	deoxycytosine
Δδ	change in chemical shift
ddNTP	dideoxynucleotide triphosphate (N = A,C,T, or G)
DEB	1,2:3,4-diepoxybutanes
DFT	2,4-difluorotoluene
dG	deoxyguanosine
DHEB	1,2-dihydroxy-3,4-epoxybutanes
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate (N = A,C,T, or G)
Dpo4	*Sulfolobus solfataricus* P2 DNA polymerase IV
DQF-COSY	double quantum filtered COSY
dT	deoxythymidine
dU	deoxyuridine
EB	1,2-epoxy-3-butene
EDTA	ethylenediamine tetraacetic acid
HMKV	hyrosxymethylvinylketone
HPLC	high performance liquid chromatography
IARC	International Agency for Cancer Research
IPTG	Isopropyl β-D-thiogalactopyranoside
LB	Luria-Bertani
MAD	multiple wavelength anomalous dispersion
MALDI-TOF	matrix assisted laser desorption ionization-time of flight
MARDIGRAS	matrix analysis of relaxation for discerning geometry of an aqueous solution
MIR	multiple isomorphous replacement
MR	molecular replacement
NMR	nuclear magnetic resonance
NOESY	nuclear Overhauser effect spectroscopy
O^6-MeG	O^6-methylguanine

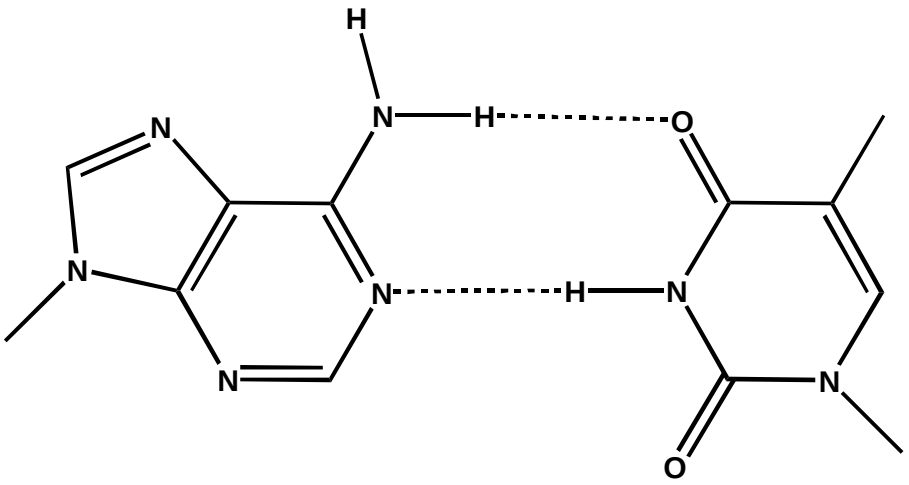
PAD	polymerase associated domain
PAH	polycyclic aromatic hydrocarbon
PCNA	proliferating cell nuclear antigen
PdG	$1,N^2$-propanodeoxyguanosine
PEG	polyethylene glycol
PEM	potential energy minimization
pol	polymerase
ppm	part per million
R_1^x	sixth root residual index
R-factor	residual factor
rMD	restrained molecular dynamics
rmsd	root mean square deviation
SBR	styrene-butadiene rubber
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
τ_m	mixing time
TOCSY	total correlation spectroscopy
TPPI	time proportional phase increment
TT	thymine dimer

CHAPTER I

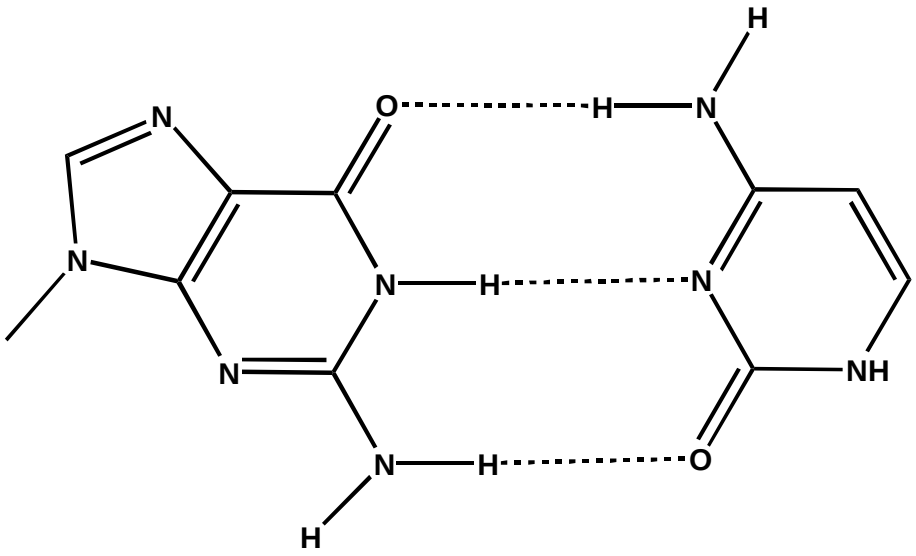
INTRODUCTION

Deoxyribonucleic Acid (DNA)

DNA is an extremely important biological macromolecule which carries genetic information for the development and functioning of all known living organisms and some viruses. Chemically, DNA is a polynucleotide made up of individual nucleotides, which join together to form a sugar-phosphate backbone from the 3' hydroxyl group on the deoxyribose sugar to the 5' phosphate group. Each nucleotide contains three components: a sugar, one of four types of nitrogen bases: adenine (A), thymine (T), guanine (G), and cytosine (C), and a phosphate group. In cells, DNA exists in a double-helix structure of two strands coiled around each other with the complementary bases hydrogen-bonded on the inside of the double-helix. The structure was first discovered by James D. Watson and Francis Crick in 1953 (*1*). Watson-Crick base pairing represents the canonical hydrogen bonding interactions in DNA, in which adenine always forms a base pair with thymine (A-T pair), as does cytosine with guanine (C-G pair) (Scheme 1-1).



Adenine **Thymine**



Guanine **Cytosine**

Scheme 1-1. Watson-Crick base pairing: A to T and G to C.

The normal right-handed double helix structure of DNA is known as the B-form, however, A-form and Z-form DNA have also been observed in organisms too (Figure 1-1)

(*2*). DNA conformations adopted in cellular systems are affected by several factors, including the sequence of the DNA, chemical modifications of the bases and solution conditions, such as the concentration of metal ions and polyamines (*3*). Geometric differences of A-, B- and Z-form DNA are listed in Table 1-1.

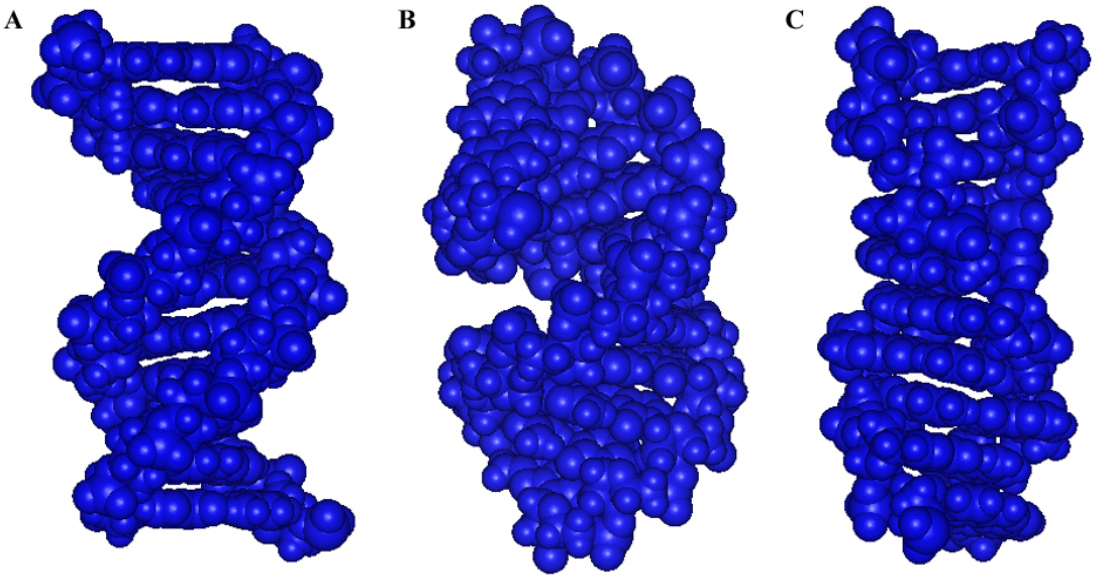


Figure 1-1. Models representing (A) B-form, (B) A-form, and (C) Z-form DNA.

B-DNA is the most common DNA conformation *in vivo* (*4*). As the phosphate backbones spiral around the outer surface of the double helix, two grooves of different widths are generated, referred to as the major and minor grooves. B-DNA has a wide major groove which is easily accessible to some DNA binding proteins, and a narrow minor groove (*5*).

Table 1-1: Comparison geometries of A-, B- and Z- form DNA (Reprinted from ref. 2. Copyright of 1993 the International Union of Crystallography by Munksgaard).

Geometry attribute	A-form	B-form	Z-form
Helix sense	right-handed	right-handed	left-handed
Repeating unit	1 bp	1 bp	2 bp
Rotation/bp	33.6°	35.9°	60°/2
Mean bp/turn	10.7	10.0	12
Inclination of bp to axis	+19°	−1.2°	−9°
Rise/bp along axis	2.3 Å (0.23 nm)	3.32 Å (0.332 nm)	3.8 Å (0.38 nm)
Rise/turn of helix	24.6 Å (2.46 nm)	33.2 Å (3.32 nm)	45.6 Å (4.56 nm)
Mean propeller twist	+18°	+16°	0°
Glycosyl angle	anti	anti	pyrimidine: anti, purine: syn
Sugar pucker	C3'-endo	C2'-endo	C: C2'-endo, G: C2'-exo
Diameter	26 Å (2.6 nm)	20 Å (2.0 nm)	18 Å (1.8 nm)

In a solution with higher salt concentrations or with alcohol added, A-form DNA is favored. The A-form is also a right-handed helix but is shorter and wider than B-form. In contrast to B-DNA, the wide, shallow minor groove is accessible to proteins, but it contains less information than the major groove. In the cell, A-DNA is the form taken by DNA-RNA hybrid double helices, as well as enzyme-DNA complexes (*6*,*7*).

The left-handed formation of DNA is referred to as Z-form DNA, which has a narrower, more elongated helix than A- or B-form DNA. The conformation of Z-DNA can be formed by regions rich in guanine-cytosine base pairs with alternating G-C sequences, due to the *anti* conformation of the C sugar compensating for the *syn* conformation of the G glycosidic bond. But the function of Z-DNA remains obscure (*8*). It has also been observed that DNA with methylated adducts may undergo a large change in conformation and adopt the Z form (*9*). This unusual conformation can be recognized

by specific Z-DNA binding proteins and may be involved in the regulation of transcription (*10*).

Given the biological function that is related to the conformation of DNA, it is important to investigate the DNA structures which will be correlated to biological activity, particularly in regard to damaged DNA.

Oncogene

An oncogene is a gene capable of causing cancer. The *ras* oncogene was originally discovered in two murine retroviruses, known as the Harvey and Kirsten murine sarcoma viruses (*11*). Nucleotide sequence analysis and *in vitro* site-specific mutagenesis studies show that changes in the *ras* protein at amino acid nucleotides 12, and 13 or 59-61 almost invariably lead to changes in oncogenic potential. Mutant *ras*61 sequences are the most frequently detectable alteration in some types of human tumors (*11*). Thus, understanding the structure of damaged DNA by chemical modification at specific nucleotides of an oncogene that has shown mutations in human cancers will be extremely significant.

DNA Damage

DNA damage by endogenous or exogenous factors, such as oxidation, hydrolysis, alkylation, radiation or toxic chemicals, occurs at a rate of 1,000 to 1,000,000 molecular lesions per cell per day (Table 1-2) (*12-14*). The high frequency of DNA damage is due to a constant assault on the DNA by genotoxic agents, nucleotide misincorporation during DNA replication, and the intrinsic biochemical instability of the DNA itself (*15*).

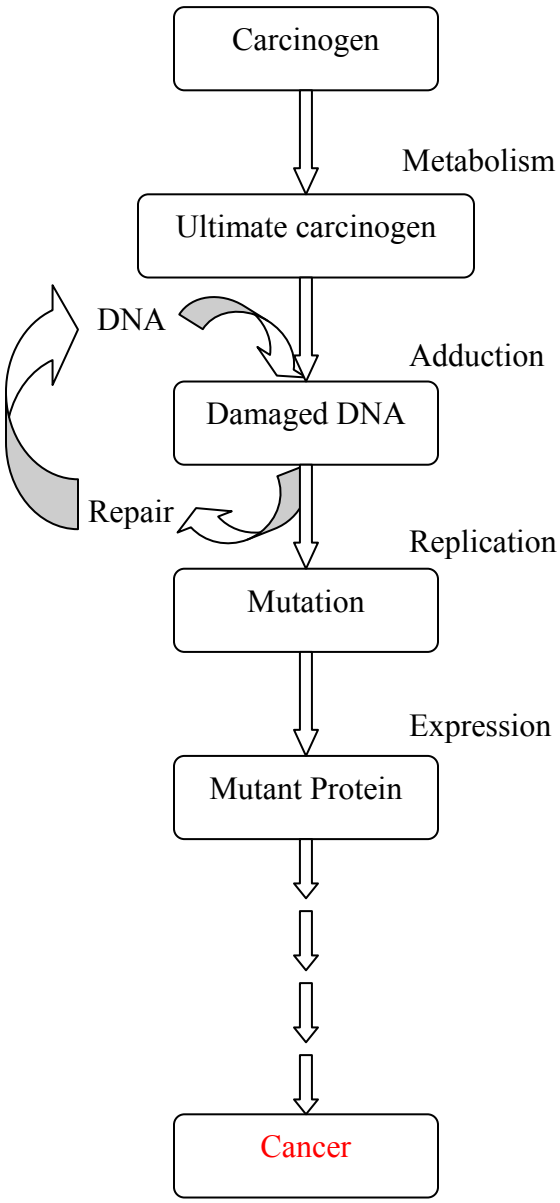
In humans, DNA damage has been shown to be involved in a variety of genetically inherited disorders, in aging (*16*), and in carcinogenesis (*17*,*18*).

Table 1-2: Sources and types of DNA damage (Reprinted from ref. 14. Copyright of 1999 Elsevier).

Source	Type	Example
Endogenous		
Chemical	DNA instability	Depurination
Biochemical	Reactive biochemicals	Aldehydes, *S*-Ado-met
	Reactive oxygen species	$HO^{\cdot}$, $NO^{\cdot}$, peroxides
	Errors at DNA replication and repair	Metal ions
Exogenous		
Hardly avoidable	Radiation	UV, cosmic and terrestric ionizing radiation
	Natural radioactive isotopes	^{222}Rn, ^{40}K
	Carcinogens in ambient air	Polynuclear aromatic hydrocarbons (PAH)
	Natural dietary carcinogens	Ethyl carbamate, estragole
Avoidable (in part)	Food pyrolysis products	Arylamines, PAH, nitroso compounds (NOC)
	Exposure at the workplace	Vinyl chloride
	Carcinogens in ambient air	Passive smoking
Avoidable (in principle)	Voluntary exposures	Smoking, some work-related exposures
	Therapeutic drugs	Tumor therapeutic agents

DNA damage caused by carcinogens may alter cellular metabolism and ultimately result in cancer (*19*). The carcinogens are typically converted to their most reactive forms, which are called ultimate carcinogens, through metabolism. DNA adducts will then be produced from the reaction between the ultimate carcinogens and DNA. The DNA adducts will be repaired by DNA repair polymerases. If not repaired, damaged DNA lesions can be replicated by incorporating of wrong bases opposite damaged ones, which is known as mutation. When the mutation occurs in critical genes, expression of damaged

DNA leads to a mutant protein. DNA damage may also result in blockages of replication and/or cellular cytotoxicity and eventually increases the risk of likelihood of cancer formation (Scheme 1-2) (*19-21*).



Scheme 1-2. Schematic representation of chemical carcinogen induced cancer development.

The modifications of DNA alter the regular helical structure of DNA by introducing non-native chemical bonds or bulky adducts that do not fit in the standard double helix, resulting in mutagenicity and DNA structure perturbation. Ongoing studies of chemical carcinogenesis and chemistry have provided new methods of investigating DNA adducts and pathways involved in cell transformation and ultimately the onset of cancer.

DNA Adducts Derived from Metabolism of Butadiene

Chemical Carcinogenesis of Butadiene

1,3-Butadiene (CAS RN 106-99-0) (BD)[1] is the major material used in the manufacture of styrene-butadiene rubber (SBR) (*22*,*23*) with several billion pounds per year produced in the United States. It is also a combustion product from automobile emissions (*24*) and cigarette smoke (*25*). The potential health hazards of butadiene were first proposed in the 1970s. Considerable efforts have been made to investigate the metabolism, mutagenesis, genotoxicity and carcinogenesis of butadiene over the past 30 years, as described in the time line in Figure 1-2 (*26*).

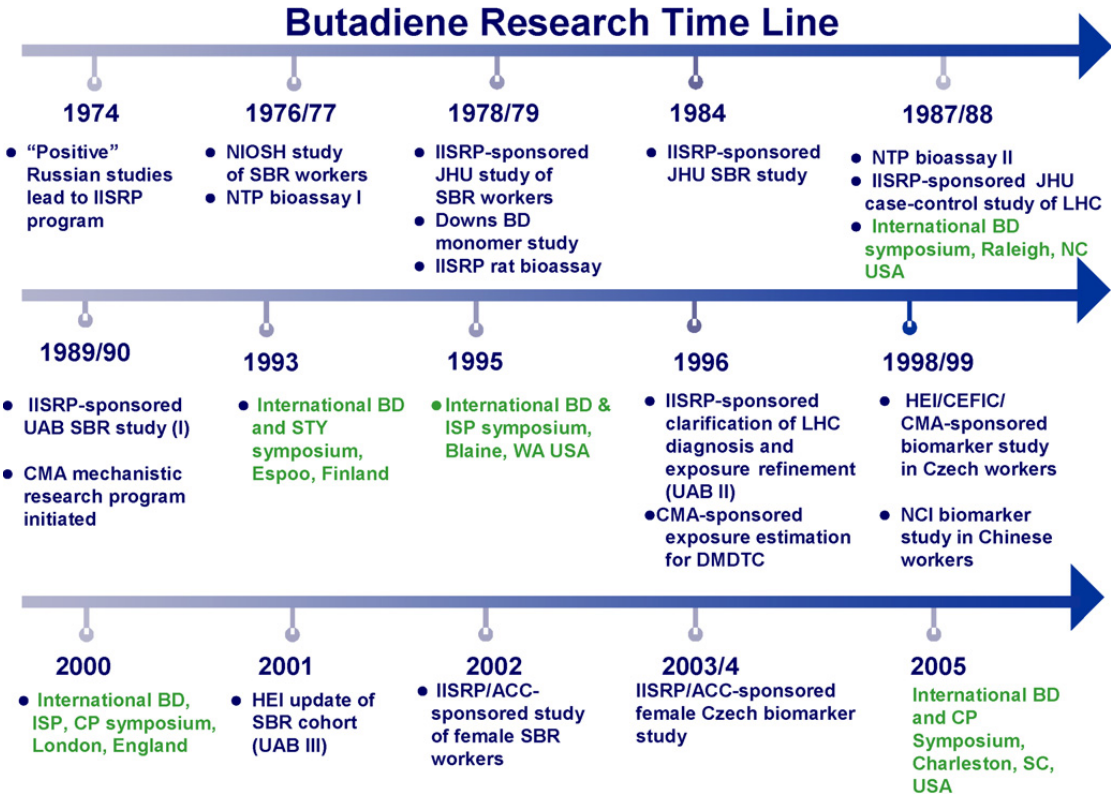
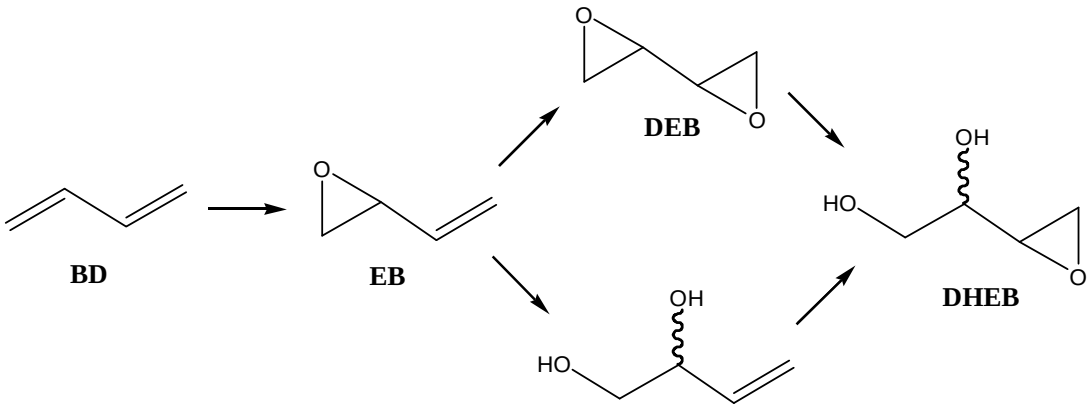


Figure 1-2. Butadiene research time line (Reprinted from ref. 26. Copyright of 2007 Elsevier).

It has been realized that BD is genotoxic and is a carcinogen in rodents, particularly in mice (*27-29*) and also in rats (*30*). BD was classified by the United States Environmental Protection Agency as "carcinogenic to humans by inhalation" (*31*). The International Agency for Research on Cancer (IARC) lists BD as a "probable human carcinogen" (Group 2A) (*23*,*26*,*32*). Chronic human exposure in the SBR industry may induce genotoxic effects (*33-35*) and is correlated with increased risk for leukemia (*22*,*36-41*). Although butadiene itself is not chemically reactive, butadiene epoxides arising from butadiene metabolism have been found to be very reactive with DNA bases.

Metabolism of Butadiene

The genotoxicity of BD has been attributed to the genotoxic and carcinogenic metabolites derived from oxidation of BD by cytochrome P450s. BD is epoxidized primarily by cytochrome P450 2E1, but also by cytochrome P450 2A6, to form 1,2-epoxy-3-butenes (EB) (*42,43*). These may be further oxidized by cytochrome P450 2E1 or 3A4 to form 1,2:3,4-diepoxybutanes (DEB) (*42,44-48*). Alternatively, hydrolysis of EB mediated by epoxide hydrolase forms 1,2-dihydroxy-3-butenes (*46,49,50*), which are metabolized by cytochrome P450 to hydroxymethylvinylketone (HMVK) (*51*). Either DEB or the 1,2-dihydroxy-3-butenes undergo cytochrome P450-mediated oxidation to form 1,2-dihydroxy-3,4-epoxybutanes (DHEB) (*46,49,52*) (Scheme 1-3). Thus, proximate electrophiles arising from BD metabolism include EB, DEB, and DHEB, and potentially, HMVK (*53*). Of all the reactive species, DEB is most highly genotoxic (*23,32,54*), probably due to its potential to form DNA-DNA (*55-58*) and DNA-protein cross-links (*59,60*), the latter which have been observed in mice (*61,62*). Mice have been shown to possess greater sensitivity to BD exposure than rats, and this is attributed to their efficient oxidation of EB to DEB (*47,63,64*), presumably facilitating DNA cross-linking. EB and DHEB are less genotoxic (*54,65,66*), however, more attention has been paid to DHEB as it is the most abundant metabolite of BD produced in humans (*67*). Butadiene genotoxicty is further enhanced in knockout mice lacking a functional microsomal epoxide hydrolase gene (*60*). Polymorphisms in the human epoxide hydrolase gene may also contribute to differences in BD genotoxicty within the human population (*68,69*).



Scheme 1-3. Reactive metabolites of 1,3-butadiene (BD).

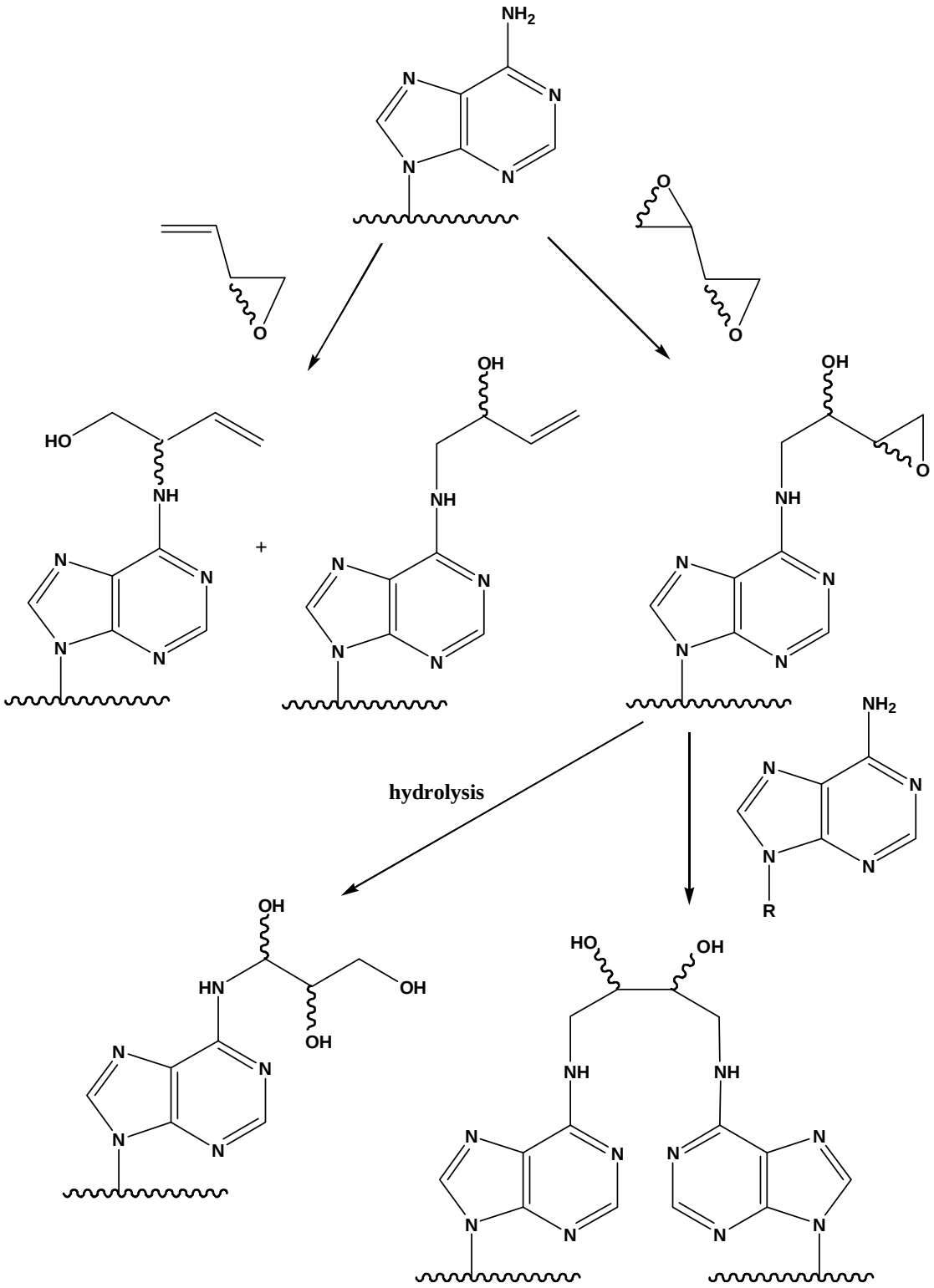
Butadiene-Derived DNA Adducts

All types of nucleoside bases of DNA have been shown to be react with butadiene epoxides leading to DNA mutation and carcinogenesis (*66*). The reaction of both EB and DEB with four nucleosides yields an abundance of DNA adducts including the N7, N^2, and N1 adducts of guanine (*70-74*), the N3 and O2 adducts of cytosine and thymine (*46,75,76*), and the N1, N^6, and N3 adducts of adenine (*70,77-80*).

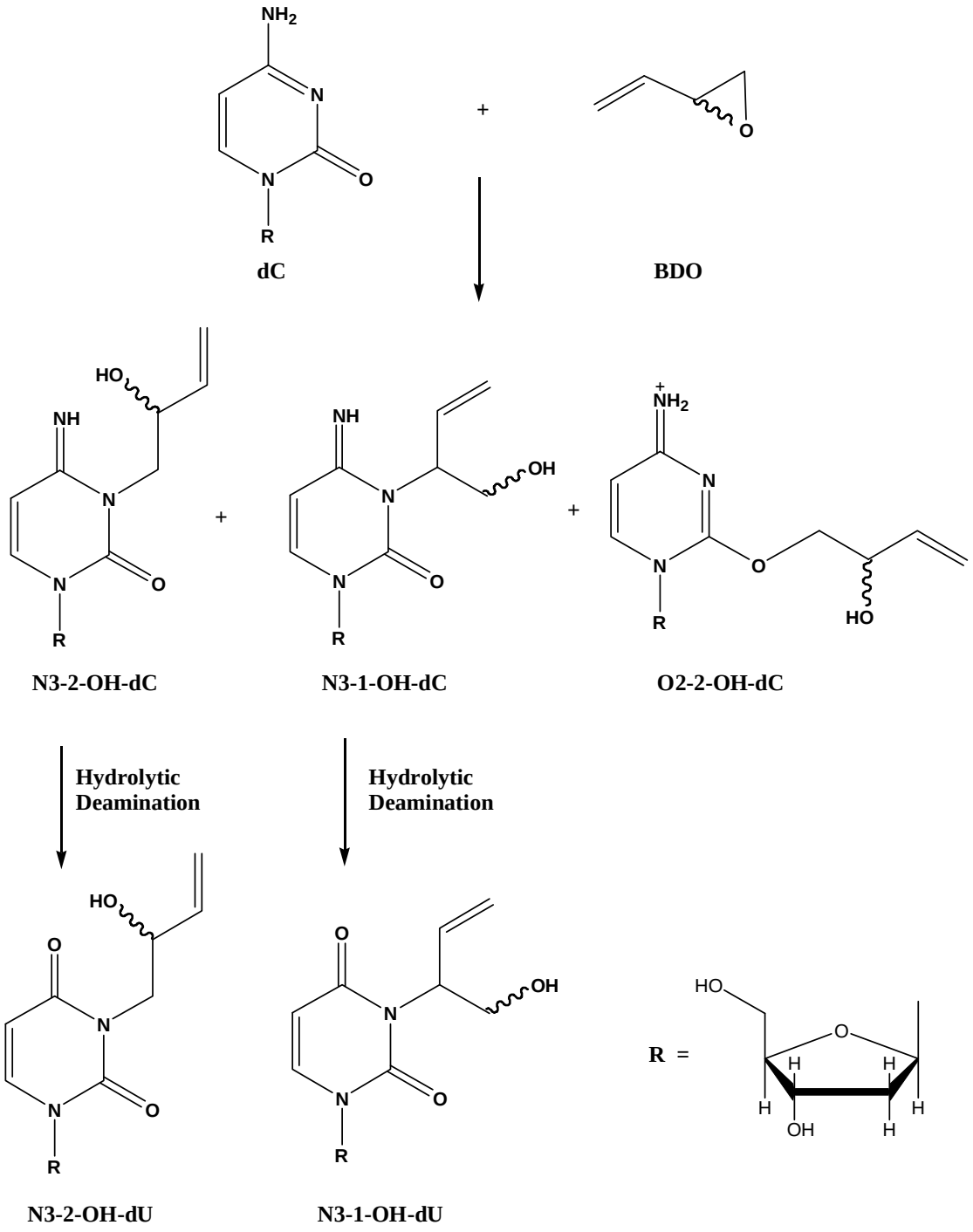
The genotoxicity of butadiene may be related to its ability to form cross-links in DNA via its diepoxide metablite DEB. The isolation of N7-N7 guanine cross-linked adducts from DEB-treated salmon sperm DNA suggested the presence of DEB-induced DNA cross-links *in vivo* (*55,56,58*). Due to the difficulty in the chemical synthesis of the N7-N7 guanine cross-linked adducts in large amounts, N^6-N^6 adenine intrastrand cross-linked adducts were made as model compounds for the structural study of BD-induced cross-links. Adenine N^6-N^6 intrastrand cross-linked adducts of butadiene can be chemically synthesized from the reaction between DEB and deoxyadenosine (Scheme 1-4) (*81*). Oxidation of BD by P450s finally leads to the (*R,R*)-, or (*S,S*)-, *meso*- diepoxide species (DEB) (*82*). All three optical isomers of DEB are genotoxic (*23,32,83*), and (*S,S*)-

DEB is the most cytotoxic and genotoxic one (*84,85*); it is believed to be the active form of the anti-tumor agent L-threitol-1,4-bismethanesulfonate (treosulfan) (*86-88*). The initial alkylation of deoxyadenosine by butadiene epoxides takes place at the N1 position. Dimroth rearrangement of the N1 adducts result in stereospecific adenine N^6 adducts, including (*S,S*)- or (*R,R*)-N^6,N^6-dA intrastrand cross-linked adducts (*89,90*).

The N3 position of cytosine is one of the highly nucleophilic sites which may react with alkylating agents. Although BD-induced N3-dC adducts are not detected in vivo, the chemically synthesized BD-derived N3-dU adducts deaminated from N3-dC adducts are determined to be strongly mutagenic (*91*). Therefore structural investigation of BD-induced N3-dU adducts may be of interest. Uridine N3 adducts of butadiene (BD-N3-dU) can be made from the reaction between EB and deoxycytidine (Scheme 1-5) (*75*). Diastereomeric pairs of O2- and N3-deoxycytidine adducts are formed initially. The O2-deoxycytidine adducts are unstable and decomposed quickly. The major adducts, N3-(2-hydroxy-3-buten-1-yl) deoxycytidine adducts, are also not stable and are deaminated to the corresponding N3-deoxyuridine adducts within a few hours. The N3-(1-hydroxy-3-buten-2-yl) deoxycytidine diastereomers are not detected, suggesting faster deamination rates compared to the N3-(2-hydroxy-3-buten-1-yl) deoxycytidine adducts. The final products of four N3-deoxyuridine adducts are extremely stable.



Scheme 1-4. Formation of N^6,N^6-dA intrastrand cross-linked adducts (Reprinted from ref. 92. Copyright of 2000 Wiley-Liss).



Scheme 1-5. Formation of BD N3-deoxyuridine adducts. (Reprinted from ref. 75. Copyright of 1997 Academic Press)

Biological Activity Studies of BD-Derived N^6 Adenine and N3 Uridine Adducts

Site-specific mutagenesis and *in vitro* replication are fundamentally important molecular biology techniques which provide direct biochemical information of mutations at a specific site in DNA and facilitate the investigation of the carcinogenetic effects of DNA mutations.

The site-specific mutagenesis experiments on the (*S,S*) and (*R,R*) stereoisomers of BD-derived N^6,N^6-deoxyadenosine intrastrand cross-linked adducts indicate that they are mutagenic in both *E. coli* and COS-7 cells (*92*). Replication of the (*S,S*) cross-linked adducted DNA in COS-7 cells yields 19% single base substitutions at the 3'-adducted base, with the primary mutation of A→G transitions (13%), followed by A→T (5.6%) and A→C transversions (0.8%). Using the (*R,R*) cross-linked adduct under the same experimental conditions, mutations at a level of 54% are detected, of which A→G transitions (40%) are predominant, followed by A→C (9%) and A→T mutations (5%). However, the level of mutations in *E. coli* is less than that observed in COS-7 cells with mutation frequencies of 3% and 8% of the (*S,S*)- and (*R,R*)- cross-linked adduct, respectively. The mutation frequency of the stereoisometric cross-linked adducts observed in both *E.coli* and COS-7 cells are illustrated in Figure 1-3 (*92*). Because N^6,N^6-dA cross-links (*92*) are significantly more mutagenic than N^6-dA monoadducts arising from EB or DHEB (*81*,*93*), their potential as rare but biologically significant lesions cannot be discounted.

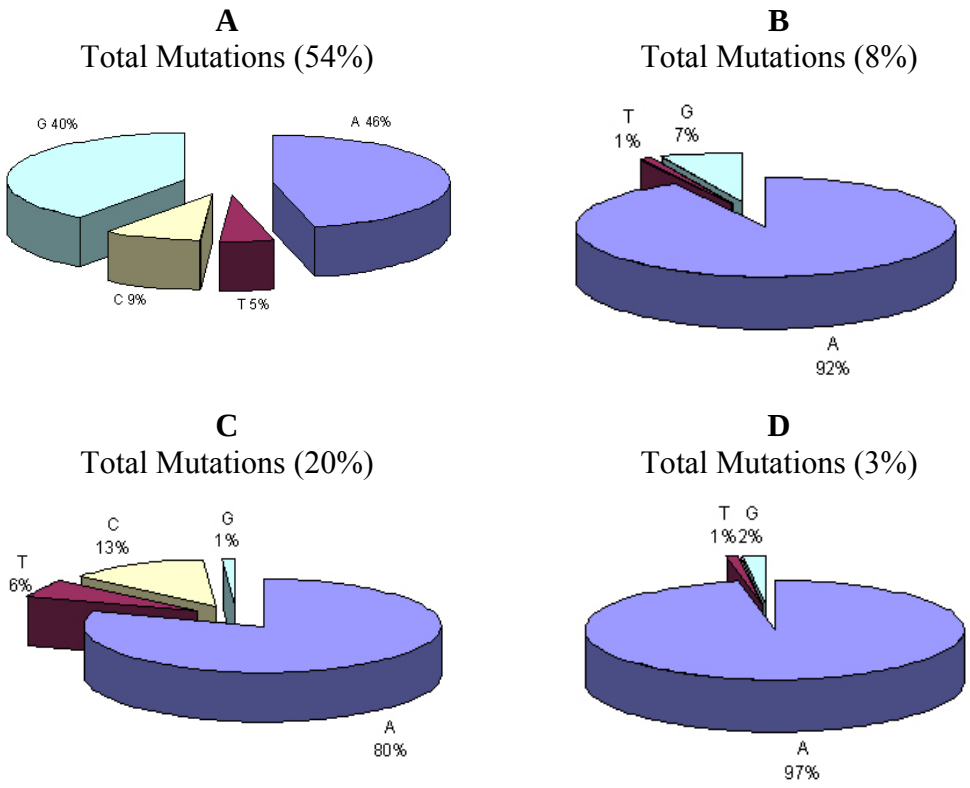


Figure 1-3. Mutation percentage of site-specifically modified ss pMS2 DNAs containing *R*,*R*- and *S*,*S*-butadiene cross-links. **A.** Mutations in *R*,*R*-butadiene replicated in COS-7 cells. **B.** Mutations in *R*,*R*-butadiene replicated in *E. coli* cells. **C.** Mutations in *S*,*S*-butadiene replicated in COS-7 cells. **D.** Mutations in *S*,*S*-butadiene replicated in *E. coli* cells. (Reprinted from ref. 93. Copyright of 2002 American Chemical Society.)

The stereoisomeric BD-derived N3-dU adducts are found to be highly mutagenic in mammalian COS-7 cells from the study of site-specific mutagenesis, with the predominant mutations of C→T transitions (53.4%) and C→A transversions (32.5%), followed by →G transversions (11%) (*91*). Fernandes et al. also performed *in vitro* replication studies, indicating that the BD-N3-dU lesions block replications of several replicative DNA polymerases significantly, such as bacterial Klenow fragment, mammalian pol δ and pol ε (*91*). The ability to bypass these adducts of Y-family trans-lesion polymerases was then examined by the same lab. Their investigations indicate

human polymerases ι, κ, and yeast pol ζ are significantly blocked by the BD-N3-dU lesions, while human pol η is able to incorporate nucleotides preferring G or A, opposite the adducts, which suggest that pol η is responsible for the induction of C→T mutations. The extension of primers containing an A incorporated opposite the adducts to full-length products can be achieved by pol η and pol ζ efficiently (*94*).

Structure Determination of Biological Macromolecules

Biological macromolecules, such as protein, DNA duplex, lipids, membranes and polysaccharides, are the principal non-aqueous components of living cells. To understand cellular processes, knowledge of three-dimensional structures of biological macromolecules is crucial. Two techniques are widely used for the structural determination of biological macromolecules at the level of distinguishing individual atoms: X-ray crystallography and nuclear magnetic resonance (NMR).

X-ray crystallography can be applied to various biological macromolecules with almost no molecular weight restriction to solve high resolution structures. This technique enables the distinction of two atoms in space as close as 2 Å apart. However, large, stable protein crystals, within which each molecule unit is lined up in a regular lattice, are required for this technique (*95*). It has also been recognized that some flexible structural parts are not able to be resolved by X-ray crystallography, whereas the development of NMR can be utilized to solve the problem. In contrast to X-ray crystallography, NMR does not require crystals and provides more detailed dynamics information of the molecule since it remains in solution during an NMR experiment. The dynamic information include interactions among large molecules and small ligands such as

substrates, cofactors, and drugs, but it is limited to small (<40 kDa), stable, soluble molecules without aggregation at the high concentrations (*96*). Due to the complementary roles in structural determination, a combination of both methods has proven to be powerful in elucidating DNA and protein structures (*97*,*98*).

Structural Studies of Oligonucleotides Using NMR Spectroscopy

The development of high magnet field NMR instruments, which currently can reach field strengths of 900 MHz, makes it increasingly significant for determining the solution structures of small size molecules such as oligonucleotides. NMR observes atomic nuclei of isotopes such as ^{1}H, ^{13}C, ^{15}N or ^{31}P. The isotopes are simulated with a strong, constant magnetic field. Then an orthogonal magnetic field was used to perturb the alignment of magnetic nuclei. Thus, the frequency of the nuclei during their relaxation to the initial state can be measured. The resonance frequency of the stimulated atomic nuclei, known as its chemical shift, depends on the distance and type of neighboring nuclei. In regards to biological molecules like DNA and proteins, NMR usually measures the spin of protons and records the chemical shifts of protons, as well as local environment information in NMR spectra. Because of many protons existing in a biological molecule, the chemical shift of each proton is hardly distinguished in 1-D NMR. Therefore, 2-D NMR has been developed to solve this problem.

^{1}H Nuclear Overhauser Effect spectroscopy (NOESY), which uses the dipolar interaction of spins for correlation of protons located within approximately 5 Å of each other, is one of the most powerful 2-D NMR experiments for the solution structure determination of oligodeoxynucleotides (Figure 1-4) (*99*). Multiple mixing times (τ_m)

ranging from 30 to 300 ms are typically used in NOESY experiments of oligodeoxynucleotides to address spin diffusion effects. When observing non-exchangeable protons, deuterated solvent is used for NOESY experiments to avoid excessive protic solvent signals. The assignment of exchangeable protons, such as amino and imino protons which relate to Watson-Crick base pairing can be obtained from a 2-D NOESY spectrum with sample dissolved in H_2O (*100-102*).

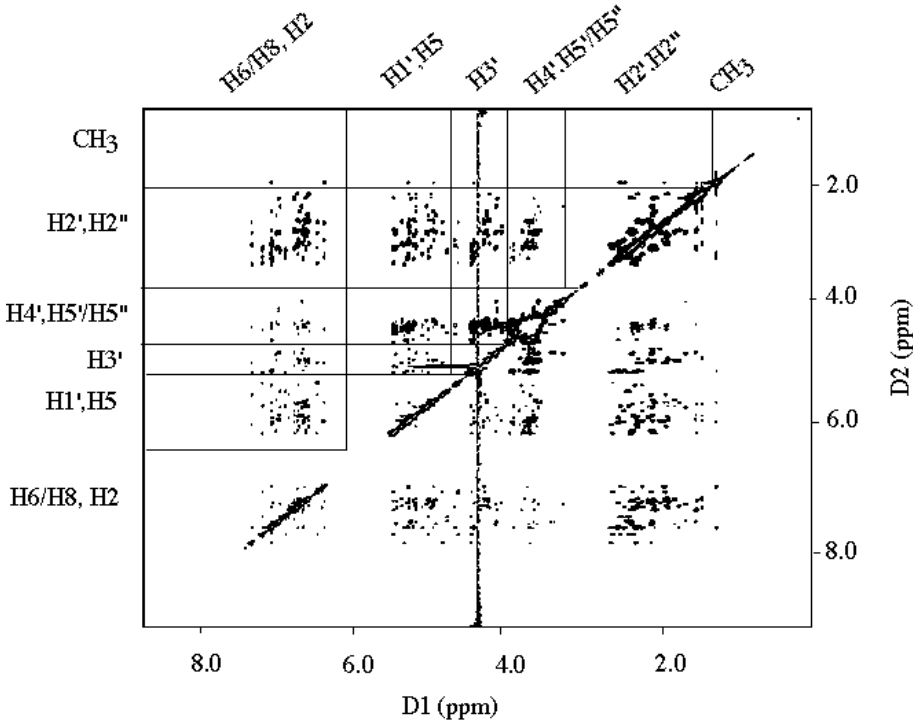
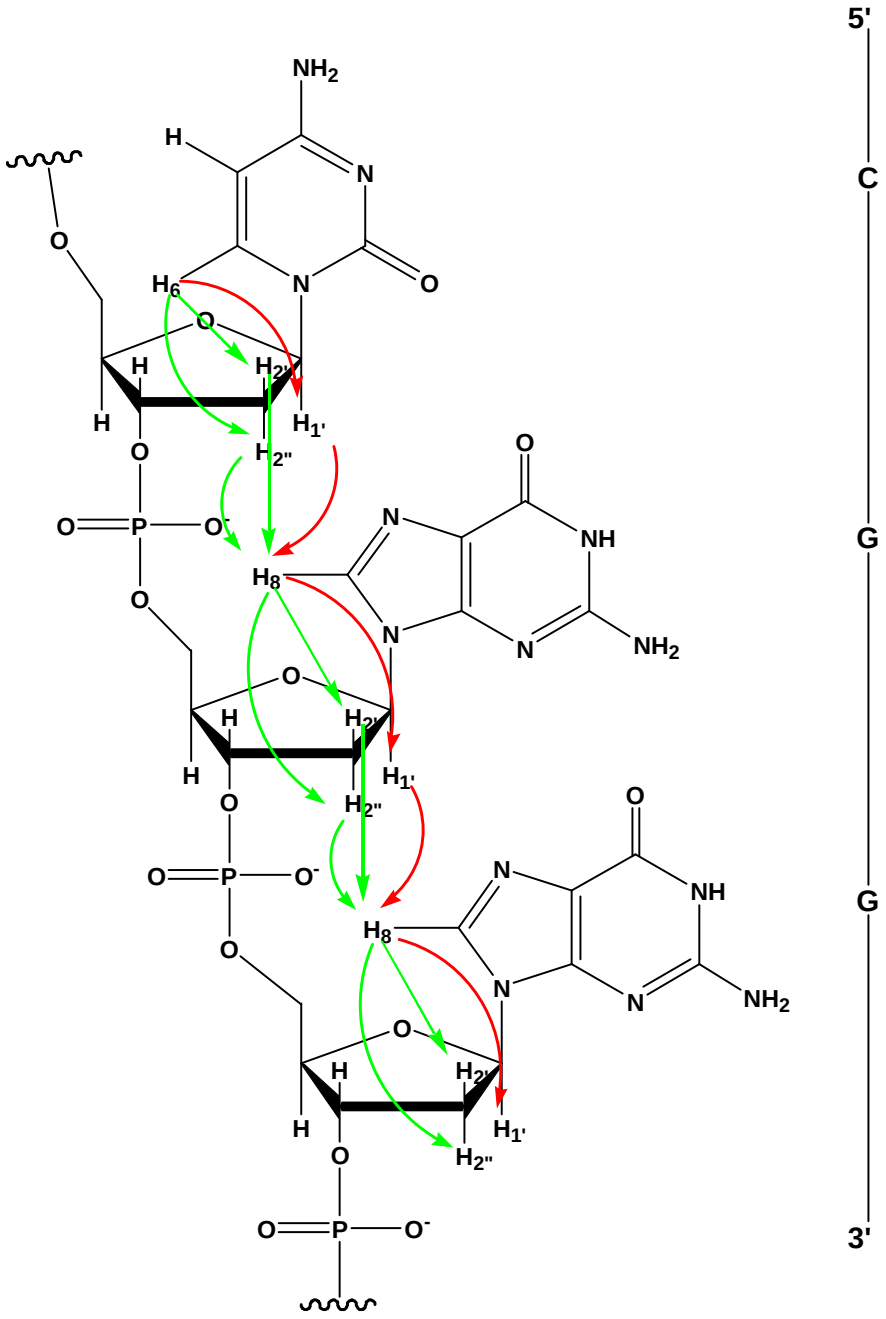


Figure 1-4. A typical ^{1}H-NOESY spectrum of oligodeoxynucleotides. Interactions, or nOes, are observed between groups of protons as noted on the left and top of the spectrum.

The sequential assignment of the non-exchangeable protons of a DNA duplex in an aqueous solution can be performed on spatially adjacent bases through the DNA strand (*103*). In the 5' to 3' direction, the intra-nucleotide ^{1}H-^{1}H distance between the aromatic H6 or H8 proton on the base of purine or pyrimidine, respectively, and its own H2'

anomeric proton allows for a cross peak. This is followed by a cross peak of the H2' proton to the H6 or H8 proton of its 3' neighboring base. In the same way, the inter-nucleotide distance between an H6 or H8 proton and its own H1' proton allows for a cross peak, followed by the H1' proton to 3'-neighboring H6 or H8 proton. This method of sequential assignment accomplished by "walking" through DNA duplex from 5' to 3' end is known as a NOESY walk (Scheme 1-6). The NOESY walk is particularly useful in observing the DNA damage-induced disruption on the sequential connectivity. It also provides chemical shift information, as well as the orientation of the damaged base. The sequential assignments of exchangeable protons can also be made from the cross peaks between the amino protons of guanine and the imino protons of thymine in a H_2O NOESY spectrum, following a similar “walk” assignment procedure as for the non-exchangeable protons.



Scheme 1-6. Sequential assignment (NOESY walk) of an oligonucleotide using NOESY spectroscopy.

In addition to NOESY spectra, 2-D correlation spectroscopy (COSY) (*104*) is also significant in determining molecular structures by correlating the chemical shift of ^{1}H nuclei which are J-coupled to one another. In a ^{1}H COSY spectrum, only the signals of protons which are less than three bonds apart are visible. In studies of the DNA duplex by COSY spectroscopy, the cross signals between the cytosine H5 and H6 protons are of special importance, as well as the methyl protons in thymine. The double quantum filtered COSY (DQF-COSY) (*105-107*) experiment has the advantage of partial cancellation of the diagonal peaks and elimination of strong signals from solvent protons that do not experience homonuclear J-coupling, as compared with a magnitude COSY spectrum. In magnitude COSY experiments, signals with small J couplings and broad lines will show huge diagonal signals, but only very small or vanishing cross peaks. However, in the DQF-COSY, both the cross and diagonal peak intensities depend on the size of the coupling constants, making it easier to observe cross peaks between signals which are close together in chemical shift. From the cross peaks assigned in a DQF-COSY spectrum, J-coupling constants can be calculated. J-coupling constants are particularly useful in generating structural restraints such as the pucker of the 5-membered sugar ring of oligodeoxynucleotides in DNA structural studies (Figure 1-5) (*108*). ^{31}P-^{1}H Hetronuclear COSY experiments (*109,110*) are also useful in the determination of DNA duplex structure, particularly in determining the ε and ζ backbone angles (Figure 1-5) from experimental data. This is accomplished by using a Karplus relationship (*111-113*), which has the form:

$$^{3}J = A \cos(\theta) + B \cos^{2}(\theta) + C$$

where A, B, and C are empirically derived constants for each type of coupling constant.

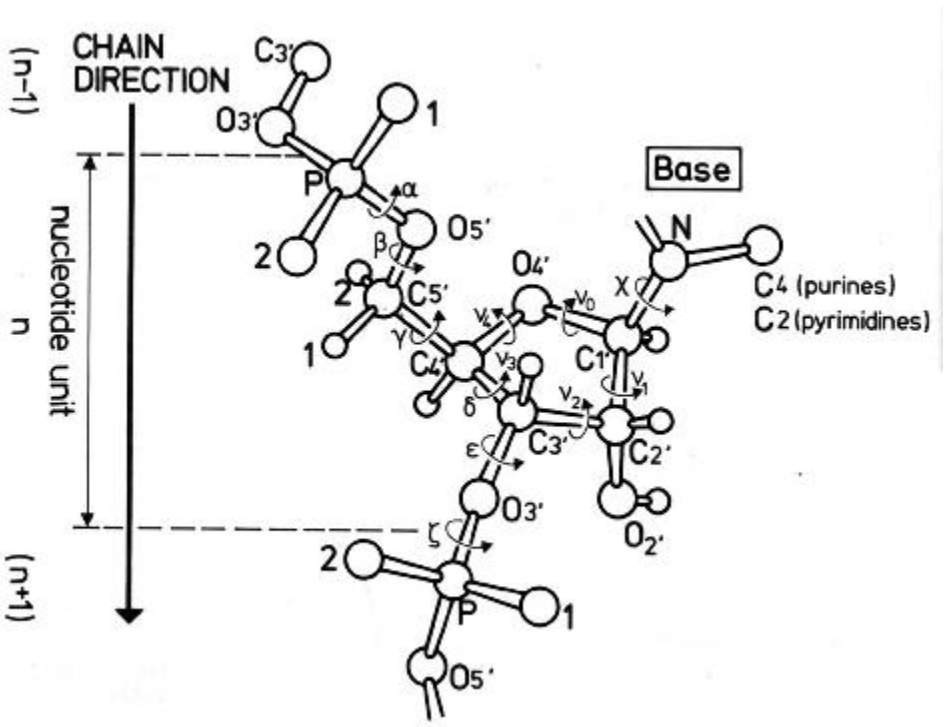


Figure 1-5. Backbone torsion angles in nucleic acid structures. (Reprinted from ref. 109. Copyright of Springer-Verlag New York, Inc. 1984)

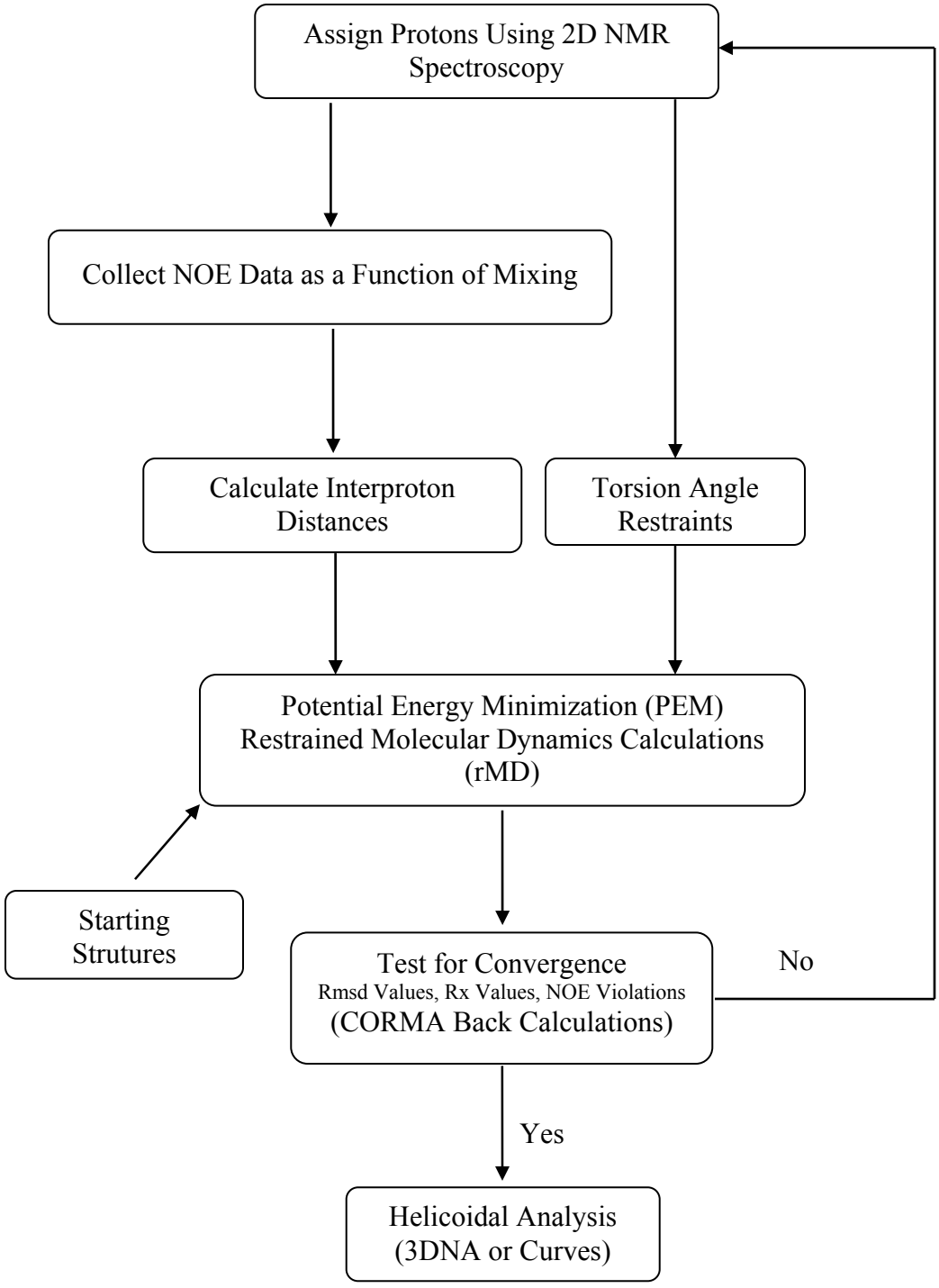
In the magnitude COSY experiment, protons more than three chemical bonds apart usually give no cross signal. Thus, the total correlation spectroscopy (TOCSY) experiments are developed for detection of the interactions of all protons with a spin system that are not directly connected via three chemical bonds, resulting in a successive scalar coupling of a complete spin system. The scalar coupling range observed from TOCSY experiments depends on the mixing time. For small mixing period, short-range proton-proton couplings can be observed. As the mixing period gets longer, correlation with longer distant protons can be observed, so the extent of correlation depends mainly on the length of the mixing period.

Once NMR data are properly collected and analyzed, solution structure of DNA duplex can be acquired by applying restrained molecular dynamics (rMD) calculations

containing proton-proton distances and conformational restraints. The typical strategy for solution structural determination of oligodeoxynucleotide is outlined in Scheme 1-7. Once the assignment of all of the cross peaks in the NOESY spectrum is completed, the program Matrix Analysis of Relaxation for Discerning Geometry of an Aqueous Solution (MARDIGRAS) (*114-116*) can be applied to calculate proton-proton distances and error bounds from cross peak intensities. The intensity matrix is converted to a relaxation rate matrix which is further used to calculate distances. The error bounds are estimated by adding noise and relative errors to the intensities converted from the final rates assuming a two-spin approximation. In order to eliminate the spin diffusion effect at long mixing times, complete relaxation matrix analysis (CORMA) is implemented to derive inter-proton distances while accounting for spin diffusion (*117*). DQF-COSY data yield dihedral angle restraints extracted from vicinal proton-proton coupling constants.

With all of the restraints applied, a three-dimensional DNA structure can be refined from potential energy minimizations (PEM) and rMD calculations. An initial A- or B-form starting structure, together with molecular topology and parameter information is refined with the PEM calculation. This relaxes the molecule until it reaches the local energy minimum point without making drastic changes in the structures. The energy minimized structure is then refined by rMD calculates using a simulated annealing method. The simulated annealing refinement method heats up the system to allow the molecule of interest to overcome energy barriers. When the system is cooled gradually, the structure reaches the most stable conformation with the lowest energy. The root mean square deviation (rmsd) of the structures represents the preciseness of the rMD calculations. The accuracy of the solution structure obtained from the rMD calculations is

determined from the residual factor, or R-factor, which is calculated by comparing the calculated NOE intensities to the experimentally determined intensities using the program CORMA (*117-119*). The helicoidal analysis of final refined structure can be carried out by using the 3DNA program (*120*) and Curves (*121,122*).



Scheme 1-7. The strategy for the NMR-generated structural refinement of the oligodeoxynucleotide.

Trans-Lesion DNA Synthesis Polymerases (Y-family)

Based on sequence homology, over 300 new DNA polymerases have now been identified in bacteria, archaea, and eukaryotes forming a new Y-family of DNA polymerases. These polymerases share significant sequence identity and similarity amongst themselves, but exhibit little sequence homology to previously identified A, B, C, D and X-families of DNA polymerases (*123-125*). Y-family polymerases are characterized by the ability of trans-lesion synthesis and low-fidelity synthesis on normal DNA with error rates of 10^{-2} to 10^{-4}, which is about 1–2 orders of magnitude higher than those of high-fidelity replicative polymerases (*123,126*).

The Y-family DNA polymerases contain 350 to 800 amino acid nucleotides, of which the N-terminal 250 to 350 amino acid nucleotides form the palm, thumb, and finger domains (*127,128*). The C-terminal nucleotides are used for nuclear localization (*129*) and for interaction with processivity factor (e.g. PCNA) (*130-132*) or other polymerases (*133,134*). Although they are in the same family, each polymerase differs in substrate specificity, that is, the type of lesions bypassed and mutation generated (*127*). For example, Pol IV and Pol κ bypass abasic and bulky DNA adduct lesions forming base-substitution and frameshift mutations (*135-138*), yeast Pol η can bypass a TT dimer and extend primer successfully (*139*), and human Pol ι violates the Watson-Crick base-pairing rule by incorporating G opposite T, leading to less efficiency in further extension (*140,141*). Thus, each organism contains several Y-family polymerases with different substrate preferences and trans-lesion synthesis efficiencies (*127*), such as pol IV and pol

V in *E. coli*, pol η and Rev1 in *S. cerevisiae*, and four polymerases of pol η, pol ι, pol κ, and Rev1 in humans (*123*,*142*).

The crystal structures of Y-family polymerases reported to date include archaeal *Sulfolobus solfataricus* P1 DNA polymerase Dbh (PDB code 1K1Q) (*134*,*143*) and P2 DNA polymerase IV (Dpo4) (PDB code 1JX4) (*144*), the N-terminal catalytic domain of *Saccharomyces cerevisiae* Pol η (PDB code 1JIH) (*145*), human Pol ι (PDB code 1T3N) (*146*) and human Pol κ (PDB code 1T94) (*147*), alone or complexed with DNA and a nucleotide substrate. These crystal structures reveal a conventional right-hand-shaped catalytic core consisting of palm, finger and thumb domains similar to replicative DNA polymerases and an additional C-terminal domain (Figure 1-6) (*124*). The palm domain contains three catalytically essential carboxylates and coordinates two metal ions to carry out the reaction between the 3'-hydroxyl group of a primer strand and the α-phosphate group of an incoming nucleotide. The unique C-terminal domain of the Y-family polymerases has been named as the 'little finger' domain (*144*), which is also known as the 'wrist' in Dbh (*143*) or the polymerase associated domain (PAD) in Pol η (*145*). The flexible little finger and thumb weaken the association of the polymerase and DNA substrate by altering the positions of the DNA substrate relative to the catalytic core, thus reducing the catalytic efficiency and increasing frameshift mutations (*148-151*).

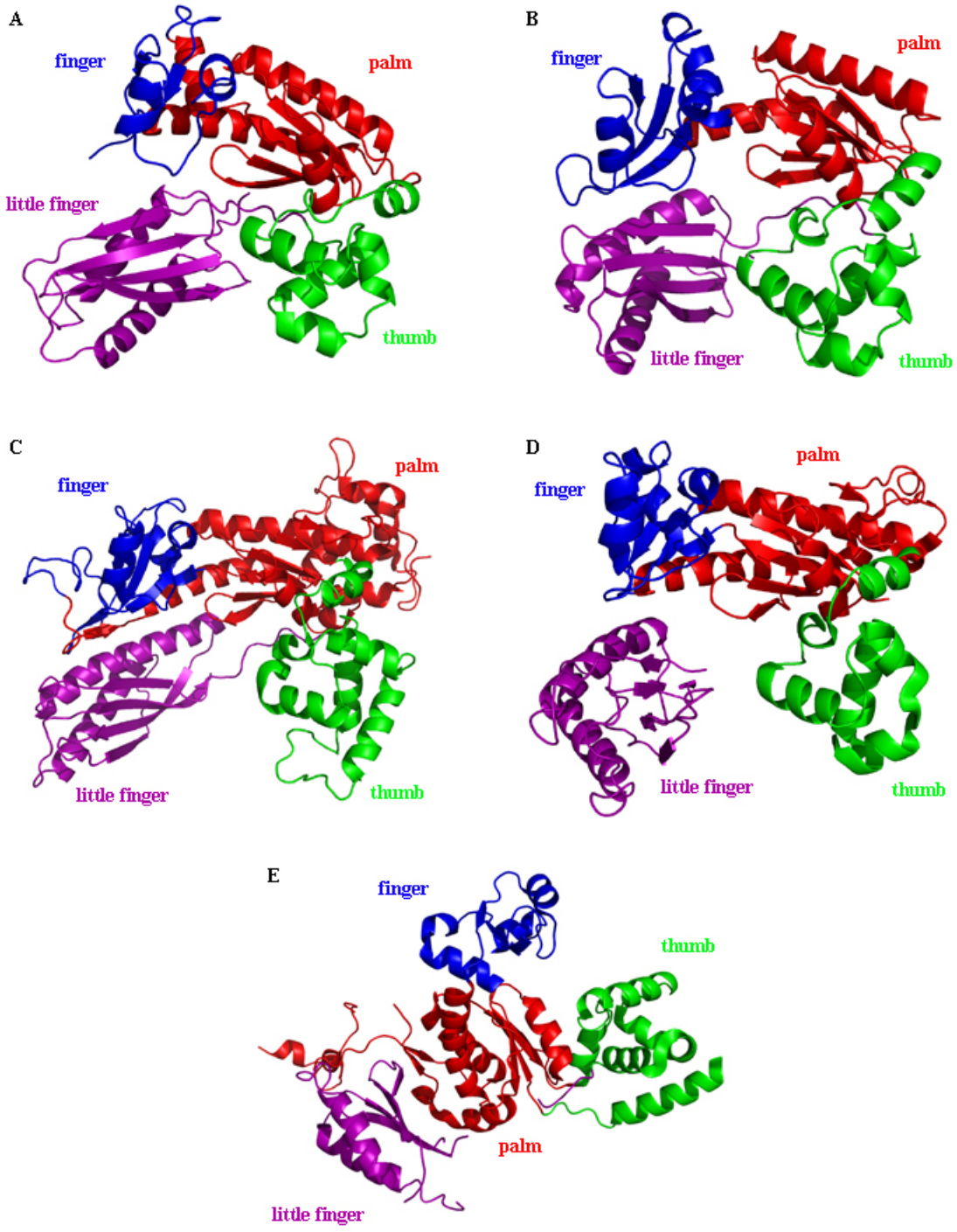


Figure 1-6. Comparison of the Y-family polymerases (A) Dhb (PDB code 1K1Q), (B) Dpo4 (PDB code 1JX4), (C) Pol η (PDB code 1JIH), (D) Pol ι (PDB code 1T3N) and (E) Pol κ (PDB code 1T94). The palm, finger, thumb and little finger domains are shown in red, blue, green and purple, respectively.

In general, Y-family polymerases have less restrictive active sites based on the smaller finger and thumb domains compared with replicative polymerases, and are therefore better able to accommodate bulky lesions and mismatched base pairs (*124*). High fidelity synthesis of the replicative DNA polymerases, such as A- and B-family polymerases, is achieved through an "induced-fit" conformational change of the finger domain to prevent a wrong incoming nucleotide so that only Watson–Crick base pairs are able to be accommodated at the active site (*152-155*). In contrast to such an "induced-fit" screening of incoming nucleotides, the flexible little finger domain and open, spacious active site of the Y-family polymerase accept a damaged or mismatched replicating base pair easily and make it difficult to align DNA, metal ions and the incoming nucleotide (*144,156*), which helps explain the high-error-rate and low-fidelity of DNA synthesis by Y-family polymerases (*157,158*).

X-ray Crystallography

X-ray crystallography has been used for structure determination of proteins and other biological molecules since the late 1950s, beginning with the first crystal structure of sperm whale myoglobin by Kendrew and Perutz, for which they were awarded the Nobel Prize in Chemistry in 1962 (*159,160*). Dorothy Hodgkin obtained the Nobel Prize in Chemistry in 1964 for her determinations of the structures of important biochemical substances, including cholesterol (1937), vitamin B_{12} (1945) and penicillin (1954) by X-ray techniques (*161*). As the development of the crystallographic technique and the introduction of sophisticated computer hardware and software, the essential features of the mechanisms of more and more enzymes, such as lysozyme (*162,163*), ribonuclease

(*152,159*), chymotrypsin (*164-167*), carboxypeptidase (*168-170*) and alcohol dehydrogenase (*171*) are becoming known in detail.

In X-ray crystallography, the growth of crystals of highly purified and soluble protein is the first step in structural determination and most likely the rate limiting step in protein crystallographic study. Over the past decade the theoretical and practical aspects of the crystallization of macromolecules, like proteins, DNA and RNA have been studied in detail (*172*). Protein crystallization experiments typically use the highly concentrated protein which is usually in a superstaturated state, and are allowed to proceed for weeks or even months. Figure 1-7 shows the pathway of the crystal growth through the phase diagram by vapor diffusion, a very useful method in crystallization (*173*). In a vapor diffusion experiment, equal volumes of precipitant and protein are mixed and added into the drop, and then equilibrated against a large reservoir of solution containing precipitant or another dehydrating agent. When the mixture of protein and precipitant reaches the nucleation zone, as the vapor diffuses from the reservoir to the drop, the first crystal nuclei are formed. Protein concentration in solution becomes more dilute when it is crystallized, and the path moves down to a metastable zone where it is suitable for the growth of crystals. Crystallization conditions have to be optimized to obtain the desired crystals, the size of which is typically between 20 and 300 μm along each edge.

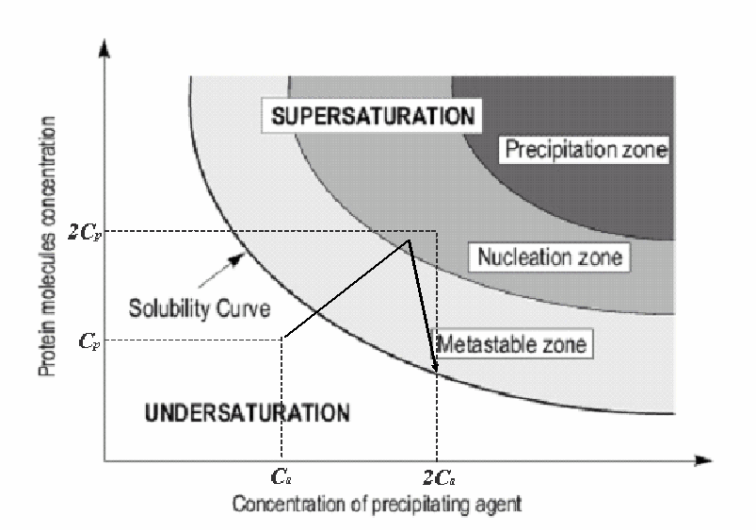


Figure 1-7. Phase diagram for vapor diffusion experiment.

When crystals of a suitable size are obtained, they are frozen in liquid nitrogen and are ready to be analyzed by X-rays. X-rays, the high energy electromagnetic radiation having a wavelength of 0.2 Å to 2.0 Å, can be generated from accelerating electrons in a synchrotron storage ring or from electrons striking a copper anode and then be focused into a beam and collimated with sets of adjustable slits to a 0.1–0.3 mm diameter to ensure that the beam is parallel.

The crystal is mounted in a device known as a goniostat which allows the crystal to be positioned appropriately and kept in the X-ray beam when rotated. Before X-ray analysis starts, the distance from the crystal to the detector is calculated and adjusted based on the desired resolution. This is because the resolution of spots collected on the detector increases with the increasing of the diffracted angles. The relationship between diffracted angle and the distance between the planes is given by Bragg's law:

$$2d \cdot \sin\theta = n\lambda$$

where *n* is an integer determined by the order given; λ is the wavelength of x-rays; *d* is the spacing between the planes; and θ is the diffracted angle between the incident ray and the scattering planes. Therefore, once the diffracted angle is determined, the distance of the detector from the crystal can be adjusted accordingly.

X-ray diffraction data can be collected from experiments either on home-source generators or at synchrotron sources. The much more highly intense X-ray beam of synchrotron sources allows shorter exposure times and a higher signal to noise ratio of the diffraction spots, which makes it more favored in crystallographic studies nowadays (*174*). The application of detectors using charged coupled device (CCD) technology also significantly reduces time for collection of a complete data set from a single crystal (*175,176*). As the diffraction spots become weaker at a higher resolution, a compromise between increased resolution and decreased diffraction quality has to be made. Once the sufficient high-resolution data are collected and processed, the unit cell dimensions, the crystal system, and the space group of the crystal can be determined. The unit cell of a crystal is represented by three lengths: a, b, and c; and three angles: α, β and γ. The crystal system is determined by the shape of the unit cell and a totally of seven crystal systems exist (Table 1-3). The space group is determined by the symmetry of the diffraction pattern. A total of 230 space groups exist, but only 65 for chiral molecules such as proteins.

Table 1-3: Seven crystal systems.

Crystal system	Conditions imposed on cell geometry	No. of space groups
Triclinic	$a \neq b \neq c$, $\alpha \neq \beta \neq \gamma$	2
Monoclinic	$a \neq b \neq c$, $\alpha = \gamma = 90$ º, $\beta \neq 90$ º	13
Orthorhombic	$a \neq b \neq c$, $\alpha = \beta = \gamma \neq 90$ º	59
Tetragonal	$a = b \neq c$, $\alpha = \beta = \gamma = 90$ º	68
Rhombohedral	$a = b = c$, $\alpha = \beta = \gamma \neq 90$ º	25
Hexagonal	$a = b \neq c$, $\alpha = \beta = 90$ º, $\gamma = 120$ º	27
Cubic	$a = b = c$, $\alpha = \beta = \gamma = 90$ º	36

For structure determination, the quantity of data required and strategy for data collection depend on many factors, including properties of the unit cell, the symmetry of the crystal (*177*). When exposed in the X-ray beam, the crystal is rotated on a spindle perpendicular to the beam, hence a series of X-ray diffraction images, each corresponding to 0.5-1 degree of rotation, can be collected to a total crystal rotation of 45 to 180 degrees. The spacing and intensities of the spots are the most significant information contained in the diffraction pattern. The spacing of the spots is dependent on the size and shape of the unit cell in the crystal. The intensity is determined by the amplitude of the diffracted waves and by the phase difference. The amplitudes can be mathematically calculated from the square root of the intensities. However, the information of the phase is lost, and is therefore referred to as the phase problem.

The three most frequently used methods devised to overcome the phase problem are molecular replacement (MR) (*178*), multiple isomorphous replacement (MIR) (*179*,*180*) and multiple wavelength anomalous dispersion (MAD) (*181*). MR is the simplest and most frequently used when a homologous molecule structure is available. It is particularly valuable for solving the structure of a drug-protein complex when the

structure of the pure protein has been previously determined. In MR, with the use of the coordinate file and phases from the known model structure and the intensities from the experimental diffraction data, the electron density map of the new molecule can be calculated. Then the map is used to determine the structure, which is used to simulate a new set of phases.

MIR is normally used in cases where no closely related protein structure is available. It requires at least three data sets: one data set from the unmodified protein and two data sets from the protein with different heavy atoms attached, such as Hg, Pt, or Au. One or a few heavy atoms are preferred to attach to the protein molecule so that it will not appreciably alter either the conformation of the protein or the unit cell dimensions. The diffraction spots originating from the heavy atoms and the changes in the spot intensities of the diffraction pattern allow determining the phases.

MAD is a more recently developed technique and is a particularly powerful and popular method for obtaining the phase information, since it requires only a single crystal. Like the MIR technique, the single protein crystal to be studied must contain some type of heavy atom, which will cause significant anomalous scattering. The most commonly used heavy atom for phase determination via MAD is selenium, since it is usually possible to replace the natural sulfur containing amino acid methionine with selenomethionine. The MAD technique utilizes different wavelengths of X-rays to adjust the strength of the anomalous scattering. Thus the heavy atoms can act as reference markers and alter the intensity of spots in the diffraction pattern, which allow the position of the heavy atoms to be determined and the phase to be assigned.

The resulting phases and originally observed intensities of diffracted spots can be converted to the density map of the molecule by performing Fourier transform. The electron density map will form the three dimensional contours where the model can be built. Once the starting model is built, the phase can be calculated. Then the calculated phase, together with the original spot intensities, can be used to rebuild an improved electron density map and a new model. This cyclic processing continues until no further significant improvements are made. The improvement of the model from each cycle is judged by comparing observed with calculated wave amplitudes, which is given as R-factor. For most proteins of typical size, an R-factor between 0.15 and 0.25 is quite satisfactory. Another factor to evaluate the quality of refinement is $R_{free,}$ which represents the difference between randomly selected 5-10% of the original data that is not used in the calculation and the calculated data derived from the refined structure. R_{free} below 0.32 is acceptable for most protein structures (*96*).

In conclusion, the strategy to obtain a crystal structure of the molecule by using X-ray crystallography is illustrated in Figure 1-8 (*96*). A purified and highly concentrated macromolecule of interest such as a protein is crystallized. The crystal is then exposed to an X-ray beam and the diffraction data are collected. The intensity of the spots and the phases are obtained by processing of the diffraction patterns. Hence, an electron density map can be calculated and the model structure can be built and refined to fit the electron density map.

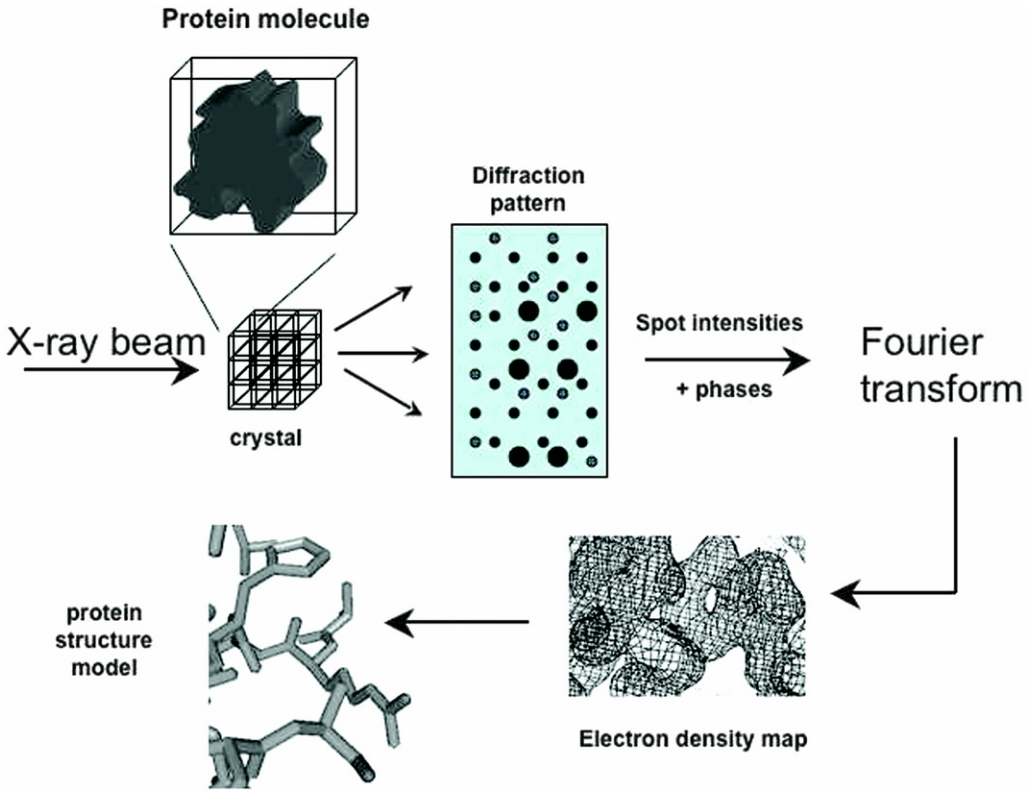


Figure 1-8. Structure determination by using X-ray crystallography. (Reprinted from ref. 97. Copyright of 2005 Lippincott Williams & Wilkins.)

Crystal Structures of Dpo4-DNA Complexes

Due to the excellent thermostabilty and high yield from purification, Dpo4 is well characterized on its structural features and serves as a model in understanding the low-fidelity and trans-lesion synthesis mechanisms of Y-family polymerases.

Crystal structures of Dpo4 in ternary complexes containing undamaged DNA and an incoming nucleotide (Figure 1-9) were first solved by the Yang group (PDB code 1JX4) (*144*). The catalytic core of Dpo4 contains the thumb, palm, and finger domains similar to high-fidelity DNA polymerases, and a unique little finger domain located at the C-terminus and tethered to the thumb domain. Two forms (Figure 1-10) of crystal structures of Dpo4 ternary complexes reported include normal replication by insertion of an

expected incoming nucleotide opposite template base, referred to as type I; and template misalignment by introduction of a -1 frameshift mutation, referred to as type II (*144*).

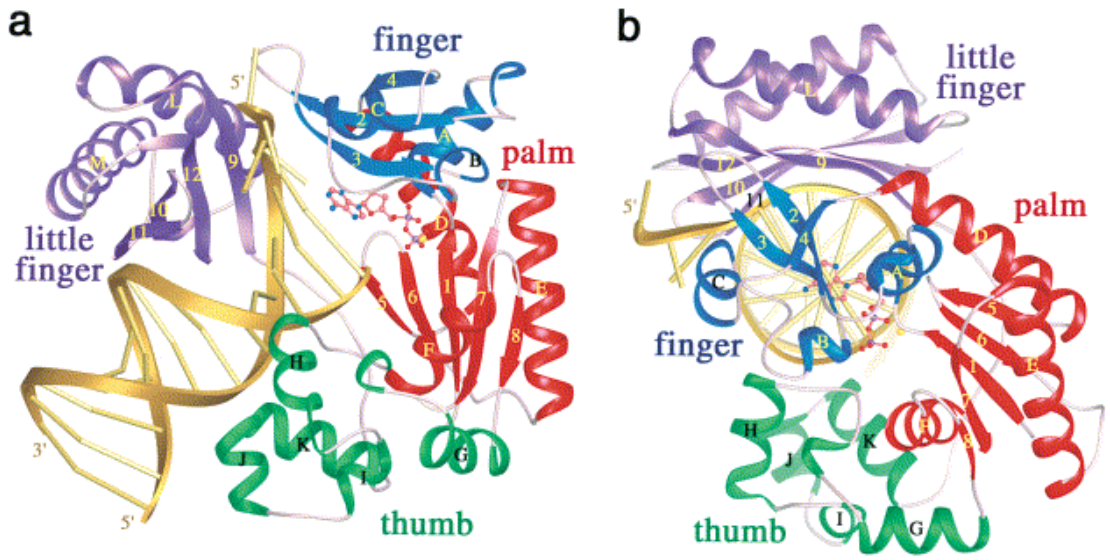


Figure 1-9. Crystal structure of Dpo4 complexes containing undamaged DNA and an incoming nucleotide. (a) A view looking into the active site with the palm, finger, thumb, and little finger domains well separated. (b) A view down the DNA-helical axis. (Reprinted from ref. 145. Copyright of 2001 Cell Press)

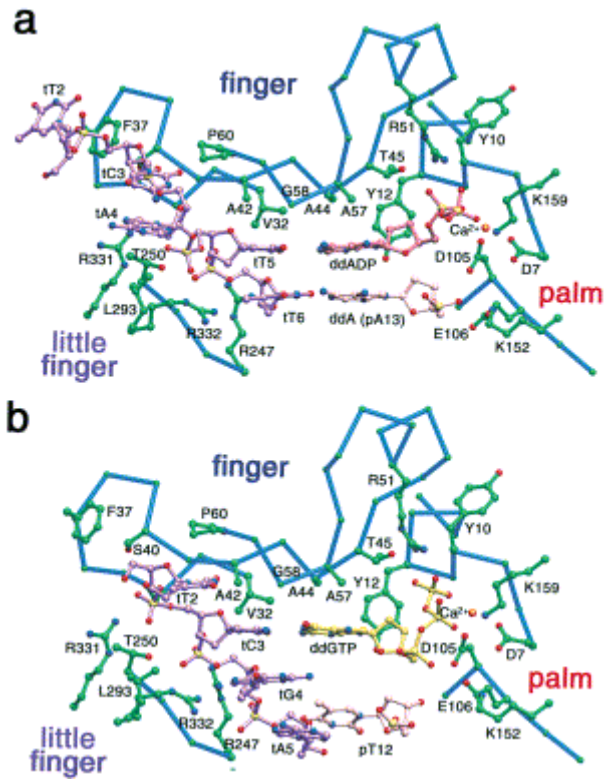


Figure 1-10. Active sites of Dpo4 complexes. (a) type I (normal) (b) type II (template misaligned). (Reprinted from ref. 145. Copyright of 2001 Cell Press)

Recently, a series of crystal structures of Dpo4-DNA complexes containing modified DNA templates have been reported, including UV cross-linking *cis-syn* thymine dimer (TT) (*148,182*); abasic lesions (*183*); benzo[*a*]pyrene diol epoxide (BPDE) adenosine adduct (*184*); 1*S*-(-)-*trans-anti*-B[*c*]Ph-N^6-dA adduct (*185*); 1,N^2-ε-guanine (*186*); 7,8-dihydro-8-oxoguanine (8-oxoG) (*149,187,188*); *N*-(deoxyguanosin-8-yl)-2-acetylaminofluorene (AAF-dG) (*189*); N^6-adenyl PAH-diol epoxide adducts (*190*); benzo[*a*]pyrene N^2-dG adducts (*191,192*); 2-amino-1-methyl-6-phenylimidazo[4,5-*b*] pyridine guanine adduct (*193*); O^6-methylguanine (O^6-MeG) (*194*); O^6-benzylguanine (O^6-BzG) (*195*); 2,4-difluorotoluene (DFT) analog of thymine (*196*) and 1,N^2-propanodeoxyguanosine (PdG) (*197*). Pre-steady-state kinetic and *in vitro* analysis have

measured fidelity and proposed possible catalytic mechanism of Dpo4 (*198-202*). The smaller finger and thumb domains, compared with the corresponding domains of the A- and B-family DNA polymerases, and the flexible little finger domain in Dpo4 result in a more solvent-accessible and less geometrically restrictive active site. Thus the interactions between active site and replicating base pair are reduced significantly. Consequently, the spacious active site and flexible little finger domain enable Dpo4 to accommodate and bypass various DNA lesions and mismatched base pairs in their active sites, causing the bypass and trans-lesion synthesis of Dpo4 to be accompanied with mutations and frameshifts (*156,203*).

Dissertation Statement

The goal of this research project is to determine the structures of site- and stereo-specific butadiene-derived adducts, including the adenine N^6-N^6 intrastrand cross-linked adduct and BD-N3-dU adducts in oligodeoxynucleotides containing the *ras*61 (5'-CXA-3') protooncogene sequence and to investigate the conformational perturbations of DNA structures induced by the butadiene adducts. The underlying goal is to correlate the structural features of butadiene-derived adducts in a biologically significant DNA sequence with mutagenesis and replication studies for better understanding of the biological processing of these adducts. The first hypothesis of this project is that the DNA cross-links derived from DEB are important mediators of the genotoxicity and, by determining the structure of a N^6-N^6 dA cross-link, it will be possible to examine how such a cross-link alters the structure of DNA. The second hypothesis is that determining the structures of stereoisomeric BD-derived N3-dU adducts will help to examine how the

adducts alter the structure of DNA. The third hypothesis is that the crystal structures of complexes between BD-N3-dU adducts and Dpo4 will help to explain how these adducts are bypassed.

Chapter II describes the materials and methods. Chapter III reports the solution structural study of the (*S*,*S*)-N^6-N^6 intrastrand dA cross-linked adduct in duplex DNA by using NMR. Chapter IV details the crystallographic study of ternary complexes containing Dpo4 and DNA modified with *R*- or *S*-BD-N3-dU adducts and a correct or incorrect incoming nucleotide to elucidate the change of the complex structures caused by the presence of the butadiene adducts. Chapter V reveals the crystallographic structures of binary complexes containing Dpo4, *R*- or *S*-BD-N3-dU adducted template and primer with A opposite the adducted lesion and ternary complexes with the addition of dGTP into the binary complexes described above. The structural difference of active sites induced by stereoisomeric butadiene adducts and different types of incoming nucleotides, as well as the implications of the lesion bypass processing of Y-family polymerases from structural studies are discussed. Chapter VI concludes the work described in this dissertation and discusses the future direction relevant to this project.

CHAPTER II

MATERIALS AND METHODS

Sample Preparation

The unmodified oligodeoxynucleotides were synthesized by the Midland Certified Reagent Co. (Midland, TX) and purified by reverse-phase high performance liquid chromatography (RP-HPLC). The 1,4-bis(2'-deoxyadenosin-N^6-yl)-2*S*,3*S*-butanediol intrastrand cross-linked oligodeoxynucleotide ((*S*,*S*)-BD-(61-2,3) cross-link) was synthesized as described (*82*,*204*). The N3-((2*R* or 2*S*)-hydroxy-3-buten-1-yl)-2'-deoxyuridine (*R*/*S*-BD-N3-dU) modified oligodeoxynucleotide was provided by the laboratory of Professor Richard P. Hodge (*91*). The concentrations of the single-stranded oligodeoxynucleotides were determined from the extinction coefficients at 254 nm (*205*). The modified oligodeoxynucleotide and its complementary strand were annealed in a buffer solution consisting of 10 mM NaH_2PO_4, 0.1 M NaCl, and 50 μM Na_2EDTA at pH 7.0. In each study, the annealed duplex oligodeoxynucleotide was eluted from a column containing DNA Grade Biogel hydroxylapatite (Bio-Rad Laboratories, Richmond, CA) with a gradient from 10 to 200 mM NaH_2PO_4 at pH 7.0. It was then lyophilized, resuspended in 1 mL of H_2O, and desalted using Sephadex G-25 (Bio-Rad Laboratories, Richmond, CA). The purity of the duplex was determined by using a PACE 5500 (Beckman Instruments, Inc., Fullerton, CA) instrument. Electrophoresis was conducted using an eCAP ssDNA 100-R kit applying 12,000 V for 30 min. The electropherogram was monitored using UV absorbance at 254 nm. MALDI-TOF mass spectra were

obtained on a Voyager-DE (PerSeptive Biosystems, Inc.) instrument in negative reflector mode. The matrix used in each of these studies contained 0.5 M 3-hydroxypicolinic acid and 0.1 M ammonium citrate. Figure 2-1 shows the MALDI Mass Spectrum of N^6-N^6 dA intrastrand crosslink duplex DNA sample. The peaks observed at 3274.1 and 3485.4 correspond to the complementary and adducted strands respectively.

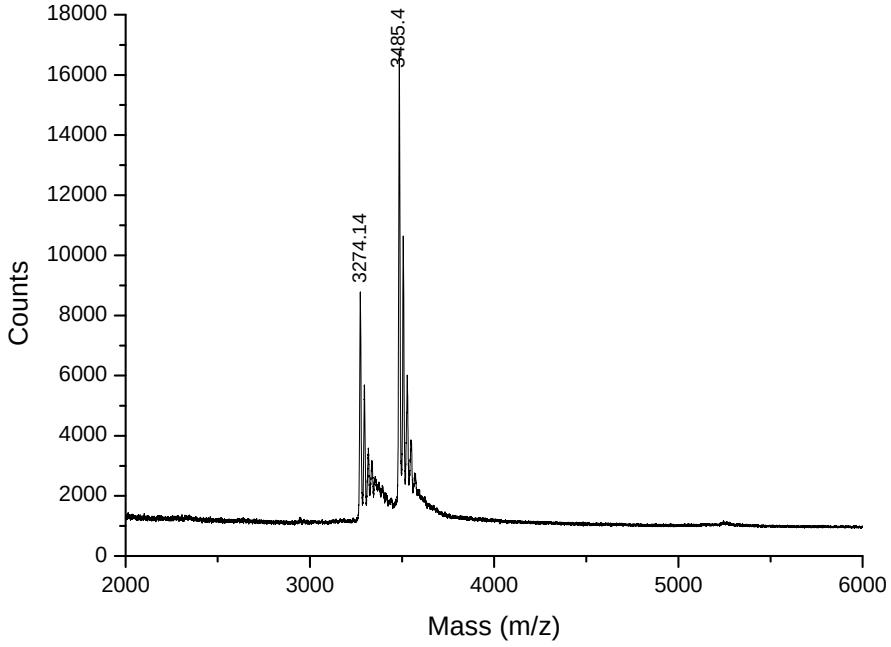


Figure 2-1. MALDI mass spectrum of N^6-N^6 dA intrastrand crosslink duplex DNA sample.

Thermal Melting Experiments

The melting temperatures were measured on a Varian Cary 4E spectrophotometer at 254 nm in a buffer consisting of 10 mM NaH_2PO_4, 50 μM Na_2EDTA, and 1 M NaCl at pH 7.0, at a temperature increasing rate of 1 °C/min from 10 to 80 °C. The melting temperatures of the unmodified and modified oligodeoxynucleotides were obtained from

the first-order derivatives of the melting curves. UV-melting curves of N^6-N^6 dA intrastrand cross-linked duplex DNA is shown in Figure 2-2. The derivative of the curves indicates the melting temperature to be around 43 °C.

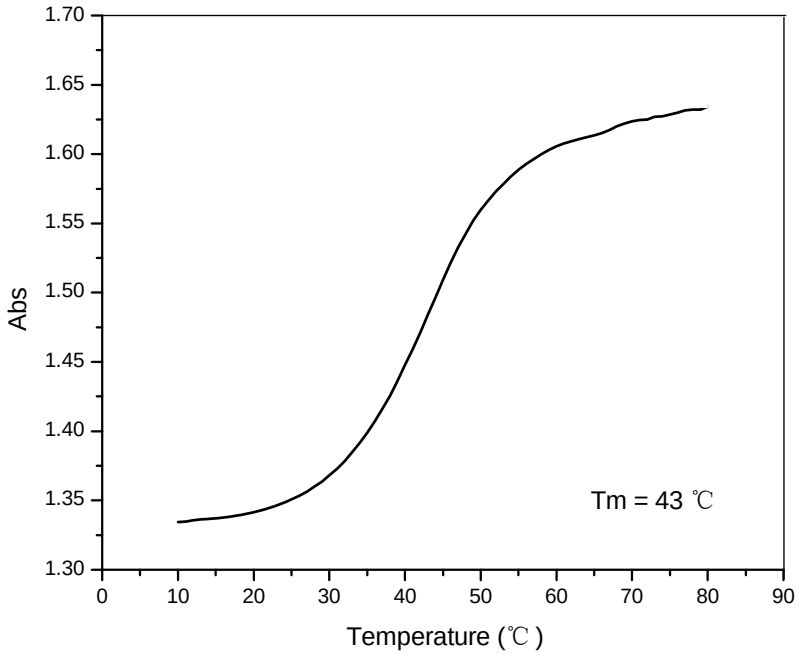


Figure 2-2. UV-melting curves of N^6-N^6 dA intrastrand crosslink duplex DNA sample.

NMR Spectroscopy

The modified duplex samples were prepared in 10 mM NaH_2PO_4, 0.1 M NaCl, and 50 μM Na_2EDTA at pH 7.0 to reach the final concentration of 2 mM. For observation of non-exchangeable protons, the samples were exchanged three times with 99.9% D_2O and dissolved in 0.5 mL of 99.99% D_2O. For observation of exchangeable protons, the samples were dissolved in 0.5 mL of 9:1 H_2O/D_2O. ^1H NMR spectra were recorded at 600.13 MHz and 800.23 MHz on Bruker spectrometers. The non-exchangeable protons

were monitored at 25 °C; the exchangeable protons were monitored at 10 °C. Chemical shifts were referenced to the water resonance. NMR data were processed using FELIX2000 (Accelyrs, Inc., San Diego, CA) on an Octane workstation (Silicon Graphics, Inc., Mountain View, CA).

NOESY spectra of the non-exchangeable protons were recorded using TPPI phase cycling with mixing times of 150, 200, and 250 ms. These spectra were acquired sequentially without removing the sample from the magnet. Spectra for the exchangeable protons were recorded using a 200 ms mixing time. These experiments were recorded with 1024 real data points in the d1 dimension and 2048 real data points in the d2 dimension. A relaxation delay of 2.0 s was used. Water suppression was performed using the WATERGATE sequence (*206*).

TOCSY experiments were performed with mixing times of 90 and 150 ms, utilizing the homonuclear Hartman-Hahn transfer with the MLEV17 (*207*) sequence for mixing. The butadiene protons were assigned via through bond magnetization transfer from the TOCSY spectra. DQF-COSY spectra were zero-filled to give a matrix of 1024 x 2048 real points. A skewed sine-bell squared apodization function with a 90° phase shift and a skew factor of 1.0 was used in both dimensions.

Distance and Torsion Angle Restraints

NOE distance restraints were obtained from NOESY spectra acquired at mixing times of 150, 200 and 250 ms. Footprints were drawn around cross-peaks obtained at a mixing time of 200 ms using FELIX2000. Identical footprints were transferred and fit to the cross-peaks obtained at the other two mixing times. The intensities of these cross-

peaks were determined by volume integration. These were combined as necessary with intensities generated from complete relaxation matrix analysis of a starting DNA structure to generate a hybrid intensity matrix (*208-210*). MARDIGRAS (*114-116*) was used to iteratively refine the hybrid intensity matrix and to optimize the agreement between the calculated and experimental NOE intensities. Calculations were initiated using isotropic correlation times of 2, 3, and 4 ns, and with both site-specific modified A-form (IniA) and B-form (IniB) starting structures and the three mixing times, yielding eighteen sets of distances. Analysis of these data yielded the experimental distance restraints used in subsequent restrained molecular dynamics calculations, and the corresponding standard deviations for the distance restraints. The distance restraints were divided into five classes, reflecting the confidence level in the experimental data.

Deoxyribose pseudorotation (*211*) was determined graphically using the sums of ^{3}J ^{1}H coupling constants measured from DQF-COSY spectra (*212*). Discrete $J_{1'2''}$ and $J_{1'2'}$ couplings were measured from active and passive couplings, respectively, of the H2" (d2) to H1' (d1) spectral region. The data were fit to curves relating the coupling constants to the deoxyribose sugar pseudo rotation angle (P), sugar pucker amplitude (φ), and the percentage S-type conformation. The sugar pseudo rotation angle and amplitude ranges were converted to the five dihedral angles ν_0 to ν_4. Coupling constants measured from ^{1}H-^{31}P HMBC spectra were applied (*109,110*) to the Karplus relationship (*111*) to determine the backbone dihedral angle ε (C4'-C3'-O3'-P), related to the H3'-C3'-O3'-P angle by a 120° shift. The ζ (C3'-O3'-P-O5') backbone angles were calculated from the correlation between ε and ζ in B-DNA (*89*).

Restrained Molecular Dynamics Calculations (rMD Calculations)

Classical A-DNA and B-DNA were used as reference structures to create starting structures for the refinement (*213*). The butadiene cross-linked adduct was constructed at the A^6 and A^7 positions using the BUILDER module of INSIGHT II (Accelrys, Inc., San Diego, CA). A-form and B-form structures of the appropriate sequence were energy-minimized by the conjugate gradients method for 200 iterations using the AMBER 7.0 force field (*214*) without experimental restraints to give starting IniA and IniB used for the subsequent relaxation matrix analysis and MD calculations.

The restraint energy function was comprised of terms describing distances and dihedral restraints, both of which were in the form of a standard square-well potential (*215*). Bond lengths involving hydrogens were fixed with the SHAKE algorithm (*216*). The generalized Born approach was used to model solvation (*217,218*). The calculations utilized a salt concentration of 0.2 mM. A series of randomly seeded rMD calculations were performed over a time course of 40 ps. These used the SANDER module of AMBER 7.0 force field, including the Parm94.dat parameter set. The atomic partial charges of BD modified lesion were calculated by using GAUSSIAN 03 (*219*) and shown in Appendix A. Restrained MD calculations were performed using AMBER in implicit solvent. The simulated annealing protocol utilized a starting temperature of 25 K. In the first 1ps the temperature was increased to 600 K and this target temperature was maintained for 4 ps, followed by cooling to 298 K over 15 ps. During the final 20 ps the temperature was reduced to 0 K. The temperature was controlled by coupling the molecules to a temperature bath. The first 1 ps of heating used a coupling of 0.4 ps, followed by the 4ps of constant temperature dynamics with a coupling of 1.0 ps. In the

following 15 ps of cooling, a heat coupling of 1.0 ps was used, followed by a value of 0.5 ps for the next 15 ps. During the final 5 ps of cooling, the coupling was ramped down to 0.01 ps. In the first 1 ps of heating, the experimental force constants were amplified by factors that ranged from 0.5 to 1.00. During the 4 ps of constant temperature dynamics and the first 15 ps of cooling, the amplification factor was increased to 1.75. In the final 20 ps of cooling, the amplification factor was reduced to a value of 1.00. Structure coordinates extracted from the final 4 ps of each rMD calculation were averaged and energy-minimized for 200 iterations using the conjugate gradient algorithm.

Back-calculation of ^{1}H NOE data was performed using CORMA (v. 4.0) (*209*). Helicodial parameters were examined using 3DNA (*120*) and Curves (*121,122*).

X-ray Crystallography

Expression and Purification of Dpo4

The vector pET22b(+)/Dpo4-NHis was a gift from the Professor F. Peter Guengerich Laboratory. The vector was overexpressed in *E. coli* strain BL21(DE3) competent cells (Stratagene, CA) and purified with a protocol adapted from the Guengerich lab (*186*). 20 ng of the plasmid containing an ampicillin resistant Dpo4 gene was added to 100 μg of BL21(DE3) cells. The transformation reaction was incubated on ice for 30 min, then heated in a 42 ºC water bath for 45 sec and incubated on ice for 2 min. Cells were added to SOC medium and incubated at 37 ºC for 30 min with shaking at 225 rpm. The cells were then concentrated by centrifugation at 200 g for 3 min and plated onto a Luria-Bertani (LB) agar plate containing 100 μg/mL ampicillin. The plate was incubated

overnight at 37 ºC and colonies grown the next morning. Figure 2-3 shows colonies grown in one such agar plate after a 16-hour incubation at 37 ºC. Each colony contains copies of the same plasmid of DNA.

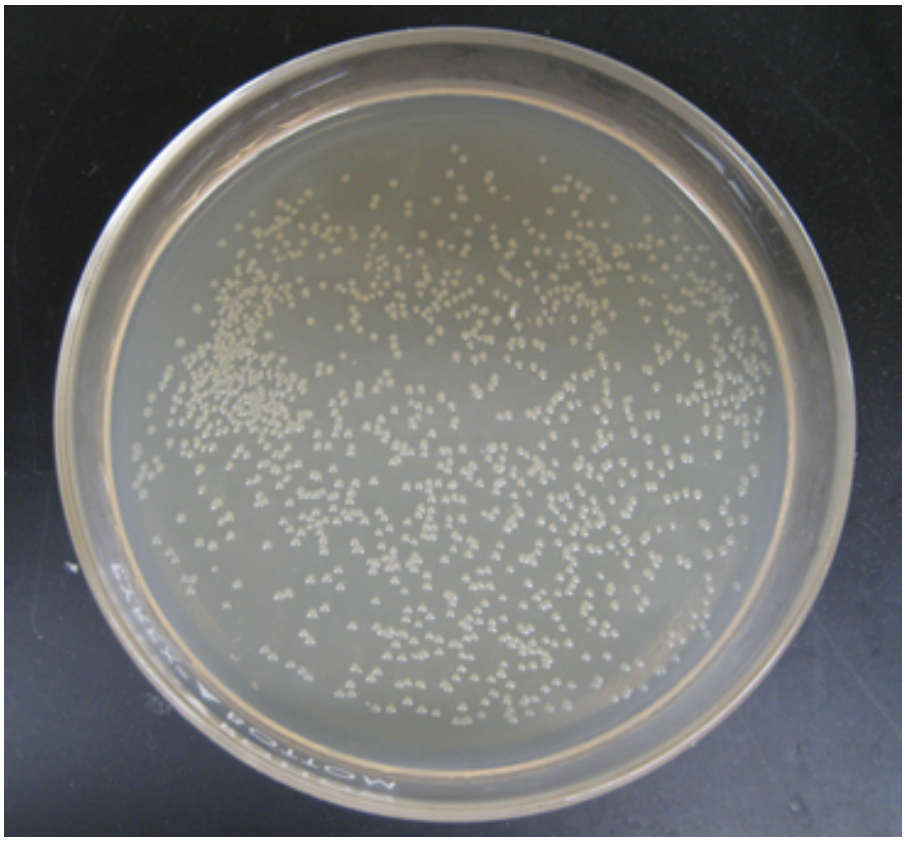


Figure 2-3. Transformation plate of the recombinant plasmid containing Dpo4 gene.

On the same day, a single colony was picked from this plate and transferred to 4 mL of sterile LB medium containing 100 μg/mL ampicillin and shaken at 225 rpm at 37 ºC overnight. The culture was added to 40 mL of fresh LB broth containing 100 μg/mL ampicillin and incubated at 225 rpm at 37 ºC overnight. The next morning, a 10 mL aliquot of culture was added to 1 L LB medium containing 100 μg/mL ampicillin. Cells

were grown at 37 ºC with shaking at 225 rpm until the absorbance at a wavelength of 600 nm (A_{600}) reached 0.6. Isopropyl β-D-thiogalactopyranoside (IPTG) was then added to a final concentration of 1 mM to induce cells for 3.5 h. The cells were harvested by centrifugation at 3,000 g for 30 min at 4 ºC (Beckman, JLA-9.100 rotor) and frozen at -80 ºC until purification.

To purify the Dpo4 protein, pellets were thawed and resuspended in lysis buffer containing 50 mM Tris-HCl, 300 mM NaCl, 10% glycerol (v/v), 5 mM β-mercaptoethanol, 1 mg/mL lysozyme, and protease inhibitor (Roche Applied Science) at pH 7.4, then cooled on ice for 30 min. The cells were then sonicated on ice for four 30 sec intervals with 30 sec of cooling between each sonication (Branson Sonifier 450, Danbury, CT). The cell debris was removed by centrifugation at 10,000 g for 45 min at 4 °C (Beckman, JL-12 rotor). The resulting supernatant was then heated at 80 °C for 10 min and centrifuged at 10,000 g for 45 min at 4 °C to remove the denatured proteins.

The supernatant was applied to a 5-mL HisTrap column (Amersham Biosciences), previously equilibrated with buffer A, which contains 50 mM Tris-HCl, 300 mM NaCl, 10% glycerol (v/v), 5 mM β-mercaptoethanol at pH 7.4. The column was washed sequentially with 15 mL of buffer A containing 20 mM imidazole, 15 mL of buffer A containing 40 mM imidazole, and 15 mL buffer A containing 60 mM imidazole. Bound His-tagged Dpo4 were eluted with 30 mL buffer A containing 400 mM imidazole, as shown in Figure 2-4. The absorbance was monitored by UV at a wavelength of 280 nm.

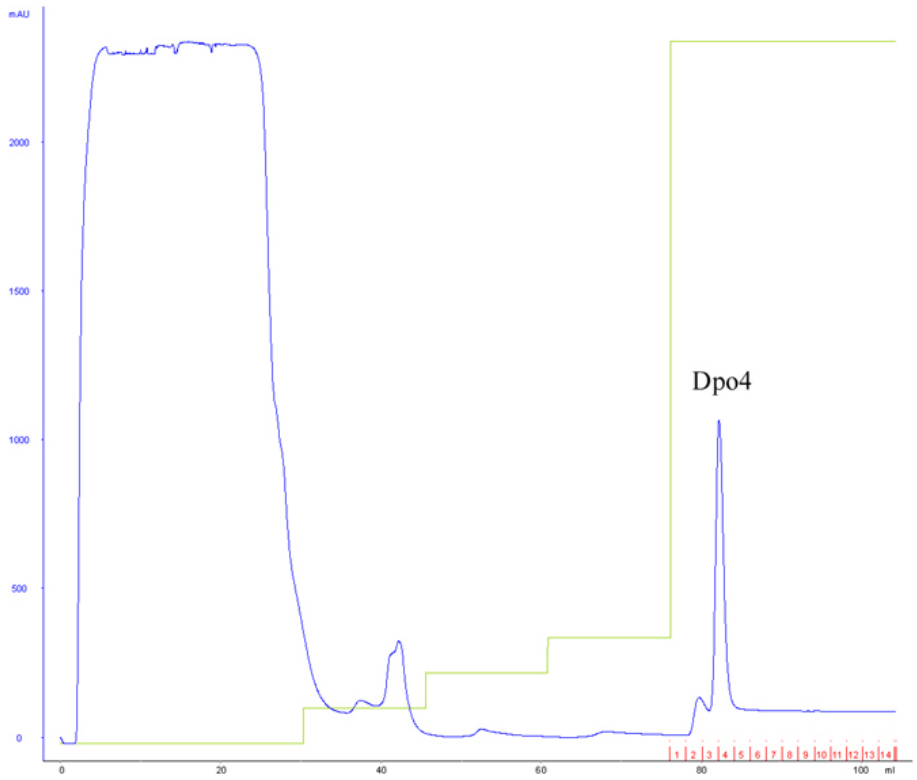


Figure 2-4. Purification of His-tagged Dpo4 using a His trap column.

Fractions containing Dpo4 were pooled together and dialyzed against buffer B containing 50 mM Tris-HCl, 0.5 mM EDTA, 10% glycerol (v/v), and 5 mM β-mercaptoethanol at pH 7.4. The protein solution was further applied to an 8-ml MonoS column (Amersham Biosciences), previously equilibrated with buffer B and eluted with a linear salt gradient of 100-500 mM NaCl in the same buffer (Figure 2-5). The purity of Dpo4 was assessed by SDS-PAGE using Pre-cast NuPage 4-12% Bis-Tris gels (Invitrogen, Carlsbad, CA) (Figure 2-6). The protein was concentrated and stored at a concentration of ~40 mg/mL for crystallographic study. Protein concentration was determined spectrophotometrically using a molar extinction coefficient of 22 mM^{-1} cm^{-1} at 280 nm (*220*).

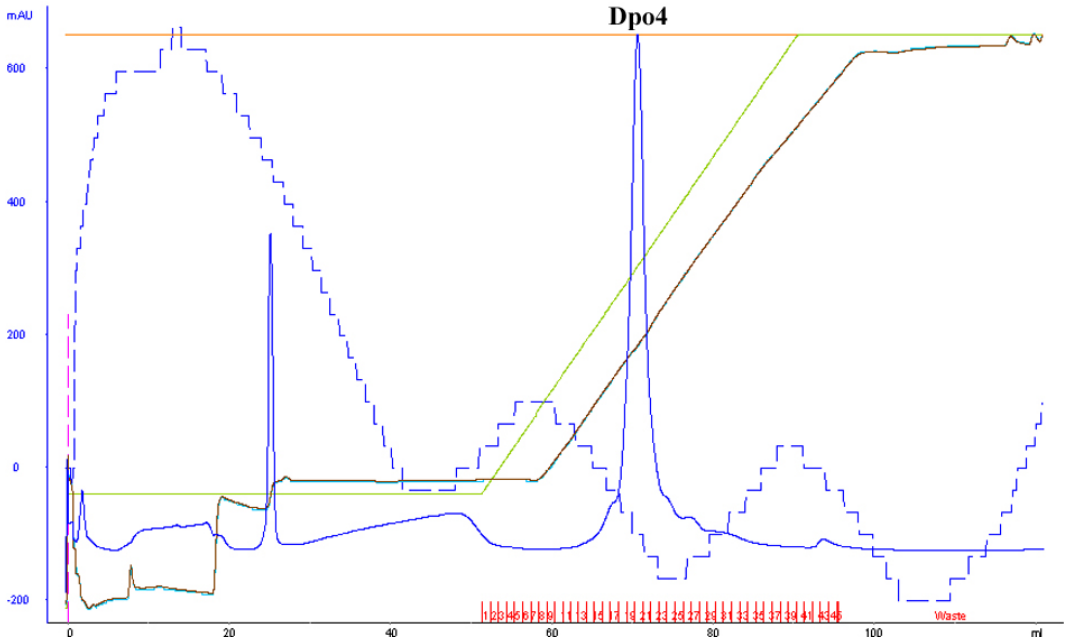


Figure 2-5. Purification of Dpo4 using a cation exchange colomn (MonoS column).

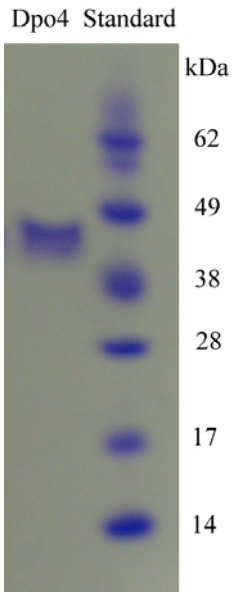


Figure 2-6. Analysis of Dpo4 by SDS-PAGE. Purified Dpo4 is shown on the left and protein size markers (Standard) are shown on the right.

Replication Bypass Experiments

The replication bypass experiments were done by Dr. Surajit Banerjee in the laboratory of Professor Michael P. Stone. These experiments utilized either a 13-mer·18-mer primer-template 5'-d(GGGGGAAGGATTT)-3'·5'-d(TCACXAAATCCTTCCCCC)-3' sequence or a 14-mer·18-mer primer-template 5'-d(GGGGGAAGGATTTA)-3'·5'-d(TCACXAAATCCTTCCC- CC)-3' sequence, where X represents *S*- or *R*-BD-N3-dU or control dC. The ^{32}P-labeled primers were annealed to either unmodified or modified templates in 1:1 molar ratios. For single nucleotide incorporation assays, the duplexes were extended in the presence of single dNTPs. Each reaction was initiated by adding 2 μL dNTPs to a preincubated DNA/Dpo4 mixture at 37 °C, yielding a total reaction volume of 10 μL. The final concentration of the reaction mixture was 25 mM Tris-HCl (pH 7.8), 5 mM DTT, 0.1 mg·mL^{-1} BSA, and 5 mM $MgCl_2$, 10 nM DNA duplex, 100 nM Dpo4, and the concentration of dNTP varied from 0 to 20 μM. The reaction mixtures were incubated over a time period of 8 min for the control experiment and 15 min for the adducted samples. Each reaction was quenched with 50 μL of 20 mM Na_2EDTA (pH 9.0) in 95% formamide (v/v) containing bromophenol and heated for 10 min at 95 °C. For full-length extension assays, the unmodified and modified primers were extended in the presence of all four dNTPs. Each reaction was initiated by adding 2 μL dNTPs to a preincubated DNA/Dpo4 mixture at 37 °C, yielding a total reaction volume of 10 μL. The final concentration of the reaction mixture was 25 mM Tris-HCl (pH 7.8), 5 mM DTT, 0.1 mg·mL^{-1} BSA, and 5 mM $MgCl_2$, 10 nM DNA duplex, 100 nM Dpo4, and the concentration of dNTP varied from 0 to 20 μM. The reaction mixtures were incubated over a time period of 15 min for the control experiment and 20 min for the adducted

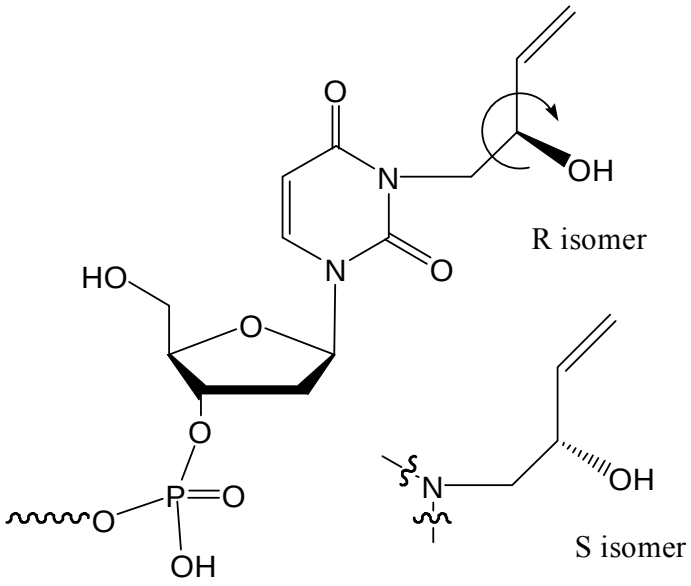
samples. Each reaction was quenched with 50 μL of 20 mM Na_2EDTA (pH 9.0) in 95% formamide (v/v) containing bromophenol and heated for 10 min at 95 °C. Electrophoresis was used to separate aliquots (5-8 μL) of the reaction on a denaturing gel containing 8.0 M urea and 16% acrylamide (w/v) (from a 19:1 acrylamide/bisacrylamide Stablized Solution, AccuGel, National Diagnostics, Atlanta, GA) with 80 mM Tris borate buffer (pH 7.8) containing 1 mM Na_2EDTA. The gel was exposed to a PhosphorImager screen (Imaging Screen K, Bio-Rad) overnight. The screen was imaged using a PhosphorImaging system (Bio-Rad, Molecular Imager FX) with the manufacturer's software Quantity One, version 4.3.0.

Crystallization of Dpo4-DNA Complexes

Two primers and one template sequence were used in the crystallizations: 5'-GGGGGAAGGATTT-3' (13-base primer), 5'-GGGGGAAGGATTTA-3' (14-base primer), and 5'-TCACXAAATCCTTCCCCC-3' (18-base template), where X represents *S*- or *R*-BD-N3-dU (Scheme 2-1). The enantiomeric purity of the *S*- and *R*- isomers was determined by enzyme digestion and capillary gel electrophoresis, as shown in Figure 2-7. The binary complexes were prepared by annealing the 18-base template and the 14-base primer, complexed with Dpo4. A series of ternary complexes were made from annealing the 18-base template and the 13-base primer, then mixing with Dpo4 and d(d)ATP, d(d)GTP, d(d)CTP or d(d)TTP, individually. Other ternary complexes were prepared by mixing annealed the 18-base template and 14-base primer with Dpo4 and d(d)GTP.

Dpo4 and the primer-template duplex were mixed in a 1:1.2 molar ratio in the solution containing 20 mM Tris-HCl, 60 mM NaCl, 5 mM $CaCl_2$, and 2.5% glycerol (v/v)

at pH 7.4. The mixture was then incubated for 10 min at room temperature after adding 1 mM d(d)ATP, d(d)GTP, d(d)CTP or d(d)TTP, individually. The final protein concentration was kept at 8 mg/ml. The reservoir solution contained 10% polyethylene glycol (PEG) 3350, 20 mM Tris-HCl (pH 7.5), and 100 mM calcium acetate. Crystals were obtained in 1-2 days using the sitting-drop vapor diffusion method at room temperature. Typical crystals of both binary and ternary complexes of Dpo4 and DNA were shown in Figure 2-8. Crystals were removed from the drop and transferred into a cryoprotectant solution containing 25% PEG 3350 (w/v) and 15% ethylene glycol (v/v) just prior to mounting on the X-ray goniometer and flash-frozen in liquid nitrogen.



***S*/*R*-BD-N3-dU**

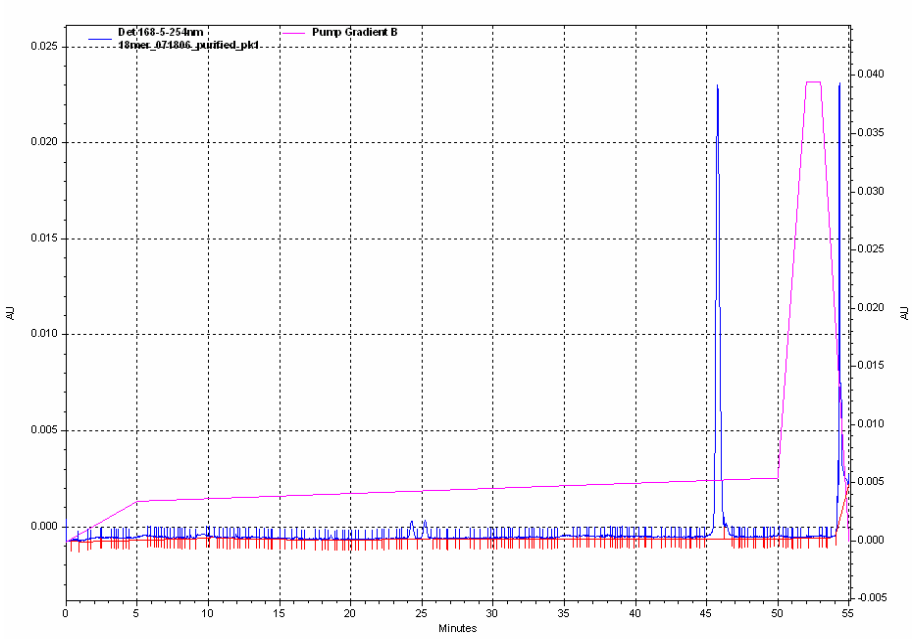
Binary Complexes	5' GGG GGA AGG ATT TA 3' 3' CCC CCT TCC TAA A**X**C ACT 5'	+ Dpo4
Ternary Complexes	5' GGG GGA AGG ATT T 3' 3' CCC CCT TCC TAA A**X**C ACT 5'	+ Dpo4 + d(d)**N**TP
	5' GGG GGA AGG ATT TA 3' 3' CCC CCT TCC TAA A**X**C ACT 5'	+ Dpo4 + d(d)GTP

X = *S*- or *R*-BD-N3-dU

N = A or T or C or G

Scheme 2-1. Template-primer duplexes used in crystallographic study.

a.



b.

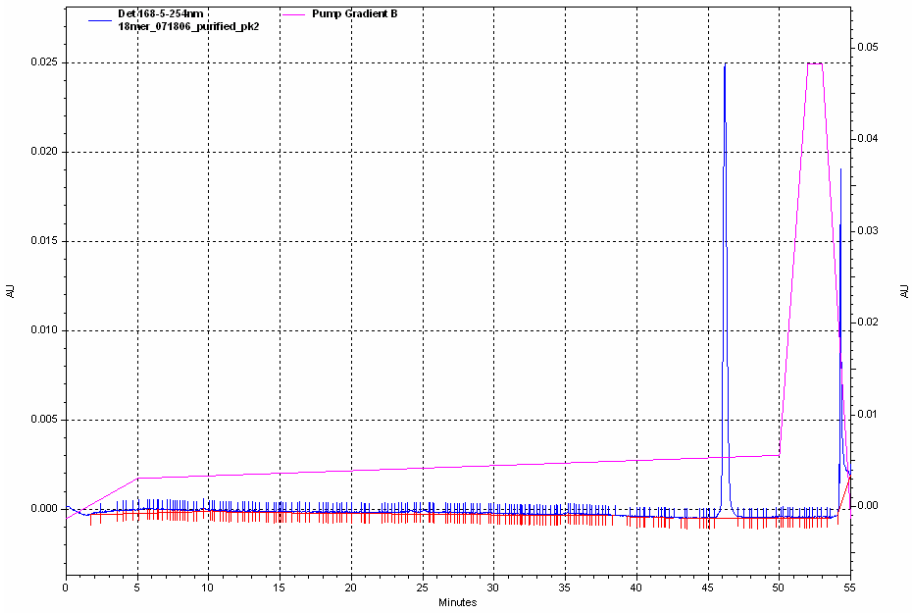


Figure 2-7. Capillary gel electrophoresis of 18-mer templates containing a) *R*-BD-N3-dU and b) *S*-BD-N3-dU.

A

B

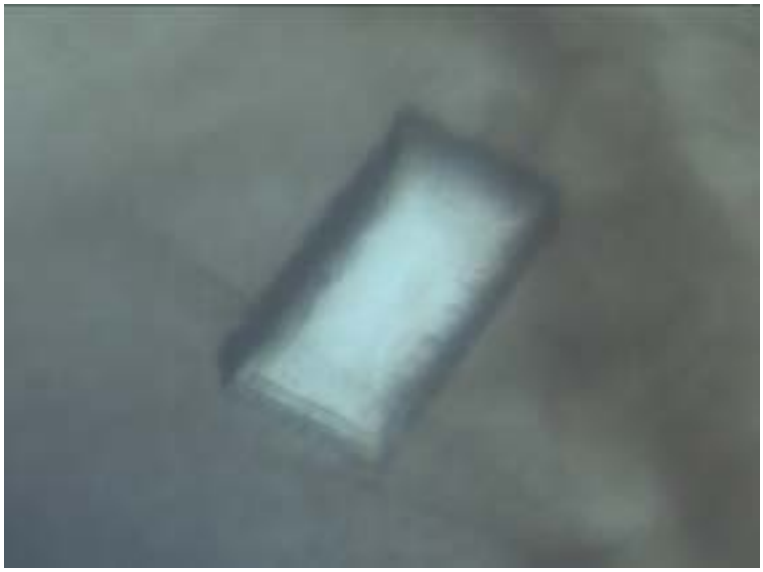


Figure 2-8. Typical crystals of Dpo4-DNA binary and ternary complexes: a) Ternary complex containing *S*-BD-N3-dU in the template and an incomging dGTP b) Binary complex containing *R*-BD-N3-dU in the template opposite A in the primer.

X-ray Diffraction Data Collection and Processing

X-ray diffraction data sets were collected at a wavelength of 1.00 Å on the ID-22 SER-CAT or BM-22 SER-CAT or ID-5 DND-CAT beam line at the Advanced Photon Source in the Argonne National Lab (Argonne, IL). For each single crystal, a series of X-ray diffraction images, each corresponding to 1 degree of rotation, were collected to a total crystal rotation of 180 degrees.

Diffraction data were autoindexed, integrated and scaled with the program HKL2000 (*221*). The orthorhombic crystal system and $P2_12_12$ space group were determined for the crystals of all complexes. Further procession, including re-index and truncate procedures of some data were performed by using CCP4 package programs (*222*) . The quality of each data set was determined by R_{merge} value.

Crystal Structure Determination and Refinement

All the structures were solved by molecular replacement with the program AMoRe (*178,223*). The first starting model was generated from the structure of a ternary Dpo4-DNA-dCTP complex (PDB code 2C22) (*187*) by changing the adduct lesion to a natural base and removing the solvent molecules. In order to place the model structure into the unit cell in exactly the same position and orientation as the new protein molecule, a cross rotation function (*178*) was carried out initially to rotate the model data to fit the new structure accurately, followed by a translation function (*224*) that moved the reoriented data through the unit cell to fit the position of the new molecule most accurately. Then the positions of the model were optimized by several rounds of rigid body refinement, with gradually increasing resolution. The model structures were further improved through

cycles of refinement. Given the same space group and similar unit cell dimensions of all crystal structures, the refined structure of the first complex was used as a starting model for the rest crystal structures.

The structures were refined using the CNS Solve package (version 1.1) (*225*), including conjugate gradient minimization, simulated annealing, restrained individual B-factor refinement and in some cases, individual unrestrained occupancy refinement. The parameters obtained from the refined structure were used to make new electron density maps, such as F_o-F_c map and $2F_o$-F_c map through Fourier transform. The annealed omit maps were used to ensure unbiased electron density maps in the vicinity of the BD adduct. Then the BD-N3-dU nucleotide was added manually to fit the annealed omit maps. The rebuilt model was then improved by correcting error based on the new electron density maps with the TURBO-FRODO program (*226*). Calcium atoms, as well as water oxygen atoms and the incoming nucleotide, were added gradually added into positive regions of the F_o - F_c electron density map after each cycle of refinement. The improvement of the structure from each cycle of refinement was judged by R and R_{free} values, the latter was made by randomly selecting 5% of the data excluded from calculations of the electron density maps. The stereochemical quality of each model was checked by analyzing the overall and nucleotide-by-nucleotide geometry in PROCHECK (*227*) and Curves (*121*,*122*). The program PYMOL (*228*) was used to prepare all the crystallographic figures.

CHAPTER III

SOLUTION STRUCTURE OF THE 1,4-BIS(2'-DEOXYADENOSIN-N^6-YL)-2*S*,3*S*-BUTANEDIOL INTRASTRAND DNA CROSS-LINK ARISING FROM BUTADIENE DIEPOXIDE IN THE HUMAN N-*RAS* CODON 61 SEQUENCE

Introduction

The 1,4-bis(2'-deoxyadenosin-N^6-yl)-2*S*,3*S*-butanediol intrastrand DNA cross-link arises from the bisalkylation of tandem N^6-dA sites in DNA by *R,R*-butadiene diepoxide (BDE). The oligodeoxynucleotide 5'-d($C^1G^2G^3A^4\underline{C^5X^6Y^7}G^8A^9A^{10}G^{11}$)- 3' · 5'-d($C^{12}T^{13}T^{14}C^{15}T^{16}T^{17}G^{18}T^{19}C^{20}C^{21}G^{22}$)-3' contains the BDE cross-link between the second and third adenines of the codon 61 sequence (underlined) of the human N-ras protooncogene and is named the (*S,S*)-BD-(61-2,3) cross-link (X,Y = cross-linked adenines) (Scheme 3-1). NMR analysis reveals that the cross-link is oriented in the major groove of duplex DNA. Watson-Crick base pairing is perturbed at base pair $X^6{\cdot}T^{17}$, whereas base pairing is intact at base pair $Y^7{\cdot}T^{16}$. The cross-link appears to exist in two conformations, in rapid exchange on the NMR time scale. In the first conformation, the β-OH is predicted to form a hydrogen bond with T^{16} O^4, whereas in the second, the β-OH is predicted to form a hydrogen bond with T^{17} O^4. In contrast to the (*R,R*)-BD-(61-2,3) cross-link in the same sequence (*229*), the anti-conformation of the two hydroxyl groups at C_β and C_γ with respect to the C_β-C_γ bond results in a decreased twist between base pairs $X^6{\cdot}T^{17}$ and $Y^7{\cdot}T^{16}$, and an approximate 10 º bending of the duplex. These conformational differences may account for the differential mutagenicity of the (*S,S*)- and

(*R*,*R*)-BD-(61-2,3) cross-links and suggest that stereochemistry plays a role in modulating biological responses to these cross-links (*92*).

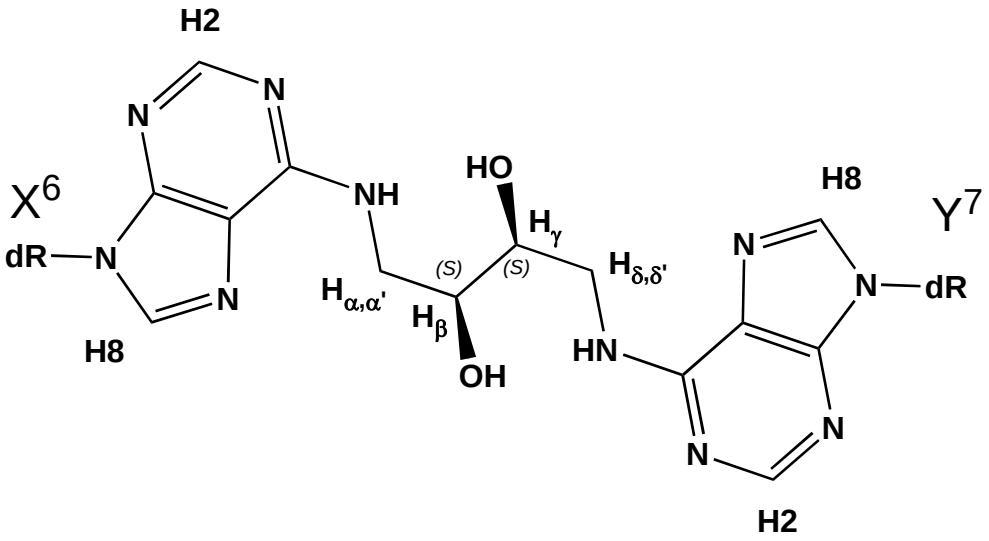
A.

<u>60</u> <u>61</u> <u>62</u>

5'-C^{1}|G^{2} G^{3} A^{4}|C^{5} X^{6} Y^{7}|G^{8} A^{9} A^{10}|G^{11}-3'

3'-G^{22}|$C^{21}$$C^{20}$$T^{19}$|$G^{18}$$T^{17}$$T^{16}$|$C^{15}$$T^{14}$$T^{13}$|$C^{12}$-5'

B.



Scheme 3-1. The *ras*61 oligodeoxynucleotide (A) and chemical structure (B) of the (2*S*,3*S*)-N^{6}-(2,3-dihydroxybutyl)-2-deoxyadenosyl cross-linked adduct and nomenclature. (Reproduced with permission from [*Chem Res Toxicol* **2007**, *20*, 187-98.] Copyright [2007] American Chemical Society)

Results

Sample Properties

The duplex oligodeoxynucleotide 5'-d($C^1G^2G^3A^4\underline{C^5A^6A^7}G^8A^9A^{10}G^{11}$)- 3' · 5'-d($C^{12}T^{13}T^{14}C^{15}T^{16}T^{17}G^{18}T^{19}C^{20}C^{21}G^{22}$)-3' utilized in the present studies contains the coding sequences for codons 60, 61 (underlined), and 62 of the human N-*ras* proto-oncogene and was engineered to be inserted into bacterial and mammalian site-specific mutagenesis systems. The *ras* oncogene was discovered in the Harvey and Kirsten murine sarcoma viruses (*230*). Mutations in codon 61 activate the oncogene. Mutations in *ras* are frequently detected alterations in human tumors (*230-235*). The identity of the duplex was verified using MALDI-TOF mass spectrometry. Mass measurement showed two signals corresponding to mass units of 3485 and 3274, which originated from the cross-linked adduct strand 5'-d(CGGACXYGAAG)-3' and the complementary strand 5'-d(CTTCTTGTCCG)-3', respectively (Figure 2-1). The ratio of two strands was determined to be 1:1 after correction for the respective absorbance coefficients by using capillary gel electrophoresis. The melting temperature of the (*S*,*S*)-BD-(61-2,3) duplex was 43 ºC (Figure 2-2), lower than the 57 ºC melting temperature of the unmodified *ras*61 duplex and the 50 ºC of the (*R*,*R*)-BD-(61-2,3) duplex (*229*). The cross-linked sample yielded excellent NMR data in the temperature range of 10 - 25 ºC.

DNA ^{1}H Resonance Assignments

Nonexchangeable Protons

The sequential assignment for the (*S*,*S*)-BD-(61-2,3) cross-linked duplex was accomplished using standard protocols (*236*,*237*). The sequential NOEs between the aromatic and anomeric protons of the (*S*,*S*)-BD-(61-2,3) cross-linked oligodeoxynucleotide duplex are displayed in Figure 3-1. For the modified strand, weak NOE connectivities were observed between X^6 H1' to Y^7 H8 and Y^7 H1' to G^8 H8. In the complementary strand, the connectivity between T^{16} H1' and T^{17} H6 was interrupted. The complete assignments of the deoxyribose H2', H2", and H3' protons were achieved, and partial assignments were made for the deoxyribose H4', H5', and H5" protons because of heavy overlap. The chemical shift assignments of non-exchangeable protons of the N^6-N^6 dA cross-link oligodeoxynucleotide duplex are listed in Appendix B.

Exchangeable Protons

A plot of the region ranging from 11 to 15 ppm of the NOESY experiment is shown in Figure 3-2. The imino cross-peaks located in this region were well-resolved, with the exception of T^{14} and T^{16}, which were overlapped. The assignments of the thymine imino resonances were confirmed by the identification of cross-peaks to adenosine H2 protons of each A·T pair, as shown in Figure 3-2B. Cross-peaks arising from the Y^7 H2 to T^{16} N3H and A^9 H2 to T^{14} N3H NOEs confirmed that both T^{14} N3H and T^{16} N3H resonated at 13.8 ppm. The sequential connectivities of the imino protons were obtained from base pairs $G^2 \cdot C^{21} \rightarrow C^5 \cdot G^{18}$ and $Y^7 \cdot T^{16} \rightarrow A^{10} \cdot T^{13}$, with breaks in connectivity observed between base pairs $C^5 \cdot G^{18} \rightarrow X^6 \cdot C^{17}$ and between $X^6 \cdot C^{17} \rightarrow Y^7 \cdot T^{16}$. The NOE between

T^{17} N3H and X^{6} H2 was of reduced intensity. The NOE between T^{16} N3H and Y^{7} H2 was also weak. The imino resonances from the terminal base pairs $C^{1}\cdot G^{22}$ and $G^{11}\cdot C^{12}$ were missing, which presumably reflected the effects of strand fraying and the resulting rapid exchange of this proton with solvent.

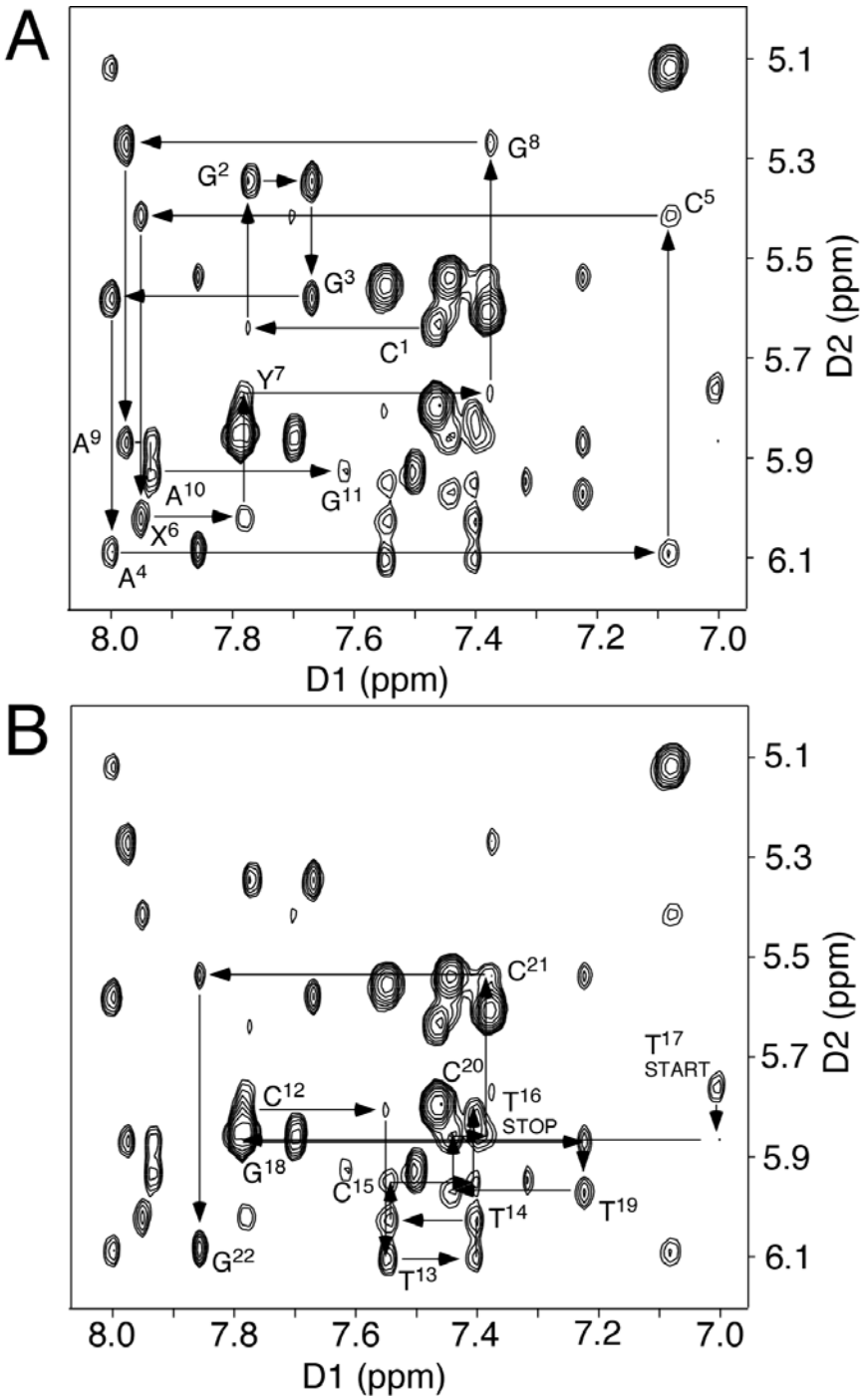


Figure 3-1. Expanded plots of a NOESY spectrum showing sequential NOE connnectivities from aromatic to anomeric protons. A. Nucleotides $C^{1}\rightarrow G^{11}$ of the modified strand of the (*S*,*S*)-BD-(61,2-3) cross-linked adduct. B. Nucleotides $C^{12}\rightarrow G^{22}$ of the complementary strand of the (*S*,*S*)-BD-(61,2-3) cross-linked adduct. The experiment was carried out at a mixing time of 250 ms and 800 MHz. (Reproduced with permission from [*Chem Res Toxicol* **2007**, *20*, 187-98.] Copyright [2007] American Chemical Society)

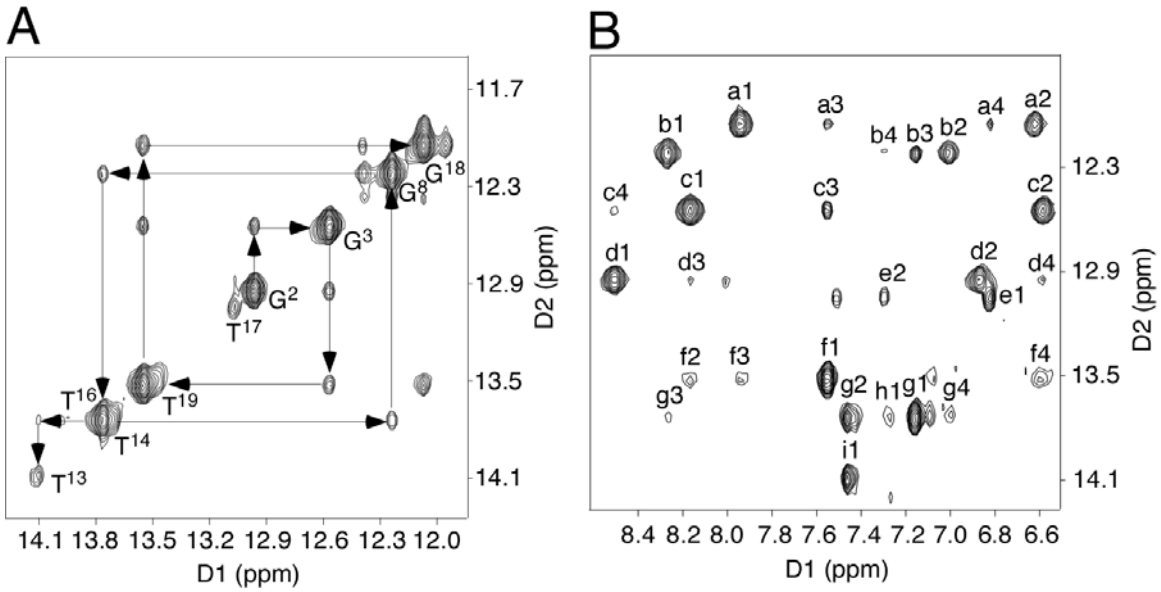


Figure 3-2. A. NOE connectivity for the imino protons for the base pairs from $G^{2}\cdot C^{21}$ to $A^{10}\cdot T^{13}$. Note the breaks in connectivity between the imino resonances for base pairs and between base pairs $C^{5}\cdot G^{18}$ and $X^{6}\cdot T^{17}$ and between $X^{6}\cdot T^{17}$ and $Y^{7}\cdot T^{16}$. The experiment was carried out at a mixing time of 250 ms and 800 MHz. **B.** NOE cross-peaks between exchangeable protons and non-exchangeable protons: a1, G^{18} N1H→G^{18} N^{2}H1; a2, G^{18} N1H→G^{18} N^{2}H2; a3, G^{18} N1H→A^{4} H2; a4, G^{18} N1H→X^{6} H2; b1, G^{8} N1H→G^{8} N^{2}H1; b2, G^{8} N1H→G^{8} N^{2}H2; b3, G^{8} N1H→A^{9} H2; b4, G^{8} N1H→Y^{7} H2; c1, G^{3} N1H→G^{3} N^{2}H1; c2, G^{3} N1H→G^{3} N^{2}H2; c3, G^{3} N1H→A^{4} H2; c4, G^{3} N1H→G^{2} N^{2}H1; d1, G^{2} N1H→G^{2} N^{2}H1; d2, G^{2} N1H→G^{2} N^{2}H2; d3, G^{2} N1H→G^{3} N^{2}H1; d4, G^{2} N1H→G^{3} N^{2}H2; e1, T^{17} N3H→X^{6} H2; e2, T^{17} N3H→Y^{7} H2; f1, T^{19} N3H→A^{4} H2; f2, T^{19} N3H→G^{3} N^{2}H1; f3, T^{19} N3H→G^{18} N^{2}H1; f4, T^{19} N3H→G^{18} N^{2}H2; g1, T^{14} N3H→A^{9} H2; g2, T^{14} N3H→A^{10} H2; g3, T^{14} N3H→G^{8} N^{2}H1; g4, T^{14} N3H→G^{8} N^{2}H2; h1, T^{16} N3H→Y^{7} H2; i1, T^{13} N3H→A^{10} H2. The experiment was carried out at 10°C. (Reproduced with permission from [*Chem Res Toxicol* **2007**, *20*, 187-98.] Copyright [2007] American Chemical Society)

Butadiene Protons

The butadiene cross-link protons were assigned by the combined use of DQF-COSY, TOCSY, and NOESY spectra. As shown in Figure 3-3, at a mixing time of 250 ms, NOEs were observed between the butadiene cross-link protons. The H_{α} proton at 2.87 ppm showed strong NOEs to $H_{\alpha'}$ at 3.32 ppm and H_{β} at 3.89 ppm, along with a weak NOE to H_{γ} at 3.95 ppm. The H_{β} proton showed NOEs to H_{α}, $H_{\alpha'}$, and H_{γ}, along with a

weak NOE to H_δ at 3.01 ppm and $H_{\delta'}$ at 3.71 ppm. The H_γ proton exhibited NOEs to H_δ and $H_{\delta'}$, along with weak NOEs to H_α and $H_{\alpha'}$.

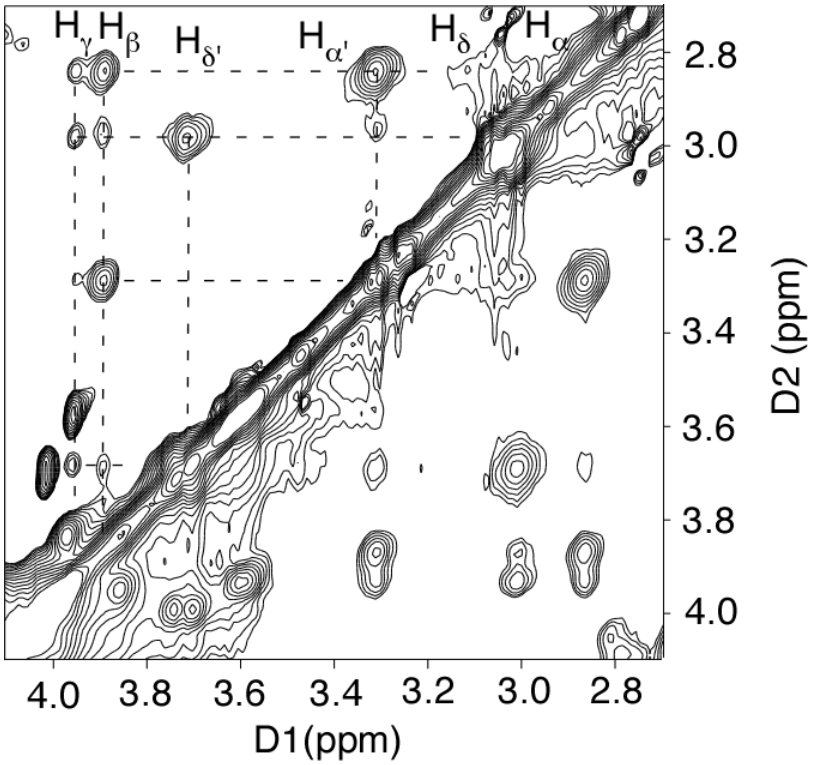


Figure 3-3. Expanded NOESY spectrum showing the assignments for the butadiene protons in the (*S*,*S*)-BD-(61,2-3) duplex. The experiment was carried out at 250 ms mixing time and 800 MHz. The temperature was 25°C. (Reproduced with permission from [*Chem Res Toxicol* **2007**, *20*, 187-98.] Copyright [2007] American Chemical Society)

Butadiene-DNA NOEs

A total of 5 NOEs were observed between the BD protons and the DNA. Some of these are shown in Figure 3-4. The H_α and $H_{\alpha'}$ protons exhibited cross-peaks to the 5'-direction, neighboring base C^5 H5. NOEs were also observed between butadiene protons

H_α, $H_{\alpha'}$, and H_β, and the exchangeable proton T^{17} N3H. These NOEs established the orientation of the butadiene cross-link in the major groove.

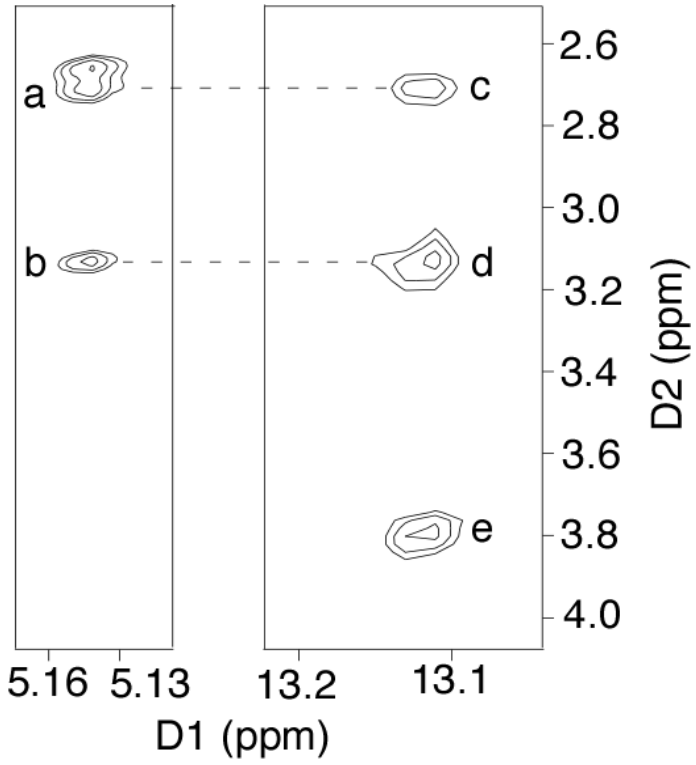


Figure 3-4. Tile plot showing NOE cross-peaks between BD and DNA of the (*S*,*S*)-BD-(61,2-3) cross-linked duplex. Cross-peaks: a, C^5 H5 → BD H_α; b, C^5 H5 → BD $H_{\alpha'}$; c, T^{17} N3H → BD H_α; d, T^{17} N3H → BD $H_{\alpha'}$; e, T^{17} N3H → BD H_β. The experiment was at 250 ms mixing time and 800 MHz. (Reproduced with permission from [*Chem Res Toxicol* **2007**, *20*, 187-98.] Copyright [2007] American Chemical Society)

Torsion Angle Measurements

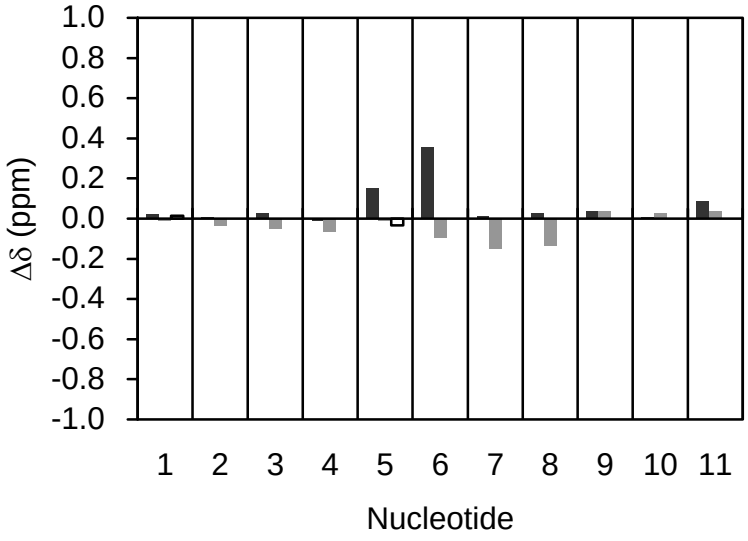
Analysis of DQF-COSY data suggested that all deoxyribose pseudorotation angles remained in the C2'-endo sugar ring conformation anticipated for the B-family DNA. The glycosyl torsion angle conformations were evaluated by inspection of ^{1}H NOESY data. The weak NOEs between purine H8 or pyrimidine H6 protons and the anomeric H1' protons of the attached deoxyribose sugars were consistent with glycosyl torsion angles

in the normal anti-conformational range. Measurements for the phosphodiester backbone torsion angles ε and ζ were obtained from a heteronuclear ^{1}H-^{31}P correlation experiment. The ^{31}P chemical shifts were observed in the expected chemical shift range for B-DNA. There were also no unusual ^{3}J ^{1}H-^{31}P couplings observed in the ^{31}P spectrum. Data obtained from ^{3}J ^{1}H-^{31}P experiments indicated that the torsion angles associated with the backbone phosphodiester linkages were not significantly perturbed by the presence of the cross-link.

Chemical Shift Effects

The ^{1}H spectrum of the cross-linked duplex exhibited significant chemical shift differences at the modified base X^{6} and its complementary base T^{17} compared to that of the unmodified *ras*61 oligodeoxynucleotide (*238*) (Figure 3-5). An upfield chemical shift of 0.36 ppm was observed for X^{6} H1', whereas a downfield shift of 0.29 ppm was observed for T^{17} H6. For the imino protons, the greatest downfield shift of 0.46 ppm was observed for T^{17} N3H. Other smaller chemical shifts were observed for a number of protons near the adduction sites.

A



B

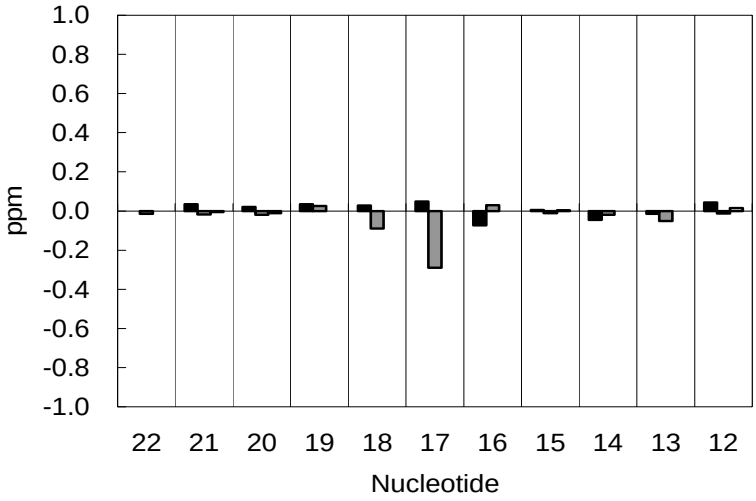


Figure 3-5. Chemical shift differences of protons of the (*S*,*S*)-BD-(61,2-3) cross-linked adduct duplex relative to the unmodified *ras61* onligodeoxynucleotide. **A**. Nucleotides $C^{1}\rightarrow G^{11}$ of the (*S*,*S*)-BD-(61,2-3) cross-linked duplex. **B**. Nucleotides $C^{12}\rightarrow G^{22}$ of the (*S*,*S*)-BD-(61,2-3) cross-linked duplex. Black bars represents the deoxyribose H1' protons; gray bars represent the purine H8 or pyrimidine H6 protons, respectively; white bars represent the cytosine H5 protons. (Reproduced with permission from [*Chem Res Toxicol* **2007**, *20*, 187-98.] Copyright [2007] American Chemical Society)

Structural Refinement

The structural refinement utilized a simulated annealing protocol for the molecular dynamics calculations. The Generalized Born model was utilized to approximate charge-charge interactions in the presence of a continuum solvent (*217*,*218*). The molecular dynamics calculations were restrained by 413 experimental distance restraints, including the 5 NOEs between the BD protons and the DNA. These were evenly distributed over the length of the oligodeoxynucleotide. In addition to the NOE-based distance restraints, 90 deoxyribose pseudorotation restraints, 104 empirical phosphodiester backbone angle measurements, and 39 hydrogen-bonding restraints were included in the calculations.

Sets of randomly seeded rMD calculations were carried out from both B-form DNA and A-form DNA starting structures. These starting structures exhibited a root-mean-square deviation (rmsd) of 6.26 Å. The structures that emerged following 40 ps of rMD were subjected to potential energy minimization and evaluated as to pairwise rmsd. Irrespective of starting structure, two ensembles of structures emerged from the calculations. In the first, designated conformation I, the butadiene C_β-hydroxyl proton was within hydrogen-bonding distance of T^{16} O^4. The maximum pairwise rmsd for this ensemble of structures was 0.87 Å. In the second, designated conformation II, the butadiene β-hydroxyl proton was within hydrogen-bonding distance of T^{17} O^4. The maximum pairwise rmsd for this ensemble of structures was 0.85 Å (Table 3-1). Stereoviews of the two different energy minimized structures that emerged from the randomly seeded rMD calculations are shown in Figure 3-6.

Table 3-1: Analysis of the rMD generated structures of the *ras*61 (*S,S*) N^6,N^6-dA cross-linked adduct (Reproduced with permission from [*Chem Res Toxicol* **2007**, *20*, 187-98.] Copyright [2007] American Chemical Society)

NMR restraints	
total no. of distance restraints	413
internucleotide distance restraints	69
intranucleotide distance restraints	339
DNA-cross-link distance restraints	5
empirical restraints	
H-bonding restraints	39
sugar pucker restraints	90
backbone torsion angle restraints	104
Structural statistics	
NMR R-factor (R_1^x)[a,b,c]	
<rMDIi>	0.0903 ± 0.0140
<rMDIIj>	0.0848 ± 0.0142
rmsdI of NOE violations (Å)	0.0200
rmsdII of NOE violations (Å)	0.0195
pairwise rmsd (Å) over all atoms	
<rMDIi> vs <rMDIav>	0.553 ± 0.13
<rMDIIj> vs <rMDIIav>	0.559 ± 0.17

[a] The mixing time was 250 ms. [b] $R_1^x = \sum \mid (a_o)_i^{1/6} - (a_c)_i^{1/6} \mid / \sum \mid (a_o)_i^{1/6} \mid$, where a_o and a_c are the intensities of observed (non-zero) and calculated NOE cross-peaks. [c]<rMDIi>, 5 converged structures of conformation I starting from randomly seeded calculations; <rMDIIj>, 7 converged structures of conformation II starting from randomly seeded calculations; <rMDIav>, average of 5 converged structures of conformation I; <rMDIIav>, average of 7 converged structures of conformation II; rmsdI, rmsd of conformation I; rmsdII, rmsd of conformation II.

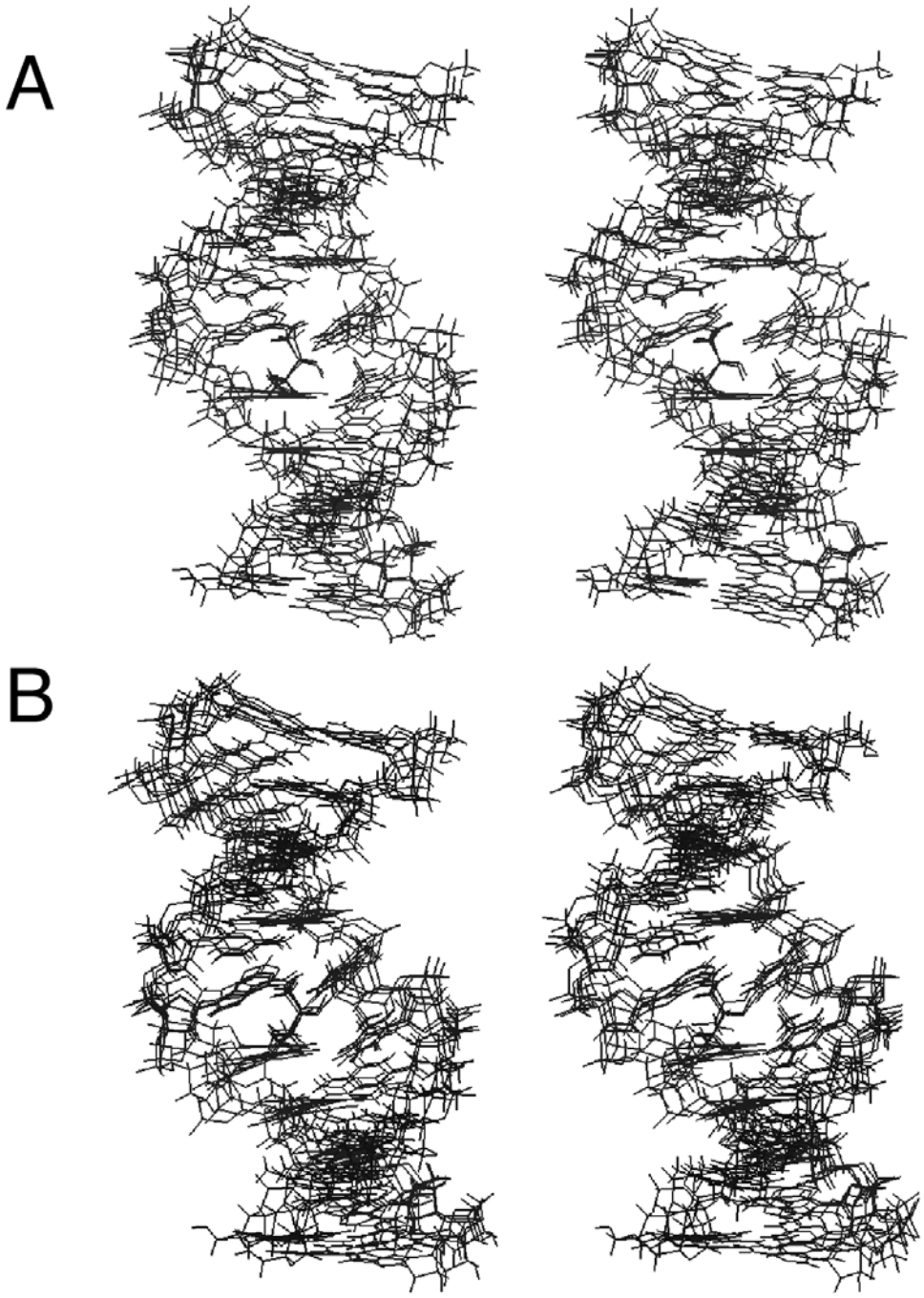


Figure 3-6. **A**. Stereoview of superimposed conformation I structures of the (*S*,*S*)-BD-(61,2-3) cross-linked duplex emergent from the simulated annealing rMD protocol. **B**. Stereoview of superimposed conformation II structures of the (*S*,*S*)-BD-(61,2-3) cross-linked duplex emergent from the simulated annealing rMD protocol. The structures resulted from randomly seeded calculations. (Reproduced with permission from [*Chem Res Toxicol* **2007**, *20*, 187-98.] Copyright [2007] American Chemical Society)

The two sets of converged structures, conformations I and II, that emerged from the rMD calculations were assessed by complete relaxation matrix calculations, using the program CORMA (v. 4.0) (*209*). These calculations yielded sixth root residuals (R_1^x values) between the NOE intensities predicted by the refined structures and the experimentally measured NOE intensities. Figure 3-7 shows R_1^x values as a function of nucleotide position. For both conformations I and II, the calculations revealed reasonable agreement between calculated intensities and experimental NOE data. This reflected the fact that the two conformations differed only in the orientation of the butadiene C_β-hydroxyl group, and the magnitudes of most or all of the 5 butadiene-DNA NOEs did not differ between the two potential conformations. The residual values were consistent over the length of the modified oligodeoxynucleotide, ranging from 0.017 to 0.152. It was concluded that either conformations I or II or an equilibrium mixture of the two could be consistent with the available NOE data, and most likely, the observed spectral data reflected the presence of an equilibrium mixture of both conformations, in rapid exchange on the NMR time scale.

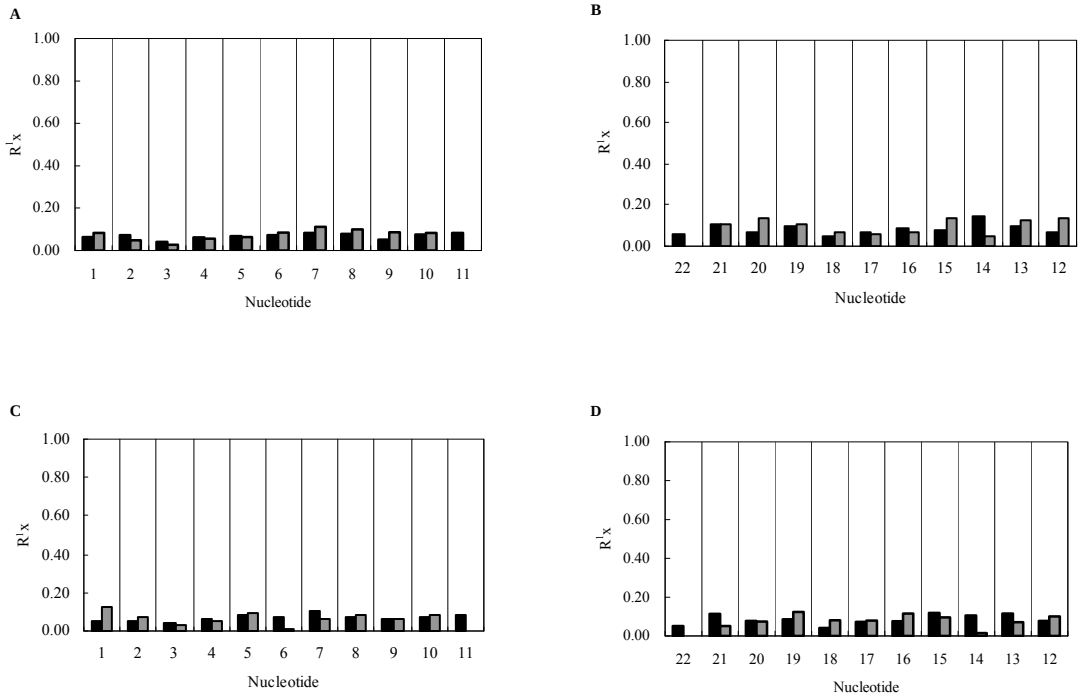


Figure 3-7. Sixth root residuals between calculated NOE intensities and experimental NOE intensities (R_1^x values) as a function of position in the (*S*,*S*)-BD-(61,2-3) adduct. **A.** Nucleotides $C^1 \rightarrow G^{11}$ of conformation I with hydrogen-bonding between β-OH and T^{16} O^4. **B.** Nucleotides $C^{12} \rightarrow G^{22}$ of conformation I. **C.** Nucleotides $C^1 \rightarrow G^{11}$ of conformation II with hydrogen-bonding between β-OH and T^{17} O^4. **D.** Nucleotides $C^{12} \rightarrow G^{22}$ of conformation II. (Reproduced with permission from [*Chem Res Toxicol* **2007**, *20*, 187-98.] Copyright [2007] American Chemical Society)

Structure of the (S,S)-BD-(61-2,3) Cross-Linked Adduct

The (*S*,*S*)-BD-(61-2,3) butadiene cross-link was oriented in the major groove, as shown in Figure 3-8. In the four-carbon butadiene-derived cross-link, the two hydroxyl groups at C_β and C_γ were in the anti-conformation with respect to the C_β-C_γ bond. This allowed the C_β-OH group to insert between T^{16} and T^{17} in the complementary strand, while allowing the C_γ-OH group to orient into the major groove and away from the DNA helix. It resulted in distortions of base pairs $X^6 \cdot T^{17}$ and $Y^7 \cdot T^{16}$ (Figure 3-9). Nevertheless,

base pairs $X^6 \cdot T^{17}$ and $Y^7 \cdot T^{16}$ remained inserted into the DNA duplex, and base stacking (Figure 3-10) was minimally perturbed compared to the corresponding stacking interactions between A^6 and A^7 in the unmodified *ras*61 oligodeoxynucleotide (*238*). Localized perturbation caused at the butadiene cross-link lesion was evident in the local base-base parameters (Figure 3-11 and 3-12) from helicoidal analysis, which suggested a 10 º decrease of the twist angle between base pairs $X^6 \cdot T^{17}$ and $Y^7 \cdot T^{16}$, resulting in an unwinding of the duplex (*120*). An approximate 10 º bending of the cross-linked duplex was calculated using the program CURVES (*121*,*122*), probably a consequence of placing the two hydroxyl groups at C_β and C_γ in the anti-conformation with respect to the C_β-C_γ bond, which did not allow the four-carbon cross-link to exist in the extended chain conformation. The presence of the C_β-OH group inserted between T^{16} and T^{17} in the complementary strand allowed it to swivel either toward T^{16} O^4 or toward T^{17} O^4. Thus, two potential conformations of the (*S*,*S*)-BD-(61-2,3) cross-linked adduct in the *ras*61 oligodeoxynucleotide differed in hydrogen-bonding interactions between the C_β-OH group of the cross-link and T^{17} O^4 (conformation I) or T^{16} O^4 (conformation II).

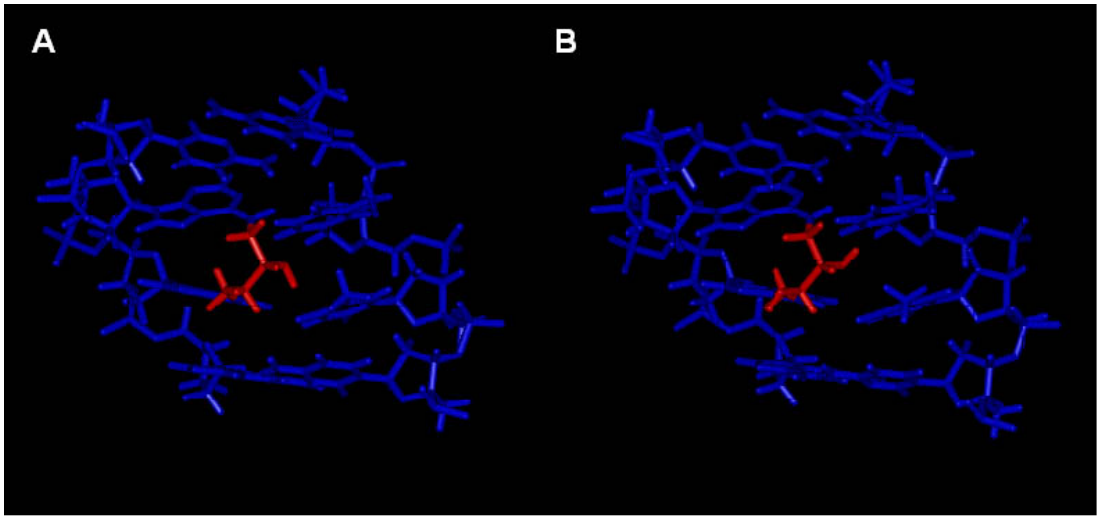


Figure 3-8. Stick models of the (*S*,*S*)-BD-(61,2-3) cross-linked duplex predicted by rMD calculations. **A**. Conformation I with hydrogen-bonding between β-OH and T^{17} O^4. **B**. Conformation II with hydrogen-bonding between β-OH and T^{16} O^4. The BD moiety is in red. (Reproduced with permission from [*Chem Res Toxicol* **2007**, *20*, 187-98.] Copyright [2007] American Chemical Society)

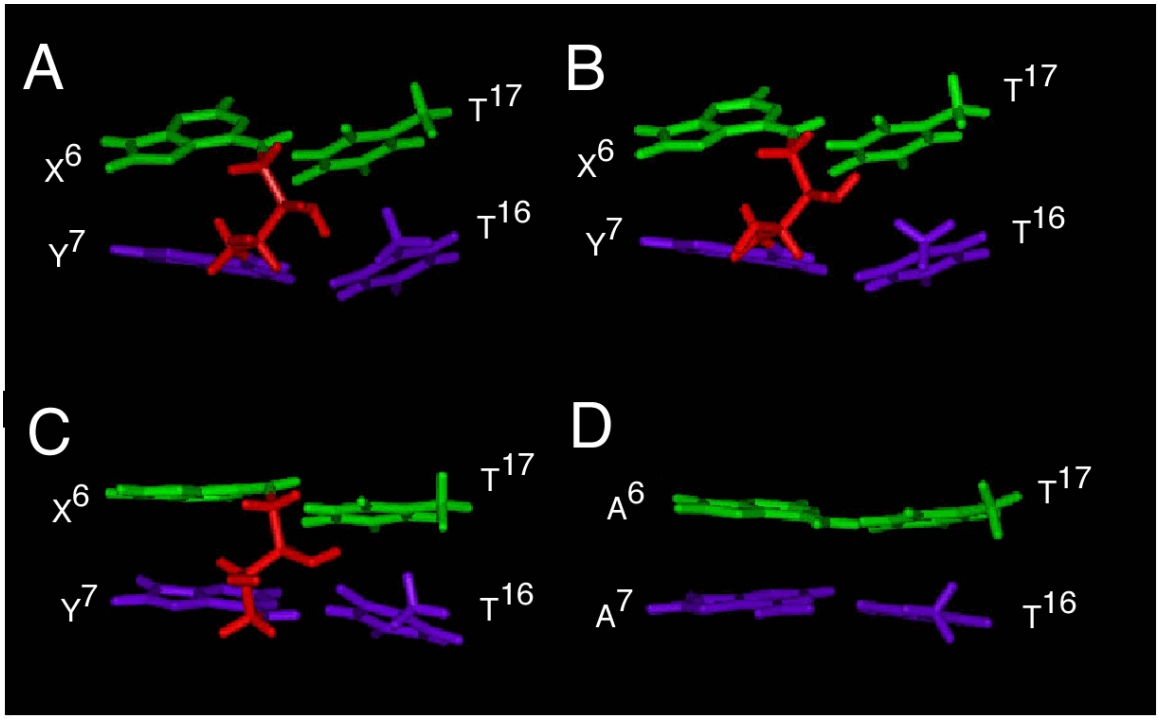


Figure 3-9. Views from the major groove of the duplex DNA of the (*S*,*S*)-BD-(61,2-3) and the (*R*,*R*)-BD-(61,2-3) cross-linked adduct duplexes at the lesion site predicted by rMD calculations, as compared to the unmodified *ras*61 oligodeoxynucleotide duplex (PDB code 1AGH). **A.** Conformation I of the (*S*,*S*)-BD-(61,2-3) cross-linked duplex detailing base pairs $X^6 \cdot T^{17}$ (green) and $Y^7 \cdot T^{16}$ (purple). **B.** Conformation II of the (*S*,*S*)-BD-(61,2-3) cross-linked duplex detailing base pairs $X^6 \cdot T^{17}$ (green) and $Y^7 \cdot T^{16}$ (purple). **C.** The (*R*,*R*)-BD-(61,2-3) cross-linked duplex detailing base pairs $X^6 \cdot T^{17}$ (green) and $Y^7 \cdot T^{16}$ (purple). **D.** The unmodified *ras*61 duplex detailing base pairs $A^6 \cdot T^{17}$ (green) and $A^7 \cdot T^{16}$ (purple). The BD moiety is in red. (Reproduced with permission from [*Chem Res Toxicol* **2007**, *20*, 187-98.] Copyright [2007] American Chemical Society)

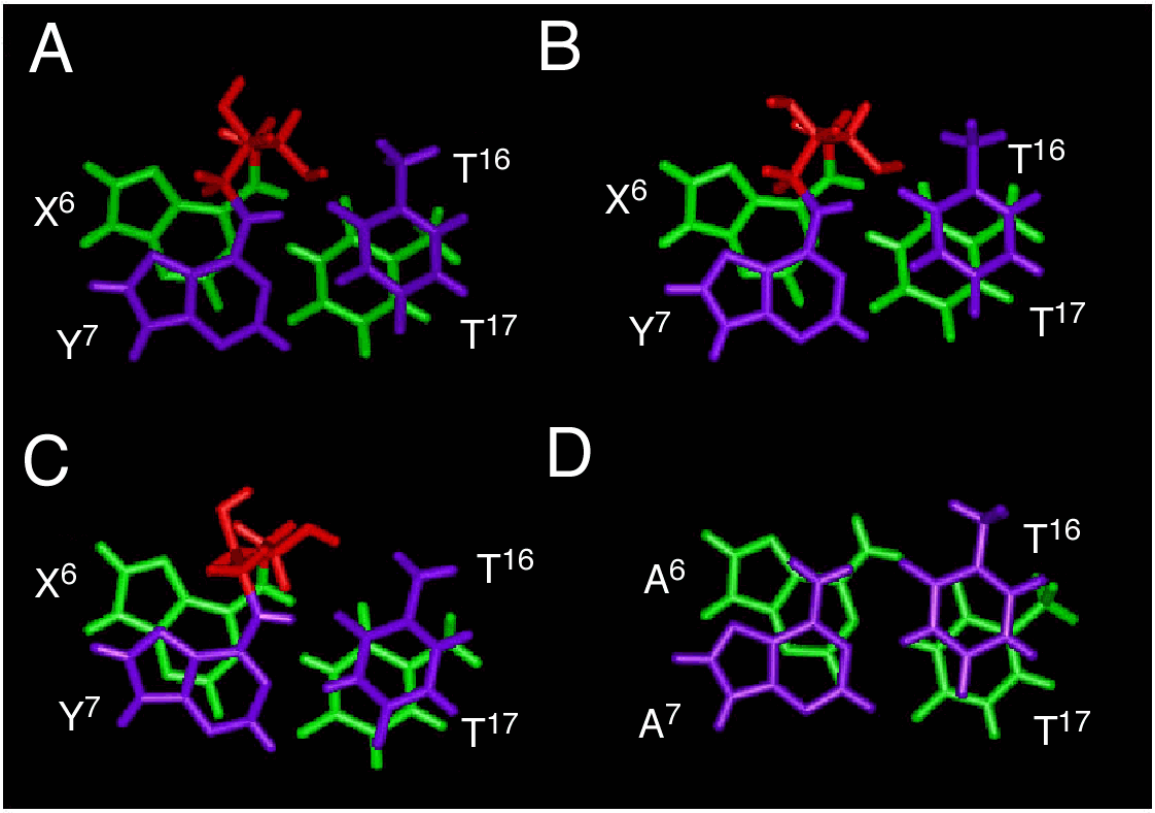


Figure 3-10. Base stacking of the (*S*,*S*)-BD-(61,2-3) and the (*R*,*R*)-BD-(61,2-3) cross-linked duplexes at the lesion site predicted by rMD calculations, as compared to the unmodified *ras61* duplex (PDB code 1AGH). **A.** Conformation I of the (*S*,*S*)-BD-(61,2-3) cross-linked duplex detailing stacking of the base pairs $X^6 \cdot T^{17}$ (green) and $Y^7 \cdot T^{16}$ (purple). **B.** Conformation II of the (*S*,*S*)-BD-(61,2-3) cross-linked duplex detailing stacking of the base pairs $X^6 \cdot T^{17}$ (green) and $Y^7 \cdot T^{16}$ (purple). **C.** The (*R*,*R*)-BD-(61,2-3) cross-linked duplex detailing stacking of the base pairs $X^6 \cdot T^{17}$ (green) and $Y^7 \cdot T^{16}$ (purple). **D.** The unmodified *ras*61 duplex detailing stacking of the base pairs $A^6 \cdot T^{17}$ (green) and $A^7 \cdot T^{16}$ (purple). The BD moiety is in red. (Reproduced with permission from [*Chem Res Toxicol* **2007**, *20*, 187-98.] Copyright [2007] American Chemical Society)

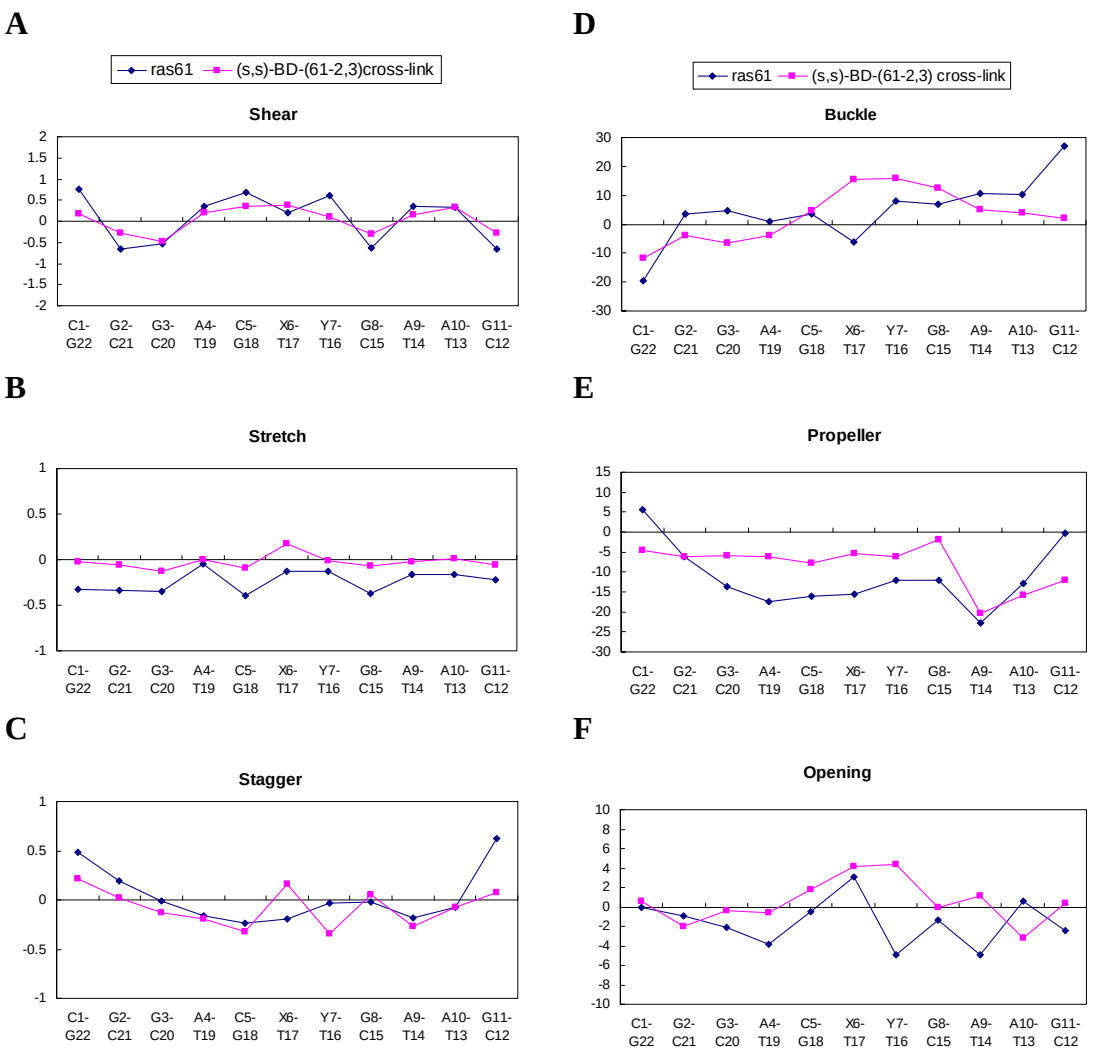


Figure 3-11. Local base-pair parameters: (A) shear, (B) stretch, (C) stagger, (D) buckle, (E) propeller, and (F) opening for the (*S*,*S*)-BD-(61,2-3) cross-linked duplexes, as compared to the unmodified *ras*61 duplex (PDB code 1AGH)

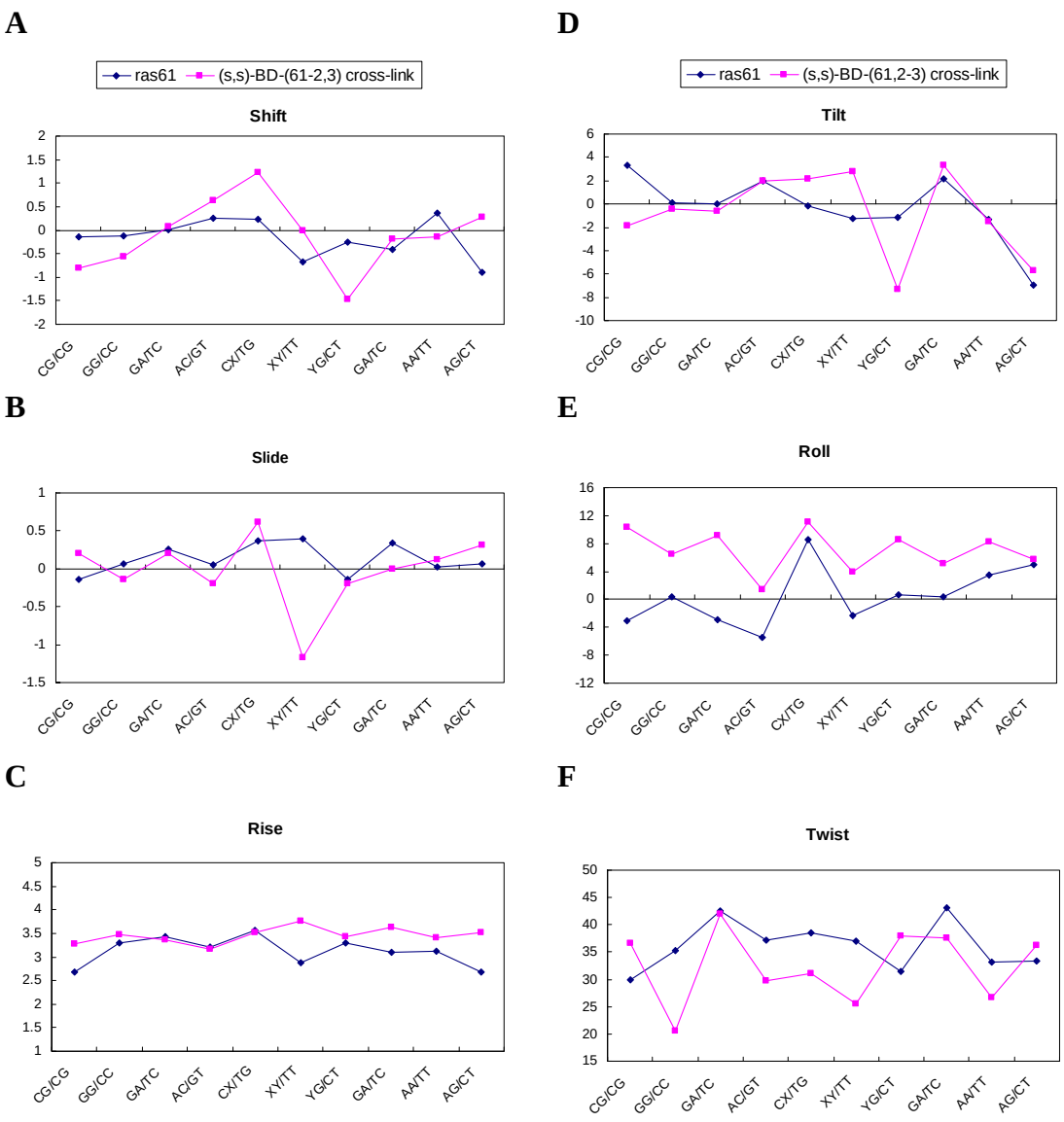


Figure 3-12. Local base-pair parameters: (A) shift, (B) slide, (C) rise, (D) tilt, (E) roll, and (F) twist for the (*S*,*S*)-BD-(61,2-3) cross-linked duplexes, as compared to the unmodified *ras*61 duplex (PDB code 1AGH)

Discussion

The genotoxicity of butadiene may be related to its ability to form cross-links in DNA via its diepoxide oxidation product. Animal studies revealed species differences between mice and rats with regard to BD genotoxicity, with mice exhibiting greater sensitivity (*29,30,239*). This was explained by the observation that in mice, the conversion of BD to DEB was more efficient than that in rats, an observation that also pointed to DEB as the key proximate electrophile in BD-mediated genotoxicity (*64*). This suspicion was confirmed by studies showing that DEB was indeed considerably more mutagenic than the mono-epoxide EB (*240*), probably due to its cross-linking ability. The primary evidence suggesting the presence of DEB-induced DNA cross-links *in vivo* comes from studies in which human lymphoblastoid TK6 and splenic T cells were exposed to low levels of DEB (*65,240,241*). These experiments revealed transitions and transversions at both G·C and A·T sites, suggesting the presence of both dG and dA adducts arising from DEB. They also revealed the presence of significant levels of deletion mutants, consistent with the presence of DNA cross-links (*65,240,241*).

Structural Analysis of the (S,S)-BD-(61-2,3) Cross-Link

The major groove orientation of the (*S*,*S*)-BD-(61-2,3) cross-link was consistent with the observation of few NOEs between the butadiene cross-link and DNA, placing the butadiene protons far from major groove nonexchangeable DNA protons. Nevertheless, the NOEs between the butadiene H_{α} and $H_{\alpha'}$ protons and the major groove C^5 H5 proton of the 5' neighboring base served to locate the cross-link in the major groove. The NOEs observed between the butadiene H_{α}, $H_{\alpha'}$, H_{β}, and the exchangeable

proton T^{17} N3H further oriented the cross-link in the major groove. The anti-conformation of the two hydroxyl groups at C_β and C_γ with respect to the C_β-C_γ bond resulted in a decrease of the twist angle between base pairs $X^6 \cdot T^{17}$ and $Y^7 \cdot T^{16}$ and an approximate 10 º bending of the cross-linked DNA duplex. This was accompanied by the disruption of Watson-Crick hydrogen-bonding at base pair $X^6 \cdot T^{17}$. The observation of a break in sequential NOE connectivity between T^{16} H6 and T^{17} H1' (Figure 3-1) and the 0.46 and 0.29 ppm upfield shifts of the T^{17} N3H and T^{17} H6 resonances, respectively (Figure 3-2), supported the conclusion that Watson-Crick base pairing was disrupted at base pair $X^6 \cdot T^{17}$. This probably also explained the reduction in the T_m of the (*S*,*S*)-BD-(61-2,3) cross-linked duplex compared to that of the unmodified *ras*61 duplex. The potential formation of a hydrogen bond between the C_γ-OH group of the cross-link and either T^{16} O^4 or T^{17} O^4, predicted by the rMD calculations, probably stabilized the conformation of the cross-link.

Comparison with the (R,R)-BD-(61-2,3) Cross-Linked Adduct

Like the (*R*,*R*)-BD-(61-2,3) cross-link, the (*S*,*S*)-BD-(61-2,3) cross-link was oriented in the major groove of DNA. For both cross-links, Watson-Crick base pairing was disrupted at base pair $X^6 \cdot T^{17}$, whereas at base pair $Y^7 \cdot T^{16}$, base pairing was less perturbed. The major conformational difference between the (*R*,*R*)-BD-(61-2,3) and (*S*,*S*)-BD-(61-2,3) cross-links was in the conformation of the four-carbon butadiene chain. As regards the (*R*,*R*)-BD-(61-2,3) cross-link, the butadiene linkage existed in an extended chain conformation, placing the two hydroxyl groups at C_β and C_γ in the syn-conformation with respect to the C_β-C_γ bond. This allowed the (*R*,*R*)-BD-(61-2,3) cross-link to be

accommodated within the major groove of the DNA with minimal structural perturbation of the duplex. In contrast, the anti-conformation of the two hydroxyl groups at C_β and C_γ with respect to the C_β-C_γ bond in the (*S*,*S*)-BD-(61-2,3) cross-link required greater structural perturbation to the DNA duplex, resulting in a decreased twist between base pairs $X^6 \cdot T^{17}$ and $Y^7 \cdot T^{16}$ and an approximate 10 º bending of the cross-linked duplex.

Structure-Activity Relationships

The present results confirm a role for stereochemistry in modulating the conformations of the diastereomeric (*R*,*R*)- and (*S*,*S*)-BD-(61-2,3) butadiene-derived cross-links. Although the (*R*,*R*)-BD-(61-2,3) cross-link existed in the extended chain conformation with minimal perturbation of the DNA duplex, the (*S*,*S*)-BD-(61-2,3) cross-link created a greater structural perturbation at base pairs $X^6 \cdot T^{17}$ and $Y^7 \cdot T^{16}$. These differential structural perturbations could influence the recognition and biological processing of these cross-links. Although both (*R*,*R*)- and (*S*,*S*)-BD-(61-2,3) cross-links were highly mutagenic in both *E. coli* and COS-7 cells, the (*S*,*S*)-BD-(61-2,3) cross-link exhibited a somewhat lower overall mutagenic frequency (20%) than that of the (*R*,*R*)-BD-(61-2,3) cross-linked adduct (54%) (*92*).

Park et al. (*58*) proposed a model in which the differences between the ability of different DEB stereoisomers to induce N7,N7-dG cross-links might be caused by the different orientations of functional groups in stereoisomeric N7-(2'-hydroxy-3',4'-epoxybut-1'-yl)-guanine intermediates. Thus, for *S*,*S*- and *R*,*R*-N7-(2'-hydroxy-3',4'-epoxybut-1'-yl)-guanine, the oxirane and the 2'-hydroxy group reside on one side of the plane formed by the carbon chain, whereas in the cross-link arising from *meso*-DEB, the

oxirane and the 2'-OH are on different sides of the plane, potentially influencing the site of the second alkylation. Molecular modeling suggested that the formation of a hydrogen bond between the 2'-hydroxy group in *S*,*S*- and *R*,*R*-N7-(2'-hydroxy- 3',4'-epoxybut-1'-yl)-guanine and N3 of the 3'-neighbor dG would favorably position the oxirane for alkylation at N7-dG in the opposite strand. In contrast, the formation of a hydrogen bond between the 2'-hydroxyl and 3'-neighbor dG in *meso*-DEB-induced *S*,*R*- and *R*,*S*-epoxy alcohol intermediates would stabilize the conformation in which the oxirane faced N7-dG in the opposite strand. As a result, the regiochemistry of cross-linking would differ. Likewise, for the *S*,*S*-N^6-(2',3',4'-trihydroxybut-1'-yl)-deoxyadenosyl adduct, hydrogen bond formation was observed between the β-OH of the side chain and the O^4 of thymine in the opposite strand, whereas for the *R*,*R* adduct, hydrogen-bonding was observed between the γ-OH and the thymine O^4 in the opposite DNA strand (*242*).

Biological Significance

Mammalian mutagenesis studies performed on B6c3F1 *lacI* transgenic mice showed point mutations at both dG and dA; the primary adenine-specific point mutations were A→T transversions (*243-247*). The site-specific mutagenesis experiments on both stereoisomers of N^6,N^6-dA intrastrand cross-linked adducts yielded primarily A→G mutations (*92*), indicating that the DEB-induced cross-linked adducts did not represent the source of the predominant A→T mutations induced by butadiene in bacterial or mammalian cells. Nonetheless, butadiene was reported to induce A→G transitions in the H-*ras* locus (*248*), to which might be attributed to DEB-induced cross-links. The N1-deoxyinosine adducts represent another potential source of A→G transitions. Site-

specific mutagenesis studies of the N1-deoxyinosine adducts arising from EB in the COS-7 cells revealed that they were strongly mutagenic with respect to A→G mutations. Consequently, the specific DNA adducts arising from exposure to BD and responsible for inducing A→T mutations remain to be determined. Mutagenesis studies of the (*S,S*)-BD-(61-2,3) cross-linked adduct suggested that A→G transitions at the 3'-adenine were the primary point mutations in both *E. coli* and COS-7 cells. The same results were observed for the (*R,R*)-BD-(61-2,3) cross-linked adduct. This observation suggested that both stereoisomeric cross-linked adducts facilitate misincorporation of dCTP opposite cross-linked nucleotide Y^7. An examination of the structure of a mismatched dC opposite the N^6, N^6-dA cross-linked adduct will be of interest. The predominant adducts formed upon the exposure of DNA to butadiene epoxides are N7-dG adducts (*80*). The N7-N7 guanine interstrand cross-linked adduct has been determined to be the major DNA-DNA cross-link induced by DEB (*37*). It will also be of considerable interest to examine the structures of N7-guanine adducts.

Summary

The (*S,S*)-BD-(61-2,3) cross-linked adduct at N^6 in the ras61 coding sequence was oriented in the major groove, resulting in the opening of base pair $X^6 \cdot T^{17}$. Two potential conformations obtained from NMR data refinement exhibited similar structures with different hydrogen-bonding interactions between the C_β-OH group of the cross-link and T^{16} O^4 or T^{17} O^4.

CHAPTER IV

INCORPORATION OF NUCLEOSIDE TRIPHOSPHATES OPPOSITE STEREOISOMERIC N3-(2*S* OR 2*R*-HYDROXY-3-BUTEN-1-YL)- 2'-DEOXYURIDINE ADDUCTS BY THE *SULFOLOBUS SOLFATARICUS* DNA POLYMERASE IV (DPO4)

Introduction

The N3-(2*S* or 2*R* -hydroxy-3-buten-1-yl)-2'-deoxyuridine adducts (*S*- or *R*- BD-N3-dU) arise from deamination of the N3-deoxycytidine adducts initially formed from the reaction between EB and deoxycytidine (*75,91*). Replication bypass and primer extension analysis demonstrated that Dpo4 was able to incorporate G or A opposite the adducted base, but the primer extension was blocked. Eight X-ray crystal structures of ternary Dpo4-DNA-d(d)NTP complexes containing either an *S*- or *R*- BD-N3-dU adducted template with all four d(d)NTPs were determined at resolutions between 1.8 and 2.5 Å. The primer-template sequence, 5'-d(TCACXAAATCCTTCCCCC)-3'·5'-d(GGGGGAAGG-ATTT)-3', where X is *S*- or *R*- BD-N3-dU, was analyzed (Scheme 4-1). The structure of each complex solved here represented the most stable conformation obtained from the structural refinement. Although the structures indicated that each complex was non-productive, the disordered incoming nucleotide suggested that multiple intermediates exist in each complex. Thus, the flexilbe dGTP and dATP are possible to reach catalytically active conformations in the complexes.

This is the first time that the ternary Dpo4-DNA-d(d)NTP complexes containing non-planar adducts in the template have been studied. The resulting crystal structures

suggest steric effect plays a significant role in the bypass of damaged DNA by trans-lesion polymerase Dpo4.

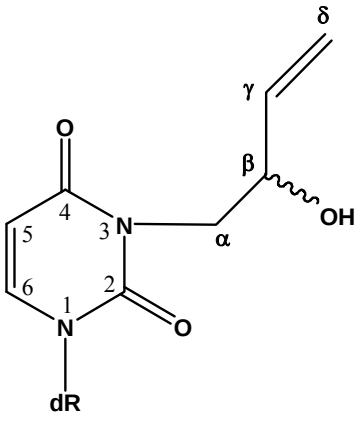
A

Primer 5' GGG GGA AGG ATT T 3'
Template 3' CCC CCT TCC TAA A**X**C ACT 5' + Dpo4 + d(d)**N**TP

X = ***S***- or ***R***-BD-N3-dU

N = **A** or **C** or **T** or **G**

B



Scheme 4-1. The primer-template sequence (A) and chemical structure (B) of the N3-(2*S* or 2*R*-hydroxy-3-buten-2-yl)-2'-deoxyuridine adducts and nomenclature.

Results

Replication and Extension of Primer by Dpo4

The replication bypass experiments were done by Dr. Surajit Banerjee in the laboratory of Professor Michael P. Stone. In the experiment containing unmodified primer-template sequence, Dpo4 inserted dGTP opposite the template C in 8 min, and extended the primer-template to full-length extension products successfully in 15 min with the existence of all four dNTPs in the reaction mixture. In the replication bypass of the template containing either the *S*- or *R*- BD-N3-dU adduct, Dpo4 inserted dATP and dGTP opposite adducted lesion in 15 min. When all four dNTPs were included in the reaction, the Dpo4 polymerase did not extend the primer strand to the full-length product in 20 min. The results of the replication and extension experiments are shown in Figure 4-1.

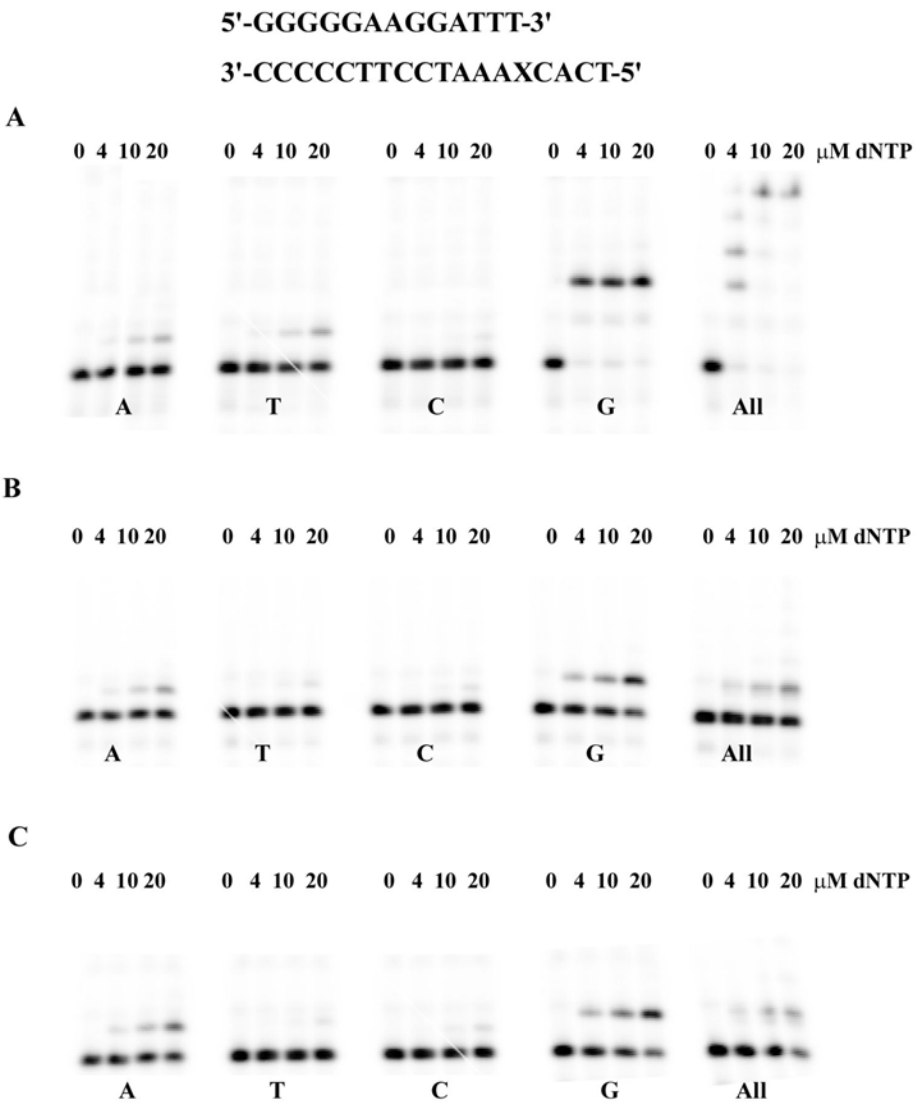


Figure 4-1. Replication bypass of (A) control, (B) *S*-BD-N3-dU adducted, (C) *R*-BD-N3-dU adducted primer-template complex with Dpo4. The reactions were performed with dNTP at the concentration of 0, 4, 10 and 20 μM. The first four panels in each assay represent insertion of a single dNTP to either the control or the adducted template. The last panel of each assay represents the inclusion of all four dNTPs in the reaction mixture and their addiction to either the control or the adducted template. The assays were carried out at 37 ºC, using 100 nM Dpo4 by Dr. Surajit Banerjee.

X-ray Diffraction Data Processing

Crystallization trials were conducted for Dpo4-DNA complexes containing either *S*- or *R*- BD-N3-dU adducted templates, in the presence of all four d(d)NTPs. Combination of Dpo4-DNA complexes containing each stereoisomeric BD adducted template with four d(d)NTPs resulted in eight ternary complexes, including the *S*-adduct with dATP (SdU/dATP), the *S*-adduct with ddCTP (SdU/ddCTP), the *S*-adduct with dTTP (SdU/dTTP), the *S*-adduct with dGTP (SdU/dGTP), the *R*-adduct with dATP (RdU/dATP), the *R*-adduct with dCTP (RdU/dCTP), the *R*-adduct with ddTTP (RdU/ddTTP), and the *R*-adduct with dGTP (RdU/ddTTP). All eight ternary complexes were crystallized and diffracted within a range of 1.8 – 2.5 Å. The statistics of data processing and data quality are summarized in Tables 4-1 and 4-2. Autoindexing and systematic absences indicated the space group $P2_12_12$ for all eight complexes. The resolution was 1.86 Å with a completeness of 99.0% for SdU/dATP, 2.25 Å with a completeness of 95.6% for SdU/ddCTP, 1.80 Å with a completeness of 94.3% for SdU/dTTP, 2.00 Å with a completeness of 96.7% for SdU/dGTP, 2.50 Å with a completeness of 97.7% for RdU/dATP, 2.45 Å with a completeness of 96.5% for RdU/dCTP, 2.40 Å with a completeness of 96.6% for RdU/ddTTP, and 2.50 Å with a completeness of 95.4% for RdU/dGTP. The resulting data sets for the ternary complexes were of good quality as revealed by R_{merge} values of 6.5% (SdU/dATP), 7.9% (SdU/ddCTP), 6.3% (SdU/dTTP), 5.8% (SdU/dGTP), 6.6% (RdU/dATP), 8.0% (RdU/dCTP), 7.2% (RdU/ddTTP), and 7.4% (RdU/dGTP), as well as signal-to-noise ratios (I/σ_I) of 11.3 (SdU/dATP), 14.3 (SdU/ddCTP), 11.5 (SdU/dTTP), 13.5

(SdU/dGTP), 14.3 (RdU/dATP), 21.9 (RdU/dCTP), 14.7 (RdU/ddTTP), and 18.5 (RdU/dGTP), respectively.

Table 4-1. Crystal data and refinement parameters for four ternary complexes of *S*-adduct.

Parameters	SdU/dATP	SdU/ddCTP	SdU/dTTP	SdU/dGTP
X-ray source	BM-22 (APS)	ID-5 (APS)	ID-22 (APS)	ID-22 (APS)
Wavelength (Å)	1.00	0.977	1.00	1.00
Temperature (K)	110	110	110	110
No. of crystals	1	1	1	1
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit cell (a,b,c)(Å)	94.53,100.52,53.47	93.52,100.75,53.22	94.49,100.71,53.20	92.51,100.70,52.92
Resolution range (Å)	27.335-1.86	29.775-2.25	44.439-1.80	45.932-2.0
Highest resolution shell	1.93-1.86	2.42-2.25	1.86-1.80	2.09-2.0
No. of measurements	659587	183470	1109692	623704
No. of unique reflections	43197	31911	45414	33122
Redundancy	6.9	5.2	6.1	5.9
Completeness (%)	99.0 (94.9)[a]	95.6 (85.7)	94.3 (57.8)	96.7 (83.2)
R_{merge} [b]	0.065	0.079	0.063	0.058
Signal/noise ($<I/\sigma^I>$)	11.3	14.3	11.5	13.5
Solvent content (%)	53.55	49.65	54.42	48.79
Model composition				
No. of amino acid nucleotides	342	342	342	342
No. of water molecules	93	127	49	104
No. of Ca^{2+} ions	3	3	3	3
No. of template nucleotides	17	17	18	18
No. of primer nucleotides	13	13	13	13
No. of dATP	1			
No. of ddCTP		1		
No. of dTTP			1	
No. of dGTP				1
R[c] (%)	24.6	22.7	25.3	23.0
R_{free} [d] (%)	27.0	27.1	27.4	27.0
Estimated coordinate error (Å)				
From Luzatti plot	0.29	0.31	0.31	0.29
From Luzatti plot (c-v)[e]	0.33	0.39	0.35	0.36
From σA plot	0.24	0.36	0.31	0.26
From σA plot (c-v)[e]	0.27	0.54	0.34	0.29
r.m.s. deviation in temp. factors				
Bonded main chain atoms ($Å^2$)	1.45	1.36	1.41	1.47
Bonded side chain atoms ($Å^2$)	2.11	1.99	1.99	2.01
r.m.s. S.D. from ideal values				
Bond lengths (Å)	0.006	0.006	0.006	0.006
Bond angles (º)	1.21	1.21	1.17	1.25
Dihedral angles (º)	21.13	21.74	20.91	21.04
Improper angles (º)	1.04	1.02	0.91	1.93

[a] Values in parentheses correspond to the highest resolution shells. [b] $R\text{merge}=\sum_{hkl}\sum_{j=1,N}|<I_{hkl}>-I_{hkl,j}|/\sum_{hkl}\sum_{j=1,N}|I_{hkl,j}|$, where the outer sum (*hkl*) is taken over the unique reflections. [c] $R_f=\sum_{hkl}\sum||F_{o,hkl}|-k|F_{c,hkl}||/\sum_{hkl}|F_{o,hkl}|$, where $|F_{o,hkl}|$ and $|F_{c,hkl}|$ are the observed and calculated structure factor amplitudes, respectively. [d] R_{free} *idem*, for the set of reflections (5% of the total) omitted from the refinement process.

Table 4-2. Crystal data and refinement parameters for four ternary complexes of *R*-adduct.

Parameters	RdU/dATP	RdU/dCTP	RdU/ddTTP	RdU/dGTP
X-ray source	BM-22(APS)	ID-22(APS)	ID-22 (APS)	ID-22(APS)
Wavelength (Å)	1.00	1.00	1.00	1.00
Temperature (K)	110	110	110	110
No. of crystals	1	1	1	1
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit cell (a,b,c)(Å)	95.41,102.06,52.57	94.48,102.37,52.53	95.90,102.23,52.82	95.48,102.22,52.68
Resolution range (Å)	29.047-2.5	45.004-2.45	45.109-2.40	45.061-2.50
Highest resolution shell	2.59-2.5	2.54-2.45	2.49-2.40	2.59-2.50
No. of measurements	653760	479099	1251882	484794
No. of unique reflections	18048	18635	20015	17410
Redundancy	6.9	6.5	6.6	6.7
Completeness (%)	97.7 (85.4)[a]	96.5 (84.6)	96.6 (77.1)	95.4 (78.5)
R_{merge} [b]	0.066	0.080	0.072	0.074
Signal/noise ($<I/\sigma^I>$)	14.3	21.9	14.7	18.5
Solvent content (%)	54.05	52.25	54.61	53.42
Model composition				
No. of amino acid nucleotides	342	343	342	343
No. of water molecules	37	77	26	37
No. of Ca^{2+} ions	3	3	2	3
No. of template nucleotides	15	17	15	17
No. of primer nucleotides	13	13	13	13
No. of dATP	1			
No. of dCTP		1		
No. of ddTTP			1	
No. of dGTP				1
R[c] (%)	23.5	23.4	23.2	23.8
R_{free} [d] (%)	28.9	27.7	26.8	28.6
Estimated coordinate error (Å)				
From Luzatti plot	0.41	0.42	0.40	0.41
From Luzatti plot (c-v)[e]	0.51	0.50	0.48	0.52
From σA plot	0.52	0.54	0.51	0.54
From σA plot (c-v)[e]	0.63	0.54	0.57	0.55
r.m.s. deviation in temp. factors				
Bonded main chain atoms ($Å^2$)	1.53	1.67	1.53	1.52
Bonded side chain atoms ($Å^2$)	1.98	2.11	2.06	1.98
r.m.s. S.D. from ideal values				
Bond lengths (Å)	0.008	0.007	0.008	0.007
Bond angles (º)	1.21	1.21	1.29	1.23
Dihedral angles (º)	21.99	22.33	21.66	22.23
Improper angles (º)	1.06	0.96	1.03	1.92

[a] Values in parentheses correspond to the highest resolution shells. [b] *R*merge=$\sum_{hkl}\sum_{j=1,N}|<I_{hkl}>-I_{hkl,j}|/\sum_{hkl}\sum_{j=1,N}|I_{hkl,j}|$, where the outer sum (*hkl*) is taken over the unique reflections. [c] $R_f=\sum_{hkl}\sum||F_{o,hkl}|-k|F_{c,hkl}||/\sum_{hkl}|F_{o,hkl}|$, where $|F_{o,hkl}|$ and $|F_{c,hkl}|$ are the observed and calculated structure factor amplitudes, respectively. [d] R_{free} *idem*, for the set of reflections (5% of the total) omitted from the refinement process.

Crystal Structures of Ternary Complexes Containing S-BD-N3-dU Adducted Template

The crystal structure of the ternary Dpo4-DNA-dATP complex containing *S*-BD-N3-dU adducted template (SdU/dATP) was determined at a resolution of 1.86 Å (Table 4-1). In the crystal structure, only the *S*-BD-N3-dU nucleotide of the template was lodged at the active site of Dpo4 (Figure 4-2 and 4-3A), which is different from the structure of Dpo4 ternary complex with native DNA in which two adjacent template bases were accommodated at the active site (*144*). The *S*-BD-N3-dU nucleotide stacked inside the duplex with the BD moiety oriented in the 5' direction (Figure 4-4A). Due to the non-planar butadiene lesion, the 5'-adjacent template C flipped out into the major groove and was directed away from the active site. The incoming dATP was inserted into the active site of Dpo4 and did not pair with the *S*-BD-N3-dU nucleotide. The electron density was complete for the entire structure except for the dATP (Figure 4-5A), indicating that the incoming dATP was disordered or that the active site for the incoming nucletide was only partially occupied. The base moieties of dATP and the 3'-terminal primer T were tilted 9.4 º and the distance between the 3'-hydroxy group of the terminal primer T and the α-phosphate group of dATP was 7.9 Å, which is out of range for the chemical reaction. The hydrogen-bonding interaction between the hydroxyl group of butadiene and O3' of the incoming dATP resulted in a twist of 55.6 º and an opening of 22.3 º between the *S*-BD-N3-dU nucleotide and dATP (Table 4-3). Three calcium ions were coordinated in this structure (Figure 4-6A), with two Ca^{2+} ions chelated by the three catalytic carboxylates - Asp^{7}, Asp^{105} and Glu^{106}. These carboxylates stabilized the phosphate moiety of the incoming dATP. However, as reversed conditions of complex structure with native DNA, one Ca^{2+} ion was closer to the γ-phosphate than the α-phosphate of dATP. The third Ca^{2+}

ion was close to two oxygens of the α-phosphate group linking the 12th and 13th primer nucleotides.

Similar structures were solved for the other three ternary complexes including SdU/ddCTP, SdU/dTTP, and SdU/dGTP at a resolution of 2.25 Å, 1.8 Å and 2.0 Å, respectively (Table 4-1). The superposition of DNA conformations at the active site of all four ternary complexes manifested similarity of conformations of the primer and template nucleotides at the active site (Figure 4-7). The active sites of all of the complexes resemble neither the "type I", nor the "type II" crystal structures of the complex between Dpo4 and native DNA (*144*). In each structure, only the *S*-BD-N3-dU nucleotide of the template was accommodated at the active site of Dpo4, stacked into the duplex with the BD moiety oriented in the 5' direction, similar to the SdU/dATP structure (Figure 4-3). Due to the non-planar butadiene lesion, the 5'-adjacent template C was looped out and rotated into the major groove. Although all four d(d)NTPs were able to be accommodated at the active site of Dpo4, the conformation of d(d)NTPs varied a lot, and they were not stacking on the 3'-teminal primer T (Figure 4-4). The conformation suggested that the spacious open site of Dpo4 was able to accommodate all kinds of incoming nucleotides with regardless of synthetic activity. The distance between the 3'-OH groups of the terminal primer nucleotide and α-phosphates of the d(d)NTP of each complex was larger than that of the active conformation of either "type I" or "type II" complex. The incomplete density of the incoming d(d)NTPs observed in all complexes (Figure 4-5) indicated that the d(d)NTPs are flexible and disordered at the active site or that the active site were only partial occupied. Moreover, the metal ions at the active site occupied similar sites in all complexes with *S*-BD-N3-dU modified templates (Figure 4-6),

attributed to stabilize the phosphate moiety of the incoming d(d)NTP despite lack of catalytic activity. Detailed analysis revealed that the altered orientations of the nucleotide triphosphates went along with different sets of hydrogen-bonding interactions with Dpo4 side chains and the BD-N3-dU nucleotide in the respective complexes, resulting in the twist and opening between the S-BD-N3-dU nucleotide and d(d)NTP (Table 4-3).

Figure 4-2. Overall structures of the Dpo4 ternary complex SdU/dATP. The primer-template sequence is shown in Scheme 4-1, with X=*S*-BD-N3-dU and N=A. the Dpo4 backbone and the primer-template are shown in ribbon diagram in cyan and gray, respectively. The *S*-BD-N3-dU nucleotide is shown in stick model colored on atom types; the dATP is shown in stick model in yellow and Ca^{2+} ions are shown as spheres in green.

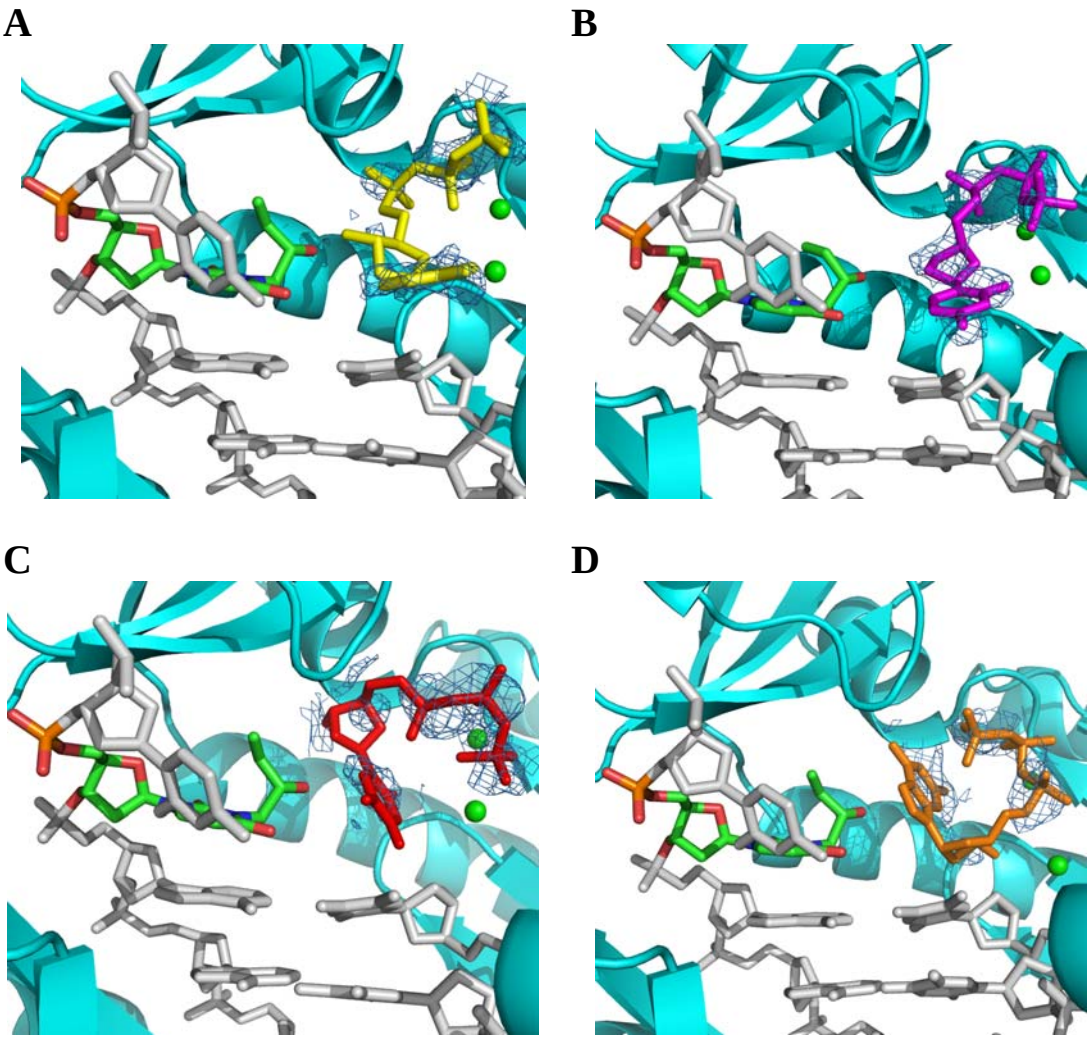


Figure 4-3. Structures of four ternary Dpo4-DNA-d(d)NTP complexes containing *S*-BD-N3-dU adducted template at the active site of Dpo4. (A) SdU/dATP, (B) SdU/ddCTP, (C) SdU/dTTP, and (D) SdU/dGTP. The Dpo4 backbone is shown in ribbon diagram in cyan. The primer-template are shown in stick model in gray, with the *S*-BD-N3-dU nucleotide colored on atom types, and the dATP, ddCTP, dTTP and dGTP are shown in stick model in yellow, purple, red and orange, respectively. Ca^{2+} ions are shown as spheres in green.

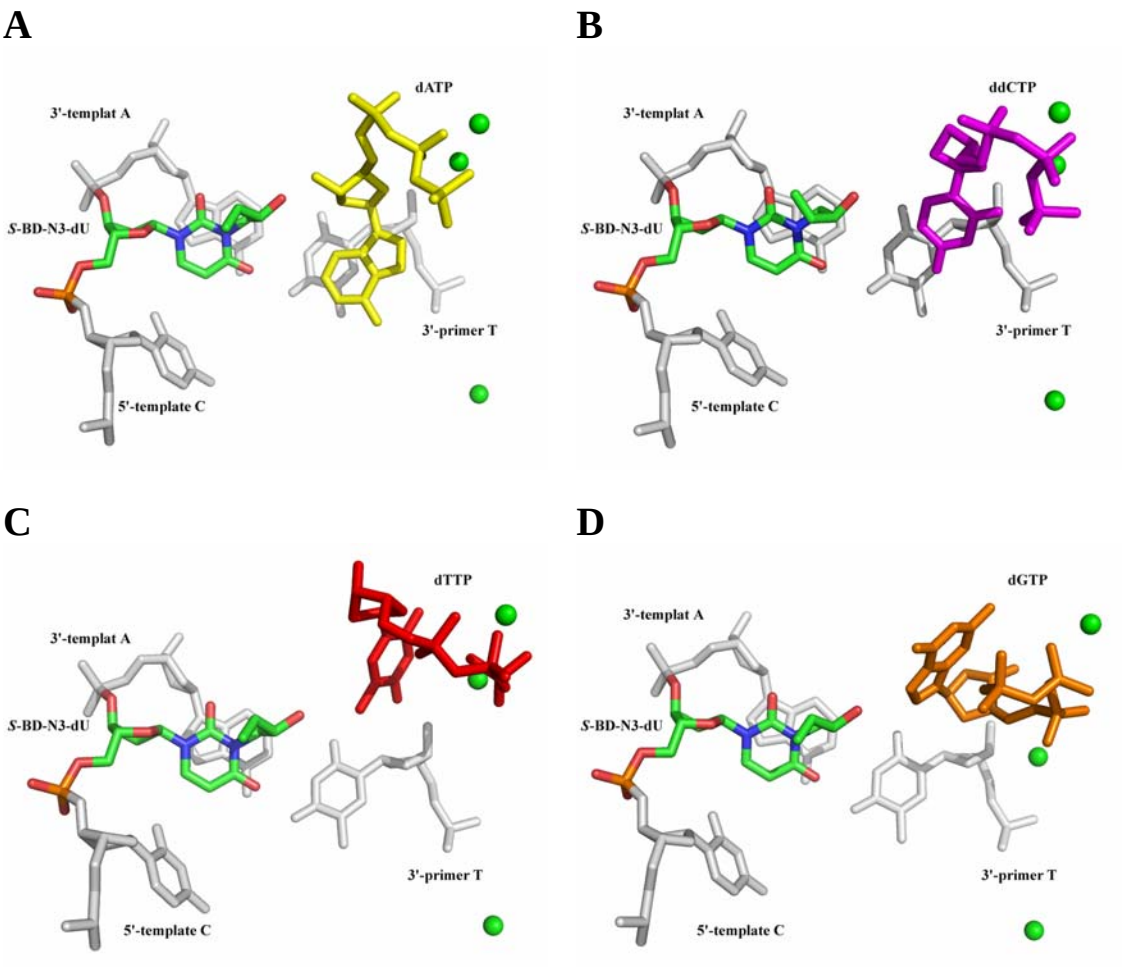


Figure 4-4. Stacking patterns of DNA base pairs at the active site of four ternary Dpo4-DNA-d(d)NTP complexes. (A) SdU/dATP, (B) SdU/ddCTP, (C) SdU/dTTP, and (D) SdU/dGTP. The primer-template are shown in stick model in gray, with the *S*-BD-N3-dU nucleotide colored on atom types, and the dATP, ddCTP, dTTP and dGTP are shown in stick model in yellow, purple, red and orange, respectively. Ca^{2+} ions are shown as spheres in green.

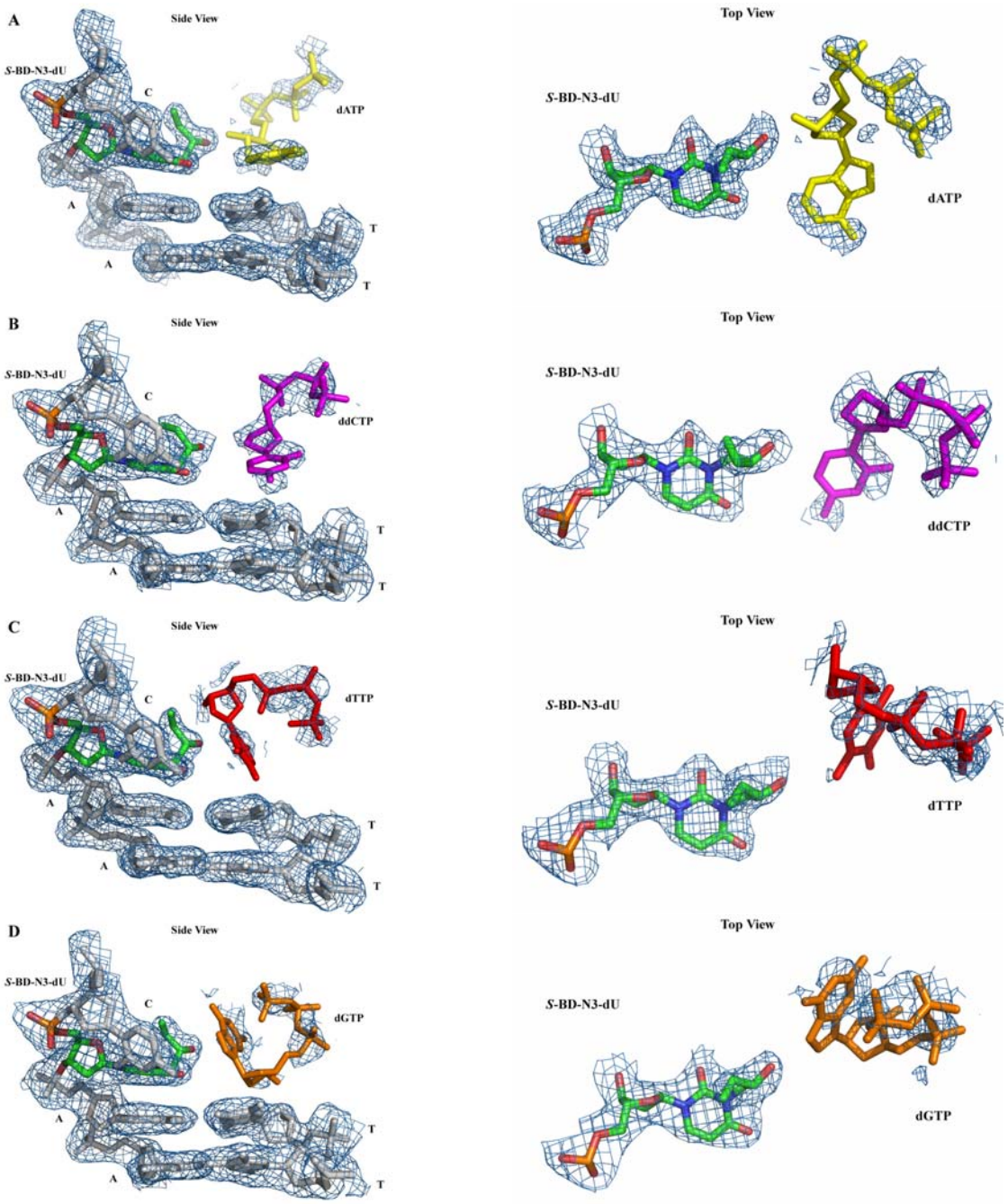


Figure 4-5. Electron density of DNA and d(d)NTPs at the active site of four ternary Dpo4-DNA-d(d)NTP complexes. (A) SdU/dATP, (B) SdU/ddCTP, (C) SdU/dTTP, and (D) SdU/dGTP. Fourier $2F_o$-F_c electron density (blue meshwork) is drawn at the 1σ level. The primer-template are shown in stick model in gray, with the *S*-BD-N3-dU nucleotide colored on atom types, and the dATP, ddCTP, dTTP and dGTP are shown in stick model in yellow, purple, red and orange, respectively.

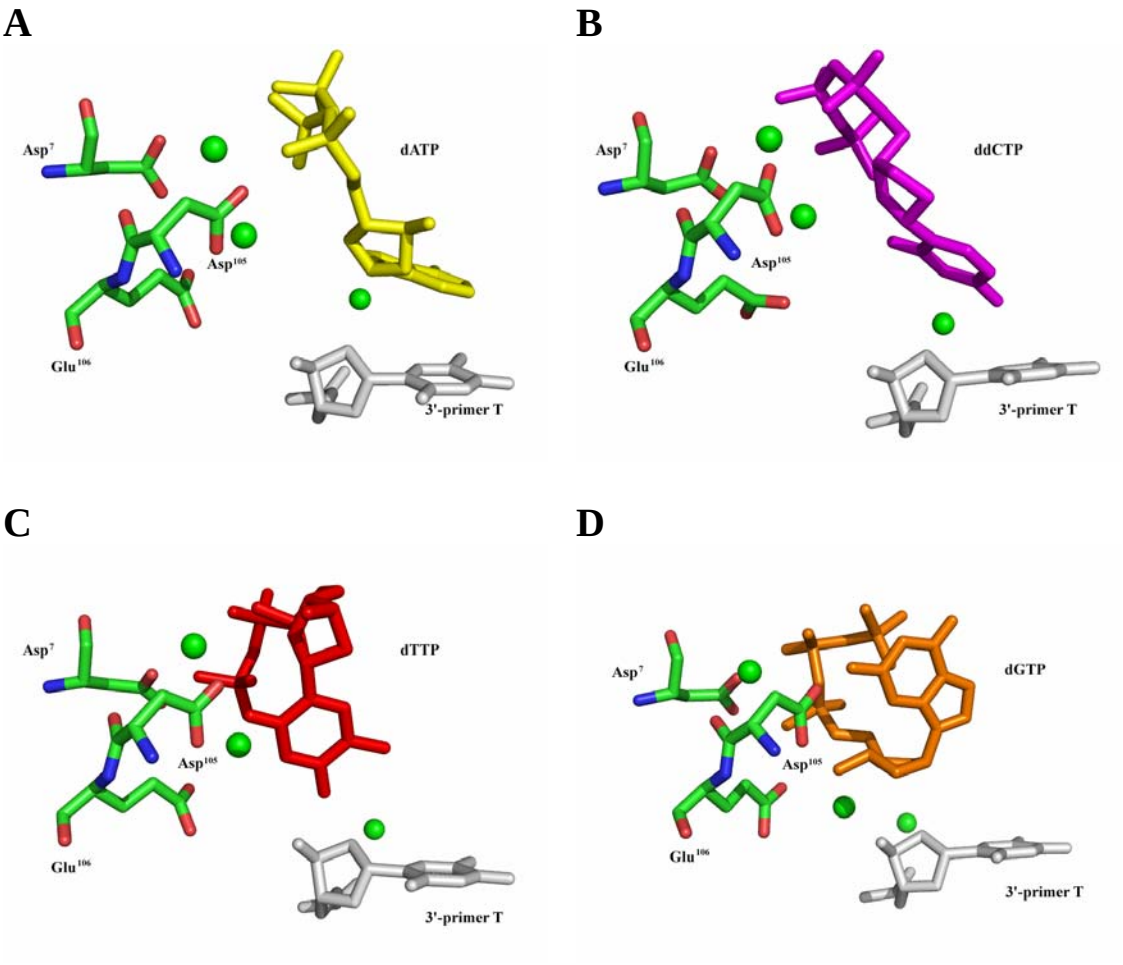
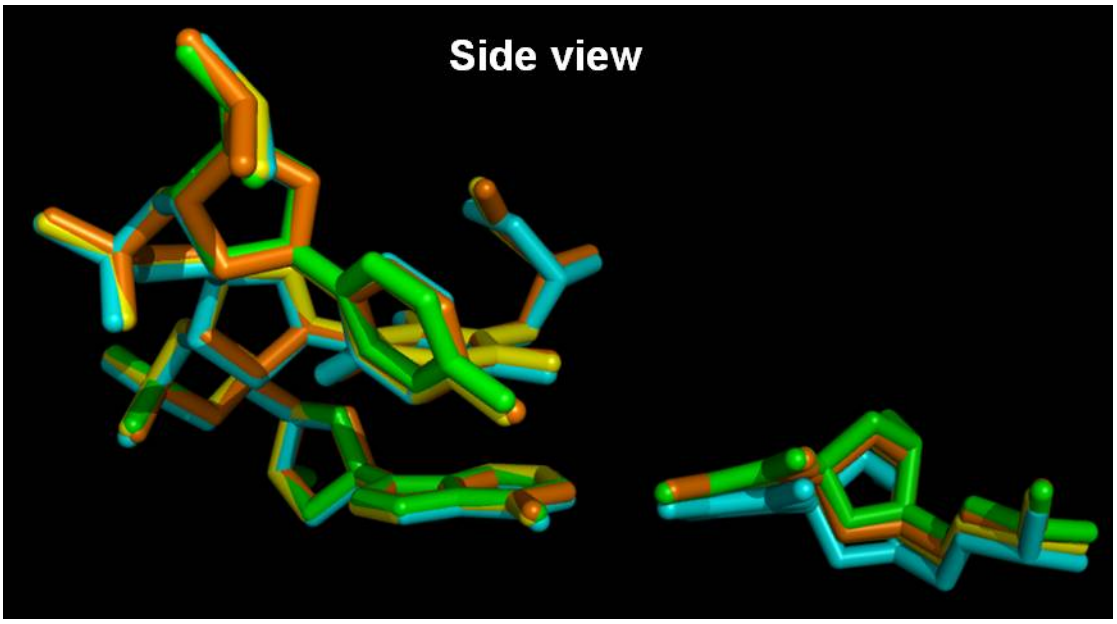


Figure 4-6. Active sites of four ternary Dpo4-DNA-d(d)NTP complexes: (A) SdU/dATP, (B) SdU/ddCTP, (C) SdU/dTTP, and (D) SdU/dGTP, showing the carboxylates Asp^{7}, Asp^{105} and Glu^{106}. The primer-template are shown in stick model in gray, the dATP, ddCTP, dTTP and dGTP are shown in stick model in yellow, purple, red and orange, respectively. Ca^{2+} ions are shown as spheres in green.

A



B

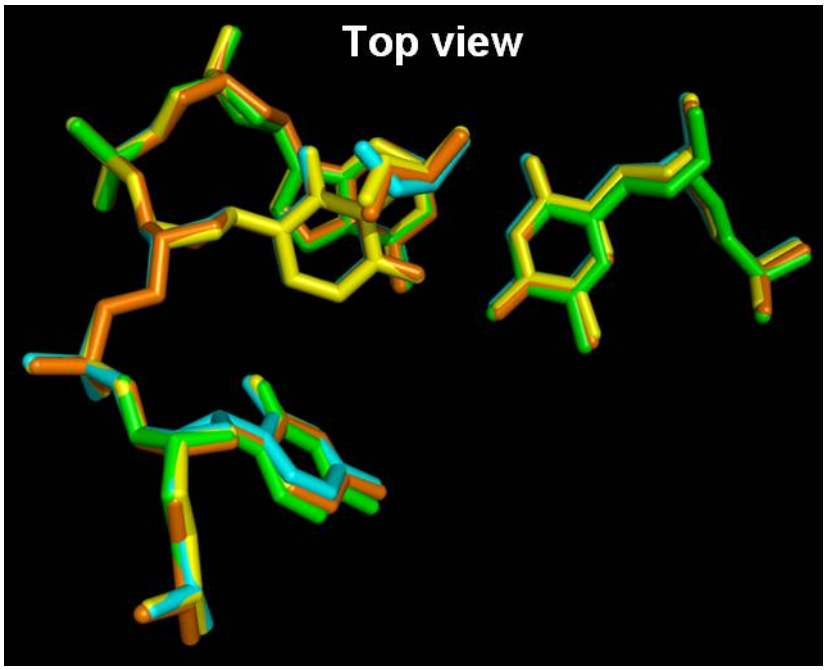


Figure 4-7. Superposition of DNA conformations at the active site of four Dpo4 ternary complex structures of SdU/dATP, SdU/dCTP, SdU/ddTTP, and SdU/dGTP: (A) side view and (B) top view. SdU/dATP, SdU/ddCTP, SdU/dTTP, and SdU/dGTP are shown in yellow, cyan, green and orange, respectively.

Table 4-3: *S*- or *R*-BD-N3-dU/d(d)NTP base pair parameters of all the ternary complexes.

Complexes	Shear (Sx)	Stretch (Sy)	Stagger (Sz)	Buckle (kappa)	Propel (omega)	Opening (sigma)
SdU/dATP	5.9	-0.6	-0.4	11.0	-34.0	22.3
SdU/ddCTP	2.2	8.8	0.1	110.3	-140.9	-38.7
SdU/dTTP	2.1	9.3	2.7	147.8	-166.3	30.3
SdU/dGTP	1.2	-0.2	7.0	-85.8	-77.9	-123.7
RdU/dATP	-4.0	-1.6	2.8	3.6	6.2	-31.0
RdU/dCTP	-2.8	-1.7	4.5	8.6	-5.7	61.0
RdU/ddTTP	-0.9	1.1	-0.2	83.6	137.1	49.6
RdU/dGTP	-2.8	-2.1	2.7	22.7	-23.5	-34.6

Base	Shift (Dx)	Slide (Dy)	Rise (Dz)	Tilt (tau)	Roll (rho)	Twist (omega)
SdU/dATP	0.5	0.4	3.7	9.4	-2.0	55.6
SdU/ddCTP	1.1	3.2	4.5	90.7	-65.5	5.9
SdU/dTTP	0.8	4.4	6.4	104.6	-74.0	45.4
SdU/dGTP	-0.2	-0.3	3.7	1.8	2.0	29.0
RdU/dATP	-1.2	-1.0	4.4	-6.4	19.6	18.3
RdU/dCTP	-0.6	-1.7	4.7	-5.5	13.6	3.6
RdU/ddTTP	0.6	-0.1	2.7	44.6	70.7	58.8
RdU/dGTP	-0.5	-1.5	3.7	-0.6	1.2	18.4

Crystal Structures of Ternary Complexes Containing R-BD-N3-dU Adducted Template

The crystal structure of the ternary Dpo4-DNA-dGTP complex containing *R*-BD-N3-dU adducted template (RdU/dGTP) was determined at a resolution of 2.5 Å (Table 4-2). This was similar to what was observed in Dpo4 ternary complexes containing the *S*-BD-N3-dU modified template, only the *R*-BD-N3-dU nucleotide of the template was accommodated at the active site of Dpo4 (Figures 4-8 and 4-9D), with the BD moiety oriented in the 5' direction. Due to the non-planar butadiene lesion, the 5'-adjacent template C was looped out of the active site and rotated into the major groove. Insertion of the incoming dGTP into the minor groove of the primer-template duplex at the active site of Dpo4 pushed the *R*-BD-N3-dU nucleotide to move toward the major groove, the stacking interaction between the *R*-BD-N3-dU nucleotide and the 3'-adjacent template A was therefore perturbed (Figure 4-10D). The electron density was complete for the whole structure except dGTP (Figure 4-11D), indicating a disordered conformation of the incoming dGTP or the active site was partially occupied. The distance between the terminal 3'-hydroxy group of the primer and the α-phosphate group of dGTP was 4.7 Å, which is out of range for the chemical reaction. The hydrogen-bonding interactions between the N1 and O^2 of the *R*-BD-N3-dU nucleotide and O^6 and N1 of dGTP, respectively, led to a twist of 18.4 º and an opening of -34.6 º between the *R*-BD-N3-dU nucleotide and dGTP (Table 4-3). Three calcium ions were coordinated in this structure, with two Ca^{2+} ions chelated by the three catalytic carboxylates -- Asp^7, Asp^{105} and Glu^{106}, stabilizing the phosphate moiety of the incoming dGTP (Figure 4-12D). However, as reversed conditions of the complex structure with native DNA, two Ca^{2+} ions were closer

to the γ-phosphate than the α-phosphate of dGTP. The third Ca^{2+} ion was close to two oxygens of the α-phosphate group linking the 12th and 13th primer nucleotides.

Crystal structures of all other ternary complexes including RdU/dATP, RdU/dCTP, and RdU/ddTTP at a resolution of 2.5 Å, 2.45 Å and 2.5 Å, respectively, displayed close similarity to the structure of RdU/dGTP (Table 4-2). All of four ternary complexes manifested only variations of conformations of 5′-adjacent template C and the incoming d(d)NTPs occupying at the active site (Figure 4-9). The active sites of the all complexes resembled neither the "type I", nor the "type II" crystal structures of the complex between Dpo4 and native DNA (*144*). In each structure, only the *R*-BD-N3-dU nucleotide of the template was accommodated at the active site of Dpo4, with the BD moiety oriented in the 5' direction, similar to that of the RdU/dGTP structure (Figure 4-9). The non-planar butadiene lesion blocked the stacking of the 5'-adjacent template C. Although all four d(d)NTPs were able to be accommodated at the active site of Dpo4, the conformations of d(d)NTPs varied a lot. They were inserted into the minor groove and perturbed the stacking interaction between the *R*-BD-N3-dU nucleotide and the 3'-adjacent template A (Figure 4-10), which implied that the spacious open site of Dpo4 was able to accommodate all kinds of incoming nucleotides regardless of synthetic activity. The distance between the 3'-hydroxyl group of the terminal primer nucleotide and the α-phosphate of d(d)NTP in each complex was larger than that in the active conformation of either the "type I" or "type II" complex. The incomplete density of the incoming d(d)NTPs observed in all complexes (Figure 4-11) implied that the d(d)NTPs were flexible and disordered or the active site was partially occupied. Moreover, three Ca^{2+} ions at the active site occupied similar sites in all complexes with *R*-BD-N3-dU modified

templates (Figure 4-12) (except for one Ca^{2+} ion missing in RdU/ddTTP complex), which attributed to stabilize the phosphate moiety of the incoming d(d)NTP, despite a lack of catalytic activity. Detailed analysis revealed that the altered orientations of the nucleotide triphosphates went along with different sets of hydrogen-bonding interactions with Dpo4 side chains and the *R*-BD-N3-dU nucleotide in the respective complexes, resulting in the twist and opening between the *R*-BD-N3-dU nucleotide and d(d)NTP (Table 4-3).

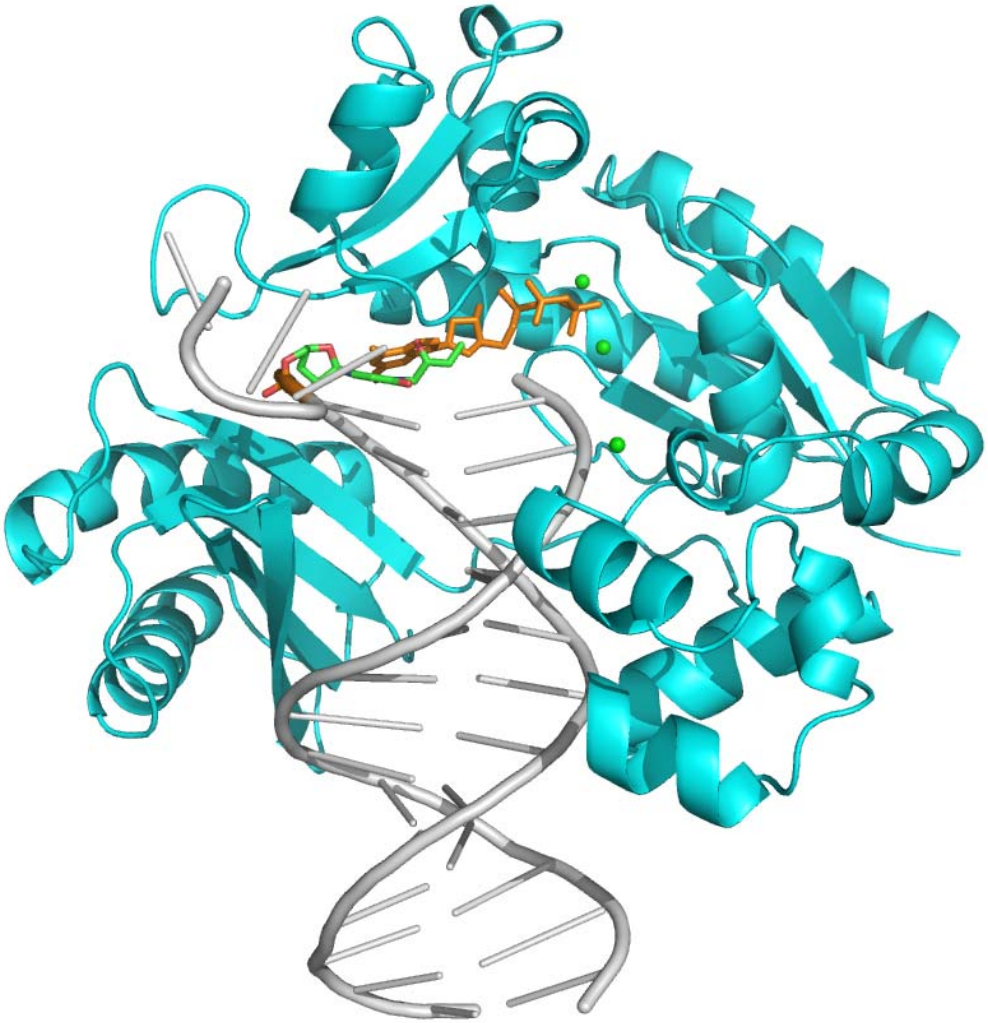


Figure 4-8. Overall structures of the Dpo4 ternary complex RdU/dGTP. The primer-template sequence is shown in Scheme 4-1, with X=*R*-BD-N3-dU and N=G. the Dpo4 backbone and the primer-template are shown in ribbon diagram in cyan and gray, respectively. The *R*-BD-N3-dU nucleotide is shown in stick model colored on atom types; the dGTP is shown in stick model in orange and Ca^{2+} ions are shown as spheres in green.

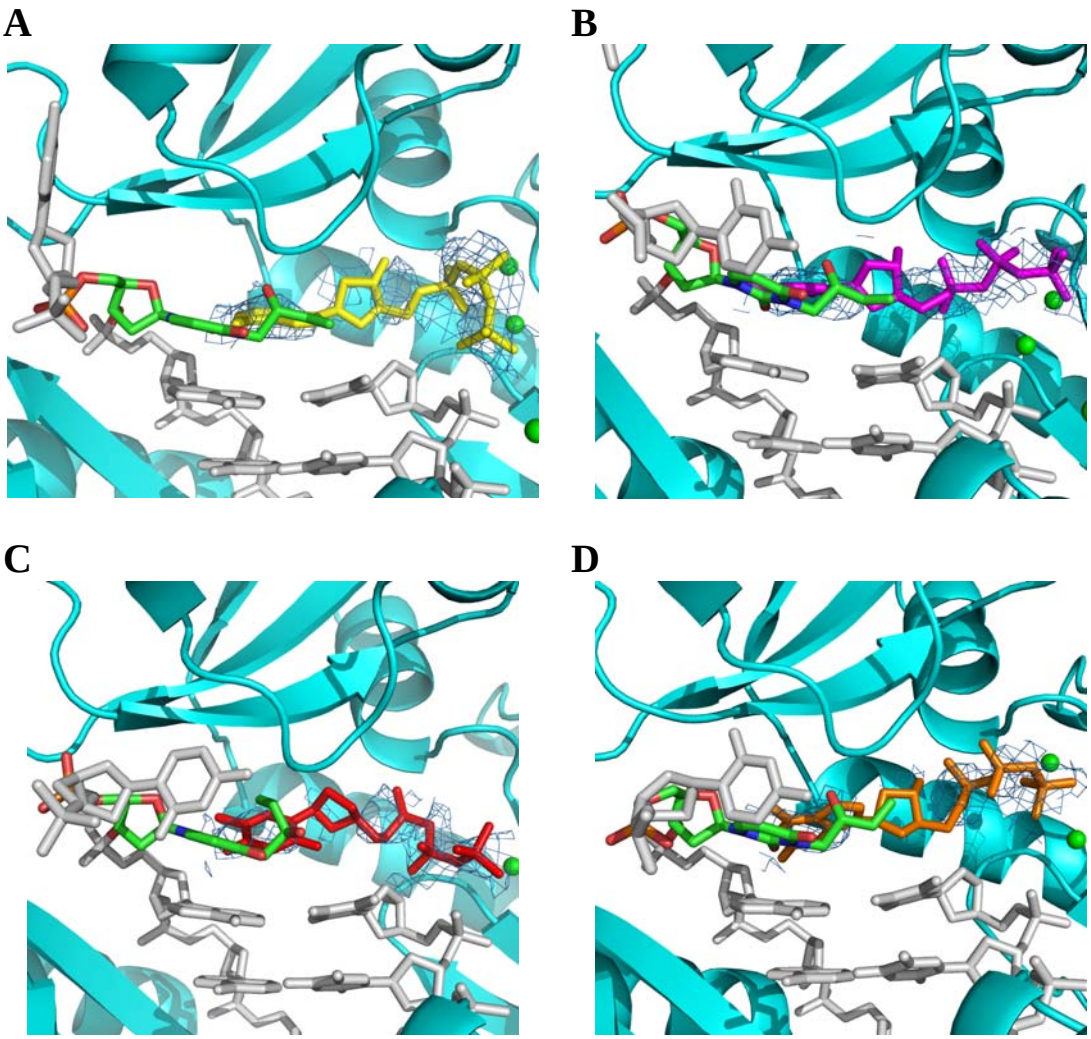


Figure 4-9. Structures of four ternary Dpo4-DNA-d(d)NTP complexes containing *R*-BD-N3-dU adducted template at the active site of Dpo4: (A) RdU/dATP, (B) RdU/dCTP, (C) RdU/ddTTP, and (D) RdU/dGTP. The Dpo4 backbone is shown in ribbon diagram in cyan. The primer-template are shown in stick model in gray, with the *R*-BD-N3-dU nucleotide colored on atom types, and the dATP, dCTP, ddTTP and dGTP are shown in stick model in yellow, purple, red and orange, respectively. Ca^{2+} ions are shown as spheres in green.

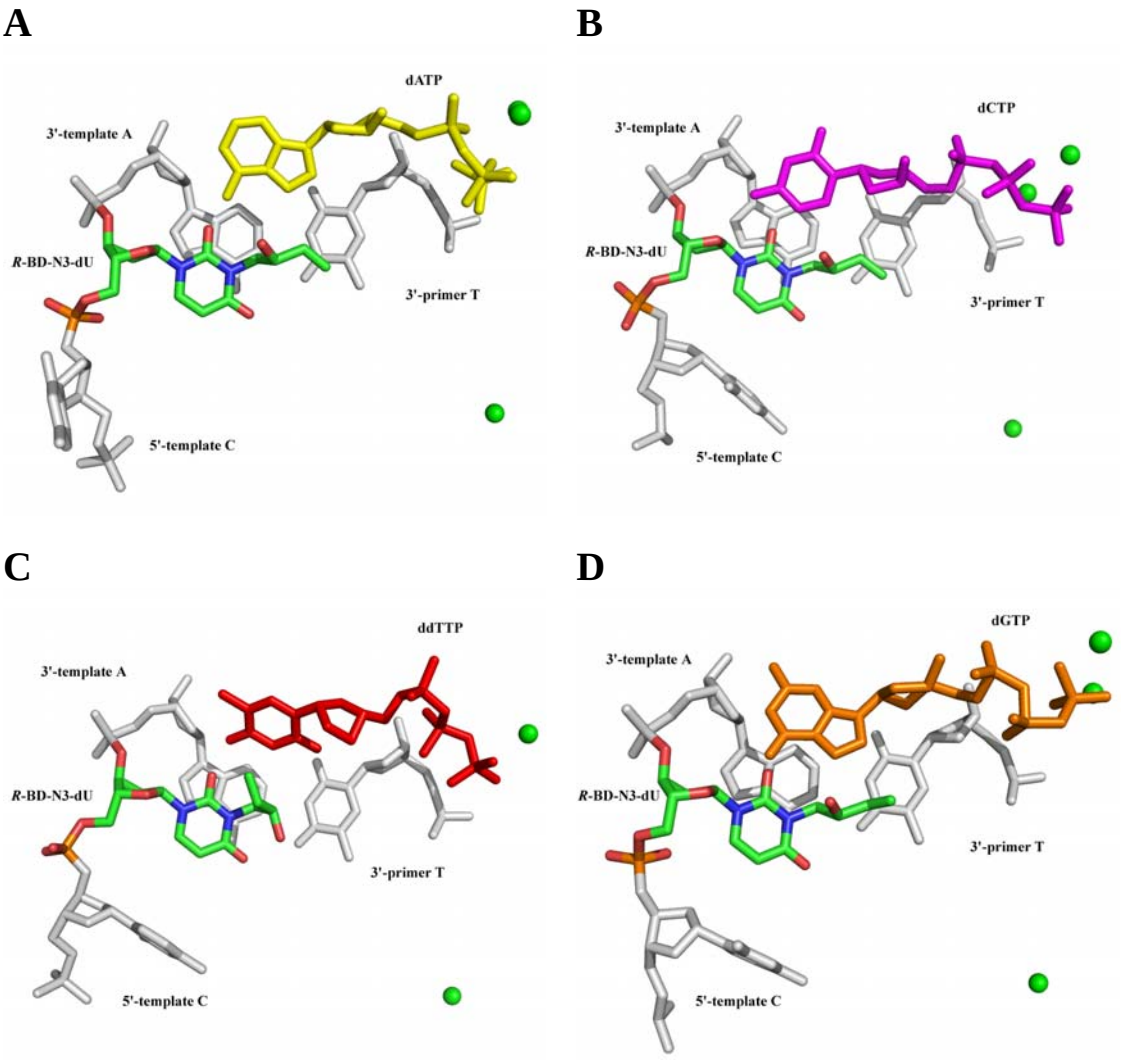


Figure 4-10. Stacking patterns of DNA base pairs at the active site of four ternary Dpo4-DNA-d(d)NTP complexes. (A) RdU/dATP, (B) RdU/dCTP, (C) RdU/ddTTP, and (D) RdU/dGTP. The primer-template are shown in stick model in gray, with the *R*-BD-N3-dU nucleotide colored on atom types, and the dATP, dCTP, ddTTP and dGTP are shown in stick model in yellow, purple, red and orange, respectively. Ca^{2+} ions are shown as spheres in green.

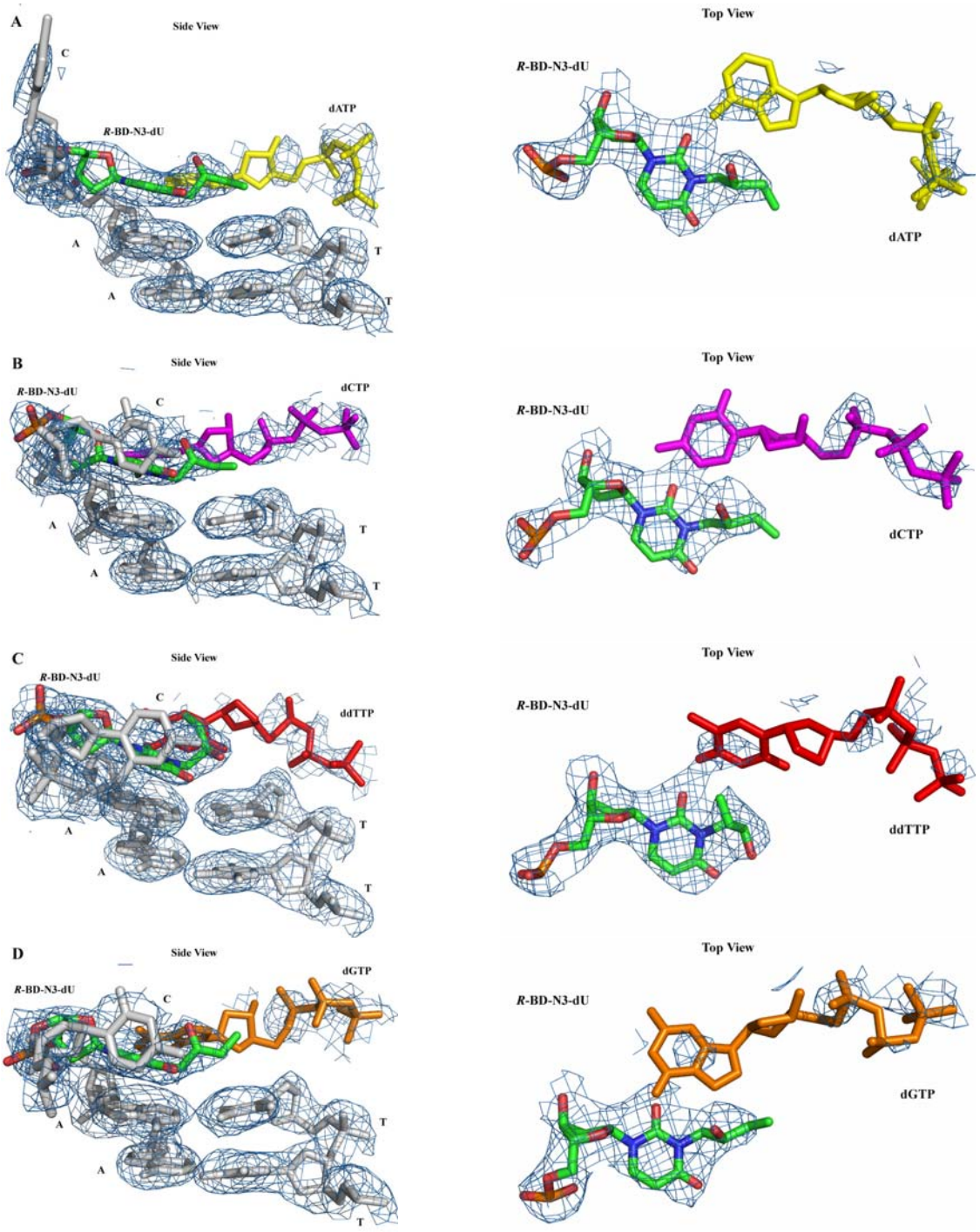


Figure 4-11. Electron density of DNA and d(d)NTPs at the active site of four ternary Dpo4-DNA-d(d)NTP complexes. (A) SdU/dATP, (B) SdU/ddCTP, (C) SdU/dTTP, and (D) SdU/dGTP. Fourier $2F_o$-F_c electron density (blue meshwork) is drawn at the 1σ level. The primer-template are shown in stick model in gray, with the *S*-BD-N3-dU nucleotide colored on atom types, and the dATP, ddCTP, dTTP and dGTP are shown in stick model in yellow, purple, red and orange, respectively.

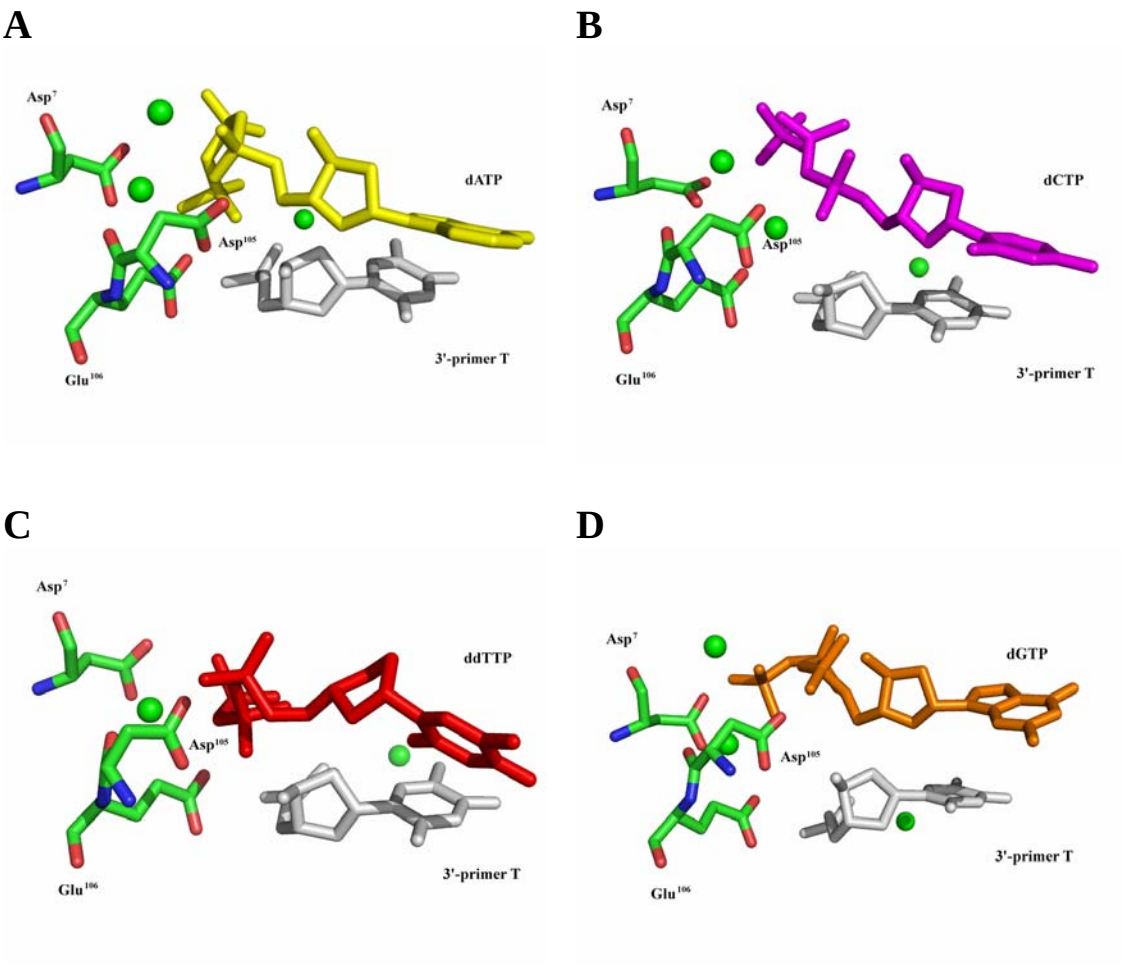


Figure 4-12. Active sites of four ternary Dpo4-DNA-d(d)NTP complexes. (A) RdU/dATP, (B) RdU/dCTP, (C) RdU/ddTTP, and (D) RdU/dGTP, showing the carboxylates Asp^{7}, Asp^{105} and Glu^{106}. The primer-template are shown in stick model in gray, the dATP, dCTP, ddTTP and dGTP are shown in stick model in yellow, purple, red and orange, respectively. Ca^{2+} ions are shown as spheres in green.

Discussion

Site-specific mutagenesis studies have shown that butadiene-derived N3-dU adducts are highly mutagenic in COS-7 cells, with predominant mutations of C → T transitions (50%) and C → A transversions (36%) (*91*). These adducts are also determined to block replication of several replicative DNA polymerases and trans-lesion polymerases significantly(*91*,*94*). Y-family polymerase human pol η is able to incorporate G or A opposite the adducts and extend primers containing an A incorporated opposite the adducts to full-length products (*94*). Dpo4, a homolog of pol η, also inserts dGTP or dATP opposite BD-N3-dU lesion but does not extend the primer to a full-length product (Figure 4-1). Thus, the structural study of the ternary complexes containing Dpo4, BD-N3-dU adducted primer-template DNA and incoming d(d)NTPs is of interest.

The "type II" crystal structures were solved in all previously reported ternary Dpo4−DNA−d(d)NTP complexes with the primer-template site-specifically modified by planar carcinogen such as 1,N^2-ε-guanine (*186*), (8-oxoG) (*149*,*187*,*188*), and PdG (*197*), as observed in Dpo4 complexes with native DNA (*144*). The correct incoming nucleotide skips the damaged lesion and is inserted opposite the 5'-adjacent template nucleotide accompanied with formation of Watson-Crick hydrogen bonds. Thus, dNTP insertion is dependent upon the identity of the 5′-adjacent template base. These crystal structures lead to the probable mechanism of the formation of a "type II" ternary complex in which Dpo4 bypasses a planar adducted nucleotide by accommodating both the damaged and the 5′-adjacent template nucleotide within its active site, and incorporates the correct incoming nucleotide utilizing the 5′-adjacent nucleotide as a template base.

However, in the ternary complexes examined here, only the non-planar BD-N3-dU nucleotides in the template reside in the active site of Dpo4, instead of two adjacent template bases admitted into the active site simultaneously in the Dpo4 complexes as reported previously. The covalent bonding of the BD moiety at the N3 position of dU destroys classic Watson-Crick base-pairing interactions with the incoming d(d)NTP, which disturbs the formation of a "type I" structure. Moreover, the steric hindrance of the non-planar BD moiety prevents the stacking of a 5'-adjacent template base at the active site, and therefore perturbs the base pairing interactions between the 5'-template C and the incoming d(d)NTP. Thus, the displaced 5'-template C, stabilized by hydrogen-bonding interactions with Dpo4, adopts a conformation that prohibits incorporation of the next incoming nucleotide, and formation of a "type II" structure is therefore impeded.

The electron density of tri-phosphates of the incoming d(d)NTPs observed at the active site of Dpo4 reveals that the incoming nucleotides are accommodated at the active site. However, the weak and incomplete electron density of sugars and bases of the incoming d(d)NTPs indicates that the d(d)NTPs are disordered. The observation that the incoming nucleotides are disordered suggests that multiple structures exist in the complexes. The crystal structures solved here represent the most stable conformations calculated from the structural refinements. The long distance observed between the 3′-hydroxyl group of the terminal primer nucleotide and the α-phosphate of the d(d)NTP in each complex indicates that the crystal structures unlikely represent catalytically active conformations. However, the disordered incoming nucleotides are possible to reach catalytically active conformation which may not be caught under the crystallization conditions.

Replication bypass studies (Figure 4-1) indicate that the Dpo4 polymerase could insert dGTP or dATP opposite BD-N3-dU nucleotides in the damaged primer-template duplex. This incorporation of dGTP or dATP opposite both the *S*- and *R*- BD-N3-dU nucleotides may correspond to the flexibility and disorder of the incoming dGTP and dATP at the active site of Dpo4, thus movements of dNTP makes it possible to form a catalytically active conformation. Site-specific mutagenesis studies performed in COS-7 cells revealed the major C→T and C→A mutations (*91*), which suggest one or a combination of mammalian polymerases in COS-7 cells are responsible for the incorporation of dATP or dTTP opposite the BD-N3-dU damaged base. The incorportation of dATP opposite the BD-N3-dU nucleotide revealed in replication bypass experiments may provide some insight into the C→T transitions in the mutagenesis study.

The conformation of the BD moiety in the template and its effect on the incoming nucleotides is remarkable. It appears to prevent either "type I" or "type II" complex structures containing either native or planar chemical adducted templates (*91,94*). This reveals the importance of steric effects for trans-lesion synthesis by Dpo4. The steric hindrance of the BD moiety disrupts the stacking of the 5'-adjacent C in the template and therefore prevents the base-paring interactions between the incoming nucleotides and the template bases. This indicates that the steric effect, other than atom electronic charge, plays a significant role in lesion bypass of damaged DNA by trans-lesion polymerase Dpo4. Thus, the structural studies of the BD-N3-dU adducts will probably be able to provide a model of the progress of related mammalian Y-family polymerases which bypass non-planar carcinogen adducted DNA.

Summary

The crystal structures of ternary Dpo4-DNA-d(d)NTP complexes, containing either an *S*- or *R*- BD-N3-dU adducted template indiacate that most stable intermediates formed under the crystallization conditions are non-productive. However, the disordered incoming nucleotides suggest that the multiple intermediates might exist in each ternary complex, so that a catalytically active conformation of the complex containing dGTP or dATP is probable to be formed. In the crystal structures, the non-planar BD lesion distorts the stacking of the 5'-adjacent template base, and therefore perturbs base-pairing hydrogen-bonding interactions between the incoming d(d)NTP and either the adducted base or the 5'-adjacent base in the template. The results indicate that the steric effect of adducts plays a significant role in the interactions between enzyme and DNA adducts. The application of these approaches to other types of DNA adducts, especially non-planar adducts, correlated with biochemical approaches, may provide insight into understanding the progress of the bypass and trans-lesion synthesis of non-planar chemical damaged DNA by Y-family polymerases.

CHAPTER V

EXTENSION OF PRIMER IN DPO4-DNA COMPLEX WHEN THE STEREOISOMERIC N3-(2*S* OR 2*R*-HYDROXY-3-BUTEN-2-YL)- 2'-DEOXYURIDINE ADDUCTS OPPOSITE PRIMER ADENINE BY THE *SULFOLOBUS SOLFATARICUS* DNA POLYMERASE DPO4

Introduction

Replication bypass and primer extension analysis demonstrated that Dpo4 was able to incorporate dGTP and extend the primer to a full-length product by utilizing a primer-template where the 3'-terminal primer placed A opposite the BD-N3-dU nucleotide in the template. X-ray crystal structures of both binary Dpo4-DNA complexes and ternary Dpo4-DNA-d(d)GTP complexes containing either an *S*- or *R*- BD-N3-dU adducted template in the primer-template sequence 5'-d(GGGGGAAGGATTTA)-3'·5'-d(TCACXAAATCCTTCCCCC)-3', where X is *S*- or *R*- BD-N3-dU were obtained. The crystal structures were determined at a resolution between 1.95 and 2.70 Å (Scheme 5-1). In all the crystal structures, only the non-planar BD-N3-dU nucleotide in the template resided at the active site and the 5'-adjacent template C flipped out to the major groove. Watson-Crick base-pairing interactions between the BD-N3-dU nucleotide and the 3'-terminal A in the primer were perturbed, and the incoming nucleotide in the ternary complexes were highly disordered.

A

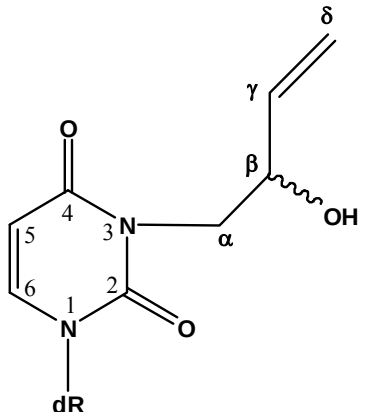
Primer 5' GGG GGA AGG ATT TA 3'
Template 3' CCC CCT TCC TAA A**X**C ACT 5' + Dpo4

B

Primer 5' GGG GGA AGG ATT TA 3'
Template 3' CCC CCT TCC TAA A**X**C ACT 5' + Dpo4 + d(d)GTP

X = *S*- or *R*-BD-N3-dU

C



Scheme 5-1. (A) The binary primer-template sequence, (B) the ternary primer-template sequence and (C) chemical structure of the N3-(2*R* or 2*S*-hydroxy-3-buten-2-yl)- 2'-deoxyuridine adducts and nomenclature.

Results

Replication and Extension of Primer by Dpo4

The replication bypass experiments were done by Dr. Surajit Banerjee in the laboratory of Professor Michael P. Stone. In the control experiment with the A·C mismatched primer-template sequence, Dpo4 inserted dGTP opposite the template C in 8 min, and extended the primer-template to full extension products successfully in 15 min with the existence of all four dNTPs in the reaction mixture. In the replication bypass of the template containing either the *S*- or *R*- BD-N3-dU adduct opposite A in the primer, Dpo4 inserted dGTP opposite the adducted lesion within 15 min and extended the primer strand to the full-length product in 20 min. The results of the replication and extension experiments are shown in Figure 5-1.

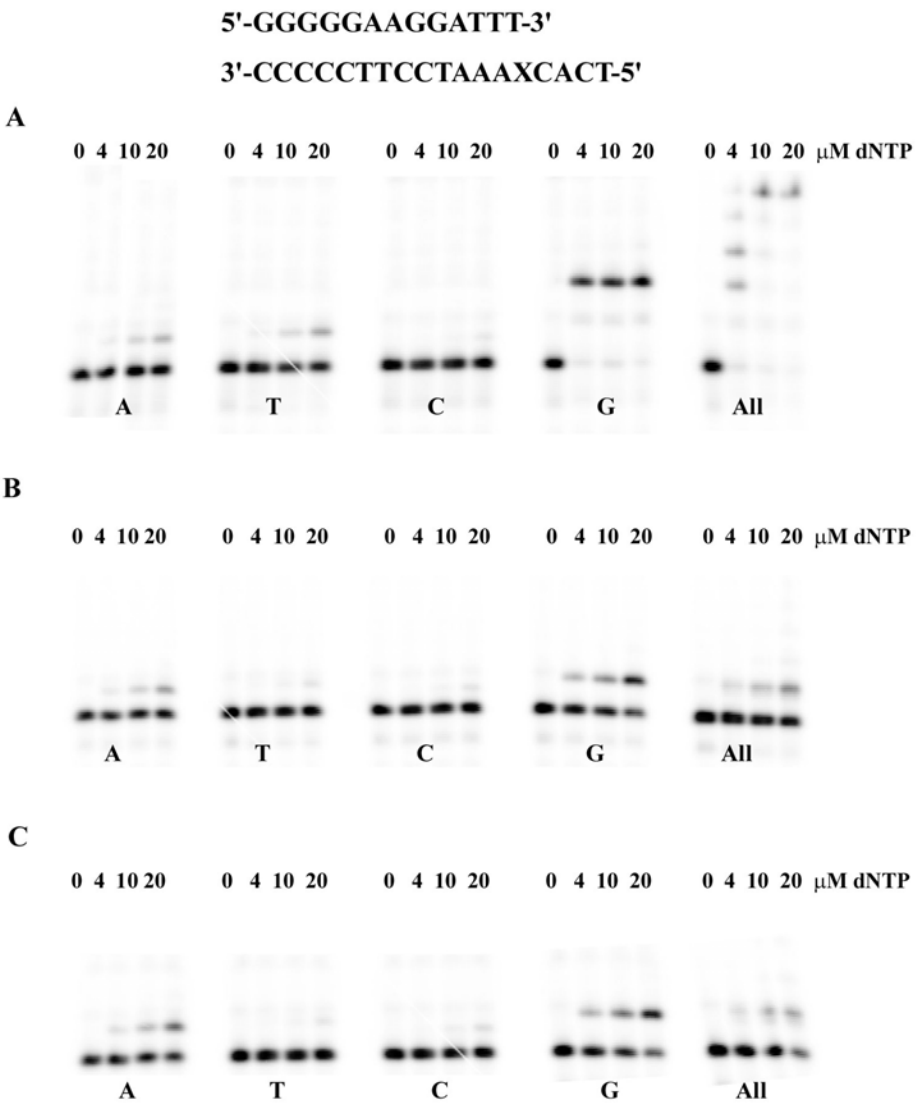


Figure 5-1. Replication bypass of (A) control, (B) *S*-BD-N3-dU adducted, (C) *R*-BD-N3-dU adducted primer-template complex with Dpo4. The reactions were performed with dNTP at the concentration of 0, 4, 10 and 20 μM. The first four panels in each assay represent insertion of a single dNTP to either the control or the adducted template. The last panel of each assay represents the inclusion of all four dNTPs in the reaction mixture and their adduction to either the control or the adducted template. The assays were carried out at 37 ºC, using 100 nM Dpo4 by Dr. Surajit Banerjee.

X-ray Diffraction Data Processing

Crystallization trials were conducted for binary Dpo4-DNA complexes and ternary Dpo4-DNA-d(d)GTP complexes containing either the *S*- or *R*-BD-N3-dU adducted template. Crystals of four complexes were obtained, referred to here as SdU:A for the binary complex of the *S*-adduct, SdU:A/dG for the ternary complex of the *S*-adduct with dGTP, RdU:A for the binary complex of the *R*-adduct, and RdU:A/ddG for the ternary complex of the *R*-adduct with ddGTP. Crystals of both the binary and ternary complexes were diffracted within a range of 1.95 - 2.70 Å. The statistics of data processing and refinement parameters are summarized in Table 5-1. Autoindexing and systematic absences indicated space group $P2_12_12$ for all four complexes, similar to the ternary complexes described in Chapter IV. The resolution was 1.95 Å with a completeness of 95.6% for SdU:A, 2.70 Å with a completeness of 92.9% for SdU:A/dG, 2.20 Å with a completeness of 93.2% for RdU:A, and 2.50 Å with a completeness of 94.4% for RdU:A/ddG. The resulting data sets for the ternary complexes were of good quality as revealed by R_{merge} values of 6.5% (SdU:A), 8.7% (SdU:A/dG), 7.9% (RdU:A), and 6.8% (RdU:A/ddG), respectively.

Table 5-1. Crystal data and refinement parameters for binary and ternary complexes.

Parameter	SdU:A	SdU:A/dG	RdU:A	RdU:A/dG
X-ray source	ID-22(APS)	ID-22(APS)	ID-22 (APS)	ID-22(APS)
Wavelength (Å)	1.00	1.00	1.00	1.00
Temperature (K)	110	110	110	110
No. of crystals	1	1	1	1
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Unit cell (a,b,c)(Å)	92.97,100.62,53.07	94.61,101.99,53.54	92.94,100.74,53.23	94.30,101.39,53.42
Resolution range (Å)	46.07-1.95	46.60-2.70	46.19-2.20	44.65-2.50
Highest resolution shell	2.02-1.95	2.80-2.70	2.28-2.20	2.59-2.50
No. of measurements	1089433	1227308	1193464	1269065
No. of unique reflections	35297	13837	24255	31700
Redundancy	6.5	6.3	5.3	6.4
Completeness (%)	95.6(65.7)[a]	92.9(64.5)	93.2(56.8)	94.4(67.0)
R_{merge} [b]	0.065	0.087	0.079	0.068
Signal/noise ($<I/\sigma^I>$)	9.5	6.8	5.9	7.0
Solvent content (%)	55.56	54.61	51.13	53.99
Model composition				
No. of amino acid nucleotides	342	342	342	342
No. of water molecules	39		51	
No. of Ca^{2+} ions	3	3	3	3
No. of template nucleotides	16	16	16	16
No. of primer nucleotides	14	14	14	14
No. of dGTP		1		
No. of ddGTP				1
R^c (%)	25.2	21.8	23.1	23.4
R_{free} [d] (%)	29.0	27.6	26.2	27.8
Estimated coordinate error (Å)				
From Luzatti plot	0.33	0.41	0.35	0.42
From Luzatti plot (c-v)[e]	0.40	0.53	0.40	0.53
From σA plot	0.32	0.63	0.42	0.56
From σA plot (c-v)[e]	0.39	0.84	0.51	0.58
r.m.s. deviation in temp. factors				
Bonded main chain atoms ($Å^2$)	1.52	1.47	1.38	1.37
Bonded side chain atoms ($Å^2$)	2.16	1.99	2.10	1.90
r.m.s. S.D. from ideal values				
Bond lengths (Å)	0.006	0.007	0.006	0.007
Bond angles (º)	1.11	1.27	1.09	1.27
Dihedral angles (º)	20.95	21.60	21.12	21.3
Improper angles (º)	0.94	1.96	0.97	1.13

[a] Values in parentheses correspond to the highest resolution shells. [a] $R_{merge}=\sum_{hkl}\sum_{j=1,N}|<I_{hkl}>-I_{hkl,j}|/\sum_{hkl}\sum_{j=1,N}|I_{hkl,j}|$, where the outer sum (*hkl*) is taken over the unique reflections. [c] $R_f=\sum_{hkl}\sum||F_{o,hkl}|-k|F_{c,hkl}||/\sum_{hkl}|F_{o,hkl}|$, where $|F_{o,hkl}|$ and $|F_{c,hkl}|$ are the observed and calculated structure factor amplitudes, respectively. [d] R_{free} *idem*, for the set of reflections (5% of the total) omitted from the refinement process.

Crystal Structure of Binary Complex SdU:A

The crystal structure of the binary Dpo4-DNA complex containing *S*-BD-N3-dU adducted template opposite A in the primer (SdU:A) was determined at a resolution of 1.95 Å (Table 5-1). Similar to the ternary complexes described in Chpater IV, only the *S*-BD-N3-dU nucleotide of the template was lodged at the active site of Dpo4 (Figures 5-2 and 5-3A). The *S*-BD-N3-dU nucleotide stacked inside the duplex with the BD moiety oriented in 5' direction (Figure 5-4A). Due to the non-planar butadiene lesion, the 5'-adjacent template C flipped out to major groove and was directed away from the active site. Three calcium ions were coordinated in this structure with two Ca^{2+} ions chelated by the three catalytic carboxylates - Asp^{7}, Asp^{105} and Glu^{106}, and the third Ca^{2+} ion close to two oxygen of the α-phosphate group linking the 12th and 13th primer nucleotides. The electron density was complete for the whole structure (Figure 5-5A). The perturbation caused by the *S*-BD-N3-dU adduct resulted in a stacking distance of 4.37 Å between the 3'-terminal A and its 5'-adjacent T in the primer, the tilt of 10.9 º and twist of 60.6 º between the terminal N3-dU:A and its 5'-directed A:T base pairs, and an opening of 18.6º between the *S*-BD-N3-dU nucleotide and the opposite A in the primer.

Crystal Structure of Ternary Complex SdU:A/dG

The crystal structure of the ternary Dpo4-DNA complex containing *S*-BD-N3-dU adducted template opposite A in the primer with the incoming dGTP (SdU:A/dG) was determined at a resolution of 2.70 Å (Table 5-1). The entire crystal structure of RdU:A/ddG is very similar to RdU:A. Only the *S*-BD-N3-dU nucleotide of the template was accommodated at the active site of Dpo4 (Figure 5-3B). The *S*-BD-N3-dU nucleotide

was stacked inside the duplex with the BD moiety oriented in 5' direction (Figure 5-4B). Due to the non-planar butadiene lesion, the 5'-adjacent template C was looped out and directed away from the active site. Therefore it could not participate in Watson-Crick base pairing with incoming dGTP. Three calcium ions were coordinated in this structure with two Ca^{2+} ions chelated by the three catalytic carboxylates - Asp^{7}, Asp^{105} and Glu^{106}, and the third Ca^{2+} ion close to two oxygen of the α-phosphate group linking the 12th and 13th primer nucleotides. The electron density was complete for the entire structure except the 3'-terminal primer A and dGTP, especially the base and sugar ring (Figure 5-5B), indicating that the 3'-terminal primer A and dGTP were highly disordered. However, the positions of the tri-phosphates could be determined base on the complete density observed. The distance between the α-phosphate of dGTP and the 3'-OH of the terminal A in the primer was 5.6 Å. The superimposed structures of SdU:A and SdU:A/dG showed the accommodation of dGTP at the active site did not disturb the whole structure significantly (Figure 5-6). The perturbation caused by the *S*-BD-N3-dU adduct resulted in the stacking distance of 6.1 Å between the 3'-terminal A and its 5'-adjacent T in the primer, the tilt of 43.0 º and twist of 37.7 º between the terminal N3-dU:A and its 5'-directed A:T base pairs.

Crystal Structure of Binary Complex RdU:A

The crystal structure of the binary Dpo4-DNA complex containing the *R*-BD-N3-dU adducted template opposite A in the primer (RdU:A) was determined at a resolution of 2.20 Å (Table 5-1). As shown in the structure, only the *S*-BD-N3-dU nucleotide of the template was lodged at the active site of Dpo4 (Figure 5-3C) and stacked inside the

duplex with the BD moiety oriented in 5' direction (Figure 5-4C). Due to the non-planar butadiene lesion, the 5'-adjacent template C flipped out to major groove and was directed away from the active site. Three calcium ions were coordinated in this structure with two Ca^{2+} ions chelated by the three catalytic carboxylates - Asp^{7}, Asp^{105} and Glu^{106}, and the third Ca^{2+} ion close to two oxygen of the α-phosphate group linking the 12th and 13th primer nucleotides. The electron density was complete for the entire structure (Figure 5-5C). The perturbation caused by the *S*-BD-N3-dU adduct resulted in a stacking distance of 4.3 Å between the 3'-terminal A and its 5'-adjacent T in the primer, the tilt of 12.6 º and twist of 62.9 º between the terminal N3-dU:A and its 5'-directed A:T base pairs, and an opening of 20.6 º between the *S*-BD-N3-dU nucleotide and the opposite A in the primer.

Crystal Structure of Ternary Complex RdU:A/ddG

The crystal structure of the ternary Dpo4-DNA complex containing *R*-BD-N3-dU adducted template opposite A in the primer with the incoming ddGTP (RdU:A/ddG) was determined at a resolution of 2.50 Å (Table 5-1). The entire crystal structure of RdU:A/dG is similar to RdU:A. Only the *R*-BD-N3-dU nucleotide of the template was accommodated at the active site of Dpo4 (Figure 5-3D). The *R*-BD-N3-dU nucleotide was stacked inside the duplex with the BD moiety oriented in 5' direction (Figure 5-4D). Due to the non-planar butadiene lesion, the 5'-adjacent template C was looped out and directed away from the active site, therefore no base pair interactions were observed between the template C and the incoming ddGTP accommodated at the active site. Three calcium ions were coordinated in this structure with two Ca^{2+} ions chelated by the three

catalytic carboxylates - Asp^{7}, Asp^{105} and Glu^{106}, and the third Ca^{2+} ion close to two oxygen of the α-phosphate group linking the 12th and 13th primer nucleotides. The electron density was complete for the entire structure except the 3'-terminal primer A and ddGTP (Figure 5-5B), indicating that the 3'-terminal primer A and dGTP were highly disordered. However, the positions of the tri-phosphates could be determined based on the complete density observed. The distance between the α-phosphate of dGTP and the 3'-OH of the terminal A in the primer was 5.6 Å. The superimposed structures of SdU:A and SdU:A/dG showed the accommodation of dGTP at the active site did not disturb the whole structure significantly (Figure 5-6). The perturbation caused by the *S*-BD-N3-dU adduct resulted in a stacking distance of 6.5 Å between the 3'-terminal A and its 5'-adjacent T in the primer, the tilt of 47.5 º and twist of 48.6 º between the terminal N3-dU:A and its 5'-directed A:T base pairs.

Figure 5-2. Overall structures of the Dpo4 binary complex SdU:A. The primer-template sequence is shown in Scheme 5-1, with X=*S*-BD-N3-dU. The Dpo4 backbone and the primer-template are shown in ribbon diagram in cyan and gray, respectively. The *S*-BD-N3-dU nucleotide is shown in stick model colored on atom types; the 3'-terminal A in the primer is shown in stick model in yellow and Ca^{2+} ions are shown as spheres in green.

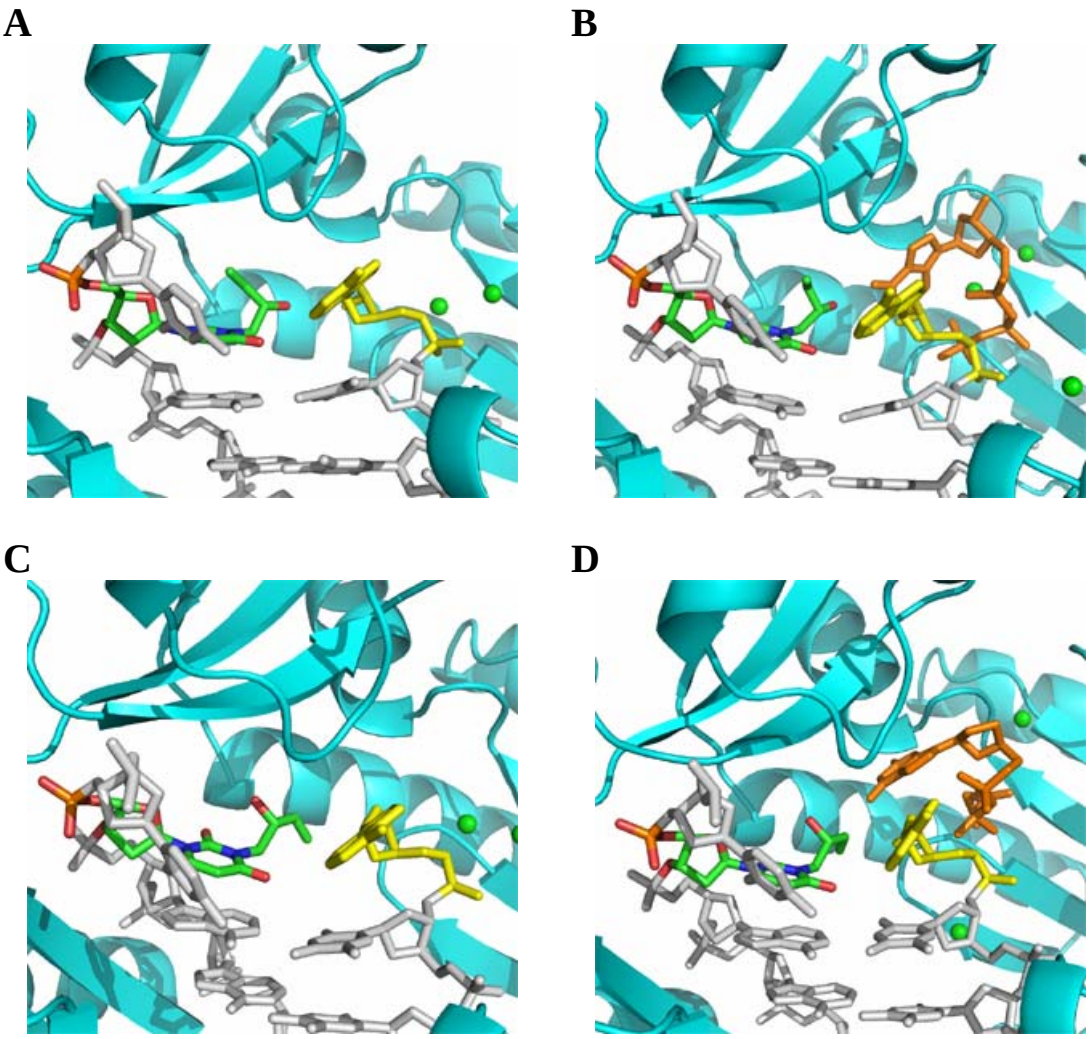


Figure 5-3. Structures of binary Dpo4-DNA and ternary Dpo4-DNA-d(d)GTP complexes containing *S*-or *R*- BD-N3-dU adducted template at the active site of Dpo4. (A) SdU:A, (B) SdU:A/dGTP, (C) RdU:A, and (D) RdU:A/ddGTP. The Dpo4 backbone is shown in ribbon diagram in cyan. The primer-template are shown in stick model in gray, with the BD-N3-dU nucleotide colored on atom types, and the 3'-terminal A in the primer and d(d)GTP are shown in stick model in yellow and orange, respectively. Ca^{2+} ions are shown as spheres in green.

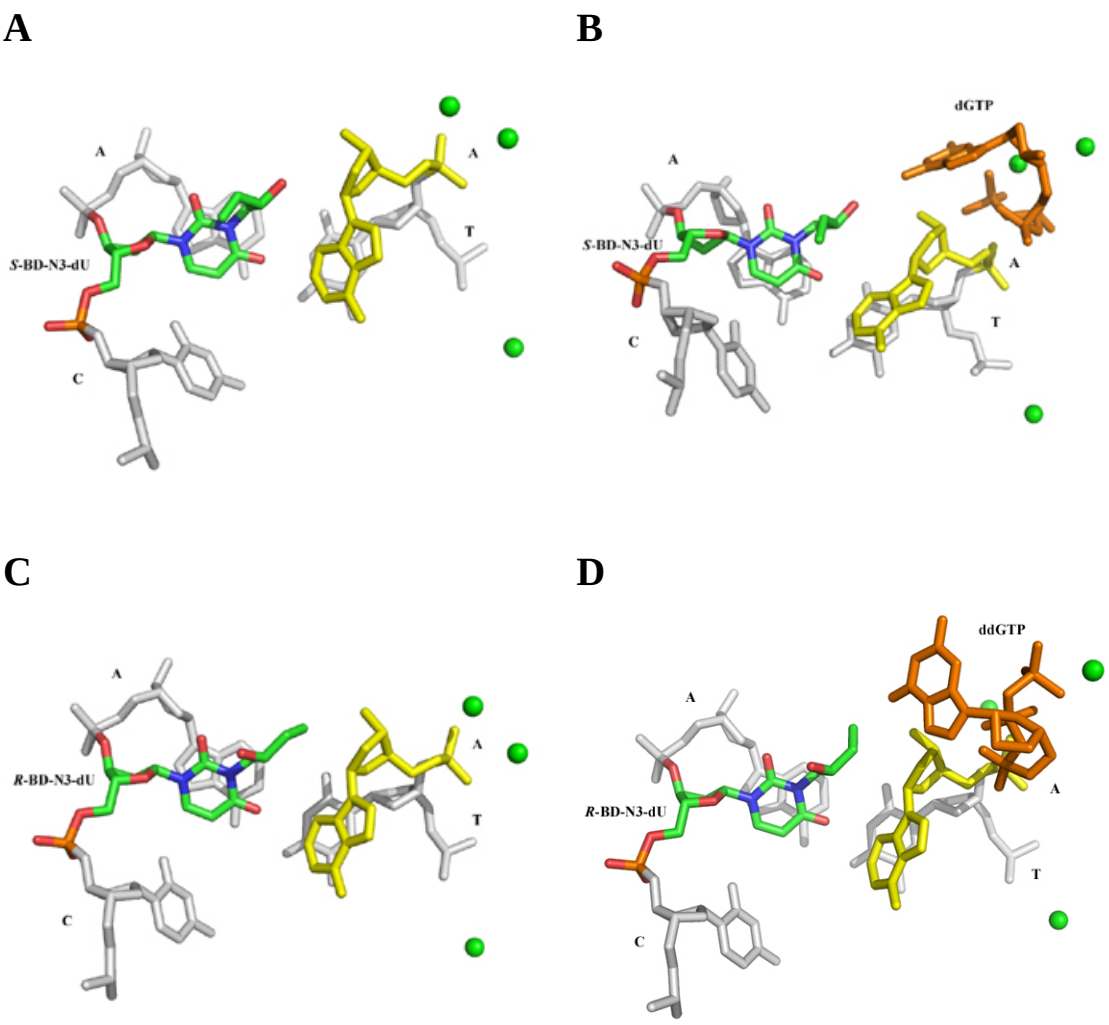


Figure 5-4. Stacking patterns of DNA base pairs at the active site of (A) SdU:A, (B) SdU:A/dGTP, (C) RdU:A, and (D) RdU:A/ddGTP. The primer-template are shown in stick model in gray, with the BD-N3-dU nucleotide colored on atom types, and the 3'-terminal A in the primer and d(d)GTP are shown in stick model in yellow and orange, respectively. Ca^{2+} ions are shown as spheres in green.

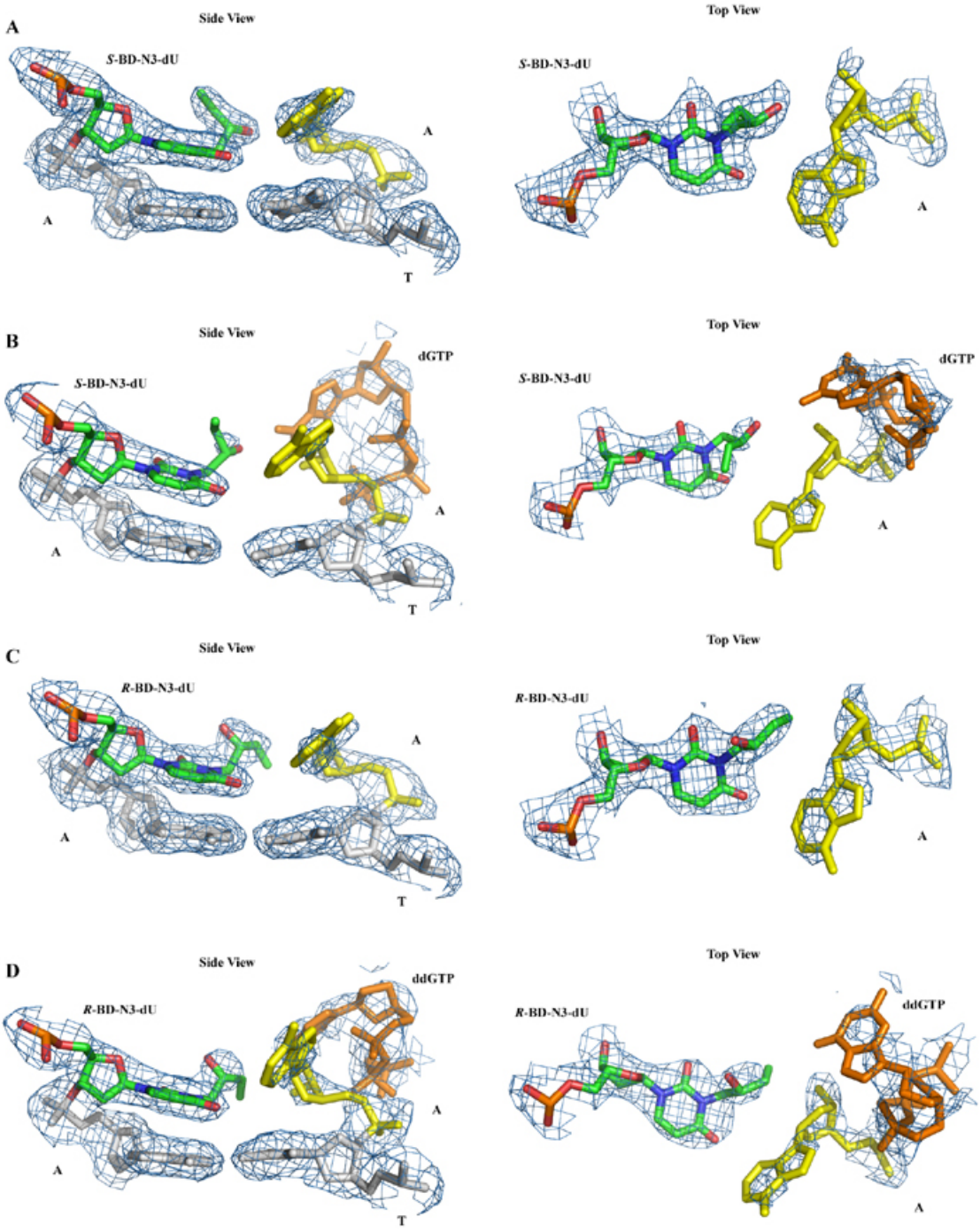


Figure 5-5. Electron density of DNA and d(d)GTP at the active site of (A) SdU:A, (B) SdU:A/dGTP, (C) RdU:A, and (D) RdU:A/ddGTP. Fourier $2F_o$-F_c electron density (blue meshwork) is drawn at the 1σ level. The primer-template are shown in stick model in gray, with the BD-N3-dU nucleotide colored on atom types, and the 3'-terminal A in the primer and d(d)GTP are shown in stick model in yellow and orange, respectively.

Figure 5-6. Superposition of overall conformations of SdU:A and SdU:A/dGTP in cyan and yellow respectively.

Figure 5-7. Superposition of overall conformations of RdU:A and RdU:A/ddGTP in cyan and yellow respectively.

Discussion

Butadiene derived N3-dU adducts cause more than 50% C → T transitions, which is the most mutations detected in COS-7 cells (*91*), indicating that dATP is preferred to insert opposite BD-N3-dU adducts. Y-family polymerases human pol η and pol ζ are able to extend the primer containing an A incorporated opposite the adducts to full-length products (*94*). Y-family polymerase Dpo4, a homolog of pol η and pol ζ, also inserts dGTP opposite BD-N3-dU lesion and extends the primer to a full-length product (Figure 5-1). Thus, the structural study of Dpo4 complexes containing the BD-N3-dU adducts opposite A in the primer, and with or without incoming d(d)GTP is of interest.

In both binary and ternary complexes investigated here, BD-N3-dU adducts in the template reside at the active site of Dpo4 and push the 5'-adjacent template nucleotide out of the active site. The covalent bonding of the BD moiety at the N3 position of dU destroys classic Watson-Crick base-pairing interactions with the 3'-terminal primer A, and leads to much higher tilt and twist angles in adducts neighboring base pairs. The stacking was perturbed between the 3'-terminal base and its 5'-adjacent base in the primer, suggesting that the 3'-terminal primer base was disordered by BD-N3-dU adducts. Moreover, the steric hindrance of the non-planar BD moiety prevents the stacking of the 5'-adjacent template base at the active site, and therefore perturbs Watson-Crick hydrogen-bonding interactions between the 5'-template base and the correct incoming nucleotide. Thus, the formation of a "type II" ternary complex where Dpo4 bypasses planar adducted nucleotides by accommodating both damaged and the 5′-adjacent template nucleotide within its active site, incorporating the correct incoming nucleotide opposite 5′-adjacent template nucleotide, is blocked here.

The crystal structures of the ternary complexes suggest that the incoming d(d)GTP is accommodated at the active site but not base-paired with the 5'-adjacent template base. Consistent with a weak and incomplete electron density, the un-base-paired and disordered d(d)GTP indicates that the ternary complexes studied here unlikely represent catalytically active conformations. Comparisons between the structures of the binary complexes and their corresponding ternary complexes reveal slight changes in the existence of the incoming d(d)GTP in the ternary complexes, suggesting that the spacious active site of Dpo4 is able to accommodate the incoming nucleotide without affecting the whole structure of the complex. The complex structures containing primer A opposite the BD-N3-dU adducts are different from the previously published "type I" or "type II" complex structures containing either native or planar chemically adducted templates, indicating the steric effect plays a significant role in lesion bypass of damaged DNA by trans-lesion polymerase Dpo4.

Crystal structures solved here are not the productive models representing the primer extension by Dpo4 when the BD-N3-dU nucleotide is opposite A in duplex structure. The non-planar BD lesion prevents the simultaneous accommodation of both the BD-N3-dU and the 5′-adjacent template nucleotide within the active site of Dpo4. Therefore, the insertion of the correct incoming nucleotide accompanied with the formation of Watson−Crick hydrogen-bonding interactions with the 5'-adjacement template nucleotide is disturbed consequently. In contrast to the crystal structures, the replication bypass analysis demonstrates that Dpo4 is able to insert a correct nucleotide opposite the 5'-adjacent template base when the BD-N3-dU adducts opposite A in the primer, and extend the primer to a full-length product. Thus, the crystal structures formed under

crystallization conditions unlikely represent the productive conformation. A possible mechanism to clarify the progress of primer extension with A in the primer opposite BD-N3-dU adducts in the template by Dpo4 remains unclear.

Summary

The BD moiety lies between the adducted uridine base and the 3'-terminal primer A, which induces the perturbation of classic Watson-Crick base-pairing interactions between the BD-N3-dU adducts and the 3'-terminal primer A. Moreover, the non-planar BD lesion prevents the stacking of the 5'-adjacent template base, and therefore perturbs base-pairing hydrogen-bonding interactions between the correct incoming nucleotide and the 5'-adjacent base in the template, leading to the formation of inactive complexes. However, replication and extension analysis indicate that Dpo4 is able to incorporate the correct nucleotide and extend the primer with the BD-derived N3-dU adducts opposite A. Crystal structures examined here display possible intermediates, but are not likely to represent the progress of primer extension. Further research remains to be carried out to understand the mechanism of the primer extension when the BD-derived N3-dU adducts are opposite A in the primer.

CHAPTER VI

CONCLUSIONS AND FUTURE DIRECTIONS

Conclusions

The structures of the site- and stereo-specific BD-derived N^6,N^6-dA intrastrand cross-linked adducts and N3-dU adducts in oligodeoxynucleotides containing the *ras*61 (5'-CAA-3') protooncogene sequence have been studied by NMR and X-ray crystallography techniques, respectively. The correlation of structure features with biochemical studies may provide an explanation of the biological processing of these particular adducts from the structural point of view.

The genotoxicity of the BD is related to its ability to form cross-links in DNA via its diepoxide oxidation product DEB, which is indeed considerably more mutagenic than the mono-epoxide EB (*240*). From *in vivo* experiments in which human lymphoblastoid TK6 and splenic T cells were exposed to low levels of DEB, transitions and transversions at both G·C and A·T sites were revealed, suggesting the presence of DEB-induced DNA cross-linked dG and dA adducts (*65*,*240*,*241*). Structural study of N^6,N^6-dA intrastrand cross-linked adducts here provide a model to understand DEB-induced DNA cross-linked adducts formed *in vivo*. NMR analysis reveals that the BD cross-link is oriented in the major groove of duplex DNA, resulting in Watson-Crick base-pair perturbation at the base pair of the adduct and its complementary bases. The cross-link appears to exist in two conformations with different hydrogen-bonding interactions between the C_β-OH

group of the cross-link and T^{16} O^4 or T^{17} O^4, in rapid exchange on the NMR time scale (*249*).

BD-N3-dU adducts are found to be highly mutagenic and block bypass and primer extension by most replicative and trans-lesion polymerases (*91,94*). Crystal structures of both binary Dpo4-DNA and ternary Dpo4-DNA-d(d)NTP complexes containing either *S*- or *R*- BD-N3-dU adducted template indicate that the steric hindrance of the non-planar BD moiety at the active site prevents the formation of catalytically active complexes. However, the disordered incoming nucleotides in the ternary complexes suggest that multiple intermediates might exist in each ternary complex, so that a catalytically active conformation of the ternary complex is probable to be formed, which may provide a possible explaination for the single nucleotides incorporation results from the replication analysis by Dpo4. The BD-N3-dU moiety is the first non-planar adduct studied in the Dpo4-DNA complexes. The resulting non-productive crystal structures suggest that steric effect plays a significant role in lesion bypass of damaged DNA by the trans-lesion polymerase Dpo4. Interestingly, replication and extension analysis indicate that Dpo4 is able to incorporate the correct nucleotide and extend the primer of the primer-template sequence with the BD-derived N3-dU adducts opposite the primer A. However, crystal structures examined here unlikely represent the progress of primer extension. Further research remains to be carried out to understand the mechanism of the primer extension when the BD-derived N3-dU adducts is opposite A in the primer. The structural results of complexes containing BD-N3-dU adducts, correlated with biochemical experiments, may provide insight into understanding the progress of the bypass and trans-lesion synthesis of non-planar chemical damaged DNA by Y-family polymerases.

Future Directions

Primary A → G mutations yielded by N^6,N^6-dA intrastrand cross-linked adducts indicate that the DEB-induced N^6,N^6-dA cross-linked adducts do not represent the source of the predominant A → T mutations induced by BD in bacterial or mammalian cells (*92*). Consequently, the specific DNA adducts arising from exposure to BD and responsible for inducing A → T mutations remain to be determined. This mutagenesis study also suggests that these cross-linked adducts facilitate misincorporation of dCTP opposite the cross-linked nucleotide adenine. Thus, structural examination of a mismatched dC opposite the N^6,N^6-dA cross-linked adduct will be of interest. The predominant adducts formed upon exposure of DNA to BD epoxides are N7-dG adducts (*80*). The N7,N7-dG interstrand cross-linked adduct has been determined to be the major DNA-DNA cross-link induced by DEB (*37*). With improvement in oligodeoxynucleotide synthesis and spectroscopic techniques, determination of the structures of *N*7-dG adducts will be expected.

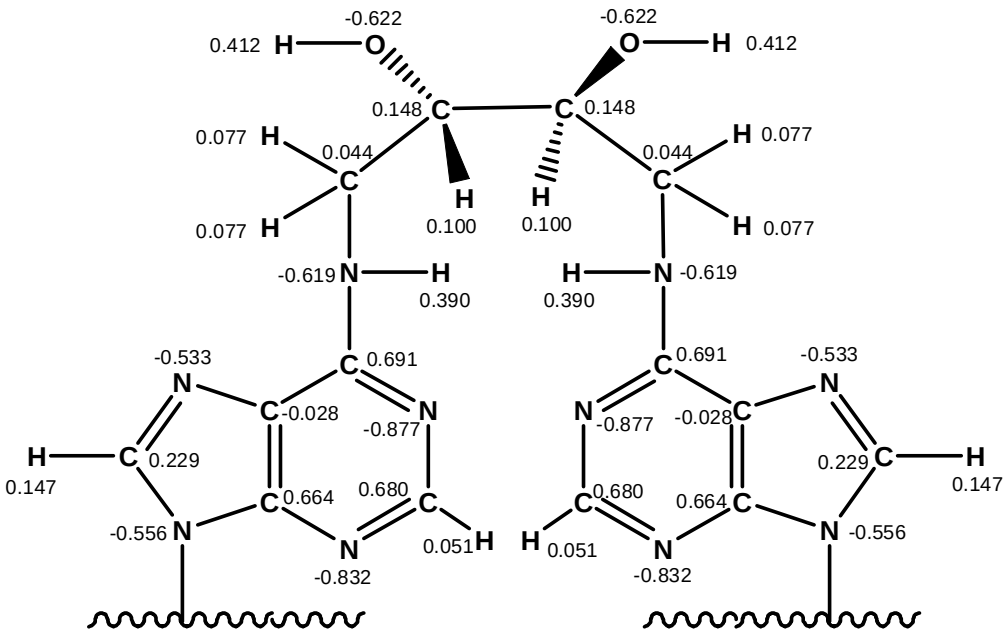
The high frequency of C → T transitions and C → A transversions resulting from the BD-N3-dU adducts in mammalian COS-7 cells suggests one or more polymerases are responsible for incorporation of dATP or dTTP opposite the adducted lesion (*91*). Human pol η is able to incorporate dGTP or dATP opposite BD-N3-dU adduct, suggesting that pol η may be one of the sources resulting in C → T mutations. The potential polymerases inducing C → A transvertions remained to be examined. Correlating with the biochemical studies, the investigation of structural properties of complexes containing the BD-N3-dU adducted template and different polymerases, such as pol η may be helpful in elucidating the biological progress. Dpo4 has been determined

to be able to incorporate the correct nucleotide and extend the primer to a full-length product with A opposite the BD-N3-dU adducts. However, the crystal structures of Dpo4 complexes containing BD-N3-dU adducts opposite A in the primer are unlikely to represent the catalytically active complexes, therefore further research remains to be carried out to explore the possible mechanism of the progress of the primer extension with BD-N3-dU adducts in the template opposite A in the primer.

APPENDIX A

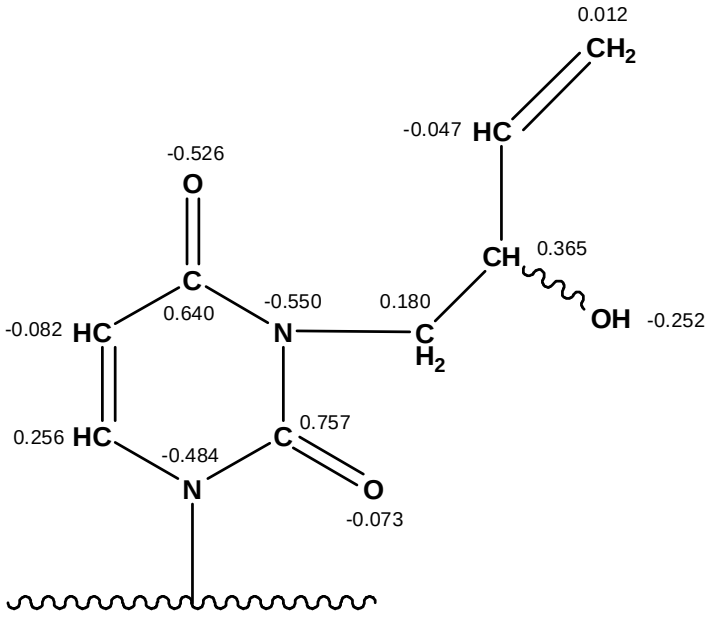
ATOM TYPE AND ATOMIC PARTIAL CHARGES

A1. Topology and parameters for the (*S*,*S*)-BD-(61,2-3) crosslinked adduct.



The atomic partial charges were calculated using GAUSSIAN 03 and the Hartree-Fock 6-31G* basis set.

A2. Topology and parameters for the BD-N3-dU adduct.



The atomic partial charges were calculated using GAUSSIAN 03 and the Hartree-Fock 6-31G* basis set.

APPENDIX B

CHEMICAL SHIFT ASSIGNMENTS

B1. Chemical shift assignments of non-exchangeable protons for the (*S*,*S*)-BD-(61,2-3) crosslinked adduct.

	H1'	H2'	H2"	H3'	H4'	H5'	H5"	H8	H5	H6
C^{1}	5.63	1.72	2.24	4.58	3.72	3.60	3.97		5.79	7.48
G^{2}	5.38	2.59	2.65	4.88	3.86	3.97	4.22	7.80		
G^{3}	5.59	2.58	2.69	4.95		4.32	4.10	7.70		
A^{4}	6.11	2.57	2.82	4.94		4.15	4.38	8.04		
C^{5}	5.43	1.85	2.26	4.93		4.07	4.17		5.15	7.11
X^{6}	6.05	2.52	2.82	4.90		4.08	4.33	7.99		
Y^{7}	5.78	2.18	2.82	4.57		3.94		7.81		
G^{8}	5.30	2.36	2.50	4.83		4.03	4.22	7.41		
A^{9}	5.90	2.53	2.81	4.97		4.08	4.35	8.00		
A^{10}	5.96	2.51	2.77	4.95		4.12	4.32	7.95		
G^{11}	5.94	2.19	2.30	4.53	4.06	4.13	4.20	7.52		
C^{12}	5.79	2.24	2.53	4.58	3.70	3.74	4.02		5.83	7.78
T^{13}	6.12	2.51	2.57	4.84		4.03	4.10			7.58
T^{14}	6.05	2.20	2.50	4.84	4.03	4.18	4.08			7.44
C^{15}	5.98	2.20	2.23	4.57		4.20	4.02		5.58	7.58
T^{16}	5.86	2.00	2.51	4.85		3.95	4.03			7.45
T^{17}	5.78	2.20	2.57	4.86		3.88	4.17			7.05
G^{18}	5.90	2.50	2.69	4.81		4.11	4.31	7.74		
T^{19}	5.99	2.11	2.47	4.81	4.20	4.14	4.09			7.25
C^{20}	5.88	2.24	2.34	4.76		4.08			5.57	7.48
C^{21}	5.57	1.91	2.05	4.75		4.02	4.07		5.63	7.40
G^{22}	6.10	2.57	2.31	4.61		3.99	4.11			7.88

H_{α}	2.87
$H_{\alpha'}$	3.32
H_{β}	3.89
H_{γ}	3.95
H_{δ}	3.00
$H_{\delta'}$	3.71

B2. Chemical shift assignments of exchangeable protons for the (*S*,*S*)-BD-(61,2-3) crosslinked adduct.

	H1	H3	NH1	NH2
C^{1}				
G^{2}	13.01		8.55	6.92
G^{3}	12.62		8.21	6.64
A^{4}				
C^{5}				
X^{6}				
Y^{7}				
G^{8}	12.29		8.31	7.04
A^{9}				
A^{10}				
G^{11}				
C^{12}				
T^{13}		14.14		
T^{14}		13.81		
C^{15}				
T^{16}		13.81		
T^{17}				
G^{18}	12.11		7.99	6.67
T^{19}		13.60		
C^{20}				
C^{21}				
G^{22}				

APPENDIX C

DISTANCE RESTRAINT FILE

Experimental distance restraints used in the rMD calculations of the (*S,S*)-BD-(61,2-3) crosslinked adduct.

```
#
# 1       CYT       H3'          1          CYT              H1'                3.19   4.24
&rst
  ixpk= 0, nxpk= 0, iat=  24,  10, r1= 2.69, r2= 3.19, r3= 4.24, r4= 4.74,
     rk2=30.0, rk3=30.0, ir6=1, ialtd=0,
&end
#
# 1       CYT       H5'          1          CYT              H1'     3.12   4.58
&rst
  ixpk= 0, nxpk= 0, iat=   4,  10, r1= 2.62, r2= 3.12, r3= 4.58, r4= 5.08,  &end
#
# 1       CYT       H5'          1          CYT              H3'     2.34   3.56
&rst
  ixpk= 0, nxpk= 0, iat=   4,  24, r1= 1.84, r2= 2.34, r3= 3.56, r4= 3.86,  &end
#
# 1       CYT       H1'          1          CYT              H5"     3.59   4.85
&rst
  ixpk= 0, nxpk= 0, iat=  10,   5, r1= 3.09, r2= 3.59, r3= 4.85, r4= 5.35,  &end
#
# 1       CYT       H3'          1          CYT              H6      3.15   4.27
&rst
  ixpk= 0, nxpk= 0, iat=  24,  13, r1= 2.65, r2= 3.15, r3= 4.27, r4= 4.77,  &end
#
# 1       CYT       H5           1          CYT              H6      1.78   2.42
&rst
  ixpk= 0, nxpk= 0, iat=  15,  13, r1= 1.28, r2= 1.78, r3= 2.42, r4= 2.92,  &end
#
# 1       CYT       H5'          1          CYT              H6      3.22   4.36
&rst
  ixpk= 0, nxpk= 0, iat=   4,  13, r1= 2.72, r2= 3.22, r3= 4.36, r4= 4.86,  &end
#
# 2       GUA       H8           1          CYT              H1'     3.66   4.7
&rst
```

```
  ixpk= 0, nxpk= 0, iat=  43,  10, r1= 3.16, r2= 3.66, r3= 4.70, r4= 5.20,  &end
#
# 2        GUA        H8          1          CYT          H2'     2.89    4.05
&rst
  ixpk= 0, nxpk= 0, iat=  43,  26, r1= 2.59, r2= 2.89, r3= 4.05, r4= 4.55,  &end
#
# 2        GUA        H3'         2          GUA          H1'     2.83    3.93
&rst
  ixpk= 0, nxpk= 0, iat=  57,  40, r1= 2.33, r2= 2.83, r3= 3.93, r4= 4.43,  &end
#
# 2        GUA        H5"         2          GUA          H1'     3.8         4.98
&rst
  ixpk= 0, nxpk= 0, iat=  35,  40, r1= 3.30, r2= 3.80, r3= 4.98, r4= 5.48,  &end
#
# 2        GUA        H8          2          GUA          H1'     2.58    3.68  change
&rst
  ixpk= 0, nxpk= 0, iat=  43,  40, r1= 2.08, r2= 2.58, r3= 3.83, r4= 4.18,  &end
#
# 2        GUA        H8          2          GUA          H4'     4.39    5.93
&rst
  ixpk= 0, nxpk= 0, iat=  43,  37, r1= 3.89, r2= 4.39, r3= 5.93, r4= 6.43,  &end
#
# 2        GUA        H5'         2          GUA          H8      3.81    5.29
&rst
  ixpk= 0, nxpk= 0, iat=  34,  43, r1= 3.31, r2= 3.81, r3= 5.29, r4= 5.79,  &end
#
# 2        GUA        H5"         2          GUA          H8      3.59    4.85
&rst
  ixpk= 0, nxpk= 0, iat=  35,  43, r1= 3.09, r2= 3.59, r3= 4.85, r4= 5.35,  &end
#
# 3        GUA        H1'         2          GUA          H1'     3.55    4.81
&rst
  ixpk= 0, nxpk= 0, iat=  73,  40, r1= 3.05, r2= 3.55, r3= 4.81, r4= 5.31,  &end
#
# 3        GUA        H3'         3          GUA          H1'     2.62    3.72
&rst
  ixpk= 0, nxpk= 0, iat=  90,  73, r1= 2.12, r2= 2.62, r3= 3.72, r4= 4.22,  &end
#
# 3        GUA        H8          3          GUA          H1'     2.19    3.8
&rst
  ixpk= 0, nxpk= 0, iat=  76,  73, r1= 1.69, r2= 2.19, r3= 3.80, r4= 4.30,  &end
#
# 3        GUA        H1'         3          GUA          H5'     2.57    4.41
&rst
  ixpk= 0, nxpk= 0, iat=  73,  67, r1= 2.07, r2= 2.57, r3= 4.41, r4= 4.91,  &end
```

```
#
# 3       GUA       H3'         3           GUA             H8      2.92    4.19
&rst
  ixpk= 0, nxpk= 0, iat=  90,  76, r1= 2.42, r2= 2.92, r3= 4.19, r4= 4.69,  &end
#
# 4       ADE       H8          3           GUA             H1'     2.67    3.81
&rst
  ixpk= 0, nxpk= 0, iat= 109,  73, r1= 2.17, r2= 2.67, r3= 3.81, r4= 4.21,  &end
#
# 4       ADE       H8          3           GUA             H8      3.83    5.17
&rst
  ixpk= 0, nxpk= 0, iat= 109,  76, r1= 3.33, r2= 3.83, r3= 5.17, r4= 5.67,  &end
#
# 4       ADE       H3'         4           ADE             H1'     3.04    3.94
&rst
  ixpk= 0, nxpk= 0, iat= 122, 106, r1= 2.54, r2= 3.04, r3= 3.94, r4= 4.44,  &end
#
# 4       ADE       H8          4           ADE             H1'     3.4     4.6
&rst
  ixpk= 0, nxpk= 0, iat= 109, 106, r1= 2.90, r2= 3.40, r3= 4.60, r4= 5.10,  &end
#
# 4       ADE       H5'         4           ADE             H8      3.31    4.47
&rst
  ixpk= 0, nxpk= 0, iat= 100, 109, r1= 2.81, r2= 3.31, r3= 4.47, r4= 4.97,  &end
#
# 5       CYT       H6          4           ADE             H1'     3.62    4.9
&rst
  ixpk= 0, nxpk= 0, iat= 141, 106, r1= 3.12, r2= 3.62, r3= 4.90, r4= 5.40,  &end
#
# 5       CYT       H5          4           ADE             H2'     3.83    5.33
&rst
  ixpk= 0, nxpk= 0, iat= 143, 124, r1= 3.33, r2= 3.83, r3= 5.33, r4= 5.83,  &end
#
# 5       CYT       H6          4           ADE             H2'     2.79    3.87
&rst
  ixpk= 0, nxpk= 0, iat= 141, 124, r1= 2.29, r2= 2.79, r3= 3.87, r4= 4.37,  &end
#
# 5       CYT       H6          4           ADE             H2"     2.64    3.84
&rst
  ixpk= 0, nxpk= 0, iat= 141, 125, r1= 2.14, r2= 2.59, r3= 3.84, r4= 4.34,  &end
#
# 5       CYT       H6          4           ADE             H3'     4.48    5.7
&rst
  ixpk= 0, nxpk= 0, iat= 141, 122, r1= 3.98, r2= 4.48, r3= 5.70, r4= 6.20,  &end
#
```

```
# 5       CYT       H5        4         ADE       H8      3.55    4.81
&rst
  ixpk= 0, nxpk= 0, iat= 143, 109, r1= 3.05, r2= 3.55, r3= 4.81, r4= 5.31,  &end
#
# 5       CYT       H6        4         ADE       H8      3.72    5.14
&rst
  ixpk= 0, nxpk= 0, iat= 141, 109, r1= 3.22, r2= 3.72, r3= 5.14, r4= 5.64,  &end
#
# 5       CYT       H2'       5         CYT       H1'     2.7     3.66
&rst
  ixpk= 0, nxpk= 0, iat= 154, 138, r1= 2.20, r2= 2.70, r3= 3.66, r4= 4.16,  &end
#
# 5       CYT       H2"       5         CYT       H1'     2.15    3.17
&rst
  ixpk= 0, nxpk= 0, iat= 155, 138, r1= 1.65, r2= 2.15, r3= 3.17, r4= 3.67,  &end
#
# 5       CYT       H6        5         CYT       H1'     3.53    4.77
&rst
  ixpk= 0, nxpk= 0, iat= 141, 138, r1= 3.03, r2= 3.53, r3= 4.77, r4= 5.27,  &end
#
# 5       CYT       H2'       5         CYT       H5      3.07    4.15 change
&rst
  ixpk= 0, nxpk= 0, iat= 154, 143, r1= 2.57, r2= 3.07, r3= 4.23, r4= 4.65,  &end
#
# 5       CYT       H3'       5         CYT       H5      2.69    4.85
&rst
  ixpk= 0, nxpk= 0, iat= 152, 143, r1= 2.19, r2= 2.69, r3= 4.85, r4= 5.35,  &end
#
# 5       CYT       H2'       5         CYT       H6      2.24    3.3
&rst
  ixpk= 0, nxpk= 0, iat= 154, 141, r1= 1.74, r2= 2.24, r3= 3.30, r4= 3.80,  &end
#
# 5       CYT       H2"       5         CYT       H6      2.73    3.89
&rst
  ixpk= 0, nxpk= 0, iat= 155, 141, r1= 2.23, r2= 2.73, r3= 3.89, r4= 4.39,  &end
#
# 5       CYT       H5        5         CYT       H6      2.07    2.81
&rst
  ixpk= 0, nxpk= 0, iat= 143, 141, r1= 1.57, r2= 2.07, r3= 2.81, r4= 3.31,  &end
#
# 5       CYT       H5'       5         CYT       H6      4.32    5.84
&rst
  ixpk= 0, nxpk= 0, iat= 132, 141, r1= 3.82, r2= 4.32, r3= 5.84, r4= 6.34,  &end
#
# 6       ABX       H8        5         CYT       H1'     4.16    5.62
```

```
&rst
  ixpk= 0, nxpk= 0, iat= 174, 138, r1= 3.66, r2= 4.16, r3= 5.62, r4= 6.12,  &end
#
# 6         ABX         H8            5             CYT               H2'      2.57    3.69
&rst
  ixpk= 0, nxpk= 0, iat= 174, 154, r1= 2.07, r2= 2.57, r3= 3.69, r4= 4.19,  &end
#
# 6         ABX         2H1X          5             CYT               H5       3.04    4.38
&rst
  ixpk= 0, nxpk= 0, iat= 182, 143, r1= 2.54, r2= 3.04, r3= 4.38, r4= 4.88,  &end
#
# 6         ABX         H8            6             ABX               H1'      2.92    3.8 change
&rst
  ixpk= 0, nxpk= 0, iat= 174, 168, r1= 2.42, r2= 2.92, r3= 4.50, r4= 5.00,  &end
#
# 6         ABX         2H1X          6             ABX               1H1X     1.71    2.59
&rst
  ixpk= 0, nxpk= 0, iat= 182, 181, r1= 1.21, r2= 1.71, r3= 2.59, r4= 3.09,  &end
#
# 6         ABX         1H4X          6             ABX               H2X      3.33    5.41
&rst
  ixpk= 0, nxpk= 0, iat= 188, 184, r1= 2.83, r2= 3.33, r3= 5.41, r4= 5.91,  &end
#
# 6         ABX         1H4X          6             ABX               H3X      2.56    4.82
&rst
  ixpk= 0, nxpk= 0, iat= 188, 186, r1= 2.06, r2= 2.56, r3= 4.82, r4= 5.32,  &end
#
# 6         ABX         2H4X          6             ABX               1H4X     1.8     2.84
&rst
  ixpk= 0, nxpk= 0, iat= 189, 188, r1= 1.30, r2= 1.80, r3= 2.84, r4= 3.34,  &end
#
# 7         ADE         H1'           8             GUA               H8       3.32    4.87
&rst
  ixpk= 0, nxpk= 0, iat= 213, 247, r1= 2.82, r2= 3.32, r3= 4.87, r4= 5.37,  &end
#
# 8         GUA         H3'           8             GUA         H1'      3.96    5.5
&rst
  ixpk= 0, nxpk= 0, iat= 261, 244, r1= 3.46, r2= 3.96, r3= 5.50, r4= 6.00,  &end
#
# 8         GUA         H5'           8             GUA               H1'      3.15    4.37
&rst
  ixpk= 0, nxpk= 0, iat= 238, 244, r1= 2.65, r2= 3.15, r3= 4.37, r4= 4.87,  &end
#
# 8         GUA         H5"           8             GUA               H1'      3.78    5.06
&rst
```

```
  ixpk= 0, nxpk= 0, iat= 239, 244, r1= 3.28, r2= 3.78, r3= 5.06, r4= 5.56,  &end
#
# 8       GUA       H8          8          GUA          H1'     2.75    3.83
&rst
  ixpk= 0, nxpk= 0, iat= 247, 244, r1= 2.25, r2= 2.75, r3= 3.83, r4= 4.33,  &end
#
# 8       GUA       H5'         8          GUA          H2'     3.37    4.69
&rst
  ixpk= 0, nxpk= 0, iat= 238, 263, r1= 2.87, r2= 3.37, r3= 4.69, r4= 5.19,  &end
#
# 9       ADE       H8          8          GUA          H1'     3.26    4.4
&rst
  ixpk= 0, nxpk= 0, iat= 280, 244, r1= 2.76, r2= 3.26, r3= 4.40, r4= 4.90,  &end
#
# 9       ADE       H3'         9          ADE          H1'     2.86    3.82
&rst
  ixpk= 0, nxpk= 0, iat= 293, 277, r1= 2.36, r2= 2.86, r3= 3.82, r4= 4.32,  &end
#
# 10      ADE       H3'         10         ADE          H1'     3.32    4.48
&rst
  ixpk= 0, nxpk= 0, iat= 325, 309, r1= 2.82, r2= 3.32, r3= 4.48, r4= 4.98,  &end
#
# 10      ADE       H5'         10         ADE          H8      3.87    5.33
&rst
  ixpk= 0, nxpk= 0, iat= 303, 312, r1= 3.37, r2= 3.87, r3= 5.33, r4= 5.83,  &end
#
# 11      GUA       H8          10         ADE          H2"     2.85    3.82
&rst
  ixpk= 0, nxpk= 0, iat= 344, 328, r1= 2.35, r2= 2.73, r3= 3.82, r4= 4.32,  &end
#
# 11      GUA       H8          10         ADE          H3'     4.1     5.54
&rst
  ixpk= 0, nxpk= 0, iat= 344, 325, r1= 3.60, r2= 4.10, r3= 5.54, r4= 6.04,  &end
#
# 11      GUA       H8          10         ADE          H8      3.45    4.77
&rst
  ixpk= 0, nxpk= 0, iat= 344, 312, r1= 2.95, r2= 3.45, r3= 4.77, r4= 5.27,  &end
#
# 11      GUA       H3'         11         GUA          H1'     3.48    4.84
&rst
  ixpk= 0, nxpk= 0, iat= 358, 341, r1= 2.98, r2= 3.48, r3= 4.84, r4= 5.34,  &end
#
# 12      CYT       H3'         12         CYT          H1'     3           4.06
&rst
  ixpk= 0, nxpk= 0, iat= 387, 373, r1= 2.50, r2= 3.00, r3= 4.06, r4= 4.56,  &end
```

```
#
# 12       CYT          H5'             12              CYT                  H3'     3            4.06
&rst
  ixpk= 0, nxpk= 0, iat= 367, 387, r1= 2.50, r2= 3.00, r3= 4.06, r4= 4.56,  &end
#
# 12       CYT          H5"             12              CYT                  H3'     2.84    3.98
&rst
  ixpk= 0, nxpk= 0, iat= 368, 387, r1= 2.34, r2= 2.84, r3= 3.98, r4= 4.48,  &end
#
# 12       CYT          H6              12              CYT                  H4'     4.25    5.75
&rst
  ixpk= 0, nxpk= 0, iat= 376, 370, r1= 3.75, r2= 4.25, r3= 5.75, r4= 6.25,  &end
#
# 12       CYT          H3'             12              CYT                  H6      2.61    3.63
&rst
  ixpk= 0, nxpk= 0, iat= 387, 376, r1= 2.11, r2= 2.61, r3= 3.63, r4= 4.13,  &end
#
# 12       CYT          H5'             12              CYT                  H6      3.26    4.4
&rst
  ixpk= 0, nxpk= 0, iat= 367, 376, r1= 2.76, r2= 3.26, r3= 4.40, r4= 4.90,  &end
#
# 12       CYT          H5"             12              CYT                  H6      3.4     4.6
&rst
  ixpk= 0, nxpk= 0, iat= 368, 376, r1= 2.90, r2= 3.40, r3= 4.60, r4= 5.10,  &end
#
# 13       THY          H7              12              CYT                  H1'     4.07    6.09
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 373, r1= 3.57, r2= 4.07, r3= 6.09, r4= 6.59,
 igr1= 409, 410, 411,
&end
#
# 13       THY          H6              12              CYT                  H1'     4.78    6.45
&rst
  ixpk= 0, nxpk= 0, iat= 406, 373, r1= 4.28, r2= 4.78, r3= 6.45, r4= 6.95,  &end
#
# 13       THY          H7              12              CYT                  H3'     1.94    4.40
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 387, r1= 1.44, r2= 1.94, r3= 4.40, r4= 4.90,
 igr1= 409, 410, 411,
&end
#
# 13       THY          H7              12              CYT                  H5      3.02    5.84
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 378, r1= 2.52, r2= 3.02, r3= 5.84, r4= 6.34,
 igr1= 409, 410, 411,
```

```
&end
#
# 13      THY         H7          12          CYT             H6      2.96    4.8
 &rst
  ixpk= 0, nxpk= 0, iat=  -1, 376, r1= 2.46, r2= 2.96, r3= 4.80, r4= 5.30,
 igr1= 409, 410, 411,
&end
#
# 13      THY         H3'         13          THY             H1'     3.42    4.62
&rst
  ixpk= 0, nxpk= 0, iat= 419, 403, r1= 2.92, r2= 3.42, r3= 4.62, r4= 5.12,  &end
#
# 13      THY         H7          13          THY             H1'     3.52    5.75 change
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 403, r1= 3.02, r2= 3.52, r3= 5.95, r4= 6.25,
 igr1= 409, 410, 411,
&end
#
# 13      THY         H5'         13          THY             H7      3.41    4.55 change
&rst
  ixpk= 0, nxpk= 0, iat= 397,  -1, r1= 2.91, r2= 3.41, r3= 5.61, r4= 5.96,
 igr2= 409, 410, 411,
&end
#
# 13      THY         H3'         13          THY             H6      2.83    3.83
&rst
  ixpk= 0, nxpk= 0, iat= 419, 406, r1= 2.33, r2= 2.83, r3= 3.83, r4= 4.33,  &end
#
# 14      THY         H7          14          THY             H1'     4.06    6.02 change
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 435, r1= 3.56, r2= 4.06, r3= 7.23, r4= 7.73,
 igr1= 441, 442, 443,
&end
#
# 15      CYT         H5'         15          CYT             H6      3.35    4.53
&rst
  ixpk= 0, nxpk= 0, iat= 461, 470, r1= 2.85, r2= 3.35, r3= 4.53, r4= 5.03,  &end
#
# 16      THY         H7          15          CYT             H3'     2.14    4.86
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 481, r1= 1.64, r2= 2.14, r3= 4.86, r4= 5.36,
 igr1= 503, 504, 505,
&end
#
# 17      THY         H3'         17          THY             H1'     3.21    4.35
```

```
&rst
  ixpk= 0, nxpk= 0, iat= 545, 529, r1= 2.71, r2= 3.21, r3= 4.35, r4= 4.85,  &end
#
# 17      THY         H6            17            THY               H1'     3.45    4.25
&rst
  ixpk= 0, nxpk= 0, iat= 532, 529, r1= 2.95, r2= 3.45, r3= 4.25, r4= 4.75,  &end
#
# 18      GUA         H5'           18            GUA               H8      3.46    4.54
&rst
  ixpk= 0, nxpk= 0, iat= 555, 564, r1= 2.96, r2= 3.46, r3= 4.54, r4= 5.04,  &end
#
# 19      THY         H7            18            GUA               H1'     3.44    5.6
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 561, r1= 2.94, r2= 3.44, r3= 5.60, r4= 6.10,
 igr1= 600, 601, 602,
&end
#
# 19      THY         H6            18            GUA               H1'     3.15    4.61
&rst
  ixpk= 0, nxpk= 0, iat= 597, 561, r1= 2.65, r2= 3.15, r3= 4.61, r4= 5.11,  &end
#
# 19      THY         H7            18            GUA               H2"     2.02    3.37
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 581, r1= 1.52, r2= 2.02, r3= 3.37, r4= 3.87,
 igr1= 600, 601, 602,
&end
#
# 19      THY         H6            18            GUA               H2"     2.55    3.27
&rst
  ixpk= 0, nxpk= 0, iat= 597, 581, r1= 2.05, r2= 2.55, r3= 3.27, r4= 3.77,  &end
#
# 19      THY         H7            18            GUA               H8      2.27    5.08
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 564, r1= 1.77, r2= 2.27, r3= 5.08, r4= 5.58,
 igr1= 600, 601, 602,
&end
#
# 19      THY         H6            19            THY               H1'     2.97    4.03
&rst
  ixpk= 0, nxpk= 0, iat= 597, 594, r1= 2.47, r2= 2.97, r3= 4.03, r4= 4.53,  &end
#
# 19      THY         H2'           19            THY               H7      2.14    4.08
&rst
  ixpk= 0, nxpk= 0, iat= 612,  -1, r1= 1.64, r2= 2.14, r3= 4.08, r4= 4.58,
 igr2= 600, 601, 602,
```

```
&end
#
# 19      THY         H2"          19          THY              H7      3.31    5.34
&rst
  ixpk= 0, nxpk= 0, iat= 613,  -1, r1= 2.81, r2= 3.31, r3= 5.34, r4= 5.84,
 igr2= 600, 601, 602,
&end
#
# 19      THY         H2'          19          THY              H6      1.24    2.02
&rst
  ixpk= 0, nxpk= 0, iat= 612, 597, r1= 0.74, r2= 1.24, r3= 2.02, r4= 2.52,  &end
#
# 19      THY         H2"          19          THY              H6      1.89    3.29
&rst
  ixpk= 0, nxpk= 0, iat= 613, 597, r1= 1.39, r2= 1.89, r3= 3.29, r4= 3.79,  &end
#
# 19      THY         H3'          19          THY              H6      2.75    3.91
&rst
  ixpk= 0, nxpk= 0, iat= 610, 597, r1= 2.25, r2= 2.75, r3= 3.91, r4= 4.41,  &end
#
# 19      THY         H7           19          THY              H6      1.73    3.01
 &rst
  ixpk= 0, nxpk= 0, iat=  -1, 597, r1= 1.23, r2= 1.73, r3= 3.01, r4= 3.51,
 igr1= 600, 601, 602,
&end
#
# 20      CYT         H5           19          THY              H7      4.87    6.23
&rst
  ixpk= 0, nxpk= 0, iat= 631,  -1, r1= 4.37, r2= 4.87, r3= 6.23, r4= 6.73,
 igr2= 600, 601, 602,
&end
#
# 20      CYT         H5           19          THY              H6      3.99    5.39
&rst
  ixpk= 0, nxpk= 0, iat= 631, 597, r1= 3.49, r2= 3.99, r3= 5.39, r4= 5.89,  &end
#
# 21      CYT         H3'          21          CYT              H2'     2.22    3.36
&rst
  ixpk= 0, nxpk= 0, iat= 670, 672, r1= 1.72, r2= 2.22, r3= 3.36, r4= 3.86,  &end
#
# 21      CYT         H2"          21          CYT              H5      2.7     3.88  change
&rst
  ixpk= 0, nxpk= 0, iat= 673, 661, r1= 2.20, r2= 2.70, r3= 4.22, r4= 4.38,  &end
#
# 22      GUA         H8           21          CYT              H1'     4.05    5.61
```

```
&rst
  ixpk= 0, nxpk= 0, iat= 689, 656, r1= 3.65, r2= 4.15, r3= 5.61, r4= 6.11,  &end
#
# 22      GUA         H8            21            CYT               H2'     3.46    5.28
&rst
  ixpk= 0, nxpk= 0, iat= 689, 672, r1= 2.96, r2= 3.46, r3= 5.28, r4= 5.78,  &end
#
# 22      GUA         H3'           22            GUA               H1'     3.54    4.78
&rst
  ixpk= 0, nxpk= 0, iat= 703, 686, r1= 3.04, r2= 3.54, r3= 4.78, r4= 5.28,  &end
#
# 22      GUA         H3'           22            GUA               H8      3.41    4.67
 &rst
  ixpk= 0, nxpk= 0, iat= 703, 689, r1= 2.91, r2= 3.41, r3= 4.67, r4= 5.17,  &end
#
# 22      GUA         H5'           22            GUA               H8      3.72    5.04
&rst
  ixpk= 0, nxpk= 0, iat= 680, 689, r1= 3.22, r2= 3.72, r3= 5.04, r4= 5.54,  &end
#
# 1       CYT         H5'           1             CYT               H5"                 1.78    2.54
&rst
  ixpk= 0, nxpk= 0, iat=   4,   5, r1= 1.28, r2= 1.78, r3= 2.54, r4= 3.04,
     rk2=25.0, rk3=25.0, ir6=1, ialtd=0,
&end
#
# 1       CYT         H2"           1             CYT               H6                  3.26    4.06
&rst
  ixpk= 0, nxpk= 0, iat=  27,  13, r1= 2.76, r2= 3.26, r3= 4.06, r4= 4.56,  &end
#
# 2       GUA         H3'           2             GUA               H8                  2.59    3.78
&rst
  ixpk= 0, nxpk= 0, iat=  57,  43, r1= 2.09, r2= 2.59, r3= 3.78, r4= 4.28,  &end
#
# 3       GUA         H8            2             GUA               H1'                 3.24    4.38
&rst
  ixpk= 0, nxpk= 0, iat=  76,  40, r1= 2.74, r2= 3.24, r3= 4.38, r4= 4.88,  &end
#
# 3       GUA         H2"           3             GUA               H1'                 1.98    2.68
&rst
  ixpk= 0, nxpk= 0, iat=  93,  73, r1= 1.48, r2= 1.98, r3= 2.68, r4= 3.18,  &end
#
# 3       GUA         H4'           3             GUA               H3'                 2.65    3.59
&rst
  ixpk= 0, nxpk= 0, iat=  70,  90, r1= 2.15, r2= 2.65, r3= 3.59, r4= 4.09,  &end
#
```

```
# 4       ADE       H2"   4       ADE       H1'                2.19   2.77
&rst
  ixpk= 0, nxpk= 0, iat= 125, 106, r1= 1.69, r2= 2.19, r3= 2.77, r4= 3.27,  &end
#
# 4       ADE       H5"         4         ADE         H2'                3.4    4.6
&rst
  ixpk= 0, nxpk= 0, iat= 101, 124, r1= 2.90, r2= 3.40, r3= 4.60, r4= 5.10,  &end
#
# 4       ADE       H5"         4         ADE         H2"                3.87   5.03
&rst
  ixpk= 0, nxpk= 0, iat= 101, 125, r1= 3.37, r2= 3.87, r3= 5.03, r4= 5.53,  &end
#
# 4       ADE       H5"         4         ADE         H8                 3.75   5.07
&rst
  ixpk= 0, nxpk= 0, iat= 101, 109, r1= 3.25, r2= 3.75, r3= 5.07, r4= 5.57,  &end
#
# 6       ABX       H8          5         CYT         H2"                3.1    4.2
&rst
  ixpk= 0, nxpk= 0, iat= 174, 155, r1= 2.60, r2= 3.10, r3= 4.20, r4= 4.70,  &end
#
# 6       ABX       H2X         6         ABX         1H1X               2.21   3.25
&rst
  ixpk= 0, nxpk= 0, iat= 184, 181, r1= 1.71, r2= 2.21, r3= 3.25, r4= 3.75,  &end
#
# 6       ABX       H3X         6         ABX         1H1X               3.63   5.17
&rst
  ixpk= 0, nxpk= 0, iat= 186, 181, r1= 3.13, r2= 3.63, r3= 5.17, r4= 5.67,  &end
#
# 6       ABX       H2X         6         ABX         2H1X               2.3    3.1
&rst
  ixpk= 0, nxpk= 0, iat= 184, 182, r1= 1.80, r2= 2.30, r3= 3.10, r4= 3.60,  &end
#
# 6       ABX       H3X         6         ABX         2H1X               3.82   4.76
&rst
  ixpk= 0, nxpk= 0, iat= 186, 182, r1= 3.32, r2= 3.82, r3= 4.76, r4= 5.26,  &end
#
# 7       ADE       H8          6         ABX         H1'                3.88   5.24
&rst
  ixpk= 0, nxpk= 0, iat= 216, 168, r1= 3.38, r2= 3.88, r3= 5.24, r4= 5.74,  &end
#
# 8       GUA       H2'         8         GUA         H1'                2.81   3.81
&rst
  ixpk= 0, nxpk= 0, iat= 263, 244, r1= 2.31, r2= 2.81, r3= 3.81, r4= 4.31,  &end
#
# 8       GUA       H2"         8         GUA         H1'                2.38   3.56
```

```
&rst
  ixpk= 0, nxpk= 0, iat= 264, 244, r1= 1.88, r2= 2.38, r3= 3.56, r4= 4.06,  &end
#
# 9       ADE       H8          8          GUA          H2'          3.29    4.45
&rst
  ixpk= 0, nxpk= 0, iat= 280, 263, r1= 2.79, r2= 3.29, r3= 4.45, r4= 4.95,  &end
#
# 9       ADE       H8          9          ADE          H1'          2.96    3.88
&rst
  ixpk= 0, nxpk= 0, iat= 280, 277, r1= 2.46, r2= 2.96, r3= 3.88, r4= 4.38,  &end
#
# 10      ADE       H3'         10         ADE          H8           3.64    4.75
&rst
  ixpk= 0, nxpk= 0, iat= 325, 312, r1= 3.14, r2= 3.64, r3= 4.75 r4= 5.25,  &end
#
# 11      GUA       H3'         11         GUA          H8           3.76    5.08
&rst
  ixpk= 0, nxpk= 0, iat= 358, 344, r1= 3.26, r2= 3.76, r3= 5.08, r4= 5.58,  &end
#
# 12      CYT       H2"         12         CYT          H1'          1.75    2.37
&rst
  ixpk= 0, nxpk= 0, iat= 390, 373, r1= 1.25, r2= 1.75, r3= 2.37, r4= 2.87,  &end
#
# 12      CYT       H5'         12         CYT          H4'          1.95    2.65
&rst
  ixpk= 0, nxpk= 0, iat= 367, 370, r1= 1.45, r2= 1.95, r3= 2.65, r4= 3.15,  &end
#
# 13      THY       H2"         13         THY          H7           5.04    5.19 change
&rst
  ixpk= 0, nxpk= 0, iat= 422,  -1, r1= 4.54, r2= 5.04, r3= 5.49, r4= 5.69,
 igr2= 409, 410, 411,
 &end
#
# 13      THY       H5'         13         THY          H6           3.5     4.74
&rst
  ixpk= 0, nxpk= 0, iat= 397, 406, r1= 3.00, r2= 3.50, r3= 4.74, r4= 5.24,  &end
#
# 13      THY       H5"         13         THY          H6           3.8     5.24
&rst
  ixpk= 0, nxpk= 0, iat= 398, 406, r1= 3.30, r2= 3.80, r3= 5.24, r4= 5.74,  &end
#
# 14      THY       H2"         14         THY          H7           2.71    6.01 change
&rst
  ixpk= 0, nxpk= 0, iat= 454,  -1, r1= 2.21, r2= 2.71, r3= 7.22, r4= 7.72,
 igr2= 441, 442, 443,
```

```
&end
#
# 14      THY         H5'          14          THY              H6                3.4    4.7
&rst
  ixpk= 0, nxpk= 0, iat= 429, 438, r1= 2.90, r2= 3.40, r3= 4.70, r4= 5.20,  &end
#
# 15      CYT         H5           14          THY              H7                3.12   5.07
&rst
  ixpk= 0, nxpk= 0, iat= 472,  -1, r1= 2.62, r2= 3.12, r3= 5.07, r4= 5.57,
 igr2= 441, 442, 443,
&end
#
# 15      CYT         H3'          15          CYT              H6                3.6    4.88
&rst
  ixpk= 0, nxpk= 0, iat= 481, 470, r1= 3.10, r2= 3.60, r3= 4.88, r4= 5.38,  &end
#
# 15      CYT         H5           15          CYT              H6                2.21   2.99
&rst
  ixpk= 0, nxpk= 0, iat= 472, 470, r1= 1.71, r2= 2.21, r3= 2.99, r4= 3.49,  &end
#
# 16      THY         H7           15          CYT              H5                2.34   3.79
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 472, r1= 1.84, r2= 2.34, r3= 3.79, r4= 4.29,
 igr1= 503, 504, 505,
&end
#
# 16      THY         H5'          16          THY              H1'               2.42   4.52
&rst
  ixpk= 0, nxpk= 0, iat= 491, 497, r1= 1.92, r2= 2.42, r3= 4.52, r4= 6.02,  &end
#
# 17      THY         H5"          17          THY              H1'               3.8    4.84
&rst
  ixpk= 0, nxpk= 0, iat= 524, 529, r1= 3.30, r2= 3.80, r3= 4.84, r4= 5.34,  &end
#
# 18      GUA         H3'          18          GUA              H2"               2.78   3.76
&rst
  ixpk= 0, nxpk= 0, iat= 578, 581, r1= 2.28, r2= 2.78, r3= 3.76, r4= 4.26,  &end
#
# 18      GUA         H3'          18          GUA              H8                3.12   4.32
&rst
  ixpk= 0, nxpk= 0, iat= 578, 564, r1= 2.62, r2= 3.12, r3= 4.32, r4= 4.82,  &end
#
# 19      THY         H3'          19          THY              H1'               3.41   4.61
&rst
  ixpk= 0, nxpk= 0, iat= 610, 594, r1= 2.91, r2= 3.41, r3= 4.61, r4= 5.11,  &end
```

```
#
# 19      THY         H5'          19          THY              H6               3.5    4.74
&rst
  ixpk= 0, nxpk= 0, iat= 588, 597, r1= 3.00, r2= 3.50, r3= 4.74, r4= 5.24,  &end
#
# 20      CYT         H3'          20          CYT              H2'              2.4    4.32
&rst
  ixpk= 0, nxpk= 0, iat= 640, 642, r1= 1.90, r2= 2.40, r3= 4.32, r4= 4.82,  &end
#
# 20      CYT         H3'          20          CYT              H2"              2.59   3.87
&rst
  ixpk= 0, nxpk= 0, iat= 640, 643, r1= 2.09, r2= 2.59, r3= 3.87, r4= 4.37,  &end
#
# 21      CYT         H2"          20          CYT              H2"              2.87   3.63
&rst
  ixpk= 0, nxpk= 0, iat= 673, 643, r1= 2.37, r2= 2.87, r3= 3.65, r4= 4.13,  &end
#
# 21      CYT         H2'          21          CYT              H1'              3.12   4.22
&rst
  ixpk= 0, nxpk= 0, iat= 672, 656, r1= 2.62, r2= 3.12, r3= 4.22, r4= 4.72,  &end
#
# 21      CYT         H5'          21          CYT              H2'              2.86   3.88
&rst
  ixpk= 0, nxpk= 0, iat= 650, 672, r1= 2.36, r2= 2.86, r3= 3.88, r4= 4.38,  &end
#
# 21      CYT         H2'          21          CYT              H5               3.59   4.38
&rst
  ixpk= 0, nxpk= 0, iat= 672, 661, r1= 3.09, r2= 3.59, r3= 4.38, r4= 4.88,  &end
#
# 21      CYT         H5'          21          CYT              H6               3.33   5.01
&rst
  ixpk= 0, nxpk= 0, iat= 650, 659, r1= 2.83, r2= 3.33, r3= 5.01, r4= 5.51,  &end
#
# 22      GUA         H8           22          GUA              H1'               2.74   3.8
&rst
  ixpk= 0, nxpk= 0, iat= 689, 686, r1= 2.24, r2= 2.74, r3= 3.80, r4= 4.30,  &end
#
# 1       CYT         H6           1           CYT              H1'              2.4    3.74
&rst
  ixpk= 0, nxpk= 0, iat=  13,  10, r1= 1.90, r2= 2.40, r3= 3.74, r4= 4.24,
     rk2=20.0, rk3=20.0, ir6=1, ialtd=0,
&end
#
# 1       CYT         H2"          1           CYT              H2'              1.73   2.35
&rst
```

```
  ixpk= 0, nxpk= 0, iat=  27,  26, r1= 1.23, r2= 1.73, r3= 2.35, r4= 2.85,  &end
#
# 2       GUA         H5'         2           GUA         H1'          3.51    4.75
&rst
  ixpk= 0, nxpk= 0, iat=  34,  40, r1= 3.01, r2= 3.51, r3= 4.75, r4= 5.25,  &end
#
# 3       GUA         H2'         3           GUA         H1'           2.9    4.8
&rst
  ixpk= 0, nxpk= 0, iat=  92,  73, r1= 2.40, r2= 2.90, r3= 4.80, r4= 5.30,  &end
#
# 3       GUA         H5"         3           GUA         H1'          4.06    5.14
&rst
  ixpk= 0, nxpk= 0, iat=  68,  73, r1= 3.56, r2= 4.06, r3= 5.14, r4= 5.64,  &end
#
# 3       GUA         H2'         3           GUA         H8           2.06    2.92
&rst
  ixpk= 0, nxpk= 0, iat=  92,  76, r1= 1.56, r2= 2.06, r3= 2.92, r4= 3.42,  &end
#
# 4       ADE         H8          3           GUA         H2"          2.38    3.76
&rst
  ixpk= 0, nxpk= 0, iat= 109,  93, r1= 1.88, r2= 2.28, r3= 3.76, r4= 4.26,  &end
#
# 4       ADE         H5'         4           ADE         H2"           3.8    5.08
&rst
  ixpk= 0, nxpk= 0, iat= 100, 125, r1= 3.30, r2= 3.80, r3= 5.08, r4= 5.58,  &end
#
# 5       CYT         H5'         5           CYT         H1'          3.57    4.47
&rst
  ixpk= 0, nxpk= 0, iat= 132, 138, r1= 3.07, r2= 3.57, r3= 4.47, r4= 4.97,  &end
#
# 6       ABX         H2"         6           ABX         H1'           2.2    3.12
&rst
  ixpk= 0, nxpk= 0, iat= 171, 168, r1= 1.70, r2= 2.20, r3= 3.12, r4= 3.62,  &end
#
# 6       ABX         H3'         6           ABX         H1'          3.25    4.39
&rst
  ixpk= 0, nxpk= 0, iat= 200, 168, r1= 2.75, r2= 3.25, r3= 4.39, r4= 4.89,  &end
#
# 6       ABX         H5"         6           ABX         H2"           2.9    4.7
&rst
  ixpk= 0, nxpk= 0, iat= 163, 171, r1= 2.40, r2= 2.90, r3= 4.70, r4= 5.20,  &end
#
# 6       ABX         H3'         6           ABX         H8           2.74    3.9 change
&rst
  ixpk= 0, nxpk= 0, iat= 200, 174, r1= 2.24, r2= 2.74, r3= 4.15, r4= 4.40,  &end
```

```
#
# 6       ABX         H5'         6           ABX             H8              3.82    5.16
&rst
  ixpk= 0, nxpk= 0, iat= 162, 174, r1= 3.32, r2= 3.82, r3= 5.16, r4= 5.66,  &end
#
# 8       GUA         H2'         8           GUA             H8              1.96    2.66
&rst
  ixpk= 0, nxpk= 0, iat= 263, 247, r1= 1.46, r2= 1.96, r3= 2.66, r4= 3.16,  &end
#
# 9       ADE         H2'         9           ADE             H1'             2.63    3.55
&rst
  ixpk= 0, nxpk= 0, iat= 295, 277, r1= 2.13, r2= 2.63, r3= 3.55, r4= 4.05,  &end
#
# 9       ADE         H2"         9           ADE             H1'             2.31    3.13
&rst
  ixpk= 0, nxpk= 0, iat= 296, 277, r1= 1.81, r2= 2.31, r3= 3.13, r4= 3.63,  &end
#
# 9       ADE         H5'         9           ADE             H1'             3.57    4.57
&rst
  ixpk= 0, nxpk= 0, iat= 271, 277, r1= 3.07, r2= 3.57, r3= 4.57, r4= 5.07,  &end
#
# 9       ADE         H3'         9           ADE             H8              2.62    4.22
&rst
  ixpk= 0, nxpk= 0, iat= 293, 280, r1= 2.12, r2= 2.62, r3= 4.22, r4= 4.72,  &end
#
# 10      ADE         H8          9           ADE             H1'             3.86    5.22
&rst
  ixpk= 0, nxpk= 0, iat= 312, 277, r1= 3.36, r2= 3.86, r3= 5.22, r4= 5.72,  &end
#
# 10      ADE         H2"         10          ADE             H1'             2.14    2.9
&rst
  ixpk= 0, nxpk= 0, iat= 328, 309, r1= 1.64, r2= 2.14, r3= 2.90, r4= 3.40,  &end
#
# 10      ADE         H5'         10          ADE             H1'             3.4     4.48
&rst
  ixpk= 0, nxpk= 0, iat= 303, 309, r1= 2.90, r2= 3.40, r3= 4.48, r4= 4.98,  &end
#
# 10      ADE         H8          10          ADE             H1'             3               4.06
&rst
  ixpk= 0, nxpk= 0, iat= 312, 309, r1= 2.50, r2= 3.00, r3= 4.06, r4= 4.56,  &end
#
# 10      ADE         H2'         10          ADE             H8              2.09    2.83
&rst
  ixpk= 0, nxpk= 0, iat= 327, 312, r1= 1.59, r2= 2.09, r3= 2.83, r4= 3.33,  &end
#
```

```
# 10     ADE      H2"      10       ADE      H8          1.96   2.76
&rst
  ixpk= 0, nxpk= 0, iat= 328, 312, r1= 1.46, r2= 1.96, r3= 2.76, r4= 3.26,  &end
#
# 11     GUA      H2'      11       GUA      H1'         2.17   3.03
&rst
  ixpk= 0, nxpk= 0, iat= 360, 341, r1= 1.67, r2= 2.17, r3= 3.03, r4= 3.53,  &end
#
# 11     GUA      H3'      11       GUA      H2'         2.42   4.44
&rst
  ixpk= 0, nxpk= 0, iat= 358, 360, r1= 1.92, r2= 2.42, r3= 4.44, r4= 4.94,  &end
#
# 11     GUA      H3'      11       GUA      H2"         2.41   3.25
&rst
  ixpk= 0, nxpk= 0, iat= 358, 361, r1= 1.91, r2= 2.41, r3= 3.25, r4= 3.75,  &end
#
# 11     GUA      H2'      11       GUA      H8          2.12   3.14
&rst
  ixpk= 0, nxpk= 0, iat= 360, 344, r1= 1.62, r2= 2.12, r3= 3.14, r4= 3.64,  &end
#
# 12     CYT      H2'      12       CYT      H1'         3.02   4.22
&rst
  ixpk= 0, nxpk= 0, iat= 389, 373, r1= 2.52, r2= 3.02, r3= 4.22, r4= 4.72,  &end
#
# 12     CYT      H5"      12       CYT      H1'         3.72   5.02
&rst
  ixpk= 0, nxpk= 0, iat= 368, 373, r1= 3.22, r2= 3.72, r3= 5.02, r4= 5.52,  &end
#
# 12     CYT      H5"      12       CYT      H4'         2.06   2.78
&rst
  ixpk= 0, nxpk= 0, iat= 368, 370, r1= 1.56, r2= 2.06, r3= 2.78, r4= 3.28,  &end
#
# 12     CYT      H5       12       CYT      H6          1.77   2.49
&rst
  ixpk= 0, nxpk= 0, iat= 378, 376, r1= 1.27, r2= 1.77, r3= 2.49, r4= 2.99,  &end
#
# 13     THY      H5"      13       THY      H1'         3.95   4.83 change
&rst
  ixpk= 0, nxpk= 0, iat= 398, 403, r1= 3.45, r2= 3.95, r3= 4.98, r4= 5.33,  &end
#
# 13     THY      H6       13       THY      H1'         3          4.4
&rst
  ixpk= 0, nxpk= 0, iat= 406, 403, r1= 2.50, r2= 3.00, r3= 4.40, r4= 4.90,  &end
#
# 14     THY      H2'      13       THY      H1'         3.42   4.62
```

```
&rst
  ixpk= 0, nxpk= 0, iat= 453, 403, r1= 2.92, r2= 3.42, r3= 4.62, r4= 5.12,  &end
#
# 14      THY         H6          13          THY             H1'             3.2     4.42
&rst
  ixpk= 0, nxpk= 0, iat= 438, 403, r1= 2.70, r2= 3.20, r3= 4.42, r4= 4.92,  &end
#
# 14      THY         H3'         14          THY             H1'             3.05    4.13
&rst
  ixpk= 0, nxpk= 0, iat= 451, 435, r1= 2.55, r2= 3.05, r3= 4.13, r4= 4.63,  &end
#
# 14      THY         H6          14          THY             H1'             3.7     5
&rst
  ixpk= 0, nxpk= 0, iat= 438, 435, r1= 3.20, r2= 3.70, r3= 5.00, r4= 5.50,  &end
#
# 15      CYT         H6          14          THY             H1'             3.14    4.24
&rst
  ixpk= 0, nxpk= 0, iat= 470, 435, r1= 2.64, r2= 3.14, r3= 4.24, r4= 4.74,  &end
#
# 15      CYT         H6          15          CYT             H1'             3.11    4.21
&rst
  ixpk= 0, nxpk= 0, iat= 470, 467, r1= 2.61, r2= 3.11, r3= 4.21, r4= 4.71,  &end
#
# 16      THY         H6          15          CYT             H1'             3.62    4.46
&rst
  ixpk= 0, nxpk= 0, iat= 500, 467, r1= 3.12, r2= 3.62, r3= 4.46, r4= 4.96,  &end
#
# 16      THY         H6          16          THY             H1'             3.55    4.81
&rst
  ixpk= 0, nxpk= 0, iat= 500, 497, r1= 3.05, r2= 3.55, r3= 4.81, r4= 5.31,  &end
#
# 18      GUA         H3'         18          GUA             H1'             3.2     3.98
&rst
  ixpk= 0, nxpk= 0, iat= 578, 561, r1= 2.70, r2= 3.20, r3= 3.98, r4= 4.48,  &end
#
# 18      GUA         H8          18          GUA             H1'             2.52    3.62 change
&rst
  ixpk= 0, nxpk= 0, iat= 564, 561, r1= 2.02, r2= 2.52, r3= 3.79, r4= 4.12,  &end
#
# 19      THY         H2'         19          THY             H1'             2.02    3.04
&rst
  ixpk= 0, nxpk= 0, iat= 612, 594, r1= 1.52, r2= 2.02, r3= 3.04, r4= 3.54,  &end
#
# 20      CYT         H6          19          THY             H1'             2.41    4.44
&rst
```

```
  ixpk= 0, nxpk= 0, iat= 629, 594, r1= 1.91, r2= 2.41, r3= 4.44, r4= 4.94,  &end
#
# 20      CYT         H2"          20          CYT              H1'              2.02    3.08
&rst
  ixpk= 0, nxpk= 0, iat= 643, 626, r1= 1.52, r2= 2.02, r3= 3.08, r4= 3.58,  &end
#
# 20      CYT         H6           20          CYT              H1'              3.27    4.43
&rst
  ixpk= 0, nxpk= 0, iat= 629, 626, r1= 2.77, r2= 3.27, r3= 4.43, r4= 4.93,  &end
#
# 20      CYT         H2'          20          CYT              H5               3.33    4.25
&rst
  ixpk= 0, nxpk= 0, iat= 642, 631, r1= 2.83, r2= 3.33, r3= 4.25, r4= 4.75,  &end
#
# 20      CYT         H3'          20          CYT              H6               3.47    5.05
&rst
  ixpk= 0, nxpk= 0, iat= 640, 629, r1= 2.97, r2= 3.47, r3= 5.05, r4= 5.55,  &end
#
# 20      CYT         H5           20          CYT              H6               2.12    2.88
&rst
  ixpk= 0, nxpk= 0, iat= 631, 629, r1= 1.62, r2= 2.12, r3= 2.88, r4= 3.38,  &end
#
# 20      CYT         H5'          20          CYT              H6               2.93    4.19
&rst
  ixpk= 0, nxpk= 0, iat= 620, 629, r1= 2.43, r2= 2.93, r3= 4.19, r4= 4.69,  &end
#
# 21      CYT         H5'          21          CYT              H1'              3.3     4.32
&rst
  ixpk= 0, nxpk= 0, iat= 650, 656, r1= 2.80, r2= 3.30, r3= 4.32, r4= 4.72,  &end
#
# 21      CYT         H3'          21          CYT              H2"              2.83    4.09
&rst
  ixpk= 0, nxpk= 0, iat= 670, 673, r1= 2.43, r2= 2.82, r3= 4.09, r4= 4.59,  &end
#
# 21      CYT         H3'          21          CYT              H6               3.55    4.79
&rst
  ixpk= 0, nxpk= 0, iat= 670, 659, r1= 3.05, r2= 3.55, r3= 4.79, r4= 5.29,  &end
#
# 21      CYT         H5           21          CYT              H6               1.95    2.65
&rst
  ixpk= 0, nxpk= 0, iat= 661, 659, r1= 1.45, r2= 1.95, r3= 2.65, r4= 3.15,  &end
#
# 22      GUA         H2"          22          GUA              H1'              1.88    2.54
&rst
  ixpk= 0, nxpk= 0, iat= 706, 686, r1= 1.38, r2= 1.88, r3= 2.54, r4= 3.04,  &end
```

```
#
# 22      GUA         H2"          22          GUA              H2'                1.77    2.67
&rst
  ixpk= 0, nxpk= 0, iat= 706, 705, r1= 1.27, r2= 1.77, r3= 2.67, r4= 3.17,  &end
#
# 22      GUA         H3'          22          GUA              H2'                2.43    3.25
&rst
  ixpk= 0, nxpk= 0, iat= 703, 705, r1= 1.93, r2= 2.43, r3= 3.25, r4= 3.75,  &end
#
# 22      GUA         H3'          22          GUA              H2"                2.83    4.09
&rst
  ixpk= 0, nxpk= 0, iat= 703, 706, r1= 2.33, r2= 2.83, r3= 4.09, r4= 4.59,  &end
#
# 22      GUA         H2'          22          GUA              H8                 2.26    3.2
&rst
  ixpk= 0, nxpk= 0, iat= 705, 689, r1= 1.76, r2= 2.26, r3= 3.20, r4= 3.70,  &end
#
# 22      GUA         H2"          22          GUA              H8                 1.76   3.42
&rst
  ixpk= 0, nxpk= 0, iat= 706, 689, r1= 1.26, r2= 1.76, r3= 3.42, r4= 3.92,  &end
#
# 1       CYT         H2'          1           CYT              H6           2.3    2.95
&rst
  ixpk= 0, nxpk= 0, iat=  26,  13, r1= 1.80, r2= 2.30, r3= 2.95, r4= 3.45,
     rk2=15.0, rk3=15.0, ir6=1, ialtd=0,
&end
#
# 2       GUA         H2'          2           GUA              H1'           3.04   4.24
&rst
  ixpk= 0, nxpk= 0, iat=  59,  40, r1= 2.54, r2= 3.04, r3= 4.24, r4= 4.74,  &end
#
# 2       GUA         H2"          2           GUA              H1'    1.92     2.6
&rst
  ixpk= 0, nxpk= 0, iat=  60,  40, r1= 1.42, r2= 1.92, r3= 2.60, r4= 3.10,  &end
#
# 2       GUA         H3'          2           GUA              H2"    2.48     3.36
&rst
  ixpk= 0, nxpk= 0, iat=  57,  60, r1= 1.98, r2= 2.48, r3= 3.36, r4= 3.86,  &end
#
# 2       GUA         H5"          2           GUA              H2"    3.44     4.66  change
&rst
  ixpk= 0, nxpk= 0, iat=  35,  60, r1= 2.94, r2= 3.44, r3= 4.80, r4= 5.16,  &end
#
# 2       GUA         H2'          2           GUA              H8     2.04    2.76
&rst
```

```
  ixpk= 0, nxpk= 0, iat=  59,  43, r1= 1.54, r2= 2.04, r3= 2.76, r4= 3.26,  &end
#
# 2       GUA         H2"          2              GUA              H8     3.11    4.21
 &rst
  ixpk= 0, nxpk= 0, iat=  60,  43, r1= 2.61, r2= 3.11, r3= 4.21, r4= 4.71,  &end
#
# 3       GUA         H2"          3              GUA              H5'    3.94    5.18
 &rst
  ixpk= 0, nxpk= 0, iat=  93,  67, r1= 3.44, r2= 3.94, r3= 5.18, r4= 5.68,  &end
#
# 3       GUA         H2"          3              GUA              H8     2.35    3.27
&rst
  ixpk= 0, nxpk= 0, iat=  93,  76, r1= 1.85, r2= 2.35, r3= 3.27, r4= 3.77,  &end
#
# 4       ADE         H2'          4              ADE               H8     2.12    3
&rst
  ixpk= 0, nxpk= 0, iat= 124, 109, r1= 1.62, r2= 2.12, r3= 3.00, r4= 3.50,  &end
#
# 4       ADE         H2"          4              ADE              H8     2.39    3.63
&rst
  ixpk= 0, nxpk= 0, iat= 125, 109, r1= 1.89, r2= 2.39, r3= 3.63, r4= 4.13,  &end
#
# 4       ADE         H3'          4              ADE              H8     3.2     4.32
&rst
  ixpk= 0, nxpk= 0, iat= 122, 109, r1= 2.70, r2= 3.20, r3= 4.32, r4= 4.82,  &end
#
# 5       CYT         H5           4              ADE              H2"    3.31    4.49
&rst
  ixpk= 0, nxpk= 0, iat= 143, 125, r1= 2.81, r2= 3.31, r3= 4.49, r4= 4.99,  &end
#
# 6       ABX         1H1X         5              CYT              H5     3.02    4.22
&rst
  ixpk= 0, nxpk= 0, iat= 181, 143, r1= 2.52, r2= 3.02, r3= 4.22, r4= 4.72,  &end
#
# 6       ABX         H5"          6              ABX              H2'    3.55    4.45
&rst
  ixpk= 0, nxpk= 0, iat= 163, 170, r1= 3.05, r2= 3.55, r3= 4.45, r4= 4.95,  &end
#
# 6       ABX         H3'          6              ABX              H2"    2.53    3.43
&rst
  ixpk= 0, nxpk= 0, iat= 200, 171, r1= 2.03, r2= 2.53, r3= 3.43, r4= 3.93,  &end
#
# 6       ABX         2H4X         6              ABX              H3X    2.1     3.45
&rst
  ixpk= 0, nxpk= 0, iat= 189, 186, r1= 1.60, r2= 2.10, r3= 3.45, r4= 3.95,  &end
```

```
#
# 7       ADE         H8          6           ABX             H2'    3.32    6.14
&rst
  ixpk= 0, nxpk= 0, iat= 216, 170, r1= 2.82, r2= 3.32, r3= 6.14, r4= 6.64,  &end
#
# 7       ADE         H1'         7           ADE             H8     3.05    3.85  change
&rst
  ixpk= 0, nxpk= 0, iat= 213, 216, r1= 2.55, r2= 3.05, r3= 4.50, r4= 5.00,  &end
#
# 7       ADE         H2'         7           ADE             H8     1.81    2.45
&rst
  ixpk= 0, nxpk= 0, iat= 230, 216, r1= 1.31, r2= 1.81, r3= 2.45, r4= 2.95,  &end
#
# 8       GUA         H3'         8           GUA             H2'    2.3           3.12
&rst
  ixpk= 0, nxpk= 0, iat= 261, 263, r1= 1.80, r2= 2.30, r3= 3.12, r4= 3.62,  &end
#
# 8       GUA         H2"         8           GUA             H8     3.12    3.86
&rst
  ixpk= 0, nxpk= 0, iat= 264, 247, r1= 2.62, r2= 3.12, r3= 3.86, r4= 4.36,  &end
#
# 9       ADE         H5'         9           ADE             H2'    2.72    3.98
&rst
  ixpk= 0, nxpk= 0, iat= 271, 295, r1= 2.22, r2= 2.72, r3= 3.98, r4= 4.48,  &end
#
# 9       ADE         H5"         9           ADE             H2"    3.74    4.88
&rst
  ixpk= 0, nxpk= 0, iat= 272, 296, r1= 3.24, r2= 3.74, r3= 4.88, r4= 5.38,  &end
#
# 9       ADE         H5'         9           ADE             H8     3.73    5.05
&rst
  ixpk= 0, nxpk= 0, iat= 271, 280, r1= 3.23, r2= 3.73, r3= 5.05, r4= 5.55,  &end
#
# 10      ADE         H5"         10          ADE             H2'    3.32    4.48
&rst
  ixpk= 0, nxpk= 0, iat= 304, 327, r1= 2.82, r2= 3.32, r3= 4.48, r4= 4.98,  &end
#
# 10      ADE         H5'         10          ADE             H2"    4.07    5.15
&rst
  ixpk= 0, nxpk= 0, iat= 303, 328, r1= 3.57, r2= 4.07, r3= 5.15, r4= 5.65,  &end
#
# 10      ADE         H5"         10          ADE             H2"    4             4.87
&rst
  ixpk= 0, nxpk= 0, iat= 304, 328, r1= 3.50, r2= 4.00, r3= 4.87, r4= 5.37,  &end
#
```

```
# 11      GUA         H8            10            ADE               H1'    2.51    3.39
&rst
  ixpk= 0, nxpk= 0, iat= 344, 309, r1= 2.01, r2= 2.51, r3= 3.39, r4= 3.89,  &end
#
# 11      GUA         H8            10            ADE               H2'    3.11    4.41
&rst
  ixpk= 0, nxpk= 0, iat= 344, 327, r1= 2.61, r2= 3.11, r3= 4.41, r4= 4.91,  &end
#
# 11      GUA         H2"           11            GUA               H1'    2.29    3.09
&rst
  ixpk= 0, nxpk= 0, iat= 361, 341, r1= 1.79, r2= 2.29, r3= 3.09, r4= 3.59,  &end
#
# 11      GUA         H8            11            GUA               H1'    3.2          4.32
&rst
  ixpk= 0, nxpk= 0, iat= 344, 341, r1= 2.70, r2= 3.20, r3= 4.32, r4= 4.82,  &end
#
# 11      GUA         H5'           11            GUA               H2'    2.93    3.97
&rst
  ixpk= 0, nxpk= 0, iat= 335, 360, r1= 2.43, r2= 2.93, r3= 3.97, r4= 4.47,  &end
#
# 11      GUA         H5'           11            GUA               H2"    2.79     4.69  change
&rst
  ixpk= 0, nxpk= 0, iat= 335, 361, r1= 2.29, r2= 2.79, r3= 4.79, r4= 5.19,  &end
#
# 11     GUA         H4'           11            GUA               H3'    2.89    4.05
&rst
  ixpk= 0, nxpk= 0, iat= 338, 358, r1= 2.39, r2= 2.79, r3= 4.05, r4= 4.55,  &end
#
# 11     GUA         H5'           11            GUA               H3'    2.87    3.89
&rst
  ixpk= 0, nxpk= 0, iat= 335, 358, r1= 2.37, r2= 2.87, r3= 3.89, r4= 4.39,  &end
#
# 11     GUA         H5"           11            GUA               H3'    3.3          4.46
&rst
  ixpk= 0, nxpk= 0, iat= 336, 358, r1= 2.80, r2= 3.18, r3= 4.46, r4= 4.96,  &end
#
# 11     GUA         H2"           11            GUA               H8     2.32    3.14  change
&rst
  ixpk= 0, nxpk= 0, iat= 361, 344, r1= 1.82, r2= 2.32, r3= 3.24, r4= 3.64,  &end
#
# 12     CYT         H3'           12            CYT               H2'    1.98    2.68
&rst
  ixpk= 0, nxpk= 0, iat= 387, 389, r1= 1.48, r2= 1.98, r3= 2.68, r4= 3.18,  &end
#
# 12     CYT         H3'           12            CYT               H2"    2.45    3.67
```

```
&rst
  ixpk= 0, nxpk= 0, iat= 387, 390, r1= 1.95, r2= 2.45, r3= 3.67, r4= 4.17,  &end
#
# 12     CYT          H3'            12             CYT                 H4'    2.72     3.68
&rst
  ixpk= 0, nxpk= 0, iat= 387, 370, r1= 2.22, r2= 2.72, r3= 3.68, r4= 4.18,  &end
#
# 12     CYT          H2'            12             CYT                 H6     2.43     3.23
&rst
  ixpk= 0, nxpk= 0, iat= 389, 376, r1= 1.93, r2= 2.43, r3= 3.23, r4= 3.73,  &end
#
# 13     THY          H7             12             CYT                 H2'    2.35      3.83
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 389, r1= 1.85, r2= 2.35, r3= 3.83, r4= 4.33,
 igr1= 409, 410, 411,
&end
#
# 13     THY          H2'            13             THY                 H7     2.01     3.68 change
 &rst
  ixpk= 0, nxpk= 0, iat= 421,  -1, r1= 1.51, r2= 2.01, r3= 4.08, r4= 4.18,
 igr2= 409, 410, 411,
&end
#
# 13     THY          H2'            13             THY                 H6     1.56     3.52
&rst
  ixpk= 0, nxpk= 0, iat= 421, 406, r1= 1.06, r2= 1.56, r3= 3.52, r4= 4.02,  &end
#
# 13     THY          H2"            13             THY                 H6     1.96     2.86 change
&rst
  ixpk= 0, nxpk= 0, iat= 422, 406, r1= 1.46, r2= 1.96, r3= 3.27, r4= 3.36,  &end
#
# 13     THY          H7             13             THY                 H6     2.12     3.43
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 406, r1= 1.62, r2= 2.12, r3= 3.43, r4= 3.93,
 igr1= 409, 410, 411,
&end
#
# 14     THY          H6             13             THY                 H7     3.1      5.62
&rst
  ixpk= 0, nxpk= 0, iat= 438,  -1, r1= 2.60, r2= 3.10, r3= 5.62, r4= 6.12,
 igr2= 409, 410, 411,
&end
#
# 14     THY          H7             13             THY                 H6     3.04     4.95
&rst
```

```
  ixpk= 0, nxpk= 0, iat=  -1, 406, r1= 2.54, r2= 3.04, r3= 4.95, r4= 5.45,
 igr1= 441, 442, 443,
&end
#
# 14     THY          H3'            14             THY                 H2'    2.35      3.31
&rst
  ixpk= 0, nxpk= 0, iat= 451, 453, r1= 1.85, r2= 2.35, r3= 3.31, r4= 3.81,  &end
#
# 14     THY          H3'            14             THY                 H6     3.29      4.19
&rst
  ixpk= 0, nxpk= 0, iat= 451, 438, r1= 2.79, r2= 3.29, r3= 4.19, r4= 4.69,  &end
#
# 14     THY          H4'            14             THY                 H6     4.19      5.31
&rst
  ixpk= 0, nxpk= 0, iat= 432, 438, r1= 3.69, r2= 4.19, r3= 5.31, r4= 5.81,  &end
#
# 14     THY          H5"            14             THY                 H6     4.24      5.58
&rst
  ixpk= 0, nxpk= 0, iat= 430, 438, r1= 3.74, r2= 4.24, r3= 5.58, r4= 6.08,  &end
#
# 15     CYT          H2'            15             CYT                 H6     1.77      2.39
&rst
  ixpk= 0, nxpk= 0, iat= 483, 470, r1= 1.27, r2= 1.77, r3= 2.39, r4= 2.89,  &end
#
# 15     CYT          H2"            15             CYT                 H6     2.27      3.37 change
&rst
  ixpk= 0, nxpk= 0, iat= 484, 470, r1= 1.77, r2= 2.27, r3= 3.57, r4= 3.87,  &end
#
# 16     THY          H2'            16             THY                 H1'    2.65      3.59
&rst
  ixpk= 0, nxpk= 0, iat= 515, 497, r1= 2.15, r2= 2.65, r3= 3.59, r4= 4.09,  &end
#
# 16     THY          H2"            16             THY                 H2'    1.79      2.59
&rst
  ixpk= 0, nxpk= 0, iat= 516, 515, r1= 1.29, r2= 1.79, r3= 2.59, r4= 3.09,  &end
#
# 16     THY          H2'            16             THY                 H6     1.79      2.43
&rst
  ixpk= 0, nxpk= 0, iat= 515, 500, r1= 1.29, r2= 1.79, r3= 2.43, r4= 2.93,  &end
#
# 17     THY          H7             16             THY                 H2'     2.03     4.14
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 515, r1= 1.53, r2= 2.03, r3= 4.14, r4= 4.64,
 igr1= 535, 536, 537,
&end
```

```
#
# 17     THY          H7             16             THY                 H6      3.34      5
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 500, r1= 2.84, r2= 3.34, r3= 5.00, r4= 5.50,
 igr1= 535, 536, 537,
&end
#
# 17     THY          H2'            17             THY                 H2"     1.65      2.23
&rst
  ixpk= 0, nxpk= 0, iat= 547, 548, r1= 1.15, r2= 1.65, r3= 2.23, r4= 2.73,  &end
#
# 17     THY          H2'            18             GUA                 H8      3.17     4.49
&rst
  ixpk= 0, nxpk= 0, iat= 547, 564, r1= 2.67, r2= 3.17, r3= 4.49, r4= 4.99,  &end
#
# 18     GUA          H5'            18             GUA                 H2"     3.44      5.24
&rst
  ixpk= 0, nxpk= 0, iat= 555, 581, r1= 2.94, r2= 3.44, r3= 5.24, r4= 5.74,  &end
#
# 18     GUA          H5"            18             GUA                 H3'     2.28      3.08
&rst
  ixpk= 0, nxpk= 0, iat= 556, 578, r1= 1.78, r2= 2.28, r3= 3.08, r4= 3.58,  &end
#
# 18     GUA          H2'            18             GUA                 H8      1.88      2.9
&rst
  ixpk= 0, nxpk= 0, iat= 580, 564, r1= 1.38, r2= 1.88, r3= 2.90, r4= 3.40,  &end
#
# 18     GUA          H2"            18             GUA                 H8      2.39      3.33
&rst
  ixpk= 0, nxpk= 0, iat= 581, 564, r1= 1.89, r2= 2.39, r3= 3.33, r4= 3.83,  &end
#
# 19     THY          H3'            19             THY                 H2'     2.13      2.89
&rst
  ixpk= 0, nxpk= 0, iat= 610, 612, r1= 1.63, r2= 2.13, r3= 2.89, r4= 3.39,  &end
#
# 19     THY          H3'            19             THY                 H2"     2.44       3.3
&rst
  ixpk= 0, nxpk= 0, iat= 610, 613, r1= 1.94, r2= 2.44, r3= 3.30, r4= 3.80,  &end
#
# 20     CYT          H5             19             THY                 H2"     2.71       3.67
&rst
  ixpk= 0, nxpk= 0, iat= 631, 613, r1= 2.21, r2= 2.71, r3= 3.67, r4= 4.17,  &end
#
# 20     CYT          H6             19             THY                 H2"     2.99       4.43
&rst
```

```
  ixpk= 0, nxpk= 0, iat= 629, 613, r1= 2.49, r2= 2.99, r3= 4.43, r4= 4.93,  &end
#
# 20     CYT          H2'            20             CYT                H1'    3.07      4.57
&rst
  ixpk= 0, nxpk= 0, iat= 642, 626, r1= 2.57, r2= 3.07, r3= 4.57, r4= 5.07,  &end
#
# 20     CYT          H5'            20             CYT                H2"    2.93      4.53  change
&rst
  ixpk= 0, nxpk= 0, iat= 620, 643, r1= 2.43, r2= 2.93, r3= 4.83, r4= 5.03,  &end
#
# 20     CYT          H2"            20             CYT                H5     3.62      5.2  change
&rst
  ixpk= 0, nxpk= 0, iat= 643, 631, r1= 3.12, r2= 3.62, r3= 5.31, r4= 5.70,  &end
#
# 20     CYT          H2'            20             CYT                H6     1.7       2.5
&rst
  ixpk= 0, nxpk= 0, iat= 642, 629, r1= 1.20, r2= 1.70, r3= 2.50, r4= 3.00,  &end
#
# 20     CYT          H2"            20             CYT                H6     1.89      2.97  change
&rst
  ixpk= 0, nxpk= 0, iat= 643, 629, r1= 1.39, r2= 1.89, r3= 3.17, r4= 3.47,  &end
#
# 21     CYT          H2"            20             CYT                H1'    2.7           4.3
&rst
  ixpk= 0, nxpk= 0, iat= 673, 626, r1= 2.20, r2= 2.70, r3= 4.30, r4= 4.80,  &end
#
# 21     CYT          H6             20             CYT                H1'    3.15      4.27
&rst
  ixpk= 0, nxpk= 0, iat= 659, 626, r1= 2.65, r2= 3.15, r3= 4.27, r4= 4.77,  &end
#
# 21     CYT          H2"            20             CYT                H2'    3.37      4.65
&rst
  ixpk= 0, nxpk= 0, iat= 673, 642, r1= 2.87, r2= 3.37, r3= 4.65, r4= 5.15,  &end
#
# 21     CYT          H6             20             CYT                H2'    1.94      2.82
&rst
  ixpk= 0, nxpk= 0, iat= 659, 642, r1= 1.44, r2= 1.94, r3= 2.82, r4= 3.32,  &end
#
# 21     CYT          H2"            21             CYT                H1'    2.49      3.61
&rst
  ixpk= 0, nxpk= 0, iat= 673, 656, r1= 1.99, r2= 2.49, r3= 3.61, r4= 4.11,  &end
#
# 21     CYT          H6             21             CYT                H1'    2.7           3.76
&rst
  ixpk= 0, nxpk= 0, iat= 659, 656, r1= 2.20, r2= 2.70, r3= 3.76, r4= 4.26,  &end
```

```
#
# 21     CYT          H5"          21           CYT                H2"   3.2          4.96
&rst
  ixpk= 0, nxpk= 0, iat= 651, 673, r1= 2.70, r2= 3.20, r3= 4.96, r4= 5.46,  &end
#
# 21     CYT          H2'          21           CYT                H6     1.73   2.51
&rst
  ixpk= 0, nxpk= 0, iat= 672, 659, r1= 1.23, r2= 1.73, r3= 2.51, r4= 3.01,  &end
#
# 21     CYT          H2"          21           CYT                H6     2.19   3.41
&rst
  ixpk= 0, nxpk= 0, iat= 673, 659, r1= 1.69, r2= 2.19, r3= 3.41, r4= 3.91,  &end
#
# 22     GUA          H5'          22           GUA                H1'    3.38    4.58
&rst
  ixpk= 0, nxpk= 0, iat= 680, 686, r1= 2.88, r2= 3.38, r3= 4.58, r4= 5.08,  &end
#
# 22     GUA          H5'          22           GUA                H3'    2.84    3.7
&rst
  ixpk= 0, nxpk= 0, iat= 680, 703, r1= 2.34, r2= 2.84, r3= 3.70, r4= 4.20,  &end
#
# 22     GUA          H5"          22           GUA                H3'    2.42    3.4
&rst
  ixpk= 0, nxpk= 0, iat= 681, 703, r1= 1.92, r2= 2.42, r3= 3.40, r4= 3.90,  &end
#
# 2      GUA          H3'          2            GUA                H2'                2.3          3.38
&rst
  ixpk= 0, nxpk= 0, iat=  57,  59, r1= 1.80, r2= 2.30, r3= 3.38, r4= 3.88,
     rk2=10.0, rk3=10.0, ir6=1, ialtd=0,
&end
#
# 2      GUA          H5'          2            GUA                H3'                2.27   3.37
&rst
  ixpk= 0, nxpk= 0, iat=  34,  57, r1= 1.77, r2= 2.27, r3= 3.37, r4= 3.87,  &end
#
# 2      GUA          H5'          2            GUA                H4'                1.95   2.63
&rst
  ixpk= 0, nxpk= 0, iat=  34,  37, r1= 1.45, r2= 1.95, r3= 2.63, r4= 3.13,  &end
#
# 2      GUA          H5"          2            GUA                H4'                2.27   3.07
&rst
  ixpk= 0, nxpk= 0, iat=  35,  37, r1= 1.77, r2= 2.27, r3= 3.07, r4= 3.57,  &end
#
# 3      GUA          H5'          2            GUA                H3'                2.66   4.0 change
&rst
```

```
  ixpk= 0, nxpk= 0, iat=  67,  57, r1= 2.16, r2= 2.66, r3= 4.40, r4= 4.50,  &end
#
# 3      GUA         H3'          3              GUA              H2'              2.14    3.02
&rst
  ixpk= 0, nxpk= 0, iat=  90,  92, r1= 1.64, r2= 2.14, r3= 3.02, r4= 3.52,  &end
#
# 3      GUA         H3'          3              GUA              H2"              2.43    3.63
&rst
  ixpk= 0, nxpk= 0, iat=  90,  93, r1= 1.93, r2= 2.43, r3= 3.63, r4= 4.13,  &end
#
# 3      GUA         H3'          3              GUA              H5'              2.9     3.68
&rst
  ixpk= 0, nxpk= 0, iat=  90,  67, r1= 2.40, r2= 2.90, r3= 3.68, r4= 4.18,  &end
#
# 4      ADE         H2'          4              ADE              H1'              2.13    2.97
&rst
  ixpk= 0, nxpk= 0, iat= 124, 106, r1= 1.63, r2= 2.13, r3= 2.97, r4= 3.47,  &end
#
# 4      ADE         H5'          4              ADE              H1'              2.44    4.34
&rst
  ixpk= 0, nxpk= 0, iat= 100, 106, r1= 1.94, r2= 2.44, r3= 4.34, r4= 4.84,  &end
#
# 4      ADE         H2"          4              ADE              H2'              1.8     2.7
&rst
  ixpk= 0, nxpk= 0, iat= 125, 124, r1= 1.30, r2= 1.80, r3= 2.70, r4= 3.20,  &end
#
# 4      ADE         H3'          4              ADE              H2"              2.3     3.46
&rst
  ixpk= 0, nxpk= 0, iat= 122, 125, r1= 1.80, r2= 2.30, r3= 3.46, r4= 3.96,  &end
#
# 4      ADE         H5'          4              ADE              H3'              2.78    3.6
&rst
  ixpk= 0, nxpk= 0, iat= 100, 122, r1= 2.28, r2= 2.78, r3= 3.60, r4= 4.10,  &end
#
# 4      ADE         H5"          4              ADE              H3'              2.56    3.46
&rst
  ixpk= 0, nxpk= 0, iat= 101, 122, r1= 2.06, r2= 2.56, r3= 3.46, r4= 3.96,  &end
#
# 5      CYT         H2"          5              CYT              H2'              1.71    2.59
&rst
  ixpk= 0, nxpk= 0, iat= 155, 154, r1= 1.21, r2= 1.71, r3= 2.59, r4= 3.09,  &end
#
# 5      CYT         H5'          5              CYT              H2"              4.12    4.96
&rst
  ixpk= 0, nxpk= 0, iat= 132, 155, r1= 3.62, r2= 4.12, r3= 4.96, r4= 5.46,  &end
```

```
#
# 6      ABX         H2'          6          ABX          H1'          2.59   3.51
&rst
  ixpk= 0, nxpk= 0, iat= 170, 168, r1= 2.09, r2= 2.59, r3= 3.51, r4= 4.01,  &end
#
# 6      ABX         H5'          6          ABX          H1'          4.29   5.09
&rst
  ixpk= 0, nxpk= 0, iat= 162, 168, r1= 3.79, r2= 4.29, r3= 5.09, r4= 5.59,  &end
#
# 6      ABX         H3'          6          ABX          H2'          2.36   3.7
&rst
  ixpk= 0, nxpk= 0, iat= 200, 170, r1= 1.86, r2= 2.36, r3= 3.70, r4= 4.20,  &end
#
# 6      ABX         H2'          6          ABX          H8           1.93   2.61
&rst
  ixpk= 0, nxpk= 0, iat= 170, 174, r1= 1.43, r2= 1.93, r3= 2.61, r4= 3.11,  &end
#
# 6      ABX         H2"          6          ABX          H8           3.03   4.11
&rst
  ixpk= 0, nxpk= 0, iat= 171, 174, r1= 2.53, r2= 3.03, r3= 4.11, r4= 4.61,  &end
#
# 8      GUA         H2"          8          GUA          H2'          1.63   2.37
&rst
  ixpk= 0, nxpk= 0, iat= 264, 263, r1= 1.13, r2= 1.63, r3= 2.37, r4= 2.87,  &end
#
# 8      GUA         H5'          8          GUA          H2"          3.75   4.79 change
&rst
  ixpk= 0, nxpk= 0, iat= 238, 264, r1= 3.25, r2= 3.75, r3= 4.99, r4= 5.29,  &end
#
# 8      GUA         H3'          8          GUA          H8           2.87   3.85 change
&rst
  ixpk= 0, nxpk= 0, iat= 261, 247, r1= 2.37, r2= 2.87, r3= 4.15, r4= 4.35,  &end
#
# 9      ADE         H3'          9          ADE          H2'          2.19   2.97
&rst
  ixpk= 0, nxpk= 0, iat= 293, 295, r1= 1.69, r2= 2.19, r3= 2.97, r4= 3.47,  &end
#
# 9      ADE         H3'          9          ADE          H2"          2.62   3.82
&rst
  ixpk= 0, nxpk= 0, iat= 293, 296, r1= 2.12, r2= 2.62, r3= 3.82, r4= 4.32,  &end
#
# 9      ADE         H2'          9          ADE          H8           2.23   3.09
&rst
  ixpk= 0, nxpk= 0, iat= 295, 280, r1= 1.73, r2= 2.23, r3= 3.09, r4= 3.59,  &end
#
```

```
# 9      ADE        H2"          9          ADE              H8               3.18   3.9
&rst
  ixpk= 0, nxpk= 0, iat= 296, 280, r1= 2.68, r2= 3.18, r3= 3.90, r4= 4.40,  &end
#
# 10     ADE        H2'          10         ADE              H1'              2.26   3.16
&rst
  ixpk= 0, nxpk= 0, iat= 327, 309, r1= 1.76, r2= 2.26, r3= 3.16, r4= 3.66,  &end
#
# 10     ADE        H2"          10         ADE              H2'              1.86   2.8
&rst
  ixpk= 0, nxpk= 0, iat= 328, 327, r1= 1.36, r2= 1.76, r3= 2.80, r4= 3.30,  &end
#
# 10     ADE        H3'          10         ADE              H2'              2.1          2.84
&rst
  ixpk= 0, nxpk= 0, iat= 325, 327, r1= 1.60, r2= 2.10, r3= 2.84, r4= 3.34,  &end
#
# 10     ADE        H3'          10         ADE              H2"              2.08   2.82
&rst
  ixpk= 0, nxpk= 0, iat= 325, 328, r1= 1.58, r2= 2.08, r3= 2.82, r4= 3.32,  &end
#
# 11     GUA        H5'          11         GUA              H1'              3.29   4.29
&rst
  ixpk= 0, nxpk= 0, iat= 335, 341, r1= 2.79, r2= 3.29, r3= 4.29, r4= 4.79,  &end
#
# 11     GUA        H2"          11         GUA              H2'              1.83   2.61
&rst
  ixpk= 0, nxpk= 0, iat= 361, 360, r1= 1.33, r2= 1.83, r3= 2.61, r4= 3.11,  &end
#
# 12     CYT        H6           12         CYT              H1'              3.12   4.22
&rst
  ixpk= 0, nxpk= 0, iat= 376, 373, r1= 2.62, r2= 3.12, r3= 4.22, r4= 4.72,  &end
#
# 12     CYT        H2'          12         CYT              H5               3.34   4.36
&rst
  ixpk= 0, nxpk= 0, iat= 389, 378, r1= 2.84, r2= 3.34, r3= 4.36, r4= 4.86,  &end
#
# 13     THY        H3'          13         THY              H2"              1.97   2.77
&rst
  ixpk= 0, nxpk= 0, iat= 419, 422, r1= 1.47, r2= 1.97, r3= 2.77, r4= 3.27,  &end
#
# 14     THY        H2'          14         THY              H1'              2.99   4.17
&rst
  ixpk= 0, nxpk= 0, iat= 453, 435, r1= 2.49, r2= 2.99, r3= 4.17, r4= 4.67,  &end
#
# 14     THY        H5"          14         THY              H1'              3.53   4.73
```

```
&rst
  ixpk= 0, nxpk= 0, iat= 430, 435, r1= 3.03, r2= 3.53, r3= 4.73, r4= 5.23,  &end
#
# 14     THY          H2"          14           THY                H2'                1.82    2.74
&rst
  ixpk= 0, nxpk= 0, iat= 454, 453, r1= 1.32, r2= 1.82, r3= 2.74, r4= 3.24,  &end
#
# 14     THY          H5'          14           THY                H2"                3.27    4.93 change
&rst
  ixpk= 0, nxpk= 0, iat= 429, 454, r1= 2.77, r2= 3.27, r3= 5.13, r4= 5.43,  &end
#
# 14     THY          H2'          14           THY                H6                 1.84    2.66
&rst
  ixpk= 0, nxpk= 0, iat= 453, 438, r1= 1.34, r2= 1.84, r3= 2.66, r4= 3.16,  &end
#
# 14     THY          H2"          14           THY                H6                 1.88    2.94 change
&rst
  ixpk= 0, nxpk= 0, iat= 454, 438, r1= 1.38, r2= 1.88, r3= 3.34, r4= 3.44,  &end
#
# 14     THY          H7           14           THY                H6                 2.31    3.76
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 438, r1= 1.81, r2= 2.31, r3= 3.76, r4= 4.26,
 igr1= 441, 442, 443,
&end
#
# 15     CYT          H2'          15           CYT                H1'                3.07    4.99
&rst
  ixpk= 0, nxpk= 0, iat= 483, 467, r1= 2.57, r2= 3.07, r3= 4.99, r4= 5.49,  &end
#
# 15     CYT          H5'          15           CYT                H1'                2.48    4.42
&rst
  ixpk= 0, nxpk= 0, iat= 461, 467, r1= 1.98, r2= 2.48, r3= 4.42, r4= 4.92,  &end
#
# 15     CYT          H5"          15           CYT                H1'                3.76    4.7
&rst
  ixpk= 0, nxpk= 0, iat= 462, 467, r1= 3.26, r2= 3.76, r3= 4.70, r4= 5.20,  &end
#
# 16     THY          H6           15           CYT                H5                 4.65    6.05
&rst
  ixpk= 0, nxpk= 0, iat= 500, 472, r1= 4.15, r2= 4.65, r3= 6.05, r4= 6.55,  &end
#
# 16     THY          H2"          16           THY                H1'                1.8             2.44
&rst
  ixpk= 0, nxpk= 0, iat= 516, 497, r1= 1.30, r2= 1.80, r3= 2.44, r4= 2.94,  &end
#
```

```
# 16     THY        H5"          16          THY              H1'              3.5          5.02
&rst
  ixpk= 0, nxpk= 0, iat= 492, 497, r1= 3.00, r2= 3.50, r3= 5.02, r4= 5.52,  &end
#
# 16     THY        H2"          16          THY              H6               2.17    3.33 change
&rst
  ixpk= 0, nxpk= 0, iat= 516, 500, r1= 1.67, r2= 2.17, r3= 3.53, r4= 3.83,  &end
#
# 16     THY        H7           16          THY              H6               2.7          4.54
&rst
  ixpk= 0, nxpk= 0, iat=  -1, 500, r1= 2.20, r2= 2.70, r3= 4.54, r4= 5.04,
 igr1= 503, 504, 505,
&end
#
# 18     GUA        H2'          18          GUA              H1'              2.13    2.97
&rst
  ixpk= 0, nxpk= 0, iat= 580, 561, r1= 1.63, r2= 2.13, r3= 2.97, r4= 3.47,  &end
#
# 18     GUA        H3'          18          GUA              H2'              2.39    3.23
&rst
  ixpk= 0, nxpk= 0, iat= 578, 580, r1= 1.89, r2= 2.39, r3= 3.23, r4= 3.73,  &end
#
# 18     GUA        H5'          18          GUA              H2'              1.95    3.75
&rst
  ixpk= 0, nxpk= 0, iat= 555, 580, r1= 1.45, r2= 1.95, r3= 3.75, r4= 4.25,  &end
#
# 19     THY        H2"          19          THY              H1'              2.31    3.13
&rst
  ixpk= 0, nxpk= 0, iat= 613, 594, r1= 1.81, r2= 2.31, r3= 3.13, r4= 3.63,  &end
#
# 19     THY        H5"          19          THY              H1'              3.22    4.72
&rst
  ixpk= 0, nxpk= 0, iat= 589, 594, r1= 2.72, r2= 3.22, r3= 4.72, r4= 5.22,  &end
#
# 19     THY        H2"          19          THY              H2'              1.89    2.69
&rst
  ixpk= 0, nxpk= 0, iat= 613, 612, r1= 1.39, r2= 1.75, r3= 2.69, r4= 3.19,  &end
#
# 19     THY        H4'          19          THY              H2"              2.74    3.92
&rst
  ixpk= 0, nxpk= 0, iat= 591, 613, r1= 2.24, r2= 2.74, r3= 3.92, r4= 4.42,  &end
#
# 20     CYT        H5'          20          CYT              H1'              2.31    4.15 change
&rst
  ixpk= 0, nxpk= 0, iat= 620, 626, r1= 1.81, r2= 2.31, r3= 4.41, r4= 4.65,  &end
```

```
#
# 20     CYT         H5'          20          CYT              H2'              2.55    3.55
&rst
  ixpk= 0, nxpk= 0, iat= 620, 642, r1= 2.05, r2= 2.55, r3= 3.55, r4= 4.05,  &end
#
# 21     CYT         H2'          20          CYT              H2'              1.61    3.39
&rst
  ixpk= 0, nxpk= 0, iat= 672, 642, r1= 1.11, r2= 1.61, r3= 3.39, r4= 3.89,  &end
#
# 21     CYT         H2"          20          CYT              H6               2.86    4.76 change
&rst
  ixpk= 0, nxpk= 0, iat= 673, 629, r1= 2.36, r2= 2.86, r3= 5.36, r4= 5.56,  &end
#
# 21     CYT         H5'          21          CYT              H3'              2.8          3.64
&rst
  ixpk= 0, nxpk= 0, iat= 650, 670, r1= 2.30, r2= 2.80, r3= 3.64, r4= 4.14,  &end
#
# 22     GUA         H2'          22          GUA              H1'              1.81    2.89
&rst
  ixpk= 0, nxpk= 0, iat= 705, 686, r1= 1.31, r2= 1.81, r3= 2.89, r4= 3.39,  &end
#
# 22     GUA         H5"          22          GUA              H2'              2.52    3.52
&rst
  ixpk= 0, nxpk= 0, iat= 681, 705, r1= 2.02, r2= 2.52, r3= 3.52, r4= 4.02,  &end
#
# 19     THY         H3           3           GUA              H1               3.3          3.7
&rst
  ixpk= 0, nxpk= 0, iat= 606,  82, r1= 2.80, r2= 3.30, r3= 3.70, r4= 4.20,
     rk2=20.0, rk3=20.0, ir6=1, ialtd=0,
&end
#
# 14     THY         H3           8           GUA              H1               3.5          3.9
&rst
  ixpk= 0, nxpk= 0, iat= 447, 253, r1= 3.00, r2= 3.50, r3= 3.90, r4= 4.40,  &end
#
# 16     THY         H3           8           GUA              H1               3.5          3.9
&rst
  ixpk= 0, nxpk= 0, iat= 509, 253, r1= 3.00, r2= 3.50, r3= 3.90, r4= 4.40,  &end
#
# 13     THY         H3           14          THY              H3               3.5          3.9
&rst
  ixpk= 0, nxpk= 0, iat= 415, 447, r1= 3.00, r2= 3.50, r3= 3.90, r4= 4.40,  &end
#
# 3      GUA         H21          3           GUA              H1               2            3.2
&rst
```

```
  ixpk= 0, nxpk= 0, iat=  85,  82, r1= 1.50, r2= 2.00, r3= 3.20, r4= 3.70,  &end
#
# 3      GUA        H21        19         THY        H3         3.5        4.4
&rst
  ixpk= 0, nxpk= 0, iat=  85, 606, r1= 3.00, r2= 3.50, r3= 4.40, r4= 4.90,  &end
#
# 4      ADE        H2         19         THY        H3         2          2.4
&rst
  ixpk= 0, nxpk= 0, iat= 118, 606, r1= 1.50, r2= 2.00, r3= 2.40, r4= 2.90,  &end
#
# 18     GUA        H22        18         GUA        H1         3.3        3.7
&rst
  ixpk= 0, nxpk= 0, iat= 574, 570, r1= 2.80, r2= 3.30, r3= 3.70, r4= 4.20,  &end
#
# 3      GUA        H22        3          GUA        H1         2          2.4
&rst
  ixpk= 0, nxpk= 0, iat=  86,  82, r1= 1.50, r2= 2.00, r3= 2.40, r4= 2.90,  &end
#
# 6      ABX        H2         18         GUA        H1         3.5        4
&rst
  ixpk= 0, nxpk= 0, iat= 196, 570, r1= 3.00, r2= 3.50, r3= 4.00, r4= 4.40,  &end
#
# 8      GUA        H22        8          GUA        H1         3.3        3.7
&rst
  ixpk= 0, nxpk= 0, iat= 257, 253, r1= 2.80, r2= 3.30, r3= 3.70, r4= 4.20,  &end
#
# 9      ADE        H2         8          GUA        H1         3.5        3.9
&rst
  ixpk= 0, nxpk= 0, iat= 289, 253, r1= 3.00, r2= 3.50, r3= 3.90, r4= 4.40,  &end
#
# 4      ADE        H2         18         GUA        H1         3.5        3.9
&rst
  ixpk= 0, nxpk= 0, iat= 118, 570, r1= 3.00, r2= 3.50, r3= 3.90, r4= 4.40,  &end
#
# 18     GUA        H21        18         GUA        H1         2          2.4
&rst
  ixpk= 0, nxpk= 0, iat= 573, 570, r1= 1.50, r2= 2.00, r3= 2.40, r4= 2.90,  &end
#
# 8      GUA        H21        8          GUA        H1         2          2.4
&rst
  ixpk= 0, nxpk= 0, iat= 256, 253, r1= 1.50, r2= 2.00, r3= 2.40, r4= 2.90,  &end
#
# 2      GUA        H21        3          GUA        H1         3.5        3.9
&rst
  ixpk= 0, nxpk= 0, iat=  52,  82, r1= 3.00, r2= 3.50, r3= 3.90, r4= 4.40,  &end
```

```
#
# 2      GUA        H21        2          GUA        H1         2          2.4
&rst
  ixpk= 0, nxpk= 0, iat=  52,  49, r1= 1.50, r2= 2.00, r3= 2.40, r4= 2.90,  &end
#
# 4      ADE        H2         3          GUA        H1         3.5        4
&rst
  ixpk= 0, nxpk= 0, iat= 118,  82, r1= 3.00, r2= 3.50, r3= 4.00, r4= 4.50,  &end
#
# 7      ADE        H2         17         THY        H3         3.5        4.5
&rst
  ixpk= 0, nxpk= 0, iat= 224, 541, r1= 3.00, r2= 3.5.0, r3= 4.50, r4= 5.00,  &end
#
# 7      ADE        H2         16         THY        H3         4.5        5.0
&rst
  ixpk= 0, nxpk= 0, iat= 224, 509, r1= 4.00, r2= 4.50, r3= 5.00, r4= 5.50,  &end
#
# 2      GUA        H22        2          GUA        H1         2          3.2
&rst
  ixpk= 0, nxpk= 0, iat=  53,  49, r1= 1.50, r2= 2.00, r3= 3.20, r4= 3.70,  &end
#
# 6      ABX        H2         17         THY        H3         3          3.5
&rst
  ixpk= 0, nxpk= 0, iat= 196, 541, r1= 2.50, r2= 3.00, r3= 3.50, r4= 4.00,  &end
#
# 3      GUA        H22        19         THY        H3         3.5        3.9
&rst
  ixpk= 0, nxpk= 0, iat=  86, 606, r1= 3.00, r2= 3.50, r3= 3.90, r4= 4.40,  &end
#
# 3      GUA        H22        2          GUA        H1         4          5
&rst
  ixpk= 0, nxpk= 0, iat=  86,  49, r1= 3.50, r2= 4.00, r3= 5.00, r4= 5.50,  &end
#
# 9      ADE        H2         14         THY        H3         2          2.5
&rst
  ixpk= 0, nxpk= 0, iat= 289, 447, r1= 1.50, r2= 2.00, r3= 2.50, r4= 3.00,  &end
#
# 10     ADE        H2         14         THY        H3         3.3        3.7
&rst
  ixpk= 0, nxpk= 0, iat= 321, 447, r1= 2.80, r2= 3.30, r3= 3.70, r4= 4.20,  &end
#
# 10     ADE        H2         13         THY        H3         2          2.4
&rst
  ixpk= 0, nxpk= 0, iat= 321, 415, r1= 1.50, r2= 2.00, r3= 2.40, r4= 2.90,  &end
#
```

```
# 18     GUA        H21          19           THY          H3                3.5          3.9
&rst
  ixpk= 0, nxpk= 0, iat= 573, 606, r1= 3.00, r2= 3.50, r3= 3.90, r4= 4.40,  &end
#
# 19     THY        H3           19           THY          H7                  3.5   3.9 change
&rst
  ixpk= 0, nxpk= 0, iat= 606,  -1, r1= 3.00, r2= 3.50, r3= 4.88, r4= 5.4,
 igr2= 600, 601, 602,
&end
#
# 14     THY        H3           14           THY          H7                  3.4   3.8 change
&rst
  ixpk= 0, nxpk= 0, iat= 447,  -1, r1= 2.90, r2= 3.40, r3= 4.76, r4= 5.30,
 igr2= 441, 442, 443,
&end
#
# 13     THY        H3           14           THY          H7                  3.5   3.9 change
&rst
  ixpk= 0, nxpk= 0, iat= 415,  -1, r1= 3.00, r2= 3.50, r3= 4.68, r4= 5.20,
 igr2= 441, 442, 443,
&end
#
# 17     THY        H3           6            ABX          1H1X              3.5          3.9
&rst
  ixpk= 0, nxpk= 0, iat= 541, 181, r1= 3.00, r2= 3.50, r3= 3.90, r4= 4.40,  &end
#
# 17     THY        H3           6            ABX          2H1X              3.5          3.9
&rst
  ixpk= 0, nxpk= 0, iat= 541, 182, r1= 3.00, r2= 3.50, r3= 3.90, r4= 4.40,  &end
#
# 17     THY        H3           6            ABX          H2X               3.5          4
&rst
  ixpk= 0, nxpk= 0, iat= 541, 184, r1= 3.00, r2= 3.50, r3= 4.00, r4= 4.50,  &end
#
# 3      GUA        H1           19           THY          H7                  3.5   3.9 change
&rst
  ixpk= 0, nxpk= 0, iat=  82,  -1, r1= 3.00, r2= 3.50, r3= 4.88, r4= 5.40,
 igr2= 600, 601, 602,
&end
#
# 8      GUA        H1           14           THY          H7                  3.5   3.9 change
&rst
  ixpk= 0, nxpk= 0, iat= 253,  -1, r1= 3.00, r2= 3.50, r3= 4.88, r4= 5.40,
 igr2= 441, 442, 443,
&end
```

```
#
# 18     GUA         H1          6           ABX             1H1X          4          5.1
&rst
  ixpk= 0, nxpk= 0, iat= 570, 181, r1= 3.50, r2= 4.00, r3= 5.10, r4= 5.60,  &end

#
# 17     THY         H6          18          GUA             H1'           4.5        5.8
&rst
  ixpk= 0, nxpk= 0, iat= 532, 561, r1= 4.00, r2= 4.50, r3= 5.80, r4= 6.30,
     rk2=20.0, rk3=20.0, ir6=1, ialtd=0,
&end
#
# 16     THY         H6          17          THY             H1'           5.5        6.5
&rst
  ixpk= 0, nxpk= 0, iat= 500, 529, r1= 5.00, r2= 5.50, r3= 6.50, r4= 7.00,  &end
#
# 16     THY         H3          17          THY             H3            5.5        6.5
&rst
  ixpk= 0, nxpk= 0, iat= 509, 541, r1= 5.00, r2= 5.50, r3= 6.50, r4= 7.00,  &end
#
# 17     THY         H3          18          GUA             H1            5.5        6.5
&rst
  ixpk= 0, nxpk= 0, iat= 541, 570, r1= 5.00, r2= 5.50, r3= 6.50, r4= 7.00,  &end
```

APPENDIX D

TORSION ANGLE RESTRAINT FILE

Torsion angle restraints used in the rMD calculations of the (*S*,*S*)-BD-(61,2-3) crosslinked adduct.

```
# 2 GUA ALPHA:  (1 DC5 O3')-(2 DG P)-(2 DG O5')-(2 DG C5') -75.0 -45.0
 &rst    iat =   28,   29,   32,   33,
       r1 = -76.0, r2 = -75.0, r3 = -45.0, r4 = -44.0,
       rk2 =  60.0, rk3 =  60.0,                                    &end

# 3 GUA ALPHA:  (2 DG O3')-(3 DG P)-(3 DG O5')-(3 DG C5') -75.0 -45.0
 &rst    iat =   61,   62,   65,   66,
       r1 = -76.0, r2 = -75.0, r3 = -45.0, r4 = -44.0,
     &end

# 4 ADE ALPHA:  (3 DG O3')-(4 DA P)-(4 DA O5')-(4 DA C5') -75.0 -45.0
 &rst    iat =   94,   95,   98,   99,
       r1 = -76.0, r2 = -75.0, r3 = -45.0, r4 = -44.0,
     &end

# 5 CYT ALPHA:  (4 DA O3')-(5 DC P)-(5 DC O5')-(5 DC C5') -75.0 -45.0
 &rst    iat =  126,  127,  130,  131,
       r1 = -81.0, r2 = -80.0, r3 = -45.0, r4 = -44.0,
     &end

# 6 ABX ALPHA:  (5 DC O3')-(6 ABX P)-(6 ABX O5')-(6 ABX C5') -95.0 -25.0
 &rst    iat =  156,  157,  160,  161,
       r1 = -96.0, r2 = -95.0, r3 = -25.0, r4 = -24.0,
     &end

# 7 ADE ALPHA:  (6 ABX O3')-(7 DA P)-(7 DA O5')-(7 DA C5') -85.0 -35.0
 &rst    iat =  201,  202,  205,  206,
       r1 = -86.0, r2 = -85.0, r3 = -35.0, r4 = -34.0,
     &end

# 8 GUA ALPHA:  (7 DA O3')-(8 DG P)-(8 DG O5')-(8 DG C5') -75.0 -45.0
 &rst    iat =  232,  233,  236,  237,
       r1 = -76.0, r2 = -75.0, r3 = -45.0, r4 = -44.0,
```

```
    &end

# 9 ADE ALPHA:  (8 DG O3')-(9 DA P)-(9 DA O5')-(9 DA C5') -75.0 -45.0
 &rst    iat =  265,  266,  269,  270,
      r1 = -76.0, r2 = -75.0, r3 = -45.0, r4 = -44.0,
    &end

# 10 ADE ALPHA:  (9 DA O3')-(10 DA P)-(10 DA O5')-(10 DA C5') -76.0 -45.0
 &rst    iat =  297,  298,  301,  302,
      r1 = -77.0, r2 = -76.0, r3 = -45.0, r4 = -44.0,
    &end

# 11 GUA ALPHA:  (10 DA O3')-(11 DG3 P)-(11 DG3 O5')-(11 DG3 C5') -75.0 -45.0
 &rst    iat =  329,  330,  333,  334,
      r1 = -77.0, r2 = -76.0, r3 = -45.0, r4 = -44.0,
    &end

# 13 THY ALPHA:  (12 DC5 O3')-(13 DT P)-(13 DT O5')-(13 DT C5') -75.0 -45.0
 &rst    iat =  391,  392,  395,  396,
      r1 = -76.0, r2 = -75.0, r3 = -45.0, r4 = -44.0,
    &end

# 14 THY ALPHA:  (13 DT O3')-(14 DT P)-(14 DT O5')-(14 DT C5') -75.0 -45.0
 &rst    iat =  423,  424,  427,  428,
      r1 = -76.0, r2 = -75.0, r3 = -45.0, r4 = -44.0,
    &end

# 15 CYT ALPHA:  (14 DT O3')-(15 DC P)-(15 DC O5')-(15 DC C5') -75.0 -45.0
 &rst    iat =  455,  456,  459,  460,
      r1 = -76.0, r2 = -75.0, r3 = -45.0, r4 = -44.0,
    &end

# 16 THY ALPHA:  (15 DC O3')-(16 DT P)-(16 DT O5')-(16 DT C5') -85.0 -35.0
 &rst    iat =  485,  486,  489,  490,
      r1 = -86.0, r2 = -85.0, r3 = -35.0, r4 = -34.0,
    &end

# 17 THY ALPHA:  (16 DT O3')-(17 DT P)-(17 DT O5')-(17 DT C5') -85.0 -35.0
 &rst    iat =  517,  518,  521,  522,
      r1 = -86.0, r2 = -85.0, r3 = -35.0, r4 = -34.0,
    &end

# 18 GUA ALPHA:  (17 DT O3')-(18 DG P)-(18 DG O5')-(18 DG C5') -75.0 -45.0
 &rst    iat =  549,  550,  553,  554,
      r1 = -76.0, r2 = -75.0, r3 = -45.0, r4 = -44.0,
```

```
    &end

# 19 THY ALPHA:  (18 DG O3')-(19 DT P)-(19 DT O5')-(19 DT C5') -75.0 -45.0
 &rst    iat =   582,   583,   586,   587,
      r1 = -76.0, r2 = -75.0, r3 = -45.0, r4 = -44.0,
    &end

# 20 CYT ALPHA:  (19 DT O3')-(20 DC P)-(20 DC O5')-(20 DC C5') -75.0 -45.0
 &rst    iat =   614,   615,   618,   619,
      r1 = -76.0, r2 = -75.0, r3 = -45.0, r4 = -44.0,
    &end

# 21 CYT ALPHA:  (20 DC O3')-(21 DC P)-(21 DC O5')-(21 DC C5') -75.0 -45.0
 &rst    iat =   644,   645,   648,   649,
      r1 = -76.0, r2 = -75.0, r3 = -45.0, r4 = -44.0,
    &end

# 22 GUA ALPHA:  (21 DC O3')-(22 DG3 P)-(22 DG3 O5')-(22 DG3 C5') -85.0 -35.0
 &rst    iat =   674,   675,   678,   679,
      r1 = -86.0, r2 = -85.0, r3 = -35.0, r4 = -34.0,
    &end

# 1 CYT BETA:  (1 DC5 P)-(1 DC5 O5')-(1 DC5 C5')-(1 DC5 C4') 60.0 300.0
 &rst    iat =    29,    2,    3,    6,
      r1 = 59.0, r2 = 60.0, r3 = 320.0, r4 = 321.0,
      rk2 =  60.0, rk3 =  60.0,                                    &end

# 2 GUA BETA:  (2 DG P)-(2 DG O5')-(2 DG C5')-(2 DG C4') 160.0 200.0
 &rst    iat =    29,    32,    33,    36,
      r1 = 159.0, r2 = 160.0, r3 = 200.0, r4 = 201.0,
    &end

# 3 GUA BETA:  (3 DG P)-(3 DG O5')-(3 DG C5')-(3 DG C4') 160.0 200.0
 &rst    iat =    62,    65,    66,    69,
      r1 = 159.0, r2 = 160.0, r3 = 200.0, r4 = 201.0,
    &end

# 4 ADE BETA:  (4 DA P)-(4 DA O5')-(4 DA C5')-(4 DA C4') 160.0 200.0
 &rst    iat =    95,    98,    99,    102,
      r1 = 149.0, r2 = 150.0, r3 = 210.0, r4 = 211.0,
    &end

# 5 CYT BETA:  (5 DC P)-(5 DC O5')-(5 DC C5')-(5 DC C4') 160.0 200.0
 &rst    iat =   127,   130,   131,   134,
      r1 = 159.0, r2 = 160.0, r3 = 200.0, r4 = 201.0,
```

```
    &end

# 6 ABX BETA:  (6 ABX P)-(6 ABX O5')-(6 ABX C5')-(6 ABX C4') 145.0 215.0
 &rst    iat =   157,   160,   161,   164,
      r1 = 144.0, r2 = 145.0, r3 = 215.0, r4 = 216.0,
    &end

# 7 ADE BETA:  (7 DA P)-(7 DA O5')-(7 DA C5')-(7 DA C4') 150.0 210.0
 &rst    iat =   202,   205,   206,   209,
      r1 = 149.0, r2 = 150.0, r3 = 210.0, r4 = 211.0,
    &end

# 8 GUA BETA:  (8 DG P)-(8 DG O5')-(8 DG C5')-(8 DG C4') 160.0 200.0
 &rst    iat =   233,   236,   237,   240,
      r1 = 159.0, r2 = 160.0, r3 = 200.0, r4 = 201.0,
    &end

# 9 ADE BETA:  (9 DA P)-(9 DA O5')-(9 DA C5')-(9 DA C4') 160.0 200.0
 &rst    iat =   266,   269,   270,   273,
      r1 = 159.0, r2 = 160.0, r3 = 200.0, r4 = 201.0,
    &end

# 10 ADE BETA:  (10 DA P)-(10 DA O5')-(10 DA C5')-(10 DA C4') 160.0 200.0
 &rst    iat =   298,   301,   302,   305,
      r1 = 159.0, r2 = 160.0, r3 = 200.0, r4 = 201.0,
    &end

# 11 GUA BETA:  (11 DG3 P)-(11 DG3 O5')-(11 DG3 C5')-(11 DG3 C4') 60.0 300.0
 &rst    iat =   330,   333,   334,   337,
      r1 = 59.0, r2 = 60.0, r3 = 300.0, r4 = 301.0,
    &end

# 12 CYT BETA:  (12 DC5 P)-(12 DC5 O5')-(12 DC5 C5')-(12 DC5 C4') 60.0 300.0
 &rst    iat =   392,   365,   366,   369,
      r1 = 59.0, r2 = 60.0, r3 = 320.0, r4 = 321.0,
    &end

# 13 THY BETA:  (13 DT P)-(13 DT O5')-(13 DT C5')-(13 DT C4') 160.0 200.0
 &rst    iat =   392,   395,   396,   399,
      r1 = 159.0, r2 = 160.0, r3 = 200.0, r4 = 201.0,
    &end

# 14 THY BETA:  (14 DT P)-(14 DT O5')-(14 DT C5')-(14 DT C4') 160.0 200.0
 &rst    iat =   424,   427,   428,   431,
      r1 = 159.0, r2 = 160.0, r3 = 200.0, r4 = 201.0,
```

```
    &end

# 15 CYT BETA:  (15 DC P)-(15 DC O5')-(15 DC C5')-(15 DC C4') 160.0 200.0
 &rst    iat =   456,   459,   460,   463,
      r1 = 156.0, r2 = 157.0, r3 = 200.0, r4 = 201.0,
    &end

# 16 THY BETA:  (16 DT P)-(16 DT O5')-(16 DT C5')-(16 DT C4') 150.0 210.0
 &rst    iat =   486,   489,   490,   493,
      r1 = 149.0, r2 = 150.0, r3 = 210.0, r4 = 211.0,
    &end

# 17 THY BETA:  (17 DT P)-(17 DT O5')-(17 DT C5')-(17 DT C4') 150.0 210.0
 &rst    iat =   518,   521,   522,   525,
      r1 = 149.0, r2 = 150.0, r3 = 210.0, r4 = 211.0,
    &end

# 18 GUA BETA:  (18 DG P)-(18 DG O5')-(18 DG C5')-(18 DG C4') 160.0 200.0
 &rst    iat =   550,   553,   554,   557,
      r1 = 159.0, r2 = 160.0, r3 = 200.0, r4 = 201.0,
    &end

# 19 THY BETA:  (19 DT P)-(19 DT O5')-(19 DT C5')-(19 DT C4') 160.0 200.0
 &rst    iat =   583,   586,   587,   590,
      r1 = 159.0, r2 = 160.0, r3 = 200.0, r4 = 201.0,
    &end

# 20 CYT BETA:  (20 DC P)-(20 DC O5')-(20 DC C5')-(20 DC C4') 160.0 200.0
 &rst    iat =   615,   618,   619,   622,
      r1 = 159.0, r2 = 160.0, r3 = 205.0, r4 = 206.0,
    &end

# 21 CYT BETA:  (21 DC P)-(21 DC O5')-(21 DC C5')-(21 DC C4') 160.0 200.0
 &rst    iat =   645,   648,   649,   652,
      r1 = 159.0, r2 = 160.0, r3 = 200.0, r4 = 201.0,
    &end

# 22 GUA BETA:  (22 DG3 P)-(22 DG3 O5')-(22 DG3 C5')-(22 DG3 C4') 150.0 210.0
 &rst    iat =   675,   678,   679,   682,
      r1 = 149.0, r2 = 150.0, r3 = 210.0, r4 = 211.0,
    &end

# 1 CYT EPSLN:  (1 DC5 C4')-(1 DC5 C3')-(1 DC5 O3')-(2 DG P) 130.0 200.0
 &rst    iat =     6,    23,    28,    29,
      r1 = 129.0, r2 = 130.0, r3 = 200.0, r4 = 201.0,
```

```
      rk2 =  60.0, rk3 =  60.0,                                  &end

# 2 GUA EPSLN:  (2 DG C4')-(2 DG C3')-(2 DG O3')-(3 DG P) 140.0 190.0
 &rst    iat =   36,   56,   61,   62,
      r1 = 139.0, r2 = 140.0, r3 = 190.0, r4 = 191.0,
    &end

# 3 GUA EPSLN:  (3 DG C4')-(3 DG C3')-(3 DG O3')-(4 DA P) 140.0 190.0
 &rst    iat =   69,   89,   94,   95,
      r1 = 139.0, r2 = 140.0, r3 = 192.0, r4 = 193.0,
    &end

# 4 ADE EPSLN:  (4 DA C4')-(4 DA C3')-(4 DA O3')-(5 DC P) 140.0 190.0
 &rst    iat =   102,   121,   126,   127,
      r1 = 139.0, r2 = 140.0, r3 = 190.0, r4 = 191.0,
    &end

# 5 CYT EPSLN:  (5 DC C4')-(5 DC C3')-(5 DC O3')-(6 ABX P) 140.0 190.0
 &rst    iat =   134,   151,   156,   157,
      r1 = 139.0, r2 = 140.0, r3 = 190.0, r4 = 191.0,
    &end

# 6 ABX EPSLN:  (6 ABX C4')-(6 ABX C3')-(6 ABX O3')-(7 DA P) 130.0 210.0
 &rst    iat =   164,   199,   201,   202,
      r1 = 129.0, r2 = 130.0, r3 = 210.0, r4 = 211.0,
    &end

# 7 ADE EPSLN:  (7 DA C4')-(7 DA C3')-(7 DA O3')-(8 DG P) 130.0 210.0
 &rst    iat =   209,   227,   232,   233,
      r1 = 129.0, r2 = 130.0, r3 = 210.0, r4 = 211.0,
    &end

# 8 GUA EPSLN:  (8 DG C4')-(8 DG C3')-(8 DG O3')-(9 DA P) 140.0 190.0
 &rst    iat =   240,   260,   265,   266,
      r1 = 139.0, r2 = 140.0, r3 = 190.0, r4 = 191.0,
    &end

# 9 ADE EPSLN:  (9 DA C4')-(9 DA C3')-(9 DA O3')-(10 DA P) 140.0 191.0
 &rst    iat =   273,   292,   297,   298,
      r1 = 139.0, r2 = 140.0, r3 = 191.0, r4 = 192.0,
    &end

# 10 ADE EPSLN:  (10 DA C4')-(10 DA C3')-(10 DA O3')-(11 DG3 P) 140.0 240.0
 &rst    iat =   305,   324,   329,   330,
      r1 = 139.0, r2 = 140.0, r3 = 282.0, r4 = 283.0,
```

```
    &end

# 12 CYT EPSLN:  (12 DC5 C4')-(12 DC5 C3')-(12 DC5 O3')-(13 DT P) 140.0 190.0
 &rst    iat =   369,   386,   391,   392,
      r1 = 139.0, r2 = 140.0, r3 = 192.0, r4 = 193.0,
    &end

# 13 THY EPSLN:  (13 DT C4')-(13 DT C3')-(13 DT O3')-(14 DT P) 140.0 190.0
 &rst    iat =   399,   418,   423,   424,
      r1 = 139.0, r2 = 140.0, r3 = 190.0, r4 = 191.0,
    &end

# 14 THY EPSLN:  (14 DT C4')-(14 DT C3')-(14 DT O3')-(15 DC P) 140.0 240.0
 &rst    iat =   431,   450,   455,   456,
      r1 = 139.0, r2 = 140.0, r3 = 240.0, r4 = 241.0,
    &end

# 15 CYT EPSLN:  (15 DC C4')-(15 DC C3')-(15 DC O3')-(16 DT P) 140.0 190.0
 &rst    iat =   463,   480,   485,   486,
      r1 = 139.0, r2 = 140.0, r3 = 190.0, r4 = 191.0,
    &end

# 16 THY EPSLN:  (16 DT C4')-(16 DT C3')-(16 DT O3')-(17 DT P) 130.0 210.0
 &rst    iat =   493,   512,   517,   518,
      r1 = 129.0, r2 = 130.0, r3 = 210.0, r4 = 211.0,
    &end

# 17 THY EPSLN:  (17 DT C4')-(17 DT C3')-(17 DT O3')-(18 DG P) 120.0 220.0
 &rst    iat =   525,   544,   549,   550,
      r1 = 119.0, r2 = 120.0, r3 = 233.0, r4 = 234.0,
    &end

# 18 GUA EPSLN:  (18 DG C4')-(18 DG C3')-(18 DG O3')-(19 DT P) 140.0 190.0
 &rst    iat =   557,   577,   582,   583,
      r1 = 139.0, r2 = 140.0, r3 = 200.0, r4 = 201.0,
    &end

# 19 THY EPSLN:  (19 DT C4')-(19 DT C3')-(19 DT O3')-(20 DC P) 140.0 190.0
 &rst    iat =   590,   609,   614,   615,
      r1 = 139.0, r2 = 140.0, r3 = 190.0, r4 = 191.0,
    &end

# 20 CYT EPSLN:  (20 DC C4')-(20 DC C3')-(20 DC O3')-(21 DC P) 140.0 190.0
 &rst    iat =   622,   639,   644,   645,
      r1 = 139.0, r2 = 140.0, r3 = 192.0, r4 = 193.0,
```

```
    &end

# 21 CYT EPSLN:  (21 DC C4')-(21 DC C3')-(21 DC O3')-(22 DG3 P) 140.0 240.0
 &rst    iat =   652,   669,   674,   675,
      r1 = 139.0, r2 = 140.0, r3 = 240.0, r4 = 241.0,
    &end

# 1 CYT GAMMA:  (1 DC5 O5')-(1 DC5 C5')-(1 DC5 C4')-(1 DC5 C3')  -35.0  85.0
 &rst    iat =    2,    3,    6,   23,
      r1 =  -36.0, r2 =  -35.0, r3 =  85.0, r4 =  86.0,
      rk2 =  60.0, rk3 =  60.0,                                    &end

# 2 GUA GAMMA:  (2 DG O5')-(2 DG C5')-(2 DG C4')-(2 DG C3')  45.0  75.0
 &rst    iat =   32,   33,   36,   56,
      r1 =  44.0, r2 =  45.0, r3 =  75.0, r4 =  76.0,
    &end

# 3 GUA GAMMA:  (3 DG O5')-(3 DG C5')-(3 DG C4')-(3 DG C3')  40.0  75.0
 &rst    iat =   65,   66,   69,   89,
      r1 =  39.0, r2 =  40.0, r3 =  75.0, r4 =  76.0,
    &end

# 4 ADE GAMMA:  (4 DA O5')-(4 DA C5')-(4 DA C4')-(4 DA C3')  44.0  75.0
 &rst    iat =   98,   99,   102,   121,
      r1 =  43.0, r2 =  44.0, r3 =  75.0, r4 =  76.0,
    &end

# 5 CYT GAMMA:  (5 DC O5')-(5 DC C5')-(5 DC C4')-(5 DC C3')  45.0  75.0
 &rst    iat =   130,   131,   134,   151,
      r1 =  44.0, r2 =  45.0, r3 =  75.0, r4 =  76.0,
    &end

# 6 ABX GAMMA:  (6 ABX O5')-(6 ABX C5')-(6 ABX C4')-(6 ABX C3')  40.0  80.0
 &rst    iat =   160,   161,   164,   199,
      r1 =  39.0, r2 =  40.0, r3 =  80.0, r4 =  81.0,
    &end

# 7 ADE GAMMA:  (7 DA O5')-(7 DA C5')-(7 DA C4')-(7 DA C3')  40.0  80.0
 &rst    iat =   205,   206,   209,   227,
      r1 =  39.0, r2 =  40.0, r3 =  80.0, r4 =  81.0,
    &end

# 8 GUA GAMMA:  (8 DG O5')-(8 DG C5')-(8 DG C4')-(8 DG C3')  45.0  75.0
 &rst    iat =   236,   237,   240,   260,
      r1 =  42.0, r2 =  43.0, r3 =  75.0, r4 =  76.0,
```

```
    &end

# 9 ADE GAMMA:  (9 DA O5')-(9 DA C5')-(9 DA C4')-(9 DA C3')  45.0  75.0
 &rst    iat =  269,  270,  273,  292,
      r1 =  44.0, r2 =  45.0, r3 =  75.0, r4 =  76.0,
    &end

# 10 ADE GAMMA:  (10 DA O5')-(10 DA C5')-(10 DA C4')-(10 DA C3')  45.0  75.0
 &rst    iat =  301,  302,  305,  324,
      r1 =  41.0, r2 =  42.0, r3 =  75.0, r4 =  76.0,
    &end

# 13 THY GAMMA:  (13 DT O5')-(13 DT C5')-(13 DT C4')-(13 DT C3')  45.0  75.0
 &rst    iat =  395,  396,  399,  418,
      r1 =  41.0, r2 =  42.0, r3 =  75.0, r4 =  76.0,
    &end

# 14 THY GAMMA:  (14 DT O5')-(14 DT C5')-(14 DT C4')-(14 DT C3')  45.0  75.0
 &rst    iat =  427,  428,  431,  450,
      r1 =  44.0, r2 =  45.0, r3 =  75.0, r4 =  76.0,
    &end

# 15 CYT GAMMA:  (15 DC O5')-(15 DC C5')-(15 DC C4')-(15 DC C3')  45.0  75.0
 &rst    iat =  459,  460,  463,  480,
      r1 =  44.0, r2 =  45.0, r3 =  75.0, r4 =  76.0,
    &end

# 16 THY GAMMA:  (16 DT O5')-(16 DT C5')-(16 DT C4')-(16 DT C3')  40.0  80.0
 &rst    iat =  489,  490,  493,  512,
      r1 =  39.0, r2 =  40.0, r3 =  80.0, r4 =  81.0,
    &end

# 17 THY GAMMA:  (17 DT O5')-(17 DT C5')-(17 DT C4')-(17 DT C3')  40.0  80.0
 &rst    iat =  521,  522,  525,  544,
      r1 =  39.0, r2 =  40.0, r3 =  80.0, r4 =  81.0,
    &end

# 18 GUA GAMMA:  (18 DG O5')-(18 DG C5')-(18 DG C4')-(18 DG C3')  45.0  75.0
 &rst    iat =  553,  554,  557,  577,
      r1 =  34.0, r2 =  35.0, r3 =  75.0, r4 =  76.0,
    &end

# 19 THY GAMMA:  (19 DT O5')-(19 DT C5')-(19 DT C4')-(19 DT C3')  45.0  75.0
 &rst    iat =  586,  587,  590,  609,
      r1 =  44.0, r2 =  45.0, r3 =  75.0, r4 =  76.0,
```

```
    &end

# 20 CYT GAMMA:  (20 DC O5')-(20 DC C5')-(20 DC C4')-(20 DC C3')  45.0  75.0
 &rst    iat =   618,   619,   622,   639,
      r1 =  34.0, r2 =  35.0, r3 =  75.0, r4 =  76.0,
    &end

# 21 CYT GAMMA:  (21 DC O5')-(21 DC C5')-(21 DC C4')-(21 DC C3')  45.0  75.0
 &rst    iat =   648,   649,   652,   669,
      r1 =  31.0, r2 =  32.0, r3 =  75.0, r4 =  76.0,
    &end

# 22 GUA GAMMA:  (22 DG3 O5')-(22 DG3 C5')-(22 DG3 C4')-(22 DG3 C3')  35.0  85.0
 &rst    iat =   678,   679,   682,   702,
      r1 =  34.0, r2 =  35.0, r3 =  85.0, r4 =  86.0,
    &end

# 708 atoms read from pdb file xlinkss_amb_m.pdb.
# 1 CYT NU0:  (1 DC5 C4')-(1 DC5 O4')-(1 DC5 C1')-(1 DC5 C2') -40.2 -10.2
 &rst    iat =     6,     8,     9,    25,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
      rk2 =  40.0, rk3 =  40.0,                                    &end

# 1 CYT NU1:  (1 DC5 O4')-(1 DC5 C1')-(1 DC5 C2')-(1 DC5 C3')  15.6  45.6
 &rst    iat =     8,     9,    25,    23,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 1 CYT NU2:  (1 DC5 C1')-(1 DC5 C2')-(1 DC5 C3')-(1 DC5 C4') -38.5   0.0
 &rst    iat =     9,    25,    23,     6,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 1 CYT NU3:  (1 DC5 C2')-(1 DC5 C3')-(1 DC5 C4')-(1 DC5 O4') -22.6  31.1
 &rst    iat =    25,    23,     6,     8,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 1 CYT NU4:  (1 DC5 C3')-(1 DC5 C4')-(1 DC5 O4')-(1 DC5 C1') -11.8  36.6
 &rst    iat =    23,     6,     8,     9,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 2 GUA NU0:  (2 DG C4')-(2 DG O4')-(2 DG C1')-(2 DG C2') -40.2 -10.2
 &rst    iat =    36,    38,    39,    58,
```

```
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 2 GUA NU1:  (2 DG O4')-(2 DG C1')-(2 DG C2')-(2 DG C3')  15.6  45.6
 &rst    iat =   38,   39,   58,   56,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 2 GUA NU2:  (2 DG C1')-(2 DG C2')-(2 DG C3')-(2 DG C4') -38.5   0.0
 &rst    iat =   39,   58,   56,   36,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 2 GUA NU3:  (2 DG C2')-(2 DG C3')-(2 DG C4')-(2 DG O4') -22.6  31.1
 &rst    iat =   58,   56,   36,   38,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 2 GUA NU4:  (2 DG C3')-(2 DG C4')-(2 DG O4')-(2 DG C1') -11.8  36.6
 &rst    iat =   56,   36,   38,   39,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 3 GUA NU0:  (3 DG C4')-(3 DG O4')-(3 DG C1')-(3 DG C2') -40.2 -10.2
 &rst    iat =   69,   71,   72,   91,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 3 GUA NU1:  (3 DG O4')-(3 DG C1')-(3 DG C2')-(3 DG C3')  15.6  45.6
 &rst    iat =   71,   72,   91,   89,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 3 GUA NU2:  (3 DG C1')-(3 DG C2')-(3 DG C3')-(3 DG C4') -38.5   0.0
 &rst    iat =   72,   91,   89,   69,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 3 GUA NU3:  (3 DG C2')-(3 DG C3')-(3 DG C4')-(3 DG O4') -22.6  31.1
 &rst    iat =   91,   89,   69,   71,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 3 GUA NU4:  (3 DG C3')-(3 DG C4')-(3 DG O4')-(3 DG C1') -11.8  36.6
 &rst    iat =   89,   69,   71,   72,
```

```
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 4 ADE NU0:  (4 DA C4')-(4 DA O4')-(4 DA C1')-(4 DA C2') -40.2 -10.2
 &rst    iat =   102,   104,   105,   123,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 4 ADE NU1:  (4 DA O4')-(4 DA C1')-(4 DA C2')-(4 DA C3')  15.6  45.6
 &rst    iat =   104,   105,   123,   121,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 4 ADE NU2:  (4 DA C1')-(4 DA C2')-(4 DA C3')-(4 DA C4') -38.5   0.0
 &rst    iat =   105,   123,   121,   102,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 4 ADE NU3:  (4 DA C2')-(4 DA C3')-(4 DA C4')-(4 DA O4') -22.6  31.1
 &rst    iat =   123,   121,   102,   104,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 4 ADE NU4:  (4 DA C3')-(4 DA C4')-(4 DA O4')-(4 DA C1') -11.8  36.6
 &rst    iat =   121,   102,   104,   105,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 5 CYT NU0:  (5 DC C4')-(5 DC O4')-(5 DC C1')-(5 DC C2') -40.2 -10.2
 &rst    iat =   134,   136,   137,   153,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 5 CYT NU1:  (5 DC O4')-(5 DC C1')-(5 DC C2')-(5 DC C3')  15.6  45.6
 &rst    iat =   136,   137,   153,   151,
      r1 = -0.4, r2 =  0.6, r3 =  45.6, r4 =  46.6,
    &end

# 5 CYT NU2:  (5 DC C1')-(5 DC C2')-(5 DC C3')-(5 DC C4') -38.5   0.0
 &rst    iat =   137,   153,   151,   134,
      r1 = -39.5, r2 = -38.5, r3 =  21.0, r4 =  22.0,
    &end

# 5 CYT NU3:  (5 DC C2')-(5 DC C3')-(5 DC C4')-(5 DC O4') -22.6  31.1
 &rst    iat =   153,   151,   134,   136,
```

```
      r1 = -36.6, r2 = -35.6, r3 =  31.1, r4 =  32.1,
    &end

# 5 CYT NU4:  (5 DC C3')-(5 DC C4')-(5 DC O4')-(5 DC C1') -11.8  36.6
 &rst    iat =   151,   134,   136,   137,
      r1 = -12.8, r2 = -11.8, r3 =  38.6, r4 =  39.6,
    &end

# 6 ABX NU0:  (6 ABX C4')-(6 ABX O4')-(6 ABX C1')-(6 ABX C2') -40.2 -10.2
 &rst    iat =   164,   166,   167,   169,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 6 ABX NU1:  (6 ABX O4')-(6 ABX C1')-(6 ABX C2')-(6 ABX C3')  15.6  45.6
 &rst    iat =   166,   167,   169,   199,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 6 ABX NU2:  (6 ABX C1')-(6 ABX C2')-(6 ABX C3')-(6 ABX C4') -38.5   0.0
 &rst    iat =   167,   169,   199,   164,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 6 ABX NU3:  (6 ABX C2')-(6 ABX C3')-(6 ABX C4')-(6 ABX O4') -22.6  31.1
 &rst    iat =   169,   199,   164,   166,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 6 ABX NU4:  (6 ABX C3')-(6 ABX C4')-(6 ABX O4')-(6 ABX C1') -11.8  36.6
 &rst    iat =   199,   164,   166,   167,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 7 ADE NU0:  (7 DA C4')-(7 DA O4')-(7 DA C1')-(7 DA C2') -40.2 -10.2
 &rst    iat =   209,   211,   212,   229,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 7 ADE NU1:  (7 DA O4')-(7 DA C1')-(7 DA C2')-(7 DA C3')  15.6  45.6
 &rst    iat =   211,   212,   229,   227,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 7 ADE NU2:  (7 DA C1')-(7 DA C2')-(7 DA C3')-(7 DA C4') -38.5   0.0
 &rst    iat =   212,   229,   227,   209,
```

```
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 7 ADE NU3:  (7 DA C2')-(7 DA C3')-(7 DA C4')-(7 DA O4') -22.6  31.1
 &rst    iat =   229,   227,   209,   211,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 7 ADE NU4:  (7 DA C3')-(7 DA C4')-(7 DA O4')-(7 DA C1') -11.8  36.6
 &rst    iat =   227,   209,   211,   212,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 8 GUA NU0:  (8 DG C4')-(8 DG O4')-(8 DG C1')-(8 DG C2') -40.2 -10.2
 &rst    iat =   240,   242,   243,   262,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 8 GUA NU1:  (8 DG O4')-(8 DG C1')-(8 DG C2')-(8 DG C3')  15.6  45.6
 &rst    iat =   242,   243,   262,   260,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 8 GUA NU2:  (8 DG C1')-(8 DG C2')-(8 DG C3')-(8 DG C4') -38.5   0.0
 &rst    iat =   243,   262,   260,   240,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 8 GUA NU3:  (8 DG C2')-(8 DG C3')-(8 DG C4')-(8 DG O4') -22.6  31.1
 &rst    iat =   262,   260,   240,   242,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 8 GUA NU4:  (8 DG C3')-(8 DG C4')-(8 DG O4')-(8 DG C1') -11.8  36.6
 &rst    iat =   260,   240,   242,   243,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 9 ADE NU0:  (9 DA C4')-(9 DA O4')-(9 DA C1')-(9 DA C2') -40.2 -10.2
 &rst    iat =   273,   275,   276,   294,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 9 ADE NU1:  (9 DA O4')-(9 DA C1')-(9 DA C2')-(9 DA C3')  15.6  45.6
 &rst    iat =   275,   276,   294,   292,
```

```
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 9 ADE NU2:  (9 DA C1')-(9 DA C2')-(9 DA C3')-(9 DA C4') -38.5   0.0
 &rst    iat =   276,   294,   292,   273,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 9 ADE NU3:  (9 DA C2')-(9 DA C3')-(9 DA C4')-(9 DA O4') -22.6  31.1
 &rst    iat =   294,   292,   273,   275,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 9 ADE NU4:  (9 DA C3')-(9 DA C4')-(9 DA O4')-(9 DA C1') -11.8  36.6
 &rst    iat =   292,   273,   275,   276,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 10 ADE NU0:  (10 DA C4')-(10 DA O4')-(10 DA C1')-(10 DA C2') -40.2 -10.2
 &rst    iat =   305,   307,   308,   326,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 10 ADE NU1:  (10 DA O4')-(10 DA C1')-(10 DA C2')-(10 DA C3')  15.6  45.6
 &rst    iat =   307,   308,   326,   324,
      r1 =  14.6, r2 =  15.6, r3 =  48.6, r4 =  49.6,
    &end

# 10 ADE NU2:  (10 DA C1')-(10 DA C2')-(10 DA C3')-(10 DA C4') -38.5   0.0
 &rst    iat =   308,   326,   324,   305,
      r1 = -42.5, r2 = -41.5, r3 =   0.0, r4 =   1.0,
    &end

# 10 ADE NU3:  (10 DA C2')-(10 DA C3')-(10 DA C4')-(10 DA O4') -22.6  31.1
 &rst    iat =   326,   324,   305,   307,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 10 ADE NU4:  (10 DA C3')-(10 DA C4')-(10 DA O4')-(10 DA C1') -11.8  36.6
 &rst    iat =   324,   305,   307,   308,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 11 GUA NU0:  (11 DG3 C4')-(11 DG3 O4')-(11 DG3 C1')-(11 DG3 C2') -40.2 -10.2
 &rst    iat =   337,   339,   340,   359,
```

```
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 11 GUA NU1:  (11 DG3 O4')-(11 DG3 C1')-(11 DG3 C2')-(11 DG3 C3')  15.6  45.6
 &rst    iat =   339,   340,   359,   357,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 11 GUA NU2:  (11 DG3 C1')-(11 DG3 C2')-(11 DG3 C3')-(11 DG3 C4') -38.5   0.0
 &rst    iat =   340,   359,   357,   337,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 11 GUA NU3:  (11 DG3 C2')-(11 DG3 C3')-(11 DG3 C4')-(11 DG3 O4') -22.6  31.1
 &rst    iat =   359,   357,   337,   339,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 11 GUA NU4:  (11 DG3 C3')-(11 DG3 C4')-(11 DG3 O4')-(11 DG3 C1') -11.8  36.6
 &rst    iat =   357,   337,   339,   340,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 12 CYT NU0:  (12 DC5 C4')-(12 DC5 O4')-(12 DC5 C1')-(12 DC5 C2') -40.2 -10.2
 &rst    iat =   369,   371,   372,   388,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 12 CYT NU1:  (12 DC5 O4')-(12 DC5 C1')-(12 DC5 C2')-(12 DC5 C3')  15.6  45.6
 &rst    iat =   371,   372,   388,   386,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 12 CYT NU2:  (12 DC5 C1')-(12 DC5 C2')-(12 DC5 C3')-(12 DC5 C4') -38.5   0.0
 &rst    iat =   372,   388,   386,   369,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 12 CYT NU3:  (12 DC5 C2')-(12 DC5 C3')-(12 DC5 C4')-(12 DC5 O4') -22.6  31.1
 &rst    iat =   388,   386,   369,   371,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 12 CYT NU4:  (12 DC5 C3')-(12 DC5 C4')-(12 DC5 O4')-(12 DC5 C1') -11.8  36.6
 &rst    iat =   386,   369,   371,   372,
```

```
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 13 THY NU0:  (13 DT C4')-(13 DT O4')-(13 DT C1')-(13 DT C2') -40.2 -10.2
 &rst    iat =   399,   401,   402,   420,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 13 THY NU1:  (13 DT O4')-(13 DT C1')-(13 DT C2')-(13 DT C3')  15.6  45.6
 &rst    iat =   401,   402,   420,   418,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 13 THY NU2:  (13 DT C1')-(13 DT C2')-(13 DT C3')-(13 DT C4') -38.5   0.0
 &rst    iat =   402,   420,   418,   399,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 13 THY NU3:  (13 DT C2')-(13 DT C3')-(13 DT C4')-(13 DT O4') -22.6  31.1
 &rst    iat =   420,   418,   399,   401,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 13 THY NU4:  (13 DT C3')-(13 DT C4')-(13 DT O4')-(13 DT C1') -11.8  36.6
 &rst    iat =   418,   399,   401,   402,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 14 THY NU0:  (14 DT C4')-(14 DT O4')-(14 DT C1')-(14 DT C2') -40.2 -10.2
 &rst    iat =   431,   433,   434,   452,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 14 THY NU1:  (14 DT O4')-(14 DT C1')-(14 DT C2')-(14 DT C3')  15.6  45.6
 &rst    iat =   433,   434,   452,   450,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 14 THY NU2:  (14 DT C1')-(14 DT C2')-(14 DT C3')-(14 DT C4') -38.5   0.0
 &rst    iat =   434,   452,   450,   431,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 14 THY NU3:  (14 DT C2')-(14 DT C3')-(14 DT C4')-(14 DT O4') -22.6  31.1
 &rst    iat =   452,   450,   431,   433,
```

```
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 14 THY NU4:  (14 DT C3')-(14 DT C4')-(14 DT O4')-(14 DT C1') -11.8  36.6
 &rst    iat =   450,   431,   433,   434,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 15 CYT NU0:  (15 DC C4')-(15 DC O4')-(15 DC C1')-(15 DC C2') -40.2 -10.2
 &rst    iat =   463,   465,   466,   482,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 15 CYT NU1:  (15 DC O4')-(15 DC C1')-(15 DC C2')-(15 DC C3')  15.6  45.6
 &rst    iat =   465,   466,   482,   480,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 15 CYT NU2:  (15 DC C1')-(15 DC C2')-(15 DC C3')-(15 DC C4') -38.5   0.0
 &rst    iat =   466,   482,   480,   463,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 15 CYT NU3:  (15 DC C2')-(15 DC C3')-(15 DC C4')-(15 DC O4') -22.6  31.1
 &rst    iat =   482,   480,   463,   465,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 15 CYT NU4:  (15 DC C3')-(15 DC C4')-(15 DC O4')-(15 DC C1') -11.8  36.6
 &rst    iat =   480,   463,   465,   466,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 16 THY NU0:  (16 DT C4')-(16 DT O4')-(16 DT C1')-(16 DT C2') -40.2 -10.2
 &rst    iat =   493,   495,   496,   514,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 16 THY NU1:  (16 DT O4')-(16 DT C1')-(16 DT C2')-(16 DT C3')  15.6  45.6
 &rst    iat =   495,   496,   514,   512,
      r1 =  11.6, r2 =  12.6, r3 =  48.6, r4 =  49.6,
    &end

# 16 THY NU2:  (16 DT C1')-(16 DT C2')-(16 DT C3')-(16 DT C4') -38.5   0.0
 &rst    iat =   496,   514,   512,   493,
```

```
      r1 = -36.5, r2 = -35.5, r3 =   5.0, r4 =   6.0,
    &end

# 16 THY NU3:  (16 DT C2')-(16 DT C3')-(16 DT C4')-(16 DT O4') -22.6  31.1
 &rst    iat =   514,   512,   493,   495,
      r1 = -26.6, r2 = -25.6, r3 =  34.1, r4 =  35.1,
    &end

# 16 THY NU4:  (16 DT C3')-(16 DT C4')-(16 DT O4')-(16 DT C1') -11.8  36.6
 &rst    iat =   512,   493,   495,   496,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 17 THY NU0:  (17 DT C4')-(17 DT O4')-(17 DT C1')-(17 DT C2') -40.2 -10.2
 &rst    iat =   525,   527,   528,   546,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 17 THY NU1:  (17 DT O4')-(17 DT C1')-(17 DT C2')-(17 DT C3')  15.6  45.6
 &rst    iat =   527,   528,   546,   544,
      r1 =  10.6, r2 =  11.6, r3 =  48.6, r4 =  49.6,
    &end

# 17 THY NU2:  (17 DT C1')-(17 DT C2')-(17 DT C3')-(17 DT C4') -38.5   0.0
 &rst    iat =   528,   546,   544,   525,
      r1 = -36.5, r2 = -35.5, r3 =   15.0, r4 =   16.0,
    &end

# 17 THY NU3:  (17 DT C2')-(17 DT C3')-(17 DT C4')-(17 DT O4') -22.6  31.1
 &rst    iat =   546,   544,   525,   527,
      r1 = -37.6, r2 = -36.6, r3 =  34.1, r4 =  35.1,
    &end

# 17 THY NU4:  (17 DT C3')-(17 DT C4')-(17 DT O4')-(17 DT C1') -11.8  36.6
 &rst    iat =   544,   525,   527,   528,
      r1 = -12.8, r2 = -11.8, r3 =  40.6, r4 =  41.6,
    &end

# 18 GUA NU0:  (18 DG C4')-(18 DG O4')-(18 DG C1')-(18 DG C2') -40.2 -10.2
 &rst    iat =   557,   559,   560,   579,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 18 GUA NU1:  (18 DG O4')-(18 DG C1')-(18 DG C2')-(18 DG C3')  15.6  45.6
 &rst    iat =   559,   560,   579,   577,
```

```
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 18 GUA NU2:  (18 DG C1')-(18 DG C2')-(18 DG C3')-(18 DG C4') -38.5   0.0
 &rst    iat =   560,   579,   577,   557,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 18 GUA NU3:  (18 DG C2')-(18 DG C3')-(18 DG C4')-(18 DG O4') -22.6  31.1
 &rst    iat =   579,   577,   557,   559,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 18 GUA NU4:  (18 DG C3')-(18 DG C4')-(18 DG O4')-(18 DG C1') -11.8  36.6
 &rst    iat =   577,   557,   559,   560,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 19 THY NU0:  (19 DT C4')-(19 DT O4')-(19 DT C1')-(19 DT C2') -40.2 -10.2
 &rst    iat =   590,   592,   593,   611,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 19 THY NU1:  (19 DT O4')-(19 DT C1')-(19 DT C2')-(19 DT C3')  15.6  45.6
 &rst    iat =   592,   593,   611,   609,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 19 THY NU2:  (19 DT C1')-(19 DT C2')-(19 DT C3')-(19 DT C4') -38.5   0.0
 &rst    iat =   593,   611,   609,   590,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 19 THY NU3:  (19 DT C2')-(19 DT C3')-(19 DT C4')-(19 DT O4') -22.6  31.1
 &rst    iat =   611,   609,   590,   592,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 19 THY NU4:  (19 DT C3')-(19 DT C4')-(19 DT O4')-(19 DT C1') -11.8  36.6
 &rst    iat =   609,   590,   592,   593,
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 20 CYT NU0:  (20 DC C4')-(20 DC O4')-(20 DC C1')-(20 DC C2') -40.2 -10.2
 &rst    iat =   622,   624,   625,   641,
```

```
     r1 = -41.2, r2 = -40.2, r3 = -5.2, r4 =  -4.2,
    &end

# 20 CYT NU1:  (20 DC O4')-(20 DC C1')-(20 DC C2')-(20 DC C3')  15.6  45.6
 &rst    iat =   624,   625,   641,   639,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 20 CYT NU2:  (20 DC C1')-(20 DC C2')-(20 DC C3')-(20 DC C4') -38.5   0.0
 &rst    iat =   625,   641,   639,   622,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 20 CYT NU3:  (20 DC C2')-(20 DC C3')-(20 DC C4')-(20 DC O4') -22.6  31.1
 &rst    iat =   641,   639,   622,   624,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 20 CYT NU4:  (20 DC C3')-(20 DC C4')-(20 DC O4')-(20 DC C1') -11.8  36.6
 &rst    iat =   639,   622,   624,   625,
      r1 = -15.8, r2 = -14.8, r3 =  36.6, r4 =  37.6,
    &end

# 21 CYT NU0:  (21 DC C4')-(21 DC O4')-(21 DC C1')-(21 DC C2') -40.2 -10.2
 &rst    iat =   652,   654,   655,   671,
      r1 = -41.2, r2 = -40.2, r3 = -10.2, r4 =  -9.2,
    &end

# 21 CYT NU1:  (21 DC O4')-(21 DC C1')-(21 DC C2')-(21 DC C3')  15.6  45.6
 &rst    iat =   654,   655,   671,   669,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 21 CYT NU2:  (21 DC C1')-(21 DC C2')-(21 DC C3')-(21 DC C4') -38.5   0.0
 &rst    iat =   655,   671,   669,   652,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 21 CYT NU3:  (21 DC C2')-(21 DC C3')-(21 DC C4')-(21 DC O4') -22.6  31.1
 &rst    iat =   671,   669,   652,   654,
      r1 = -23.6, r2 = -22.6, r3 =  31.1, r4 =  32.1,
    &end

# 21 CYT NU4:  (21 DC C3')-(21 DC C4')-(21 DC O4')-(21 DC C1') -11.8  36.6
 &rst    iat =   669,   652,   654,   655,
```

```
      r1 = -12.8, r2 = -11.8, r3 =  36.6, r4 =  37.6,
    &end

# 22 GUA NU0:  (22 DG3 C4')-(22 DG3 O4')-(22 DG3 C1')-(22 DG3 C2') -40.2 -10.2
 &rst    iat =   682,   684,   685,   704,
      r1 = -41.2, r2 = -40.2, r3 = -5.2, r4 =  -4.2,
    &end

# 22 GUA NU1:  (22 DG3 O4')-(22 DG3 C1')-(22 DG3 C2')-(22 DG3 C3')  15.6  45.6
 &rst    iat =   684,   685,   704,   702,
      r1 =  14.6, r2 =  15.6, r3 =  45.6, r4 =  46.6,
    &end

# 22 GUA NU2:  (22 DG3 C1')-(22 DG3 C2')-(22 DG3 C3')-(22 DG3 C4') -38.5   0.0
 &rst    iat =   685,   704,   702,   682,
      r1 = -39.5, r2 = -38.5, r3 =   0.0, r4 =   1.0,
    &end

# 22 GUA NU3:  (22 DG3 C2')-(22 DG3 C3')-(22 DG3 C4')-(22 DG3 O4') -22.6  31.1
 &rst    iat =   704,   702,   682,   684,
      r1 = -23.6, r2 = -22.6, r3 =  33.1, r4 =  34.1,
    &end

# 22 GUA NU4:  (22 DG3 C3')-(22 DG3 C4')-(22 DG3 O4')-(22 DG3 C1') -11.8  36.6
 &rst    iat =   702,   682,   684,   685,
      r1 = -16.8, r2 = -15.8, r3 =  36.6, r4 =  37.6,
    &end
```

APPENDIX E

MOLECULAR DYNAMICS FILES

File D1. Energy minimization protocol (initial minimisation prior to MD GB model)

```
 &cntrl
  imin   = 1,
  maxcyc = 500,
  ncyc   = 250,
  ntb    = 0,
  igb    = 1,
  cut    = 12
 /
```

File D2. Simulated annealing protocol, 40 ps

```
 &cntrl
    nstlim=400000, pencut=-0.001, nmropt=1,
    ntpr=1000, ntt=1, ntwx=1000,
    cut=10.0, ntb=0, vlimit=10,
    tempi=25.0, ntx=1, igb=1, saltcon=0.2, offset=0.13, ig=776400,
 &end
 &ewald
 &end
#
#Simple simulated annealing algorithm: J.Smith's ppt .in file
#
#from steps 0 to 1000: heat the system to 1000K with slower T cooling
#from steps 1001-3100: equilibrium with increasing constraints
#from steps 3101-5000: equilibrium with full constraints
#from steps 5000-19000: recool to low temperature
#from steps 19000-20000: final cooling with short tautp
 &wt type='TEMP0', istep1=0,istep2=10000,value1=25.0,
          value2=600.0,   &end
 &wt type='TEMP0', istep1=10001, istep2=50000, value1=600.0,
          value2=600.0,     &end
 &wt type='TEMP0', istep1=50001, istep2=200000, value1=600.0,
          value2=298.0,     &end
 &wt type='TEMP0', istep1=200001, istep2=400000, value1=298.0,
          value2=0.0,     &end

 &wt type='TAUTP', istep1=0,istep2=10000,value1=0.4,
          value2=0.4,     &end
 &wt type='TAUTP', istep1=10001,istep2=50000,value1=1.0,
          value2=1.0,     &end
```

```
 &wt type='TAUTP', istep1=50001,istep2=200000,value1=1.0,
          value2=0.5,     &end
 &wt type='TAUTP', istep1=200001,istep2=350000,value1=0.5,
          value2=0.5,     &end
 &wt type='TAUTP', istep1=350001,istep2=370000,value1=0.5,
          value2=0.05,     &end
 &wt type='TAUTP', istep1=370001,istep2=400000,value1=0.05,
          value2=0.01,     &end

 &wt type='REST', istep1=0,istep2=10000,value1=0.5,
          value2=1.00,  &end
 &wt type='REST', istep1=10001,istep2=200000,value1=1.00,
          value2=1.75,  &end
 &wt type='REST', istep1=200001,istep2=400000,value1=1.75,
value2=1.0,  &end

 &wt type='END'  &end
LISTOUT=POUT
```

References

(1) Watson, J. D.; Crick, F. H. *Nature* **2003**, *421*, 397-8; discussion 396.

(2) Ghosh, A.; Bansal, M. *Acta Crystallogr D Biol Crystallogr* **2003**, *59*, 620-6.

(3) Basu, H. S.; Feuerstein, B. G.; Zarling, D. A.; Shafer, R. H.; Marton, L. J. *J Biomol Struct Dyn* **1988**, *6*, 299-309.

(4) Leslie, A. G.; Arnott, S.; Chandrasekaran, R.; Ratliff, R. L. *J Mol Biol* **1980**, *143*, 49-72.

(5) Dickerson, R. E.; Drew, H. R.; Conner, B. N.; Wing, R. M.; Fratini, A. V.; Kopka, M. L. *Science* **1982**, *216*, 475-85.

(6) Lu, X. J.; Shakked, Z.; Olson, W. K. *J Mol Biol* **2000**, *300*, 819-40.

(7) Wahl, M. C.; Sundaralingam, M. *Biopolymers* **1997**, *44*, 45-63.

(8) Crawford, J. L.; Kolpak, F. J.; Wang, A. H.; Quigley, G. J.; van Boom, J. H.; van der Marel, G.; Rich, A. *Proc Natl Acad Sci U S A* **1980**, *77*, 4016-20.

(9) Rothenburg, S.; Koch-Nolte, F.; Haag, F. *Immunol Rev* **2001**, *184*, 286-98.

(10) Oh, D. B.; Kim, Y. G.; Rich, A. *Proc Natl Acad Sci U S A* **2002**, *99*, 16666-71.

(11) Franks, L. M., Teich, N.M. *Introduction to the Cellular and Molecular Biology of Cancer*; 3rd ed. ed.; Oxford University Press: New York., 1997.

(12) Lodish H, B. A., Matsudaira P, Kaiser CA, Krieger M, Scott MP, Zipursky SL, Darnell J. *Molecular Biology of the Cell*; 5th ed.; WH Freeman: New York, 2004.

(13) Altieri, F.; Grillo, C.; Maceroni, M.; Chichiarelli, S. *Antioxid Redox Signal* **2008**, *10*, 891-937.

(14) Gupta, R. C.; Lutz, W. K. *Mutat Res* **1999**, *424*, 1-8.

(15) Lindahl, T. *Nature* **1993**, *362*, 709-15.

(16) Finkel, T.; Holbrook, N. J. *Nature* **2000**, *408*, 239-47.

(17) Hoeijmakers, J. H. *Nature* **2001**, *411*, 366-74.

(18) Peltomaki, P. *Mutat Res* **2001**, *488*, 77-85.

(19) Cooper, G. M. *The Cell: A Molecular Approach*; 1st ed.; The American Society for Microbiology: Washingotn, DC., 1997; Vol. 1.

(20) Miller, J. A.; Miller, E. C. *Prog Exp Tumor Res* **1969**, *11*, 273-301.

(21) Sancar, A.; Lindsey-Boltz, L. A.; Unsal-Kacmaz, K.; Linn, S. *Annu Rev Biochem* **2004**, *73*, 39-85.

(22) Himmelstein, M. W.; Acquavella, J. F.; Recio, L.; Medinsky, M. A.; Bond, J. A. *Crit Rev Toxicol* **1997**, *27*, 1-108.

(23) Jackson, M. A.; Stack, H. F.; Rice, J. M.; Waters, M. D. *Mutat Res* **2000**, *463*, 181-213.

(24) Pelz, N.; Dempster, N. M.; Shore, P. R. *J Chromatogr Sci* **1990**, *28*, 230-5.

(25) Brunnemann, K. D.; Kagan, M. R.; Cox, J. E.; Hoffmann, D. *Carcinogenesis* **1990**, *11*, 1863-8.

(26) Swenberg, J. A.; Boysen, G.; Georgieva, N.; Bird, M. G.; Lewis, R. J. *Chem Biol Interact* **2007**, *166*, 78-83.

(27) Huff, J. E.; Melnick, R. L.; Solleveld, H. A.; Haseman, J. K.; Powers, M.; Miller, R. A. *Science* **1985**, *227*, 548-9.

(28) Melnick, R. L.; Huff, J.; Chou, B. J.; Miller, R. A. *Cancer Res* **1990**, *50*, 6592-9.

(29) Melnick, R. L.; Huff, J. E.; Roycroft, J. H.; Chou, B. J.; Miller, R. A. *Environ Health Perspect* **1990**, *86*, 27-36.

(30) Owen, P. E.; Glaister, J. R.; Gaunt, I. F.; Pullinger, D. H. *Am Ind Hyg Assoc J* **1987**, *48*, 407-13.

(31) Agency U. S. E. P., 2002.

(32) IARC *IARC Sci Publ* **1999**, *71*, 109-125.

(33) Ward, J. B., Jr.; Ammenheuser, M. M.; Bechtold, W. E.; Whorton, E. B., Jr.; Legator, M. S. *Environ Health Perspect* **1994**, *102 Suppl 9*, 79-85.

(34) Ward, J. B., Jr.; Ammenheuser, M. M.; Whorton, E. B., Jr.; Bechtold, W. E.; Kelsey, K. T.; Legator, M. S. *Toxicology* **1996**, *113*, 84-90.

(35) Sram, R. J.; Rossner, P.; Peltonen, K.; Podrazilova, K.; Mrackova, G.; Demopoulos, N. A.; Stephanou, G.; Vlachodimitropoulos, D.; Darroudi, F.; Tates, A. D. *Mutat Res* **1998**, *419*, 145-54.

(36) Delzell, E.; Sathiakumar, N.; Hovinga, M.; Macaluso, M.; Julian, J.; Larson, R.; Cole, P.; Muir, D. C. *Toxicology* **1996**, *113*, 182-9.

(37) Meinhardt, T. J.; Lemen, R. A.; Crandall, M. S.; Young, R. J. *Scand J Work Environ Health* **1982**, *8*, 250-9.

(38) Matanoski, G. M.; Schwartz, L. *J Occup Med* **1987**, *29*, 675-80.

(39) Santos-Burgoa, C.; Matanoski, G. M.; Zeger, S.; Schwartz, L. *Am J Epidemiol* **1992 Oct 1**, *136. 136*, 843-54.

(40) Santos-Burgoa, C.; Eden-Wynter, R. A.; Riojas-Rodriguez, H.; Matanoski, G. M. *Ann N Y Acad Sci* **1997 Dec 26**, *837. 837:176-88*, 176-88.

(41) Albertini, R.; Clewell, H.; Himmelstein, M. W.; Morinello, E.; Olin, S.; Preston, J.; Scarano, L.; Smith, M. T.; Swenberg, J.; Tice, R.; Travis, C. *Regul Toxicol Pharmacol* **2003**, *37*, 105-32.

(42) Csanady, G. A.; Filser, J. G. *Arch Toxicol* **1993**, *67*, 227-30.

(43) Duescher, R. J.; Elfarra, A. A. *Arch Biochem Biophys* **1994 Jun**, *311. 311*, 342-9.

(44) Malvoisin, E.; Evrard, E.; Roberfroid, M.; Mercier, M. *J Chromatogr* **1979**, *186*, 81-7.

(45) Seaton, M. J.; Follansbee, M. H.; Bond, J. A. *Carcinogenesis* **1995**, *16*, 2287-93.

(46) Malvoisin, E.; Roberfroid, M. *Xenobiotica* **1982**, *12*, 137-44.

(47) Himmelstein, M. W.; Turner, M. J.; Asgharian, B.; Bond, J. A. *Carcinogenesis* **1994**, *15*, 1479-86.

(48) Himmelstein, M. W.; Asgharian, B.; Bond, J. A. *Toxicol Appl Pharmacol* **1995**, *132*, 281-8.

(49) Cheng, X.; Ruth, J. A. *Drug Metab Dispos* **1993**, *21*, 121-4.

(50) Nauhaus, S. K.; Fennell, T. R.; Asgharian, B.; Bond, J. A.; Sumner, S. C. **1996 Jun**, *9*, 764-73.

(51) Elfarra, A. A.; Moll, T. S.; Krause, R. J.; Kemper, R. A.; Selzer, R. R. *Adv Exp Med Biol* **2001**, *500*, 93-103.

(52) Boogaard, P. J.; Bond, J. A. *Toxicol Appl Pharmacol* **1996**, *141*, 617-27.

(53) Powley, M. W.; Jayaraj, K.; Gold, A.; Ball, L. M.; Swenberg, J. A. *Chem Res Toxicol* **2003**, *16*, 1448-54.

(54) Cochrane, J. E.; Skopek, T. R. *Carcinogenesis* **1994**, *15*, 713-7.

(55) Brookes, P.; Lawley, P. D. *Biochem J* **1961**, *80*, 496-503.

(56) Lawley, P. D.; Brookes, P. *J Mol Biol* **1967**, *25*, 143-60.

(57) Millard, J. T.; White, M. M. *Biochemistry* **1993**, *32*, 2120-4.

(58) Park, S.; Hodge, J.; Anderson, C.; Tretyakova, N. *Chem Res Toxicol* **2004**, *17*, 1638-51.

(59) Valadez, J. G.; Liu, L.; Loktionova, N. A.; Pegg, A. E.; Guengerich, F. P. *Chem Res Toxicol* **2004**, *17*, 972-82.

(60) Goggin, M.; Loeber, R.; Park, S.; Walker, V.; Wickliffe, J.; Tretyakova, N. *Chem Res Toxicol* **2007**, *20*, 839-47.

(61) Vangala, R. R.; Laib, R. J.; Bolt, H. M. *Arch Toxicol* **1993**, *67*, 34-8.

(62) Ristau, C.; Deutschmann, S.; Laib, R. J.; Ottenwalder, H. *Arch Toxicol* **1990**, *64*, 343-4.

(63) Thornton-Manning, J. R.; Dahl, A. R.; Bechtold, W. E.; Griffith, W. C., Jr.; Henderson, R. F. *Toxicology* **1997**, *123*, 125-34.

(64) Thornton-Manning, J. R.; Dahl, A. R.; Bechtold, W. E.; Griffith, W. C., Jr.; Henderson, R. F. *Carcinogenesis* **1995**, *16*, 1723-31.

(65) Cochrane, J. E.; Skopek, T. R. *IARC Sci Publ* **1993**, 195-204.

(66) Cochrane, J. E.; Skopek, T. R. *Carcinogenesis* **1994**, *15*, 719-23.

(67) Koc, H.; Tretyakova, N. Y.; Walker, V. E.; Henderson, R. F.; Swenberg, J. A. *Chem Res Toxicol* **1999**, *12*, 566-74.

(68) Abdel-Rahman, S. Z.; Ammenheuser, M. M.; Ward, J. B., Jr. *Carcinogenesis* **2001**, *22*, 415-23.

(69) Abdel-Rahman, S. Z.; El-Zein, R. A.; Ammenheuser, M. M.; Yang, Z.; Stock, T. H.; Morandi, M.; Ward, J. B., Jr. *Environ Mol Mutagen* **2003**, *41*, 140-6.

(70) Tretyakova, N.; Lin, Y.; Sangaiah, R.; Upton, P. B.; Swenberg, J. A. *Carcinogenesis* **1997**, *18*, 137-47.

(71) Selzer, R. R.; Elfarra, A. A. *Chem Res Toxicol* **1996**, *9*, 126-32.

(72) Tretyakova, N.; Sangaiah, R.; Yen, T. Y.; Swenberg, J. A. *Chem Res Toxicol* **1997**, *10*, 779-85.

(73) Zhang, X. Y.; Elfarra, A. A. *Chem Res Toxicol* **2004**, *17*, 521-8.

(74) Powley, M. W.; Li, Y.; Upton, P. B.; Walker, V. E.; Swenberg, J. A. *Carcinogenesis* **2005**, *26*, 1573-80.

(75) Selzer, R. R.; Elfarra, A. A. *Arch Biochem Biophys* **1997**, *343*, 63-72.

(76) Selzer, R. R.; Elfarra, A. A. *Carcinogenesis* **1997**, *18*, 1993-8.

(77) Tretyakova, N.; Sangaiah, R.; Yen, T. Y.; Gold, A.; Swenberg, J. A. *Chem Res Toxicol* **1997**, *10*, 1171-9.

(78) Selzer, R. R.; Elfarra, A. A. *Chem Res Toxicol* **1996**, *9*, 875-81.

(79) Selzer, R. R.; Elfarra, A. A. *Carcinogenesis* **1999**, *20*, 285-92.

(80) Tretyakova, N.; Chiang, S. Y.; Walker, V. E.; Swenberg, J. A. In *J Mass Spectrom* 1998; Vol. 33, p 363-76.

(81) Carmical, J. R.; Nechev, L. V.; Harris, C. M.; Harris, T. M.; Lloyd, R. S. *Environ Mol Mutagen* **2000**, *35*, 48-56.

(82) Nechev, L. V.; Zhang, M.; Tsarouhtsis, D.; Tamura, P. J.; Wilkinson, A. S.; Harris, C. M.; Harris, T. M. *Chem Res Toxicol* **2001**, *14*, 379-88.

(83) Cochrane, J. E.; Skopek, T. R. *Carcinogenesis* **1994 Apr**, *15. 15*, 713-7.

(84) Verly, W. G.; Brakier, L.; Feit, P. W. *Biochimica et Biophysica Acta* **1971**, *228*, 400-406.

(85) Matagne, R. *Mutat Res* **1969**, *7*, 241-7.

(86) Schmidmaier, R.; Oellerich, M.; Baumgart, J.; Emmerich, B.; Meinhardt, G. *Exp Hematol* **2004**, *32*, 76-86.

(87) Hartley, J. A.; O'Hare, C. C.; Baumgart, J. *Br J Cancer* **1999**, *79*, 264-6.

(88) Runowicz, C. D.; Fields, A. L.; Goldberg, G. L. *Cancer* **1995**, *76*, 2028-33.

(89) Qian, C.; Dipple, A. *Chem Res Toxicol* **1995**, *8*, 389-95.

(90) Kim, H. Y.; Finneman, J. I.; Harris, C. M.; Harris, T. M. *Chem Res Toxicol* **2000 Jul**, *13*, 625-37.

(91) Fernandes, P. H.; Hackfeld, L. C.; Kozekov, I. D.; Hodge, R. P.; Lloyd, R. S. *Chem Res Toxicol* **2006**, *19*, 968-76.

(92) Kanuri, M.; Nechev, L. V.; Tamura, P. J.; Harris, C. M.; Harris, T. M.; Lloyd, R. S. *Chem Res Toxicol* **2002**, *15*, 1572-80.

(93) Carmical, J. R.; Zhang, M.; Nechev, L.; Harris, C. M.; Harris, T. M.; Lloyd, R. S. *Chem Res Toxicol* **2000**, *13*, 430.

(94) Fernandes, P. H.; Lloyd, R. S. *Mutat Res* **2007**, *625*, 40-9.

(95) Drenth, J. *Principles of Protein X-Ray Crystallography*; 2nd ed.; Springer-Verlag New York, Inc: New York, 1999.

(96) Yaffe, M. B. *Crit Care Med* **2005**, *33*, S435-40.

(97) Snyder, D. A.; Chen, Y.; Denissova, N. G.; Acton, T.; Aramini, J. M.; Ciano, M.; Karlin, R.; Liu, J.; Manor, P.; Rajan, P. A.; Rossi, P.; Swapna, G. V.; Xiao, R.; Rost, B.; Hunt, J.; Montelione, G. T. *J Am Chem Soc* **2005**, *127*, 16505-11.

(98) Yee, A. A.; Savchenko, A.; Ignachenko, A.; Lukin, J.; Xu, X.; Skarina, T.; Evdokimova, E.; Liu, C. S.; Semesi, A.; Guido, V.; Edwards, A. M.; Arrowsmith, C. H. *J Am Chem Soc* **2005**, *127*, 16512-7.

(99) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York 1986.

(100) Patel, D. J.; Pardi, A.; Itakura, K. *Science* **1982**, *216*, 581-590.

(101) Roy, S.; Redfield, A. G. *Nucleic Acids Res.* **1981**, *9*, 7073-7083.

(102) Patel, D. J. *Proc.Natl.Acad.Sci.USA* **1982**, *79*, 6424-6428.

(103) Broido, M. S.; Zon, G.; James, T. L. *Biochem Biophys Res Commun* **1984**, *119*, 663-70.

(104) Marion, D.; Wuthrich, K. *Biochem Biophys Res Commun* **1983**, *113*, 967-74.

(105) Rance, M.; Sorensen, O. W.; Bodenhausen, G.; Wagner, G.; Ernst, R. R.; Wuthrich, K. *Biochem Biophys Res Commun* **1983**, *117*, 479-85.

(106) Zhou, N.; Manogaran, S.; Zon, G.; James, T. L. *Biochemistry* **1988**, *27*, 6013-20.

(107) Piantini, U.; Sorensen, O. W.; Ernst, R. R. *Journal of the American Chemical Society* **1982**, *104*, 6800-6801.

(108) Saenger, W. *Principles of nucleic acid structure*; Springer-Verlag: New York 1984.

(109) Wang, H.; Zuiderweg, E. R. P.; Glick, G. D. *Journal of the American Chemical Society* **1995**, *117*, 2981-2991.

(110) Geen, H.; Freeman, R. *Journal of Magnetic Resonance* **1991**, *93*, 93-141.

(111) Lankhorst, P. P.; Haasnoot, C. A. G.; Erkelens, C.; Altona, C. *Nucleic Acids Research* **1984**, *12*, 5419-5428.

(112) Roongta, V. A.; Jones, C. R.; Gorenstein, D. G. *Biochemistry* **1990**, *29*, 5245-5258.

(113) Wu, A. A.; Crerner, D. *International Journal of Molecular Sciences* **2003**, *4*, 158-192.

(114) Borgias, B. A.; James, T. L. *Methods Enzymol* **1989**, *176*, 169-83.

(115) Borgias, B. A.; James, T. L. *Journal of Magnetic Resonance* **1990**, *87*, 475-487.

(116) Liu, H.; Spielmann, H. P.; Ulyanov, N. B.; Wemmer, D. E.; James, T. L. *J Biomol NMR* **1995**, *6*, 390-402.

(117) James, T. L.; Basus, V. J. *Annu Rev Phys Chem* **1991**, *42*, 501-42.

(118) Schmitz, U.; Pearlman, D. A.; James, T. L. *J.Mol.Biol.* **1991**, *221*, 271-292.

(119) Madrid, M.; Llinas, E.; Llinas, M. *J.Magn.Reson.* **1991**, *93*, 329-346.

(120) Lu, X. J.; Olson, W. K. *Nucleic Acids Res* **2003**, *31*, 5108-21.

(121) Lavery, R.; Sklenar, H. *J Biomol Struct Dyn* **1988**, *6*, 63-91.

(122) Stofer, E.; Lavery, R. *Biopolymers* **1994**, *34*, 337-46.

(123) Ohmori, H.; Friedberg, E. C.; Fuchs, R. P.; Goodman, M. F.; Hanaoka, F.; Hinkle, D.; Kunkel, T. A.; Lawrence, C. W.; Livneh, Z.; Nohmi, T.; Prakash, L.; Prakash, S.; Todo, T.; Walker, G. C.; Wang, Z.; Woodgate, R. *Mol Cell* **2001**, *8*, 7-8.

(124) Yang, W. *Curr Opin Struct Biol* **2003**, *13*, 23-30.

(125) Wang, Y.; Arora, K.; Schlick, T. *Protein Sci* **2006**, *15*, 135-51.

(126) Bebenek, K.; Kunkel, T. A. *Adv Protein Chem* **2004**, *69*, 137-65.

(127) Yang, W.; Woodgate, R. *Proc Natl Acad Sci U S A* **2007**, *104*, 15591-8.

(128) Kulaeva, O. I.; Koonin, E. V.; McDonald, J. P.; Randall, S. K.; Rabinovich, N.; Connaughton, J. F.; Levine, A. S.; Woodgate, R. *Mutat Res* **1996**, *357*, 245-53.

(129) Kannouche, P.; Broughton, B. C.; Volker, M.; Hanaoka, F.; Mullenders, L. H.; Lehmann, A. R. *Genes Dev* **2001**, *15*, 158-72.

(130) Wagner, J.; Fujii, S.; Gruz, P.; Nohmi, T.; Fuchs, R. P. *EMBO Rep* **2000**, *1*, 484-8.

(131) Kannouche, P. L.; Wing, J.; Lehmann, A. R. *Mol Cell* **2004**, *14*, 491-500.

(132) Vidal, A. E.; Kannouche, P.; Podust, V. N.; Yang, W.; Lehmann, A. R.; Woodgate, R. *J Biol Chem* **2004**, *279*, 48360-8.

(133) Kannouche, P.; Fernandez de Henestrosa, A. R.; Coull, B.; Vidal, A. E.; Gray, C.; Zicha, D.; Woodgate, R.; Lehmann, A. R. *Embo J* **2003**, *22*, 1223-33.

(134) Zhou, B. L.; Pata, J. D.; Steitz, T. A. *Mol Cell* **2001**, *8*, 427-37.

(135) Wagner, J.; Gruz, P.; Kim, S. R.; Yamada, M.; Matsui, K.; Fuchs, R. P.; Nohmi, T. *Mol Cell* **1999**, *4*, 281-6.

(136) Ohashi, E.; Ogi, T.; Kusumoto, R.; Iwai, S.; Masutani, C.; Hanaoka, F.; Ohmori, H. *Genes Dev* **2000**, *14*, 1589-94.

(137) Strauss, B. S.; Roberts, R.; Francis, L.; Pouryazdanparast, P. *J Bacteriol* **2000**, *182*, 6742-50.

(138) Zhang, Y.; Yuan, F.; Wu, X.; Wang, M.; Rechkoblit, O.; Taylor, J. S.; Geacintov, N. E.; Wang, Z. *Nucleic Acids Res* **2000**, *28*, 4138-46.

(139) Washington, M. T.; Johnson, R. E.; Prakash, L.; Prakash, S. *Proc Natl Acad Sci U S A* **2001**, *98*, 8355-60.

(140) Tissier, A.; McDonald, J. P.; Frank, E. G.; Woodgate, R. *Genes Dev* **2000**, *14*, 1642-50.

(141) Zhang, Y.; Yuan, F.; Wu, X.; Wang, Z. *Mol Cell Biol* **2000**, *20*, 7099-108.

(142) Goodman, M. F. *Annu Rev Biochem* **2002**, *71*, 17-50.

(143) Silvian, L. F.; Toth, E. A.; Pham, P.; Goodman, M. F.; Ellenberger, T. *Nat Struct Biol* **2001**, *8*, 984-9.

(144) Ling, H.; Boudsocq, F.; Woodgate, R.; Yang, W. *Cell* **2001**, *107*, 91-102.

(145) Trincao, J.; Johnson, R. E.; Escalante, C. R.; Prakash, S.; Prakash, L.; Aggarwal, A. K. *Mol Cell* **2001**, *8*, 417-26.

(146) Nair, D. T.; Johnson, R. E.; Prakash, S.; Prakash, L.; Aggarwal, A. K. *Nature* **2004**, *430*, 377-80.

(147) Uljon, S. N.; Johnson, R. E.; Edwards, T. A.; Prakash, S.; Prakash, L.; Aggarwal, A. K. *Structure* **2004**, *12*, 1395-404.

(148) Ling, H.; Boudsocq, F.; Plosky, B. S.; Woodgate, R.; Yang, W. *Nature* **2003**, *424*, 1083-7.

(149) Rechkoblit, O.; Malinina, L.; Cheng, Y.; Kuryavyi, V.; Broyde, S.; Geacintov, N. E.; Patel, D. J. *PLoS Biol* **2006**, *4*, e11.

(150) Sawaya, M. R.; Prasad, R.; Wilson, S. H.; Kraut, J.; Pelletier, H. *Biochemistry* **1997**, *36*, 11205-15.

(151) Beard, W. A.; Shock, D. D.; Vande Berg, B. J.; Wilson, S. H. *J Biol Chem* **2002**, *277*, 47393-8.

(152) Doublie, S.; Tabor, S.; Long, A. M.; Richardson, C. C.; Ellenberger, T. *Nature* **1998**, *391*, 251-8.

(153) Franklin, M. C.; Wang, J.; Steitz, T. A. *Cell* **2001**, *105*, 657-67.

(154) Kiefer, J. R.; Mao, C.; Braman, J. C.; Beese, L. S. *Nature* **1998**, *391*, 304-7.

(155) Li, Y.; Korolev, S.; Waksman, G. *Embo J* **1998**, *17*, 7514-25.

(156) Vaisman, A.; Ling, H.; Woodgate, R.; Yang, W. *Embo J* **2005**, *24*, 2957-67.

(157) Mizukami, S.; Kim, T. W.; Helquist, S. A.; Kool, E. T. *Biochemistry* **2006**, *45*, 2772-8.

(158) Potapova, O.; Chan, C.; DeLucia, A. M.; Helquist, S. A.; Kool, E. T.; Grindley, N. D.; Joyce, C. M. *Biochemistry* **2006**, *45*, 890-8.

(159) Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C. *Nature* **1958**, *181*, 662-6.

(160) Kendrew, J. C.; Dickerson, R. E.; Strandberg, B. E.; Hart, R. G.; Davies, D. R.; Phillips, D. C.; Shore, V. C. *Nature* **1960**, *185*, 422-7.

(161) Hodgkin, D. C. *Science* **1965**, *150*, 979-88.

(162) Blake, C. C.; Johnson, L. N.; Mair, G. A.; North, A. C.; Phillips, D. C.; Sarma, V. R. *Proc R Soc Lond B Biol Sci* **1967**, *167*, 378-88.

(163) Blake, C. C.; Mair, G. A.; North, A. C.; Phillips, D. C.; Sarma, V. R. *Proc R Soc Lond B Biol Sci* **1967**, *167*, 365-77.

(164) Blow, D. M.; Birktoft, J. J.; Hartley, B. S. *Nature* **1969**, *221*, 337-40.

(165) Tsukada, H.; Blow, D. M. *J Mol Biol* **1985**, *184*, 703-11.

(166) Blow, D. M.; Wright, C. S.; Kukla, D.; Ruhlmann, A.; Steigemann, W.; Huber, R. *J Mol Biol* **1972**, *69*, 137-44.

(167) Blow, D. M.; Steitz, T. A. *Annu Rev Biochem* **1970**, *39*, 63-100.

(168) Rees, D. C.; Lewis, M.; Lipscomb, W. N. *J Mol Biol* **1983**, *168*, 367-87.

(169) Lipscomb, W. N. *Proc Natl Acad Sci U S A* **1980**, *77*, 3875-8.

(170) Lipscomb, W. N.; Reeke, G. N., Jr.; Hartsuck, J. A.; Quiocho, F. A.; Bethge, P. H. *Philos Trans R Soc Lond B Biol Sci* **1970**, *257*, 177-214.

(171) Eklund, H.; Nordstrom, B.; Zeppezauer, E.; Soderlund, G.; Ohlsson, I.; Boiwe, T.; Branden, C. I. *FEBS Lett* **1974**, *44*, 200-4.

(172) McPherson, A. *Crystallization of biological macromolecules* Cold Spring Harbor: New York, 1999.

(173) Giegé, A. D. a. R. *Crystallization of nucleic acids and proteins : a practical approach*; 2nd ed.; Oxford, UK New York 1999.

(174) Ealick, S. E.; Walter, R. L. *Current Opinion in Structural Biology* **1993**, *3*, 725-736.

(175) Moy, J. P. *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment* **1994**, *348*, 641-644.

(176) Gruner, S. M. *Current Opinion in Structural Biology* **1994**, *4*, 765-769.

(177) Dauter, Z. *Macromolecular Crystallography, Pt A* **1997**, *276*, 326-344.

(178) Rossmann, M. G. *Acta Crystallogr A* **1990**, *46 (Pt 2)*, 73-82.

(179) Ke, H. M. *Macromolecular Crystallography, Pt A* **1997**, *276*, 448-461.

(180) Watenpaugh, K. D. *Methods in Enzymology* **1985**, *115*, 3-15.

(181) Walsh, M. A.; Evans, G.; Sanishvili, R.; Dementieva, I.; Joachimiak, A. *Acta Crystallogr D Biol Crystallogr* **1999**, *55*, 1726-32.

(182) Johnson, R. E.; Prakash, L.; Prakash, S. *Proc Natl Acad Sci U S A* **2005**, *102*, 12359-64.

(183) Ling, H.; Boudsocq, F.; Woodgate, R.; Yang, W. *Mol Cell* **2004**, *13*, 751-62.

(184) Ling, H.; Sayer, J. M.; Plosky, B. S.; Yagi, H.; Boudsocq, F.; Woodgate, R.; Jerina, D. M.; Yang, W. *Proc Natl Acad Sci U S A* **2004**, *101*, 2265-9.

(185) Wang, L.; Wu, M.; Yan, S. F.; Patel, D. J.; Geacintov, N. E.; Broyde, S. *Chem Res Toxicol* **2005**, *18*, 441-56.

(186) Zang, H.; Goodenough, A. K.; Choi, J. Y.; Irimia, A.; Loukachevitch, L. V.; Kozekov, I. D.; Angel, K. C.; Rizzo, C. J.; Egli, M.; Guengerich, F. P. *J Biol Chem* **2005**, *280*, 29750-64.

(187) Zang, H.; Irimia, A.; Choi, J. Y.; Angel, K. C.; Loukachevitch, L. V.; Egli, M.; Guengerich, F. P. *J Biol Chem* **2006**, *281*, 2358-72.

(188) Eoff, R. L.; Irimia, A.; Angel, K. C.; Egli, M.; Guengerich, F. P. *J Biol Chem* **2007**, *282*, 19831-43.

(189) Wang, L.; Broyde, S. *Nucleic Acids Res* **2006**, *34*, 785-95.

(190) Zang, H.; Chowdhury, G.; Angel, K. C.; Harris, T. M.; Guengerich, F. P. *Chem Res Toxicol* **2006**, *19*, 859-67.

(191) Chandani, S.; Loechler, E. L. *J Mol Graph Model* **2007**, *25*, 658-70.

(192) Bauer, J.; Xing, G.; Yagi, H.; Sayer, J. M.; Jerina, D. M.; Ling, H. *Proc Natl Acad Sci U S A* **2007**, *104*, 14905-10.

(193) Zhang, L.; Rechkoblit, O.; Wang, L.; Patel, D. J.; Shapiro, R.; Broyde, S. *Nucleic Acids Res* **2006**, *34*, 3326-37.

(194) Eoff, R. L.; Irimia, A.; Egli, M.; Guengerich, F. P. *J Biol Chem* **2007**, *282*, 1456-67.

(195) Eoff, R. L.; Angel, K. C.; Egli, M.; Guengerich, F. P. *J Biol Chem* **2007**, *282*, 13573-84.

(196) Irimia, A.; Eoff, R. L.; Pallan, P. S.; Guengerich, F. P.; Egli, M. *J Biol Chem* **2007**, *282*, 36421-33.

(197) Wang, Y.; Musser, S. K.; Saleh, S.; Marnett, L. J.; Egli, M.; Stone, M. P. *Biochemistry* **2008**, *47*, 7322-34.

(198) Boudsocq, F.; Kokoska, R. J.; Plosky, B. S.; Vaisman, A.; Ling, H.; Kunkel, T. A.; Yang, W.; Woodgate, R. *J Biol Chem* **2004**, *279*, 32932-40.

(199) Fiala, K. A.; Suo, Z. *Biochemistry* **2004**, *43*, 2116-25.

(200) Fiala, K. A.; Suo, Z. *Biochemistry* **2004**, *43*, 2106-15.

(201) Fiala, K. A.; Suo, Z. *J Biol Chem* **2007**, *282*, 8199-206.

(202) Fiala, K. A.; Hypes, C. D.; Suo, Z. *J Biol Chem* **2007**, *282*, 8188-98.

(203) Yang, W.; Lee, J. Y.; Nowotny, M. *Mol Cell* **2006**, *22*, 5-13.

(204) Leuratti, C.; Jones, N. J.; Marafante, E.; Kostiainen, R.; Peltonen, K.; Waters, R. *Carcinogenesis* **1994**, *15*, 1903-10.

(205) Borer, P. N.; Kan, L. S.; Ts'o, P. O. *Biochemistry* **1975**, *14*, 4847-63.

(206) Piotto, M.; Saudek, V.; Sklenar, V. *J Biomol NMR* **1992**, *2*, 661-5.

(207) Bax, A.; Davis, D. G. *J.Magn.Reson.* **1985**, *65*, 355-360.

(208) James, T. L. *Current Opinion Strut.Biol.* **1991**, *1*, 1042-1053.

(209) Keepers, J.; Kollman, P. A.; James, T. L. *Biopolymers* **1984**, *23*, 2499-511.

(210) Keepers, J. W.; James, T. L. *J.Magn.Reson.* **1984**, *57*, 404-426.

(211) Millard, J. T.; Hanly, T. C.; Murphy, K.; Tretyakova, N. *Chem Res Toxicol* **2006**, *19*, 16-9.

(212) Schmidt, J. M. *Molecular Physics* **1998**, *95*, 809-826.

(213) Arnott, S.; Hukins, D. W. *Biochem Biophys Res Commun* **1972**, *47*, 1504-9.

(214) Case, D. A., Pearlman, D.A., Caldwell, J.W., Cheatham, T.E. III, Wang, J., Ross, W.S., Simmerling, C.L., Darden, T.A., Merz, K.M., Stanton, R.V., Cheng, A.L., Vincent, JJ., Crowley, M., Tsui, V., Gohlke, H., Radmer, R.J., Duan, Y., Pitera, J., Massova, I., Seibel, G.L., Singh, U.C., Weiner, P.K., Kollman, P.A.; 7.0 ed.; University of California: San Francisco, CA, 2002.

(215) Clore, G. M.; Brunger, A. T.; Karplus, M.; Gronenborn, A. M. *J.Mol.Biol.* **1986**, *191*, 523-551.

(216) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J.Comp.Phys.* **1977**, *23*, 327-341.

(217) Bashford, D.; Case, D. A. *Annu Rev Phys Chem* **2000**, *51*, 129-52.

(218) Tsui, V.; Case, D. A. *Biopolymers* **2000**, *56*, 275-91.

(219) Gaussian; Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.

(220) Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. *Protein Sci* **1995**, *4*, 2411-23.

(221) Otwinowski, Z.; Minor, W. *Macromolecular Crystallography, Pt A* **1997**, *276*, 307-326.

(222) *Acta Crystallogr D Biol Crystallogr* **1994**, *50*, 760-3.

(223) Navaza, J. *Acta Crystallographica Section A* **1994**, *50*, 157-163.

(224) Fujinaga, M.; Read, R. J. *Journal of Applied Crystallography* **1987**, *20*, 517-521.

(225) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice,

L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr D Biol Crystallogr* **1998**, *54*, 905-21.

(226) Cambillau, C., and Roussel, A; TURBO-FRODO, Version OpenGL.1, University Aix-Marseille II, Marseille, France.

(227) Laskowski, R. A.; Macarthur, M. W.; Moss, D. S.; Thornton, J. M. *Journal of Applied Crystallography* **1993**, *26*, 283-291.

(228) DeLano, W. L.; The PyMOL Molecular Graphics System. DeLano Scientific LLC, San Carlos, CA, USA. http://www.pymol.org.

(229) Merritt, W. K.; Nechev, L. V.; Scholdberg, T. A.; Dean, S. M.; Kiehna, S. E.; Chang, J. C.; Harris, T. M.; Harris, C. M.; Lloyd, R. S.; Stone, M. P. *Biochemistry* **2005**, *44*, 10081-92.

(230) Malumbres, M.; Barbacid, M. *Nat Rev Cancer* **2003**, *3*, 459-65.

(231) Bos, J. L. *Cancer Res.* **1989**, *49*, 4682-4689.

(232) Almoguera, C.; Shibata, D.; Forrester, K.; Martin, J.; Arnheim, N.; Perucho, M. *Cell* **1988**, *53*, 549-54.

(233) Rodenhuis, S.; van de Wetering, M. L.; Mooi, W. J.; Evers, S. G.; van Zandwijk, N.; Bos, J. L. *N Engl J Med* **1987**, *317*, 929-35.

(234) Forrester, K.; Almoguera, C.; Han, K.; Grizzle, W. E.; Perucho, M. *Nature* **1987**, *327*, 298-303.

(235) Bos, J. L.; Fearon, E. R.; Hamilton, S. R.; Verlaan-de Vries, M.; van Boom, J. H.; van der Eb, A. J.; Vogelstein, B. *Nature* **1987**, *327*, 293-7.

(236) Patel, D. J.; Shapiro, L.; Hare, D. *Q Rev Biophys* **1987**, *20*, 35-112.

(237) Reid, B. R. *Q Rev Biophys* **1987**, *20*, 1-34.

(238) Feng, B.; Stone, M. P. *Chem Res Toxicol* **1995**, *8*, 821-32.

(239) Jelitto, B.; Vangala, R. R.; Laib, R. J. *Arch Toxicol* **1989**, *Suppl. 13*, 246-249.

(240) Cochrane, J. E.; Skopek, T. R. *Carcinogenesis* **1994**, *15*, 713-717.

(241) Cochrane, J. E.; Skopek, T. R. *Carcinogenesis* **1994**, *15*, 719-723.

(242) Merritt, W. K.; Scholdberg, T. A.; Nechev, L. V.; Harris, T. M.; Harris, C. M.; Lloyd, R. S.; Stone, M. P. *Chem Res Toxicol* **2004**, *17*, 1007-19.

(243) Recio, L.; Saranko, C. J.; Steen, A. M. *Res Rep Health Eff Inst* **2000**, 49-87; discussion 141-9.

(244) Recio, L.; Steen, A. M.; Pluta, L. J.; Meyer, K. G.; Saranko, C. J. *Chem Biol Interact* **2001**, *135-136*, 325-41.

(245) Sisk, S. C.; Pluta, L. J.; Bond, J. A.; Recio, L. *Carcinogenesis* **1994**, *15*, 471-7.

(246) Recio, L.; Meyer, K. G. *Environ Mol Mutagen* **1995**, *26*, 1-8.

(247) Steen, A. M.; Meyer, K. G.; Recio, L. *Mutagenesis* **1997**, *12*, 61-7.

(248) Goodrow, T. L.; Nichols, W. W.; Storer, R. D.; Anderson, M. W.; Maronpot, R. R. *Carcinogenesis* **1994**, *15*, 2665-7.

(249) Xu, W.; Merritt, W. K.; Nechev, L. V.; Harris, T. M.; Harris, C. M.; Lloyd, R. S.; Stone, M. P. *Chem Res Toxicol* **2007**, *20*, 187-98.