OXIDATION OF CYTOSINE IN DNA

By

MARTA WIESŁAWA SZULIK

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

May 2014

Nashville, TN

Approved:

Professor Michael P. Stone, Ph.D.

Professor Martin Egli, Ph.D.

Professor Terry P. Lybrand, Ph.D.

Professor Carmelo Rizzo, Ph.D.

To my family,

lovely and wonderful Mom, Dad, Brother

and to my beloved Sunshine

ACKNOWLEDGEMENTS

I would like to thankfully acknowledge my advisor Dr. Michael P. Stone, for giving me the greatest opportunity to perform research in his laboratory. For his guidance, help and the patience during this journey. He made me grow through this voyage to the scientist I am today. Thank you so much for encouragement, support and many great words of advice.

I am also very grateful to members of my PhD committee. Dr. Martin Egli, Dr. Carmelo Rizzo and Dr. Terry Lybrand you were so generous with time and advice both during and outside our formal committee meetings, and I greatly appreciate the different perspectives you have brought to this work and my scientific development.

Throughout my graduate career I had the great pleasure to work with a number of great collaborators. My NMR guru – Dr. Markus Voehler, Director of the Biomolecular NMR Facility, who not only taught me almost everything I know today about the biomolecular NMR, but always had the time and patience to help me with the experimental setup. Markus, I so appreciate your advice, expertise, and all discussions we had, I could not have done a huge part of this work without your unending help. I would also like to thank Dr. Don Stec for help and guidance in the NMR area.

I would like to thank Dr. Pradeep Pallan for guiding me in crystallization setup, data collection and refinement. Special thanks go to the Structural Biology Center team under the supervision of Dr. Andrzej Joachimiak at Argonne National Laboratory: Dr. Jurek Osipiuk, Dr. Robcio Jedrzejczak and special thanks to Dr. Bogi Nocek

for the help and mentoring during X-ray data collection and crystal structure refinement, as well as long and always exciting scientific brain-storming discussions.

This work would not have been accomplished without the tremendous help and engagement of Dr. Surajit Banerjee. I would like to acknowledge Dr. Barry Gold, Dr. Luis A. Marky, and members of their laboratories for providing samples and additional data results and analysis. I would like to thank Dr. Brandt Eichman and Sonja Brooks for helpful discussions during the development of this project. I would also like to thank Dr. Carmelo Rizzo for the opportunity to work in his laboratory and learn oligo synthesis techniques from his group members: Edward Hawkins, and Dr. Plamen Christov.

It has been a tremendous pleasure and fun to work with the Stone Lab during the past five years. A special thank goes to Ewa Kowal, who has been a great friend and wonderful colleague. I am so grateful to past and current members of the Stone group Kyle Brown, Surajit Banerjee, Hai Huang, Sarah Musser, Liang Li, Kallie Stavros, Drew Kellum for their help, advice and friendship. I also had the privilege to work with very talented undergraduate student – Patrick S. Donahue, thank you so much for your help and understanding. I would also like to acknowledge Ms. Sandra Ford for all the help, unending patience, support and friendship.

This work would not have been possible without the financial support of the National Institutes of Health, the Center in Molecular Toxicology, Center for Structural Biology, and Department of Chemistry at Vanderbilt University. I would like to thank the Vanderbilt Graduate School for providing travel grants to attend scientific conferences.

To all my friends, old and new, especially wszyscy Wolscy, Beierowie, pani Bożenka, ciocia Ania z Wiednia z Hansem, Agą oraz Alą, Klawisz, Ewa Karolak, Marta Szelążek, Ola Jankowiak, Agnieszka Bodzioch, Inka Krzyżanowska, Jacek Pecyna, Kasia z Hubertem, Ewa z Michałem, Becky and Kuba for all your friendships, greatest support and cheering me up constantly. Thank you so much for being such great friends.

I would not be here if it were not for the support of my family. Thank you Dad. Thank you Mom, I am so grateful for being so inspiring to me, for your love, support and encouragement. You have always been there for me. Marcin, you are the best brother I could have and anyone could hope for. Thank you so much for being so wonderful, I am so proud to be your sister. Thank you my in-laws, Mom Renia and Dad Antek, Piotr, for believing in me and support.

Most importantly, special thanks and love go to my Sunshine, to my husband. No one has been more important to me in this journey. Słonko, thank you so so much for you love, understanding, patience and support. For giving me all the strength. It would not be possible without you. You are the best husband, partner and friend ever.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

1D	One dimensional
2D	Two-dimensional
3D	Three-dimensional
5caC	5-carboxyl-2'-deoxycytidine
5fC	5-formyl-2'-deoxycytidine
5hmC	5-hydroxymethyl-2'-deoxycytidine
5mC	5-methyl-2'-deoxycytidine
Å	Angstrom
A	Adenine
AE-HPLC	Anionic-exchange high performance liquid chromatography
AID	Activation-induced deaminase
AlkB	Oxidative demethylase of N1-methyladenine or N3-methylcytosine DNA lesions
AP	Apurinic/apyrimidinic/abasic site
APOBEC	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like
BER	Base excision repair
C	Cytosine
C	Celsius
CH_3CN	Acetonitrile
COSY	Correlation spectroscopy
D_2O	Deuterated/heavy water
dA	Deoxyadenosine

dC	Deoxycytidine
DDD	Dickerson Drew Dodecamer
DFT	Density functional theory
dG	Deoxyguanosine
DMSO	Dimethyl sulfoxide
DNA	Deoxyrybonucleic acid
DNMT	DNA methyl transferase
DQF-COSY	Double-quantum filter correlation spectroscopy
DSC	Differential scanning calorimetry
dT	Deoxythymidine
eq.	Equation
EM	Electromagnetic
ES	Embryonic stem
EX1	Exchange regime 1
EX2	Exchange regime 2
G	Guanine
H_2O	Water
K	Kelvin
k_0	Rate of exchange from the open state in the absence of ammonia
k_B	Rate constant for exchange catalysis
kcal	kilocalorie
k_{cl}	Rate for base-pair closing

k_D	Bimolecular collision rate between the imino group and proton acceptor in the open state
k_{ex}	Chemical Exchange rate
k_{op}	Rate for base-pair opening
LSQ	Least squares quadratic
MAD	Multiple anomalous wavelength dispersion
MALDI-TOF	Matrix assisted laser desorption ionization time-of-flight mass spectrometry
MeCP2	Methyl CpG binding protein 2
mg	Miligram
MHz	Megahertz
min	Minute
mL	Mililiter
mM	Milimolar
mol	Mole
MPD	2-methyl-2,4-pentadienol
MR	Molecular replacement
mRNA	Messenger ribonucleic acid
Na_2EDTA	Sodium ethylenediaminetetraacetate
Na_2HPO_4	Sodium phosphate dibasic
NaCl	Sodium chloride
NaH_2PO_4	Sodium phosphate monobasic
NaN_3	Sodium azide
NaPi	Sodium phosphate buffer

NE-CAT	Northeastern Collaborative Access Team
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect Spectroscopy
PDB	Protein Data Bank
pK_a	Acid dissociation constant
Pol	Polymerase
ppm	Parts per million
rMD	Restrained molecular dynamic
rmsd	Root-mean-square deviation
RNA	Ribonucleic Acid
RP-HPLC	Reversed-phase high performance liquid chromatography
s	Second
SAD	Single anomalous wavelength dispersion
SAM	S-adenosylmethionine
s.d.	Standard deviation
T	Thymine
T_1	Longitudinal relaxation
TDG	Thymine DNA glycosylase
TEOA	Triethanolamine
TET	Ten-eleven translocation enzyme
T_M	Melting temperature
TOCSY	Total correlation spectroscopy

U	Uracil
UNG	Uracil DNA glycosylase
UV	Ultraviolet
VIS	Visible
αK_{op}	Equilibrium constant for base pair opening
ΔC_p	Heat capacity
ΔG	Gibbs free energy
ΔH	Enthalpy
ΔS	Entropy
µg	Microgram
µL	Microliter
µM	Micromolar
σ_m	Electronic substituent constant

CHAPTER 1

INTRODUCTION

DNA structure

DNA, deoxyribonucleic acid, is the most important biomacromolecule as it carries genetic information for the development and functioning of all living organisms. It was first observed by the Swiss chemist Friedrich Miescher in the late 1800s (*1*). But nearly a century passed from that discovery until researchers unraveled the structure of the DNA molecule and realized its central importance to biology. In the decades following Miescher's discovery, other scientists in 1944, Avery, MacLeord and McCarty, published their discovery that DNA is the “fundamental unit” of genetic information, based on the results from the experiments on the pathogenicity of pneumococci (*2*). Soon after that, Phoebus Levene and Erwin Chargaff carried out a series of research efforts that revealed additional details about the DNA molecule, including its primary chemical components and the ways in which they are joined with one another (*3-5*). For many years, there was a debate among scientists, about the molecule that carries life's biological instructions. Most neglected DNA as too simple a molecule to play such a critical role. Instead, they argued that proteins were more likely to carry out this vital function because of their greater complexity and wider variety of forms. The importance

of DNA became clear in 1953 when James Watson and Francis Crick revealed the double helix structure of DNA (*6*), after studying Maurice Wilkins's and Rosalind Franklin's X-ray diffraction patterns of DNA fibers (*5, 7, 8*). A few years later, in an presentation in 1957, Francis Crick laid out the central dogma of molecular biology, which predicted the relationship between DNA, RNA, and proteins, and articulated the "adaptor hypothesis" (*9*). Final confirmation of the replication mechanism that was implied by the double-helical structure followed in 1958 through the Meselson–Stahl experiment (*10*). Further work by Crick and coworkers showed that the genetic code was based on non-overlapping triplets of bases, called codons, allowing Har Gobind Khorana, Robert W. Holley and Marshall Warren Nirenberg to decode the genetic code (*11-13*) These findings represent the birth of molecular biology.

Watson and Crick revealed the structure of DNA based on x-ray diffraction data of DNA fibers collected by Rosalind Franklin and Maurice Wilkins (*4*). This turned out to be the B type structure, and Watson and Crick proposed that a complete turn of the B-DNA molecule occurs once every ten bases. However, it was not until 1980 when a first crystal structure of this complete DNA turn was published (*3*). The sequence studied was a self-complementary dodecamer $[5'\text{-}(CGCGAATTCGCG)\text{-}3']_2$ and it seem to crystallize very well with "unusual ease and rapidity" compared to other DNA oligonucleotides.

Chemically, DNA is a polyanion at neutral pH. It usually appears as a double-stranded helix composed of two biopolymers that are made of nucleotides. Nucleotides comprise: purines Guanine (G) and Adenine (A), and pyrimidines Cytosine (C) and Thymine (T). The nitrogen bases are connected via deoxyribose ring and phosphate

backbone. The two strands are held together by hydrogen bonding between the bases as shown in Figure 1.1.

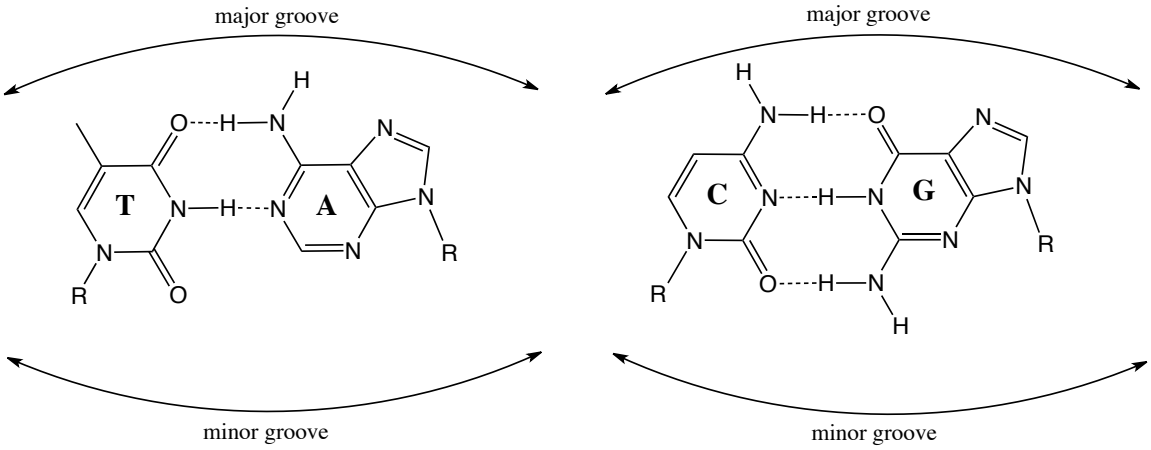


Figure 1.1. Watson-Crick pairing of A:T base pair (left) and a G:C base pair (right).

Bases fit into the double helical model when pyrimidines on one strand are always paired with purine on the other and vice versa. Chargaff's rule has its origin in the specific pairing of A with T and G with C. This pairs a keto base (G or T) with an amino base (C or A), and a purine with a pyrimidine. Two H-bonds can form between an A and T pair, and three can form between a G and C pair. These are the complementary base pairs. The base-pairing scheme immediately suggests a way to replicate and copy the genetic information.

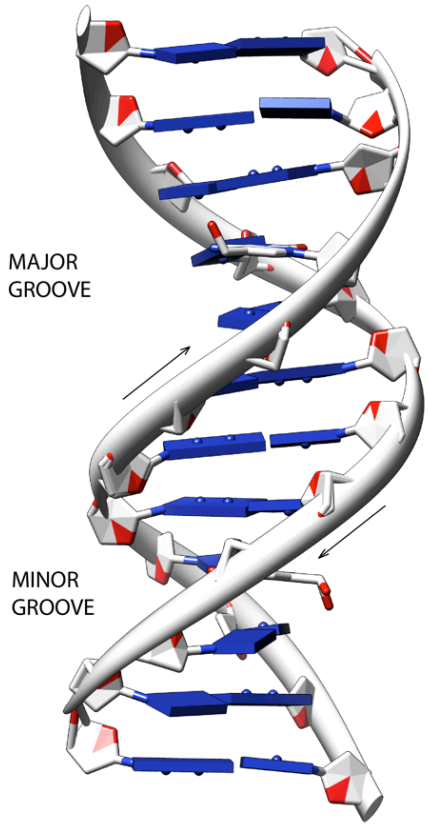


Figure 1.2. Antiparallel plectonemically coiled DNA strands. The arrows are pointed 5' to 3', and they illustrate the antiparallel nature of the duplex. The nucleotides arrayed in a 5' to 3' orientation on one strand align with complementary nucleotides in the 3' to 5' orientation of the opposite strand.

The two DNA strands are coiled around the same helical axis and are intertwined with themselves (which is referred to as a plectonemic coil). One consequence of this intertwining is that the two strands cannot be separated without the DNA rotating, one turn of the DNA for every "untwisting" of the two strands. The intertwined strands make two grooves of different widths, referred to as the major groove and the minor groove. The major groove is wider than the minor groove in DNA (Figure 1.2), and many sequence specific proteins interact in the major groove. The N7 and C6 groups of purines and the C4 and C5 groups of pyrimidines face into the major groove, where they can

make specific contacts with amino acids in DNA-binding proteins. Thus specific amino acids serve as H−bond donors and acceptors to form H-bonds with specific nucleotides in the DNA. H−bond donors and acceptors are also present in the minor groove, and indeed some proteins bind specifically in the minor groove.

Three different duplex have been described (*14*) (Figure 1.3). The most common form present in most DNA at neutral pH and physiological salt concentrations is B-form. That is the classic, right-handed double helical structure with a turn every 3.4 nm, such that the distance between two neighboring base pairs is 0.34 nm. Hence, there are about 10 pairs per turn. In a solution with higher salt concentrations or with alcohol added, the DNA structure may change to an A form, which is still right-handed, but every 2.3 nm makes a turn and there are 11 base pairs per turn. The A-form is adopted by RNA-DNA duplexes and RNA-RNA duplexes. Another DNA structure is called the Z form, because backbone phosphates are arranged in a zigzag pattern. Z DNA is left-handed. One turn spans 4.6 nm, comprising 12 base pairs. DNA molecules with alternating G-C sequences in alcohol or high salt solution tends to have such structure.

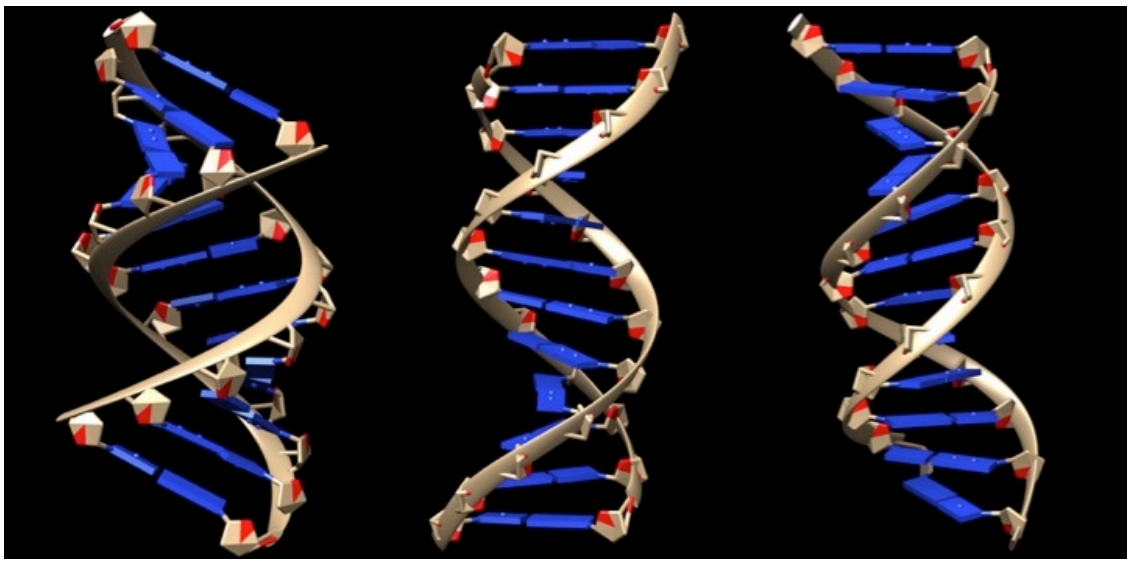


Figure 1.3. Models representing A-form (left), B-form (middle) and Z-form DNA (right).

Even classic B-DNA is not completely uniform in its structure. X-ray diffraction analysis of crystals of double-helical oligonucleotides shows that a given sequence will adopt a unique structure. These variations in B-DNA may differ in the propeller twist (between bases within a pair) to optimize base stacking, or in the three ways that two successive base pairs can be oriented relative to each other: twist, roll, or slide.

The strength of hydrogen bonds between base pairs contributes to the stability of the DNA helix. Appropriate geometrical correspondence of hydrogen bond donors and acceptors allows only for the complementary base to form stable base pairs. DNA with high CG content is more stable than DNA with low CG content, but hydrogen bonds do not stabilize the DNA themselves. Stabilization is mainly due to stacking interactions. Stacking interactions between bases are due to dispersion attractions, short-range exchange repulsions, and electrostatic interactions. Most favorable are GC stacking interactions with adjacent bases (*15*).

DNA in the form of the duplex is stable at room temperature, but the two strands separate above a melting point that is determined by the length of the molecule and CG content. Higher GC contents result in higher melting temperatures (*15*).

Epigenetic regulation of the genome

In the genome, DNA is constructed from the four nucleosides: 2'-deoxyadenosine (dA), 2'-deoxyguanosine (dG), thymidine (dT), and 2'-deoxycytidine (dC). These building blocks are assembled inside the DNA double helix (*6, 15*). In multicellular organisms all cells possess the same genetic content; however, cells might perform very different functions. For example, neuronal cells are designed to conduct electrical signals along the axons, muscle cell are involved in performing contractive motions. The same genetic material manifests itself in different functions of different cells, which is only possible because cells differ in the nature and number of active genes. During cellular development, cells switch unneeded genes on and off on their way to specialization. Chemically, long-term gene silencing is achieved by the methylation of base cytosines (dC) at position C5 in special places in the genome rich in CG repeats, called CpG islands (*16, 17*).

In duplex DNA, the C5 and C6 positions of cytosine are positioned in the major groove, held by Watson-Crick interactions with complementary guanines. The electrophilic character of the C6 position makes it a key target of modifying enzymes. For example, DNA methyltransferase (DNMTs) transiently modify C6 by attack of an active site cysteine. Methylation results from the concerted addition

of a methyl group derived from S-adenosylmethionine (SAM) to the C5 position (*18, 19*). The covalent intermediate breaks down, releasing the enzyme and generating genomic 5-methylcytosine (5mC). It was also shown that in the absence of SAM, DNMTs can also catalyze deamination at C4 (*20, 21*), or the addition of aldehydes to C5 (*22*), which raises the interesting question about the relevance of these non-classical functions *in vivo*.

The methylation pattern is a crucial part in the epigenetic information and a critical marker that distinguishes cells. The underlying 5-methylcytosine is often considered to be the fifth base of the genome. After fertilization and at certain points of the embryonic development, a large number of the methylation marks are erased (*23, 24*). This allows embryonic stem cells to differentiate into any possible specialized cell. In certain cases this genome-wide demethylation occurs without cell division and thus without the synthesis of new DNA. The methyl group then has to be actively removed by enzymes. The mechanism underlying this active demethylation is of high interest, because it is speculated that some cells might have the possibility to actively demethylate their genetic material in order to re-differentiate (*25*).

DNA Demethylation mechanisms

DNA methylation can be passive or active. Passive DNA demethylation occurs in dividing cells. While DNMT is responsible for active methylation during cell replication, its inhibition or malfunction will result in unmethylated cytosines and reduced overall methylation levels. Active DNA demethylation can occur in both dividing and nondividing cells. This process requires enzymatic reactions to revert

the 5mC back to unmodified cytosine (*26-29*). As of today, there is still no known mechanism in mammalian cells that can cleave the strong covalent C-C bond that connects cytosine to the methyl group. Instead, demethylation appears to involve a series of enzymatic reactions where 5mC is further modified, by either deamination or oxidation to a sequence of modifications that are recognized by the base excision repair (BER) pathway to remove the modified base, and replace it with unmodified cytosine. Several mechanisms of active DNA demethylation have been proposed (Figure. 1.4):

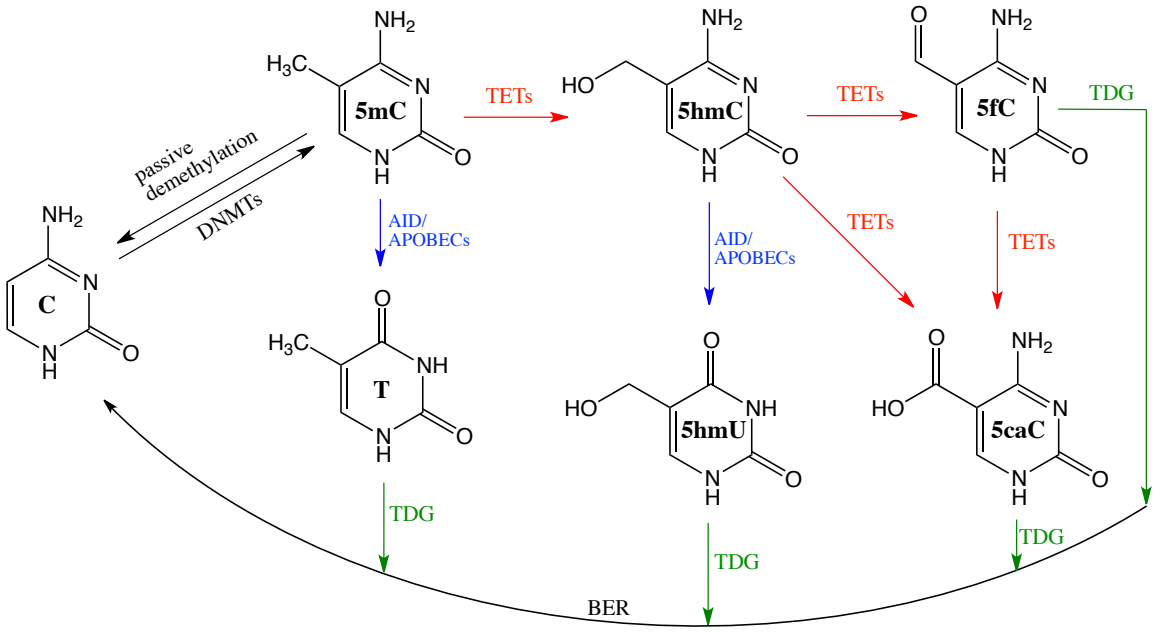


Figure 1.4. DNA demethylation pathways.

5-methylcytosine (5mC) can be chemically modified at two sites. At the C4 amino group of 5mC, which can be deaminated by AID/APOBEC (activation-induced cytidine deaminase/apolipoprotein B mRNA-editing enzyme complex), thereby

converting 5mC into thymine. This results in a G/T mismatch and induces the BER pathway to correct the base to G/C (*30-32*).

Recently, the search for alternative pathways of active demethylation led to discovery of 5-hydroxymethyl-cytosine (5hmC). Two independent research teams showed that mammalian genomic DNA contains not only 5mC, but also 5hmC, which is now considered to be the sixth base of the genome (*33, 34*). The 5hmC modified base is formed post-replicatively by enzymatic oxidation (*35*). The oxidation of 5mC is facilitated by the TET (ten eleven translocation) family of enzymes, which belong to the Fe(II)/α-ketoglutarate-dependent oxygenase family that includes histone demethylases and the DNA damage repair enzyme AlkB (*34, 36*). TET enzymes were discovered based on homology to a trypanosome enzyme, which catalyzes oxidation of the exocyclic methyl group of thymine. The oxidation product of 5mC, 5hmC, was found in developed brains at the levels of 0.3-0.7%, which is approximately ten times lower than the natural abundance of 5mC (*33, 37*). Once 5hmC is formed, two separate mechanisms can convert 5hmC back to unmodified cytosine. In the first, AID/APOBEC can deaminate 5hmC to 5-(hydroxymethyl)-uracil (5hmU) which is a substrate for thymine DNA glycosylase (TDG) and excised via a base excision mechanism (BER) (*38*). In another pathway for 5hmC, iterative oxidation by TETs further oxidizes 5hmC to form 5-formylcytosine (5fC) and 5-carboxycytosine (5caC) (*39*). Both modifications are stably detectable intermediates in genomic DNA from embryonic stem (ES) cells (*39, 40*).

The function of 5hmC is still unclear. On the one hand, it may serve as an intermediate in DNA demethylation, but on the other it may also regulate gene

expression. In support of this theory, the conversion of 5mC to 5hmC impairs the binding of the repressive methyl-binding protein MeCP2 (*41*). Moreover, 5hmC was found *in vivo* in mammalian tissue, is tissue specific, and may play an important role in regulating DNA demethylation and gene expression.

In all the mentioned pathways of active DNA demethylation, the base excision repair pathway uses thymine DNA glycosylase (TDG) to excise the modified residue – thymine, 5hmU, 5fC, 5caC and replace them with an unmodified cytosine (*42, 43*). TDG is essential for DNA demethylation and is required for normal development.

DNA repair

Specific base pairing is essential for preserving the information content of the genome. However, the structural properties of DNA bases, and their pairing properties are often modified by reactions with toxins or endogenous metabolic products. As a result mismatches can be formed, and some of them may lead to certain types of cancer (*44*). Spontaneous deamination of 5mC at CpG sites in human DNA leads to C/T transition and therefore G/T mismatches. Mutations such as G/T mismatches can be very harmful to the genetic code, in terms of the information that is stored in the DNA, and epigenetic methylation patterns. If these mismatches are not repaired by the cellular repair mechanism known as base excision repair (BER), they can lead to DNA abnormalities and carcinogenesis (*44*). Base excision repair is the main mechanism by which cells correct various types of damaged DNA bases.

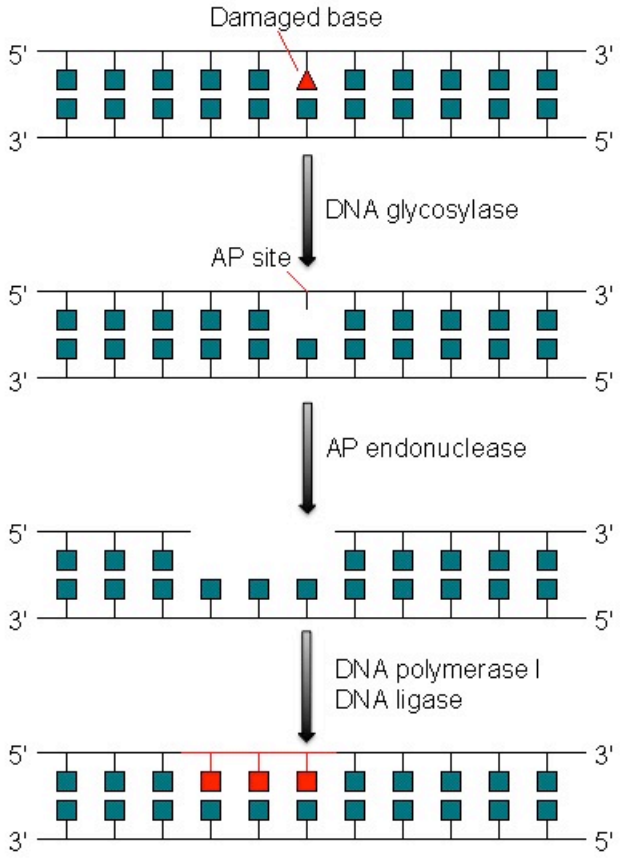


Figure 1.5. Base excision repair of DNA.

The first step in BER is the recognition of a wrong base by a glycosylase enzyme, which then excises the damaged base. This results in the formation of an apurinic or an apyrimidinic (AP) site. These AP sites are then removed by an AP endonuclease that causes nicks in the DNA as a result of the hydrolysis of the phosphodiester bond that is 5' to the AP site. These free ends are degraded by endonucleases, and the gap is then filled by a DNA polymerase enzyme (*44, 45*). DNA repair is then completed by ligation in the presence of ligases (Figure 1.5).

The recognition and detection of damage or modified bases in DNA by glycosylase is an essential step during DNA repair in all cells. This initial step

involves a specific repair enzyme finding its substrate in the midst of a vast excess of canonical bases that in many cases are structurally similar to the damaged base. A base-flipping repair protein may have to search through millions of base pairs to find a single point of damage. How these proteins efficiently locate base lesions has long been an interesting question. Assuming that repair protein first binds to DNA nonspecifically and then moves in one dimension, there are three potential damage-locating mechanisms: in the first, every base is actively flipped and inserted into the active site by the protein until the damage is located; in the second, repair protein selectively detects an unstable or non-Watson- Crick base pair that contains the damaged base; and in the third, a transiently extrahelical base lesion is captured by the repair protein (*46*). In the process of enzymatic base flipping, the DNA undergoes deformation, which is unfavorable in terms of energy. One of the most dramatic substrate conformational changes in site-specific nucleic acid recognition is the complete rotation of a base and its attached sugar from the base stack through either the major or minor groove. The unfavorable energetic events during the reaction include: nucleotide rotation, breaking of the Watson-Crick hydrogen bonds, and disruption of the base stacking and DNA bending. The binding repair enzyme has to pay for these costs through use of a favorable bending energy, which would increase in this case (Figure 1.6) (*47*).

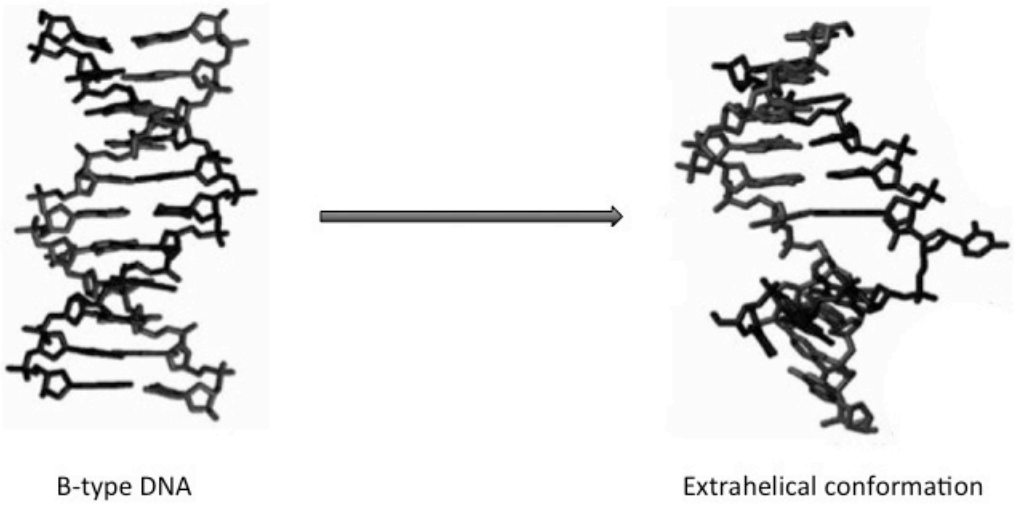


Figure 1.6. Base flipping occurring on the DNA molecule during the process of enzymatic flipping of the damage DNA base. (Figure adopted from (*47*)).

The most common experimental method to measure the flipping or base opening is imino exchange by nuclear magnetic resonance (NMR) spectroscopy. This technique indirectly measures the rates and the equilibrium for base pair opening and closing, by monitoring the exchange of the imino protons on G or T with solvent protons that have been magnetically labeled by selective inversion (*48-52*). The key assumption is that exchange occurs only in the open state. However, when the exchange of the imino proton is much faster than the opening of the base, then the latter becomes the rate-limiting step, and the observed exchange rate is equal to base pair opening; in other words, exchange occurs at each opening event. Since the exchanging imino protons are on either T or G, it is assumed, that within the whole base pair, these two bases will be flipping out from the stack in order to exchange their imino proton. However, as proposed based on computational studies, it is possible that for T exchange, the opposite adenine flips out, leading to exposure of the T imino proton without requiring T to leave the base stack. Likewise, imino proton exchange on G could occur

by flipping the opposite cytosine base (*53*). The increased flexibility or dynamic of a damaged or modified base pair could guide the repair enzyme to the site. The studies with uracil DNA glycosylase (UDG) have provided a precedent for this mechanism (*54-56*). In these studies, the DNA duplexes containing a single U:X or T:X base pair were used, where X was an adenine analog able to form a single hydrogen bond, double or triple hydrogen bonds with the opposite base. Progressive destabilization of the studied base pair by removal of a hydrogen bond would be expected to favor base pair opening and enzyme binding, because less binding energy would be required to open the destabilized base pair. The results with a U:X base pair indicated that base pairs with stronger hydrogen bonds inhibits uracil flipping into the enzyme binding site. The second result with a T:X base pair showed that a base pairs with larger equilibrium constants for opening lead to enhanced binding. The energetic correlations between the intrinsic stability and dynamics of the base pair and the enzyme binding to both uracil and thymine base pairs revealed that the initial damage recognition might rely on the intrinsic dynamic and physical properties of the base pair (*54-56*).

Another example of a damage-recognition mechanism involves selective detection of a non-Watson-Crick base pair that contains the damaged bases by the repair enzyme. An example of this type of mispairing is G pairing either with T or U in the Wobble base pair configuration (Figure 1.7). In this unique base pair G, which is normally capable of forming three hydrogen bonds, forms only two hydrogen bonds with the opposite T or U, due to a slight movement of G into the minor groove. This unusual structural feature is recognized by thymine DNA glycosylase and allows for efficient removal of T or U from the mismatched DNA (*57*).

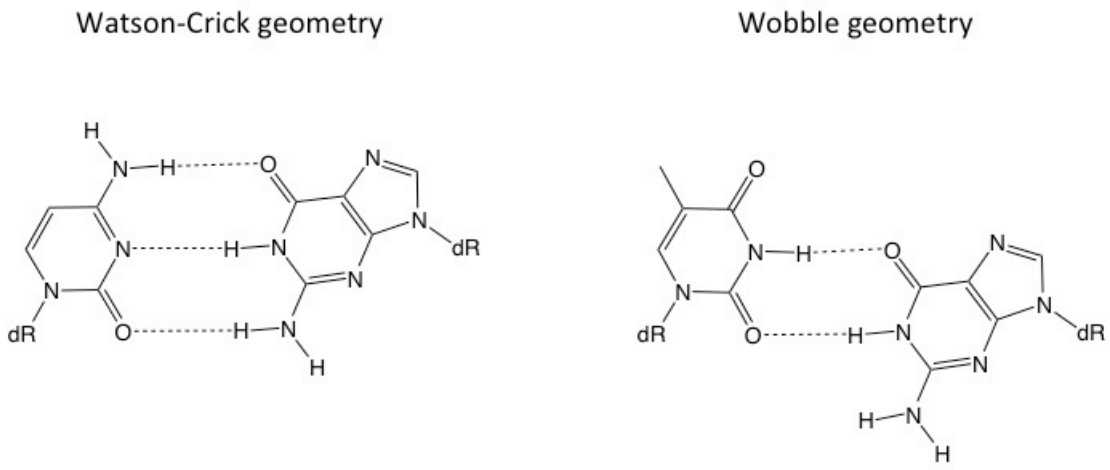


Figure 1.7. Different base pairing geometries.

Experimental methods for structural and dynamic analysis

Biological macromolecules play an important role in all cellular processes. Structural biology studies the relationship between the structure of biological macromolecules and their intrinsic function in order to understand chemical reactions, that are crucial to life. Two techniques are widely used for the structure determination of biological macromolecules at the level of distinguishing individual atoms: X-ray crystallography and nuclear magnetic resonance spectroscopy (NMR).

Nuclear Magnetic Resonance Spectroscopy

Nuclear Magnetic Resonance (NMR) spectroscopy has become a prominent tool in structural biology, as it has the ability to determine the atomic resolution structures of biomacromolecules in a native solution environment. NMR provides the means

for studying critical biological phenomena including nucleic acid and protein structure, and dynamics and it serves as a practical approach to drug design (*8*). The most basic physical principles that describe this biophysical technique as well as NMR experiments that are most relevant to structural biology are described below.

The NMR era started in the late 1930s when Rabi *et al.* measured the intrinsic magnetic moment of each spin (*58*). Bloch (*59*) and Purcell (*60*) then worked on transition effects caused by the resonance during the radiofrequency sweep in the static magnetic field, for which they were awarded with the Nobel Prize in 1962. Only three years later, Ernst and Anderson introduced pulsed excitation and signal averaging with the application of Fourier transformation of the time domain data. This allowed for averaging of multiple scans resulting in improved signals (*61*). That was the time when detection in two dimensions started, with pioneering contributions by Jeener (*62*), Earnst and studied by others (*63*). Kurt Wüthrich was first to apply these techniques to protein analysis (*64*), when he published the first NMR derived assignments of a protein leading to structure determination by NMR in 1982. Since then the analysis of biomacromolecules by NMR (*65*) at the 3D structural level has expanded, and it continues to be one of the major techniques in structural biology.

Nuclear magnetic resonance relies upon manipulation of the fundamental quantum mechanical properties of atomic nuclei to provide unique information on their electronic environment (*9*). The quantum mechanical property of interest is the nuclear spin angular momentum. Placement of the nuclei with the magnetically active spins (where the spin quantum number is $I = ½$) within a static magnetic field of sufficient strength induces the spin angular momentum vector to rotate about an axis parallel

to the magnetic field (Figure 1.8) (*10-12*). The frequency of rotational motion, known as the resonance frequency, depends upon the electronic environment surrounding a given nucleus, i.e. whether it forms part of a covalent or non-covalent bond, or is in the proximity to other atoms. Subsequent excitation of these processing spins by an oscillating electromagnetic (EM) field, which transmits energies corresponding to their resonance frequencies, rotates the spins to the plane perpendicular to the static field (Figure 1.8). As they continue their precession in the transverse plane, a time-dependent signal that encodes the collective rotational frequencies of the spins is collected (Figure 1.8). Once processed by Fourier transformation, this signal provides a 'frequency map' that reflects the unique electronic surroundings of each spins (Figure 1.8).

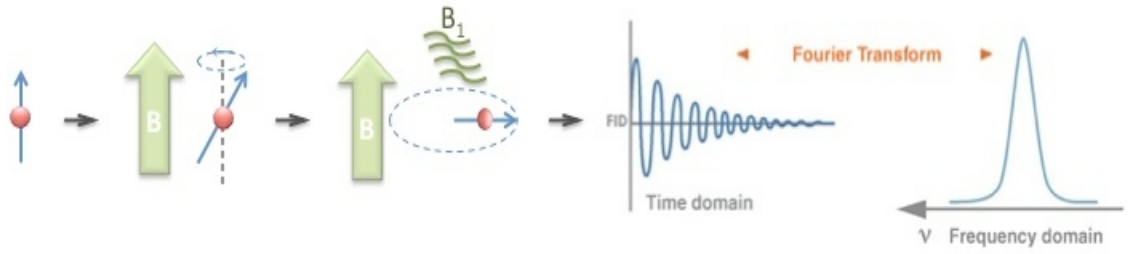


Figure 1.8. Nuclear magnetic resonance spectroscopy, and a basic outline of the NMR experiment. Nuclear spins (red sphere and blue arrow) undergo precessive motion at a characteristic resonance frequency when placed in a static magnetic field B (green thick arrow). Application of an oscillating magnetic pulse (B_1) rotates the spin vector to the perpendicular plane, where it continues its precessional motion and induces an oscillating voltage signal. Fourier transform of this time-domain signal generates a frequency spectrum with the higher intensity peak centered at the spin's resonance frequency.

Without the presence of the excitatory EM field, transverse spin precession gradually returns to the original parallel precession extant prior to excitation. This return of excited spins to their ground energy state is known as 'relaxation', and provides a powerful

means for accessing global and local molecular motions, as the timescales of motion influence the rapidity with which spins relax (*12, 13, 66, 67*).

In terms of NMR, biomacromolecules are complex assemblies of nuclear spins. Although one dimentional ^{1}H NMR exploits the high natural abundance of the ^{1}H isotope in macromolecules, the large number of protons present in a macromolecule (as compared to a small molecule) often leads to extensive overlap within the resulting 1D ^{1}H NMR spectrum. Therefore, 2D NMR has been developed to overcome this problem.

One of the most powerful 2D NMR experiments for the solution structure determination of oligodeoxynucleotides is ^{1}H-^{1}H Nuclear Overhauser Effect Spectroscopy (NOESY). This experiment uses the dipolar interaction of spins for correlation of protons located within approximately 5 Å apart in space. With the application of different mixing times raging from 30 to 300 ms NOESY experiments can address spin diffusion effects. For the detection of non-exchangeable protons, deuterated solvent is used for NOESY in order to avoid excessive solvent signals. For the assignment of exchangeable protons, such as amino and imino protons, which relates to Watson-Crick base pairing, the NOESY spectrum can be obtained with the sample in H_2O solution.

The assignment of non-exchangeable protons in a DNA duplex can be performed on adjacent bases throughout the DNA strand (*68*). In the 5' to 3' direction, the intra-nucleotide, ^{1}H-^{1}H distance between the aromatic H6 or H8 proton on the aromatic base purine or pyrimidine, respectively, and its own H1' deoxyribose sugar anomeric proton

allows for a cross-peak. This is followed by a cross peak of the H1' sugar proton to the H6 or H8 proton of its 3' neighboring base (Figure 1.9).

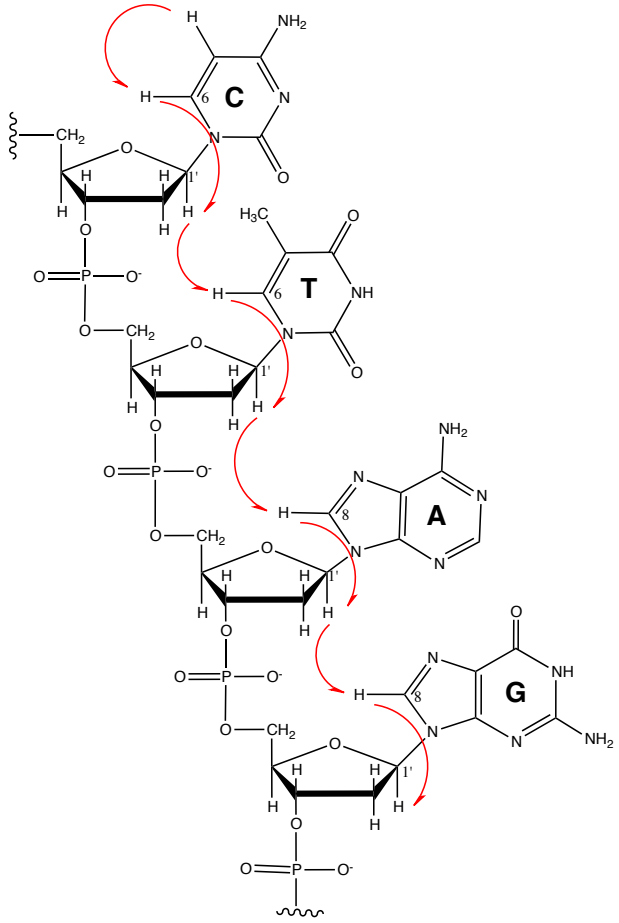


Figure 1.9. Sequential walk between aromatic H6/H8 aromatic base protons and anomeric H1' sugar protons.

In the same way, the inter-nucleotide distance between the H6 or H8 proton and its own H3' sugar proton allows for a cross peak, followed by an H3' sugar interaction with the 3'-neighboring H6 or H8 proton. This method of sequential assignment accomplished by “walking” through the DNA duplex in a 3' to 5' direction is known as a NOESY walk (Figure 1.9).

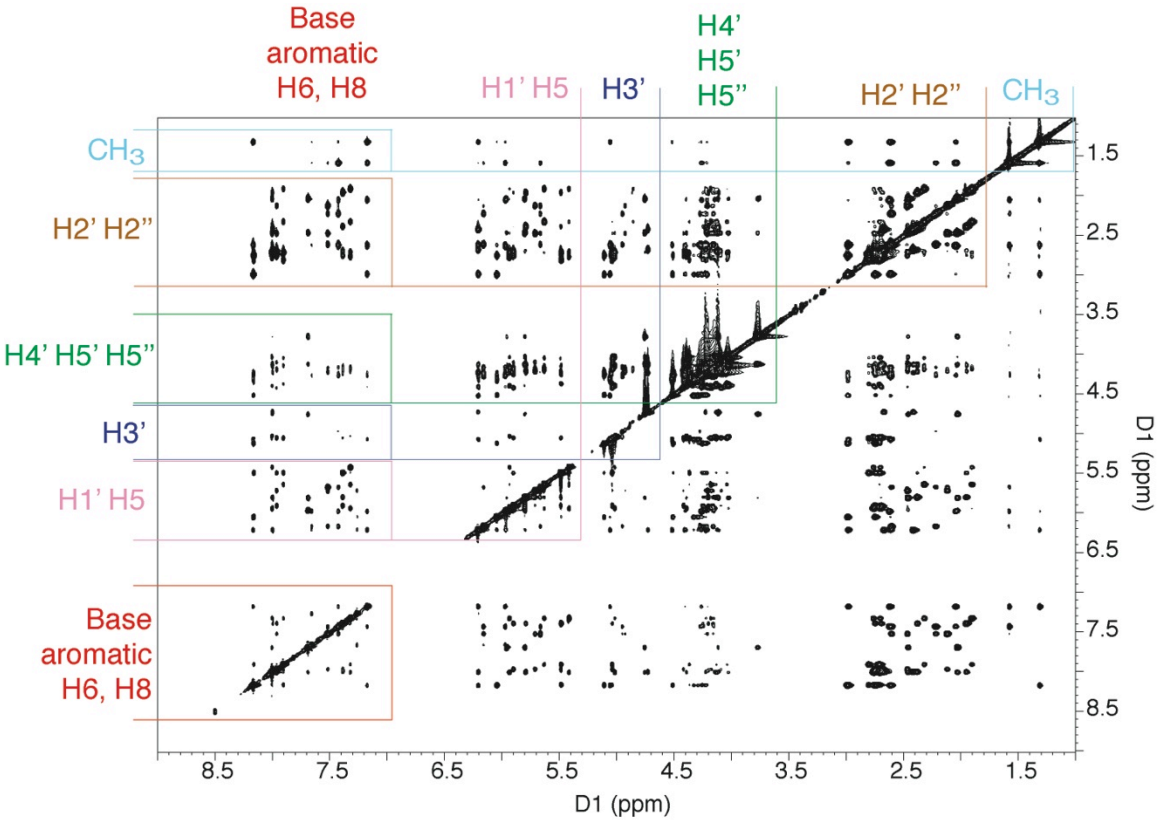


Figure 1.10. A typical ^{1}H-^{1}H NOESY spectrum of an oligodeoxynucleotide. Interactions of NOE's are observed between peaks at the specific regions.

The NOESY spectrum (Figure 1.10) is very useful for observing DNA damage and modifications–induced disruption of the sequential connectivity. It also provides chemical shift information, as well as insight into the orientation of the modified base. A NOESY spectrum measured for a sample in H_2O solvent as the allows for detection of exchangeable protons. In this way one can detect imino protons, and their sequential intra-nucleotide connectivities between guanines and thymines throughout the whole DNA duplex (*69*). Interactions between imino protons as well as imino and amino protons provide information about Watson-Crick base pairing and stacking interaction in DNA.

Another widely used 2D experiment for determining molecular structures is correlation spectroscopy (COSY). It provides additional information by correlating the chemical shift of ^{1}H nuclei which are *J*-coupled to one another (*70*). With ^{1}H-^{1}H COSY one can detect signals from geminal or vicinal hydrogens. In this way cross-peaks between cytosine H5 and H6 can be located, as well as the methyl protons of thymine. The Double Quantum Filtered COSY (DQF-COSY) (*66, 67, 71*) allows for partial cancellation of the diagonal peaks and elimination of strong signals coming from solvent protons that do not experience homonuclear *J*-coupling, as compared with magnitude COSY. In the latter, signals with small *J*-coupling constants and broad lines will show enormous diagonal signals, but only very small or vanishing cross-peaks. However, only DQF-COSY, where both cross and diagonal peak intensities depend on the size of the coupling constant, allows for detection of cross-peaks of the protons close in chemical shift and calculation of *J*-coupling constants. These values are useful in generating structural restraints such as the pucker of the deoxyribose sugar ring in DNA structural studies (*72*).

With Total Correlation Spectroscopy (TOCSY) one can detect the interactions of all protons with a spin system that are not directly connected by chemical bonds (*73*). As a result scalar coupling of a complete spin system can be determined. The scalar coupling range depends on the mixing time. For the short mixing time period, short range intraproton interactions are observed, and once the mixing time gets longer, then the correlation with longer distance protons can be detected. With the complete set of NMR 2D data a solution structure of DNA duplex can be obtained by applying restraints molecular dynamics (rMD) calculations.

X-ray crystallography

The era of 3D crystal structure determination of biomacromolecules started in the late 1950s, with the first crystal structure of sperm whale myoglobin by Kendrew and Perutz (*74, 75*), for which they were awarded the Nobel Prize in Chemistry in 1962. Thirty years later Richard Wing and co-authors reported the first single-crystal structure of more than a complete turn of right-handed B-DNA, based on the self-complementary dodecamer sequence $[5'\text{-}(CGCGAATTCGCG)\text{-}3']_2$ (*76*). The sequence crystallized in the $P2_12_12_1$ space group and the duplex was extensively characterized by Dickerson and Drew in 1981 (*5-7*). It was found to have an overall bend of 19° due to crystal packing forces, indicating intrusive flexibility of the DNA molecule. The duplex exhibited a rise of 3.4 Å per base pair, and there were 10.1 base pairs per DNA turn. Since then the Dickerson Drew Dodecamer (DDD) sequence has been extensively studied by crystallography, due to its ease of crystallization, and also by NMR. The crystal packing allows for a stable conformation of the DNA and well diffracting crystals. The DDD duplex has specific cation binding sites where Mg^{2+} or Ca^{2+} interact in the grooves of DNA. In addition a characteristic inner and outer spines of hydration are observed in the minor groove of the DDD (*77*). Cations and water molecules play an important role in crystal formation and stabilization of DNA duplex.

X-ray crystallography is used to determine the arrangement of atoms in a crystal. To analyze DNA by X-ray crystallography, it is necessary to grow crystals of a purified DNA sample. Crystals are formed as the conditions in a supersaturated solution slowly change. One of the methods for growing biomacromolecule crystals is the hanging drop vapor diffusion method, where a drop of protein solution is suspended over a reservoir

containing buffer and precipitant. Water diffuses from the drop to the solution leaving the drop with optimal crystal growth conditions. Crystallization conditions have to be optimized to obtain the desired crystals, the size of which is typically between 10-300 μm along each edge. Once crystals of good quality have been grown, of a good quality, they are looped out and flash frozen in liquid nitrogen and ready for X-ray analysis.

X-rays represent high-energy electromagnetic radiation and can be generated from accelerated electrons. Electrons are produced by a cathode and then accelerated to 99.9999% of the speed of light in a linear accelerator. The electrons are then injected into the booster synchrotron, where they are sent around an oval racetrack of electromagnets, providing further acceleration. Upon reaching the 99.9999% speed of light the electrons are injected into a storage ring. Once there, the electrons are focused into a beam and collimated with the sets of adjustable slits to ensure a parallel beam that is available for use in experimentation. X-rays are directed at the crystal when electrons diffract the X-rays, thus resulting in a diffraction pattern. Using the Fourier transformation these patterns can be converted into electron density maps. These maps show contour lines of electron density. Since electrons more or less surround atoms uniformly, it is possible to determine where atoms are located. Unfortunately since hydrogen has only one electron, it is difficult to map hydrogens. To get a three dimensional picture, the crystal is rotated while a computerized detector produces two-dimensional electron density maps for each angle of rotation. The third dimension comes from comparing the rotation of the crystal with the series of images. Computer programs use this method to come up with three dimensional spatial coordinates.

Using a synchrotron sources ensures a high intensity X-ray beam and allows shorter exposure times and a higher signal to noise ratios of the diffraction spots, which is required for crystallographic studies of biomacromolecules. With the application of CCD or Pilates detectors, the time of data collection of a complete data set from a single crystal is now rather short. As the diffraction spots become weaker at higher resolution, a compromise between increased resolution and decreased diffraction quality has to be made. When sufficiently high-resolution data are collected and processed, the unit cell dimensions, the crystal system, and the space group of the crystal can be determined.

A key factor in a crystal structure determination is the quantity of the collected data. A specific strategies have to be followed, that is dependent on the properties of the unit cell and the symmetry of the crystal. While exposed to the X-ray beam, the crystal is rotated and a series of diffraction patterns is collected. 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, and the intensity is determined by the amplitude of the diffracted waves and by the phase difference. Since the amplitudes can be recalculated, the information about the phase is lost; therefore it is referred as the phase problem. There are two methods most frequently used to overcome the phase problem: molecular replacement (MR), and multiple wavelength anomalous dispersion (MAD). The most common technique used is MR. When a homologous model is available, then coordinates and phases can be used with the experimental diffraction data to calculate an electron density map of the new molecule. In the next step the map is used to determine

the structure, which forms the basis for the refined set of structures. The MAD technique requires a single crystal to contain a heavy atom, which will cause anomalous scattering. The most common heavy metals used here are strontium, barium or selenium. With the MAD technique, the wavelengths of X-rays will be adjusted to strengthen the anomalous scattering. Thus, the heavy atoms can act as a reference marker and alter the intensity of spots in the diffraction pattern, which allows for the position of the heavy atoms to be determined and phases to be assigned.

With the solved phases and observed intensities of diffracted spots, Fourier transformation can be applied to obtain electron density map. This map forms a 3D contour into which the model can be built. Once the starting model is obtained, the phases can then be recalculated. In the next cycle, the calculated phase together with the original spot intensities can be used to rebuild and improve the electron density map and thus the new model. The improvement of the model from each cycle is called refinement, and is verified by comparing observed with calculated wave amplitudes, which is given by the R-factor. For DNA dodecamers, a satisfactory R-factor is between 0.15-0.25. Another validating factor, which evaluates the quality of the refinement is R-free, and it represents the difference between randomly selected 5-10% of the original data that are not used in the refinement cycle and the calculated data derived from the refined structure. The R-free should be below 20% for DNA structures.

Scope of this work

The main topic of this thesis concerns oxidation products of 5-methylcytosine (5mC), such as 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) with a specific focus on their structural and dynamic effects on the DNA duplex. Maiti and Drohat (*78*) showed that active demethylation of 5mC requires further oxidation of 5hmC to the downstream products 5fC and 5caC, allowing TDG to restore the epigenetically unmodified G:C pair. In their work they showed that both 5fC and 5caC are substrates for thymine DNA glycosylase (TDG), but not 5hmC. There are two major hypotheses in this research project. From studies with thymine DNA glycosylase (TDG), it was proposed that the enzyme recognizes a mismatched base of G/T or G/U mispairs in the DNA duplex, whereby the mismatch base pair forms a Wobble geometry (*57*). This unique structural geometry is recognized extrahelicaly by the enzyme, and results in excision of the mispaired base. Mechanistically, TDG employs an extrahelical base flipping mechanism (*79-81*) to position substrates into its active site for catalysis.

According to studies done by Stivers (*47, 82*), DNA stability and dynamics play a crucial role in the detection of damaged bases by a repair enzyme. The assumption is that the base pair with lower stability will have kinetically higher opening rates. In order to characterize the kinetics of base pair opening in nucleic acid duplexes, high-resolution proton exchange NMR spectroscopy has been used. The rates of exchange of imino protons with solvent protons were measured by magnetization transfer from

water protons for each DNA duplex as a function of the concentration of exchange catalyst.

The first goal of this research project is to determine the structures of 5hmC, 5fC and 5caC in the DDD sequence and to investigate the conformational perturbations of DNA induced by the presence of the modified bases, with a specific interest in base pairing and base stacking in the DNA double helix. The second goal is to study how the presence of the modifying groups in the major groove of DNA affects the stability of the double helix. The third and last goal is to determine the kinetics of base pair opening in the modified DNA double helix in order to understand dynamics around the modification site.

Chapter 2 describes the materials and methods. Chapter 3 reports the detailed thermodynamic analysis of 5hmC in duplex DNA, high-resolution crystal structure and dynamics of base pair opening. In chapter 4, the crystallographic results of 5fC in DDD are reported, as well as the kinetic characterization of base pair dynamics. Chapter 5 reveals the crystal structure of 5caC in DNA duplex, along with the kinetic results obtained by high-resolution NMR spectroscopy. Finally, chapter 6 provides conclusions driven from the work described in this dissertation and discusses future directions relevant to this project.

CHAPTER 2

MATERIALS AND METHODS

Oligodeoxynucletide synthesis

Unmodified (DDD) and modified oligodeoxynucleotides with 5mC, 5fC or 5caC (DDD^{m}, DDD^{f}, and DDD^{ca}, respectively) were synthesized by Midland Certified Reagent Co. (Midland, TX) and purified by anion-exchange high performance liquid chromatography (AE-HPLC).

The 5hmC oligodeoxynucleotide (DDD^{hm}) was synthesized on an Expedite 8909 DNA synthesizer (PerSeptive Biosystems) on a 1-μmol scale using the UltraMild line of phosphoramidites (ethylcyanide-protected 5-hydroxymethyl-dC, phenoxyacetyl-protected dA, 4-isopropyl-phenoxyacetyl-protected dG, acetyl-protected dC and dT phosphoramidites) and solid supports from Glen Research (Sterling, VA). The manufacturer's standard synthesis protocol was followed except in the case of the incorporation of the modified phosphoramidite, which was accomplished manually, off-line. At this point, the column was removed from the instrument and sealed with two syringes, one of which contained 250-300 μL of the manufacturer's 1H-tetrazole activator solution (1.9-4.0% in CH_3CN, v/v) and the other contained 250 μL of the modified phosphoramidite solution (15 mg in anhydrous CH_3CN). The 1H-tetrazole and the phosphoramidite solutions were sequentially drawn through the column (1H-tetrazole first), and this procedure was repeated periodically over 30 min. After this

time, the column was washed with manufacturer's grade anhydrous CH_3CN and returned to the instrument for the capping, oxidation, and detritylation steps. After the synthesis the deprotection reaction was run with 30% ammonium hydroxide for 17 hours at 75 °C.

Oligodeoxynucleotide purification

The unmodified and modified oligodeoxynucleotides were purified by semi-preparative reversed-phase high performance liquid chromatography (RP-HPLC). The purification was conducted on a Beckman HPLC system (32 Karat software) with a diode array UV detector measuring at 260 nm. The semi-preparative column was used (Atlantis by Waters, C18, 5 µm, 250 mm x 10.0 mm), and was equilibrated either with 30 mM phosphate buffer at pH 7.0 (for DDD, DDD^m, DDD^{hm}, and DDD^{ca}) or 0.1 M ammonium formate at pH 6.5 (for DDD^f) and acetonitrile. The purification method was set up with the following gradient: 1-15% acetonitrile over 20 min., 15-80% acetonitrile over 5 min., and then to 1% acetonitrile over 5 min., with a flow rate of 4.5 mL/min. The collected fractions were incubated at -80 ºC followed by liophylization. Figure 2.1 shows an example of RP-HPLC chromatogram of DDD^{hm} after purification. Dried oligonucleotides were resuspended in 0.5 mL of deionized water and desalted on a G-25 Sephadex column.

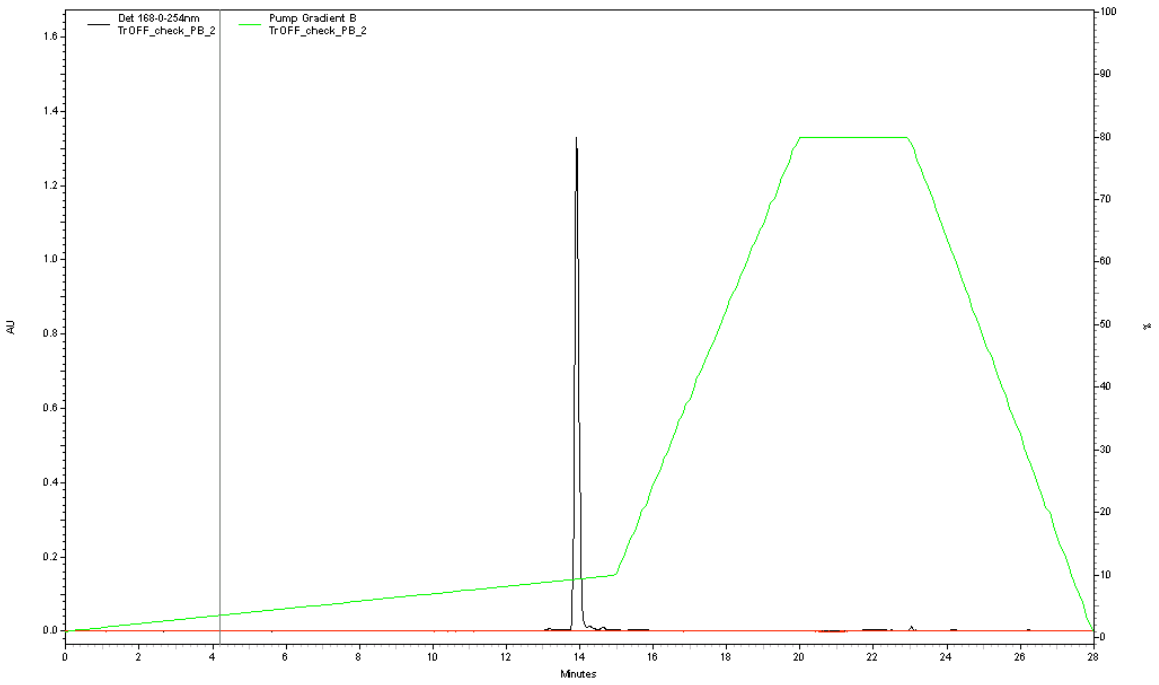


Figure 2.1. Example of an RP-HPLC chromatogram showing pure oligodeoxynucleotide.

Matrix-assisted laser desorption/ionization mass spectrometry

The purity of the oligonucleotides was confirmed by MALDI-TOF mass spectrometry. The mass analysis was performed using a Voyager-DE (PerSeptive Biosystems, Inc., Foster City, CA) spectrometer. Samples were suspended in a matrix containing 0.5 M 3-hydroxypicolinic acid in 1:1 CH_3CN:H_2O, 0.1 M ammonium citrate and spotted onto sample plates. Mass spectra were recorded in a negative reflector and averaged from 512 scans. Calculated mass for DDD $[M-H]^-$ *m/z* 3646.4, found *m/z* 3647.8; for DDD^m calculated mass of $[M-H]^-$ *m/z* 3660.5, found 3663.4; for DDD^{hm} calculated mass of $[M-H]^-$ *m/z* 3675.5, found 3679.7; for DDD^f calculated mass of $[M-H]^-$ *m/z* 3674.4, found 3673.2; for DDD^{ca} calculated mass of $[M-H]^-$ *m/z* 3690.4, found 3693.1.

Sample preparation

The lyophilized and desalted oligodeoxynucleotides were dissolved in buffer containing 10 mM Na_2HPO_4, 100 mM NaCl, 50 μM Na_2EDTA at pH 7.0, and annealed by heating the solution to 85 ºC for 15 min, and allowing it to slowly cool to room temperature. Duplex concentrations were determined by UV absorbance using extinction coefficients determined at 260 nm.

Thermodynamic measurements

Temperature-Dependent UV Spectroscopy (UV Melting Curves).

UV melting curves (absorption versus temperature profiles) for the helix−coil transition of DDD, DDD^f and DDD^{ca} duplexes were obtained with a thermoelectrically controlled Cary 100 Bio UV-VIS spectrophotometer (Varian, Inc. Palo Alto, CA), interfaced to a PC computer with Cary WinUV Thermal Application (v. 2.0) for data acquisition and analysis. The samples containing 0.20 A_{260} unit of duplex were dissolved in 1 mL of buffer containing 10 mM Na_2HPO_4/NaH_2PO_4 at pH 7.0, 100 mM NaCl, 50 μM Na_2EDTA. The UV absorbance measurements were performed for unmodified and modified duplexes, at 260 nm, at 1 min intervals with a 1 °C/min temperature gradient. The temperature was cycled between 10 and 85 °C. The first derivative of the melting curve was used to establish the T_M values.

For the samples of DDD, DDD^m and DDD^{hm} a complete thermodynamic analysis was performed in collaboration with the Marky Lab from the University of Nebraska

Medical Center. The detailed methodology is described below. The UV melting experiments were performed on Lambda-10 Perkin-Elmer or AVIV spectrophotometers, equipped with a thermoelectrically controlled sample holder. The UV absorbances at 260 nm and 275 nm were measured in the temperature range of 0–90 °C with a heating rate of 0.6 °C /min. The melting temperature, T_M, was determined from the midpoints of the UV melting curves (Figure 2.2). The van't Hoff enthalpy, ΔH_{vH}, was determined from the slopes of the linear plots of the experimentally measured values of $1/T_M$ vs. ln (C_T), according to the relationship:

$$\frac{1}{T_M} = \left(\frac{R}{\Delta H_{vH}}\right) lnC_T + \left(\frac{\Delta S_{vH}}{\Delta H_{vH}}\right) \quad (1)$$

where C_T is the total concentration of DNA strands. The nature of complex formation of each oligonucleotide was studied by performing UV melts as a function of their total strand concentration (3–330 μM). To understand the molecular changes accompanying the unfolding transitions, the differential binding of counterions (Δn_{Na+}) was determined by performing UV melts in the presence of salt (10–200 mM NaCl). This parameter is calculated using the following thermodynamic relationship (*83, 84*):

$$\Delta n_{Na+} = 0.483 \left(\frac{\Delta H_{cal}}{RT_M^2}\right) \left(\frac{\partial T_M}{\partial \log [Na^+]}\right) \quad (2)$$

where R is the molar gas constant. The first bracketed term in is experimentally determined from differential scanning calorimetric experiments, while the second bracketed term is determined by measuring the T_M at different NaCl concentrations (*83*).

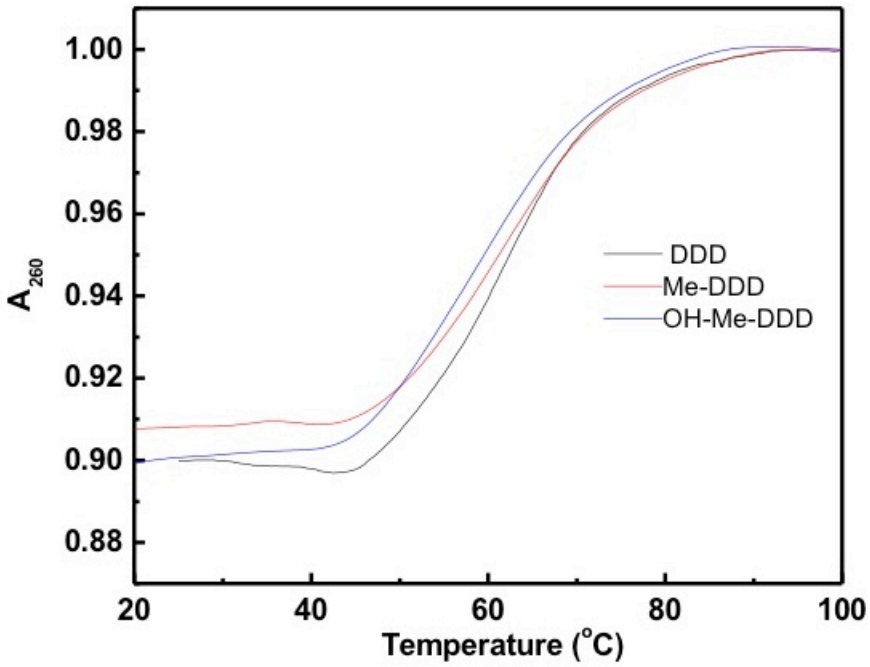


Figure 2.2. The example of the UV melting curve.

Differential Scanning Calorimetry (DSC).

The changes in heat capacity (ΔC_P) as a function of temperature were determined with a VP-DSC differential scanning calorimeter (MicroCal, LLC, Northampton, MA, USA). In a typical DSC experiment, ~200 μM on total strand concentration of the oligonucleotide solution in the sample cell (0.506 mL cell volume) was scanned from 0–90 °C at a rate of 45 °C/h with buffer in the reference cell. A buffer versus buffer scan was also done under similar conditions and subtracted from the sample runs. The resulting thermograms were plotted as ΔCp versus T profiles using the Origin 7.0 software provided with the instrument. Analysis of the integrated plots of the anomalous ΔC_p versus temperature curves ($\int \Delta C_P dT$) and normalization for the number of moles, yields the molar calorimetric enthalpy (ΔH_{cal}). The molar calorimetric entropy (ΔS_{cal})

was calculated from integration of the ΔC_P/T versus T curves [$\int(\Delta C_P/T)$dT]. The Gibbs equation was used to calculate the Gibbs free energy at 20 °C:

$$\Delta G_{(20)} = \Delta H_{cal} - T\Delta S_{cal} \quad (3)$$

Furthermore, heat capacity measurements for the unfolding of each duplex were measured indirectly in DSC experiments at different salt concentrations, this information is obtained from the slope of the ΔH_{cal} versus T_M plots.

Circular dichroism (CD) Spectroscopy.

CD measurements were conducted on an Aviv (model 202SF) CD spectrometer (Lakewood, NJ, USA). The spectrum of each duplex was obtained using a strain-free 1 cm quartz cell at low temperatures to ensure 100% duplex formation. Typically, 1 OD of a duplex sample was dissolved in 1 mL of 10 mM sodium phosphate buffer at pH 7.0. The reported spectra correspond to an average of three scans from 220 to 350 nm with a wavelength step of 1 nm.

Nuclear Magnetic Resonance Spectroscopy

The modified and unmodified samples were diluted to a duplex concentration of 0.25 mM in 180 μL of 10 mM NaH_2PO_4, 100 mM NaCl, 0.011 M NaN_3, 50 μM Na_2EDTA buffer. The samples were exchanged with D_2O and dissolved in 180 μL of 99.99% D_2O to observe nonexchangeable protons in the spectra. For the observation

of exchangeable protons, the samples were dissolved in 180 μL of 9:1 H_2O:D_2O. The NOESY and DQF-COSY spectra of samples in D_2O were collected at 15 °C on a Bruker Avance spectrometer operating at 900 MHz with 5 mm CPTCI cryoprobe. For assignment of exchangeable protons, NOESY experiments with mixing times of 150, 200 and 250 ms and TPPI quadrature detection was conducted. These data were recorded with 2048 real data points in the *t2* dimension and 1024 data points in the *t1* dimension. The relaxation delay was 2.0 s. The data in the *t1* and t2 dimension were zero-filled to give the matrix of 2K×2K real points. The NMR spectra for the exchangeable protons were recorded at 5, 15, 25, 35, 45, 55 and 65 °C. The NOESY spectra of unmodified and modified samples in H_2O were collected at 5 °C with 70 and 250 ms mixing times and relaxation delay of 2.0 s. Water suppression was achieved by a gradient Watergate pulse sequence. Chemical shifts were referenced to water. NMR data were processed using TOPSPIN software (2.0.b.6, Bruker Biospin Inc., Billerica, MA).

Imino Exchange Measurements

Characteristic of base pair opening processes in DNA relies upon the exchange properties of imino protons, i.e. N1H in guanines and N3H in thymines. The opening of individual base pairs in DNA is generally characterized from the exchange of imino protons (Figure 2.3) with solvent protons. In the native DNA double helix, the imino protons are not accessible to solvent due to their location in the center of the structure and their participation in hydrogen bonds. For the exchange to occur, the base pairs must open up. In this opening reaction, the hydrogen bond holding the imino proton

is transiently broken and the proton is moved into a solvent accessible state, where it can be transferred to acceptors present in solution (*48*).

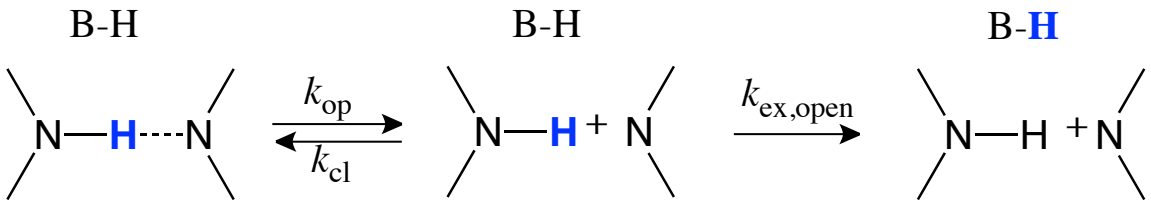


Figure 2.3. Base pair opening process and imino proton exchange.

The exchange rate observed experimentally depends upon the kinetic parameters of the opening reaction as:

$$k_{ex} = \frac{k_{op} \cdot k_{ex.open}}{k_{cl} + k_{ex,open}} \qquad (4)$$

where k_{op} and k_{cl} are the rates of opening and closing of the base pair, and $k_{ex,open}$ is the rate of exchange from the open state. To determine the opening and closing rates, one varies the rate of exchange from the open state, $k_{ex,open}$, by adding to the DNA solution increasing concentrations of a proton acceptor. In the present work ammonia was used as a proton acceptor due to its small size and lack of charge, and to minimize catalysis due to presence of OH^- ions (*85*). Increasing concentrations of ammonia base catalyst were obtained by adding to the sample small aliquots of an ammonia stock solution at the desired pH. The DNA samples contained 1mM TEOA (triethanolamine), which was used to monitor the pH of the sample directly in the NMR tube during ammonia titration. This was done by measuring the difference between the chemical

shifts of the resonances of the two methylene groups of TEOA, according to previous reports (*50*). The pH was measured after each ammonia titration, and its experimental range was found to be 8.8-9.0. A pK_a value 9.218 for ammonia at 15 °C was used. The final concentration of the active form of the ammonia base catalyst was calculated from the total ammonia concentration (c_0) and pH as follows:

$$[B] = c_0 \cdot \frac{10^{-pK}}{10^{-pH}+10^{-pK}} \quad (5)$$

The rate of exchange of the imino proton from the open state of the base pair depends on the concentration of proton acceptor B:

$$k_{ex,open} = k_0 + \alpha k_B[B] \quad (6)$$

where k_0 is the rate of exchange from the open state in the absence of ammonia, k_B is the rate constant for transfer of the imino proton to base catalyst in isolated nucleotides, and α is a factor that accounts for any differences in the rate of proton transfer between isolated nucleotides and open DNA base pairs, e.g., restricted accessibility of the proton acceptor to the imino proton in the open base. Previous investigations have shown that, for ammonia base, α is close to unity (*85*). k_B is the rate constant for exchange catalysis and is calculated from (*50*):

$$k_B = \frac{k_D}{1+10^{pK_a^{Nu}-pK_a^B}} \quad (7)$$

in which k_D is the bimolecular collision rate between the imino group and proton acceptor in the open state (1.0×10^9 $M^{-1}s^{-1}$ for ammonia at 15°C), pK_a^{Nu} and pK_a^{B} are the pK_a values for the imino proton of interest and the ammonia base catalyst, respectively. The final equation for the exchange rate is obtained by inserting eq 6 into eq 4 (with $\alpha = 1$):

$$k_{ex} = \frac{k_{op}(k_0 + k_B[B])}{k_{cl} + k_0 + k_B[B]} \quad (8)$$

Two kinetic regimes can be distinguished depending on how the rate of exchange from the open state compares with the rate of closing of the base pair (*50*). The EX2 regime occurs when the concentration of proton acceptor is low such that $k_{ex,open} << k_{cl}$. In this case, the observed rate of exchange is proportional to the rate of exchange from the open state (eq 4), and thus to the concentration of proton acceptor B:

$$k_{ex} = \frac{k_{op}}{k_{cl}} \cdot k_{ex,open} = K_{op} \cdot (k_0 + k_B[B]) \quad (9)$$

where $K_{op} = k_{op}/k_{cl}$ is the equilibrium constant for opening of the base that contains the imino proton. The other regime, called EX1 regime, occurs at high concentrations of proton acceptor, when $k_{ex,open} >> k_{cl}$. In this regime, the exchange occurs in each opening event and $k_{ex} = k_{op}$. For the determination of the base pair opening (k_{op}), imino exchange rates are measured as a function of titrated ammonia base catalyst and are fit to equation 8.

For the imino exchange experiments NMR data were collected at 15 °C at 500 MHz using a Bruker AV-III spectrometer equipped with 5 mm CPQCI probe. The samples were dissolved in 180 mL of 90% H_2O, 10% D_2O solution containing 10 mM phosphate buffer, 100 mM NaCl, 0.05 mM EDTA, 0.011 M NaN_3, 1 mM TEOA (pH 8.9). Magnetization transfer from water to the imino protons was followed by observation of the imino protons after a variable mixing time. For selective spin inversion of the water protons, a 2 ms, a 180° Sinc pulse with 1000 points was used. To minimize effects of radiation damping during the mixing time, a 0.1 $G \cdot cm^{-1}$ gradient was used. Water suppression was achieved by a binominal 1-1 echo sequence, jump and return, with flanking 1 ms smooth square shape gradients, 15 $G \cdot cm^{-1}$. Sixteen values for the variable delay ranging form 1 ms to 15 s were used for each experiment. All data were processed and analyzed in TOPSPIN.

The exchange rates were calculated from the following equation:

$$\frac{I_z(t_{mix})}{I_{z,eq}} = 1 + Ek_{ex}(e^{-R_{1i}t_{mix}} - e^{-R_{1w}t_{mix}}) \qquad (10)$$

in which I_z (t_{mix}) and $I_{z,\,eq}$ are intensities of the imino proton peaks at a given value of t_{mix} and at equilibrium, respectively; k_{ex} is the chemical exchange rate, R_{1w} is the longitudinal relaxation rate of water (3.15 sec as determined separately under the same conditions), R_{1i} is the sum of the imino proton relaxation rate and k_{ex}, and E is the efficiency of water inversion (value of -2 as given elsewhere (*86, 87*)).

The data analysis and fits were performed using PRISM software (v. 6.0b, GraphPad Software, Inc.).

Gaussian calculations

GAUSSIAN09 (*88*) was used for the calculation of bond length, angle and torsion angle values for non-standard residues (i.e. 5hmC, 5fC, 5caC). The restrained electrostatic potential charges were calculated using the B3LYP Density Functional Theory (DFT) method with a 6-31G* basis set.

X-ray Crystallography

Crystallization of DNA

Crystals of the DDD^{hm}, or DDD^{f} or DDD^{ca}, were grown at 18 °C by the hanging-drop vapor-diffusion method, using the Nucleic Acid Mini-screen (Hampton Research, Aliso Viejo, CA). Droplets of 2 μL volume containing 1.2 mM modified DNA duplex, and precipitating solution, were equilibrated against 0.75 μL of 35% MPD. Crystals grew within 8-16 days. Precise crystallization conditions and solution compositions of particular oligodeoxynucleotides are summarized in Table 2.1. Single crystals from each DNA sample were mounted in nylon loops and flash-frozen in liquid nitrogen.

Condition	**DDD^{hm}**	**DDD^{f}**	**DDD^{ca}**
pH	7.0	6.0	6.0
Buffer	40 mM Na Cacodylate	40 mM Na Cacodylate	40 mM Na Cacodylate
Salts	80 mM NaCl, 20 mM $MgCl_2$	80 mM NaCl	80 mM $SrCl_2$
Additives	12 mM spermine 4HCl	12 mM spermine 4HCl	12 mM spermine 4HCl
2-methyl-2,4-pentadienol (MPD)	10 % (v/v)	10 % (v/v)	10 % (v/v)

Table 2.1. Crystallization conditions.

X-ray Diffraction Data Collection and Processing

For the DDD^{hm} crystal diffraction data were collected on the 19-ID beamline of the Structural Biology Center at the Advanced Photon Source of Argonne National Laboratory (Argonne, Il) (*89*). The wavelength was 0.9794 Å. Initial indexing and scaling of recorded diffraction images, together with further reflection merging, were done using HKL3000 (*90*). In order to ensure high completeness of data two separate passes were collected.

For the DDD^{f} and DDD^{ca} crystals, data were collected on the 24-IDE and 24-IDC beamlines, respectively, of the Northeastern Collaborative Access Team (NE-CAT) at the Advanced Photon Source of Argonne National Laboratory (Argonne, IL). The wavelength was 0.97920 Å. Initial indexing and scaling of recorded diffraction images, together with further reflection merging, were done using XDS (*91, 92*), and SCALA (*93*) in the CCP4 (*94*) suite as part of the RAPD data-collection strategy at NE-CAT. Data collection details are shown in Table 2.2:

Parameter	DDD^{hm}	DDD^{f}	DDD^{ca}
Crystal data			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell			
a [Å]	25.6	25.3	24.2
b [Å]	41.3	41.5	41.3
c [Å]	64.3	65.6	66.4
Data collection			
Resolution range [Å]	40-1.00	35-1.74	26-1.95
Reflections			
Observed	338726	8305	20462
unique	37637	3296	5113
Completeness (%)	99.6	99.4	99.5
In the outer shell (%)	98.5	94.8	97.5
R_{merge} [a]	0.044	0.034	0.045
In the outer shell [b]	0.979	0.747	0.619
I/ σI	52	20	16
In the outer shell [b]	1.65	2.64	2.82
Structure refinement			
Resolution range [Å]	40-1.02	35-1.74	26-1.09
R-work	0.15	0.23	0.22
R-free	0.17	0.29	0.26
RMS deviation			
Bond lengths [Å]	0.018	0.009	0.008
Angle distances [Å]	1.015	1.747	2.209
Number of ions	1 Mg^{2+}	-	-
Number of ligands	3	-	-
Number of water molecules	178	56	13

Table 2.2. Crystal Data, Data Collection, and Refinement Statistics for DDD^{hm}, DDD^{f} and DDD^{ca}.

[a] $R_{merge}=\Sigma_{hkl}\Sigma_i \left| Ii - < I > \right| / \Sigma_{hkl}\Sigma_I \left| < I > \right|$, where *Ii* is the intensity for the i_{th} measurement of an equivalent reflection with indices *h*, *k*, and *l*.
[b] Numbers in parentheses are values for the highest-resolution bin.

Crystal Structure Determination and Refinement

The structures of the modified DDDs were determined by molecular replacement using the Dickerson-Drew dodecamer as the search model (PDB code 436D (*77*). Molecular replacement searches were completed with program MOLREP of the CCP4 suite (*94-96*). An initial model was subsequently manually checked and rebuilt in program COOT (*97*). The final model was further rebuilt and refined using REFMAC 5.6 (*98, 99*). The final models were refined against all reflections, except for 5% randomly selected reflections that were used for monitoring R_{free}. The final refinement statistics are presented in Table 2.2.

Helicoidal Analysis

The calculations of helicoidal parameters, including backbone torsion angles, intra- and interbase translations and rotations, was performed using Curves+ (*100*). For crystal structures, the last refined structure was used, prior to validation.

CHAPTER 3

CHARACTERIZATION OF 5-HYDROXYMETHYLCYTOSINE IN DNA: COMPARISON WITH 5-METHYLCYTOSINE

Introduction

The methylation of cytosine by DNA methyltransferases (*101, 102*) to form 5-methylcytosine (5mC), plays an important role in the epigenetic regulation of the eukaryotic genome (*103, 104*). The reverse process of active demethylation has been of considerable interest (*105, 106*) and has led to the important discovery of 5-hydroxymethylcytosine (5hmC) in mammalian DNA (*107, 108*). The 5hmC base is generated from 5mC through the action of ten-eleven translocation (TET) deoxygenases (*109-111*). It is also formed in response to oxidative stress as a consequence of UV radiation (*112*). It is thought that 5hmC is itself an important epigenetic marker and transcriptional regulator. Support for this hypothesis comes from the observation that altered levels of 5mC and 5hmC are observed in early embryonic development, embryonic stem cell differentiation, and tumors (*109, 113*). Cellular levels of 5hmC are tissue-specific, with the highest levels found in the central nervous system (*114*). 5hmC values increase during brain development, suggesting a role of this modification in brain maturation and neuronal development (*108*). The balance between 5mC and 5hmC at gene promoters and CpG islands appears to be linked to pluripotency of the cell (*110*).

Further oxidation of 5hmC by TET deoxygenases leads to the formation of 5-formylcytosine (5fC) and 5-carboxycytosine (5caC) (*109-111*). However, mass

spectrometric analyses have led to estimates that some hundreds of pmols of 5hmC are present in mammalian tissues, while the levels of 5fC and 5caC are much lower (*114*) suggesting that the latter oxidation products are more efficiently removed from DNA than is 5hmC. Indeed, Maiti and Drohat (*115*) showed that both 5fC and 5caC are substrates for thymine DNA glycosylase (TDG), but not 5hmC. This suggested that active demethylation of 5mC requires further oxidation of 5hmC to the downstream products 5fC and 5caC, allowing TDG to restore the epigenetically unmodified G:C pair.

Mechanistically, TDG employs the familiar extrahelical base flipping mechanism (*79-81*) to position its substrates into its active site for catalysis. Recently, Renciuk *et al*. (*116*) obtained a structure of the Dickerson dodecamer, in which either a 5mC or 5hmC modification was placed site-specifically either at the C^9:G^{16} base pair or the C^3:G^{22} base pair. They concluded that the presence of either 5mC or 5hmC did not influence the overall DNA structure. The hydroxyl group of the 5hmC base was oriented toward the 3' end of the duplex, away from the phosphate backbone. They also reported the thermodynamic effects of these modified bases using a combination of CD and UV spectroscopy. The presence of the 5hmC modification resulted in a slight destabilization of the duplex. They proposed that the cytosine C5 carbon provides an ideal location to encode epigenetic information. Since polymerases may not be able to distinguish 5hmC from 5mC and unmodified cytosine, these modifications might be anticipated to be non-mutagenic.

Maiti and Drohat obtained the structure of the hTDG catalytic domain complexed with DNA containing an abasic site, (*117*) which revealed interactions promoting the specificity requirement of guanine vs. adenine as the pairing partner of the target base and additional protein-DNA interactions associated with the specificity for CpG sites. Further studies showed that the conserved Asn^{140} residue was implicated in the chemical step, whereas the conserved Arg^{275} was implicated in nucleotide flipping into the active site (*118*). Subsequently, Hashimoto *et al*. (*119*) determined the structure of a post-

reactive complex of hTDG with a DNA containing a G:5hmU mismatch, showing that the glycosylase had flipped the 5hmU nucleotide from the DNA, and suggesting that TDG allows hydrogen-bonding interactions to both 5hmU and 5caC. They proposed that amino-imino tautomerization of the substrate base may explain how TDG discriminates against 5hmU and 5caC.

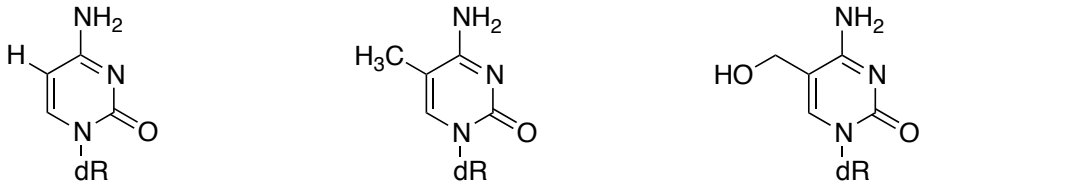
To further characterize the conformational consequences of 5hmC inserted into a DNA duplex, we collected NMR data on the Dickerson-Drew dodecamer (DDD) (*120, 121*) site-specifically modified with 5hmC, $[5'\text{-d}(C^1G^2C^3G^4A^5A^6T^7T^8\underline{Z}^9G^{10}C^{11}G^{12})\text{-}3']_2$, $\underline{Z}^9$=5hmC (Figure 3.1). We used high-resolution 2D NMR data to further characterize the 5hmC-modified duplex, and compared it to the 5mC duplex. We have also characterized the kinetics of base-pair opening in 5hmC-modified DNA by measuring the rates of exchange of imino protons with solvent protons by magnetization transfer from water, as a function of the concentration of exchange catalyst. Our results show that the equilibrium constant for opening of 5hmC is low. We gathered differential scanning calorimtery (DSC), and additional CD and UV spectroscopy data to characterize the thermodynamic effects of these 5-substituted cytosines in DNA. Finally, we determined at 1 Å resolution resolution the structure of the DDD with 5hmC at base pair C^9:G^{16}.

Results

5-hydroxymethyl-cytosine in comparison with 5-methyl-cytosine

The DNA duplexes used in these studies are shown in Figure 3.1. The sequence we chose for these studies is well-established, self complementary Dickerson Drew dodecamer. The nineth residue in this sequence is replaced by either 5mC or 5hmC.

A



Cytosine; C 5-Methyl-cytosine; 5meC 5-Hydroxymethyl-cytosine; 5hmC

B

DDD
5'-C^{1} G^{2} C^{3} G^{4} A^{5} A^{6} T^{7} T^{8} C^{9} G^{10}C^{11}G^{12}-3'
3'-G^{24}C^{23}G^{22}C^{21}T^{20}T^{19}A^{18}A^{17}G^{16}C^{15}G^{14}C^{13}-5'

DDD^{m}
5'-C^{1} G^{2} C^{3} G^{4} A^{5} A^{6} T^{7} T^{8} Y^{9} G^{10}C^{11}G^{12}-3'
3'-G^{12}C^{11}G^{10}Y^{9} T^{8} T^{7} A^{6} A^{5} G^{4} C^{3} G^{2} C^{1} -5'

DDD^{hm} (NMR)
5'-C^{1} G^{2} C^{3} G^{4} A^{5} A^{6} T^{7} T^{8} Z^{9} G^{10}C^{11}G^{12}-3'
3'-G^{12}C^{11}G^{10}Z^{9} T^{8} T^{7} A^{6} A^{5} G^{4} C^{3} G^{2} C^{1} -5'

DDD^{hm} X-ray
5'-C^{1} G^{2} C^{3} G^{4} A^{5} A^{6} T^{7} T^{8} Z^{9} G^{10}C^{11}G^{12}-3'
3'-G^{24}C^{23}G^{22}Z^{21}T^{20}T^{19}A^{18}A^{17}G^{16}C^{15}G^{14}C^{13}-5'

Figure 3.1. A. Structure of dC, 5mC and 5hmC. B. Sequences and numbering of the nucleotides for unmodified DDD, DDD^{m}, DDD^{hm} (NMR) and DDD^{hm} (X-ray) duplexes. In solution, the two strands of the DDD exhibit pseudo-dyad symmetry and the NMR resonances of symmetry-related nucleotides in the two strands are not individually observed. In the crystal, electron density for symmetry-related nucleotides in the two strands is observed and the nucleotides are numbered individually.

Unfolding Thermodynamic of Oligonucleotide Duplexes.

Typical UV melting curves of each duplex at 260nm and 275nm are compared in Figure 3.2.

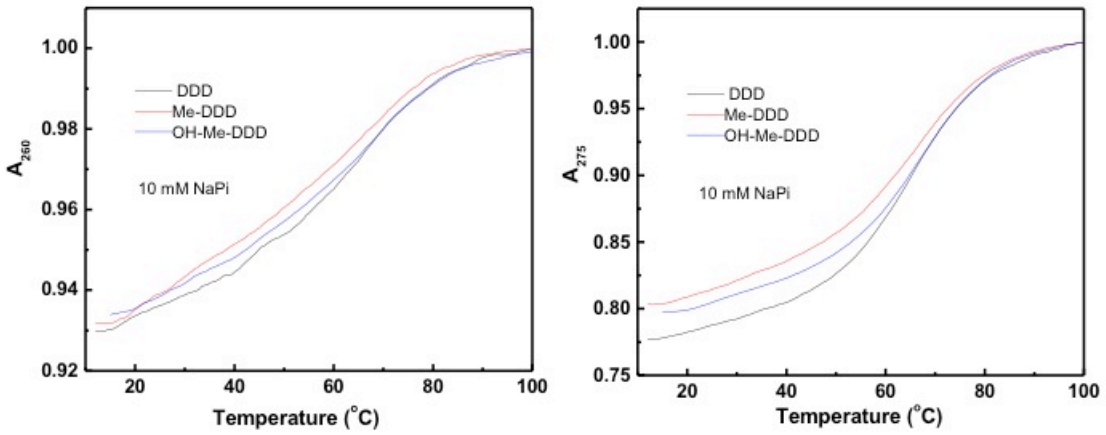


Figure 3.2. UV melting curves of DDD, DDD^{m} and DDD^{hm} at 260 nm and 275 nm, in 10 mM NaP_i buffer.

These curves show that at the lowest temperatures, below ~ 20 °C, each molecule is in the duplex state, while at higher temperatures, all curves are sigmoidal in shape which is characteristic of the unfolding of DNA duplexes. The inclusion of a methyl or hydroxymethyl groups into cytosine causes a weak and minimal thermal destabilization of the DNA. Their thermal stability is in the order: DDD^{hm} < DDD^{m} < DDD. Analysis of the melting curves as a function of C_T shows that the T_M values increase with increasing DNA concentration, Figure 3.3, as expected for the unfolding of self-complementary duplexes into two single strands.

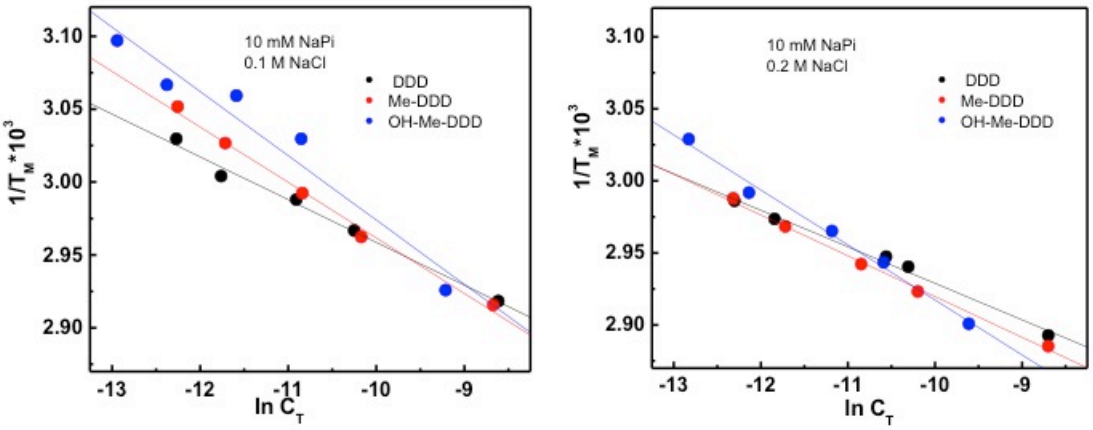


Figure 3.3. T_M-dependences on strand concentration of DDD, DDD^{m} and DDD^{hm}.

The CD spectra of each duplex at 20 °C show a positive Cotton effect at 280 nm and a strong negative Cotton effect around 250 nm, which is characteristic of a right-handed helix in the B-conformation. The intensities of both positive and negative bands are similar in magnitude, the intensity of the negative band correlates with the extent of base-pair stacking contributions. Thus, the incorporation of 5mC and 5hmC contributes equally to the extent of base-pair stacking interactions among these three duplexes. The CD spectra at 90 °C correspond to the characteristic spectra of DNA random coils.

The DSC melting curves for all three duplexes in 0.1 M and 0.2 M NaCl buffered solution are compared in Figure 3.4.

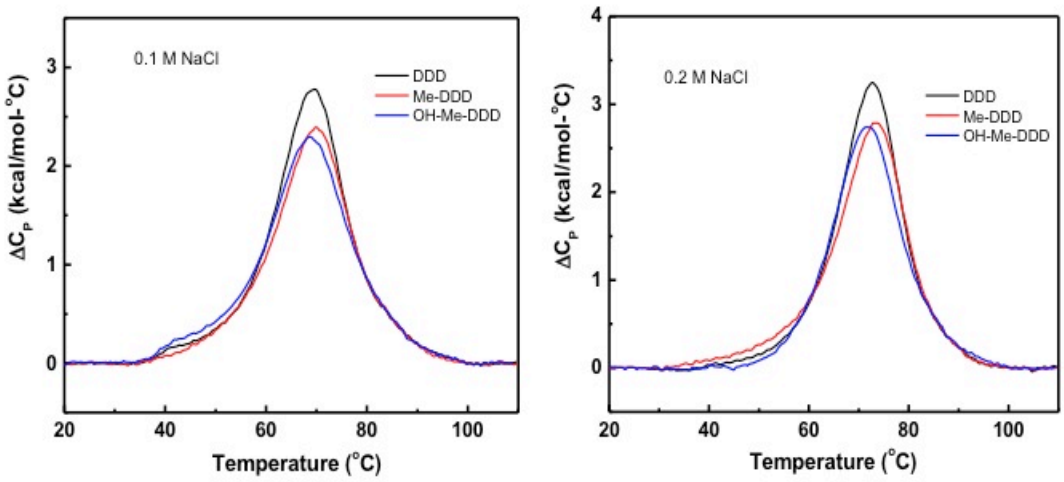


Figure 3.4. DSC melting curves of DDD, DDD^{m} and DDD^{hm} in 10 mM NaP_i buffer at pH 7 and 0.1 M NaCl (left) or 0.2 M NaCl (right).

The thermodynamic profiles at 0.1 M and 0.2 M NaCl are summarized in Table 3.1. The unfolding of each duplex is monophasic, the increased in salt shifts the duplex transition curves to higher temperatures, and this is due to a higher screening by salt of the duplex phosphates. Analysis of these curves revealed a decreased average endothermic enthalpy of 5.5 kcal/mol for the modified duplexes relative to DDD in the 0.1-0.2 M NaCl range (Table 3.1). This decrease in enthalpy, due to the incorporation of modified cytosines, is within the experimental error of the enthalpy determination. The lowering of the enthalpies at lowered salts suggests the presence of heat capacity effects. The heat capacity values were 0.6 kcal/K·mol (DDD), 1.7 kcal/K·mol (DDD^{m}) and 1.1 kcal/K·mol (DDD^{hm}). The presence of heat capacity effects has been attributed to the exposure of non-polar groups to the solvent and/or to changes in structural hydration between the unfolded and folded states (*83*).

Sample	T_M (°C)	ΔH_{cal} (kcal/mol)	$\Delta G°_{(20)}$ (kcal/mol)	$T\Delta S$ (kcal/mol)	ΔH_{vH}* (kcal/mol)	Δn_{Na+} (mol $_{Na+}$/ mol)	ΔC_p (kcal/mol-°C)
DDD	69.6 (72.6)	-108.4 (-110.2)	-15.7 (-16.8)	-92.7 (-93.4)	68 (78)	-2.8	-0.6
DDD^m	69.9 (73.5)	-99.6 (-105.8)	-14.5 (-16.3)	-85.1 (-89.5)	52 (70)	-2.6	-1.7
DDD^{hm}	68.7 (71.6)	-102.2 (-105.3)	-15.7 (-15.0)	-89.5 (-85.0)	45 (52)	-2.5	-1.1

Table 3.1. Folding thermodynamic profiles for DDD, DDD^m and DDD^{hm}. All experiments were performed in 10 mM NaP_i buffer at pH 7.0 and 0.1 M NaCl or 0.2 M NaCl in parenthesis. *The ΔH_{vH} is calculated from the dependences of the T_M on strand concentration.

Furthermore, the average values of the ΔH_{vH} /ΔH_{cal} ratios at these salt concentrations are: 0.67 (DDD), 0.59 (DDD^m) and 0.46 (DDD^{hm}), indicating all three duplexes unfolds in non two-state transitions. The incorporation of the modified bases actually lowers the size of the cooperative melting unit.

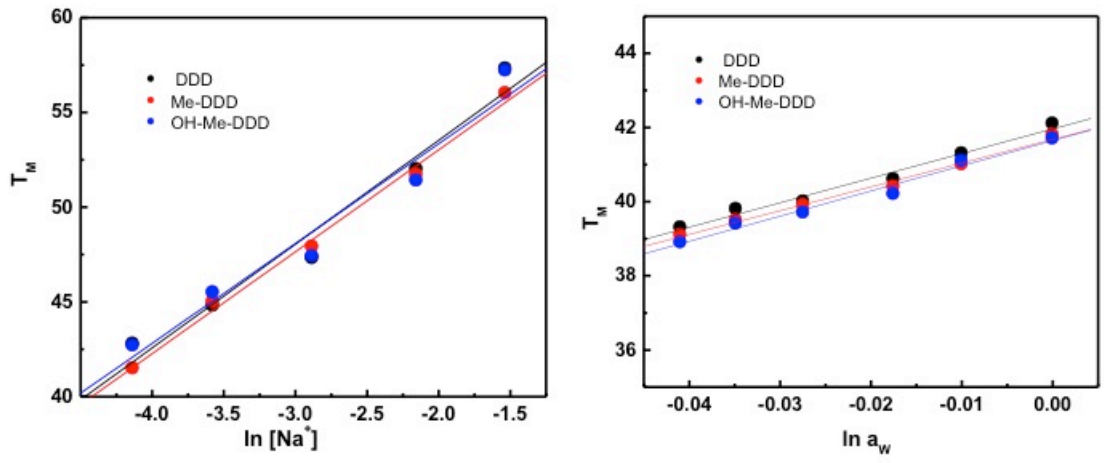


Figure 3.5. T_M-dependences of DDD, DDD^m and DDD^{hm} on salt concentration and water activity.

UV melting curves were measured at different salt concentrations, 16-216 mM NaCl, to determine the differential binding of counterions, Δn_{Na^+}. The T_M values increase linearly as a function of *ln* [Na+] as shown in Figure 3.5., consistent with the observation that salt favors the duplex state with a higher charge density parameter. We obtained Δn_{Na^+} values (Table 3.1.) of 2.8 mol Na^+/mol DNA (DDD), 2.6 kcal/K·mol (DDD^m) and 2.5 kcal/K·mol (DDD^{hm}).

The thermodynamic profiles for the folding of each duplex are listed in Table 3.1. Analysis of the data shows that a favorable Gibbs free energy accompanies the stable formation of each duplex, which results from compensation of favorable enthalpy-unfavorable entropy contributions. The favorable enthalpies arise from the formation of base-pairs and base pair stacks. Unfavorable entropy terms includes the ordering of two strands to form a duplex and condensation of counterions.

Relative to the unmodified dodecamer duplex, the DDD^m and DDD^{hm} modified oligomers were destabilized at 0.1 M and 0.2 M NaCl concentrations. Specifically, the inclusion of two 5mC or two 5hmC modifications in DDD yielded an average decrease in $\Delta G^\circ_{(20)}$ of 0.9 kcal/mol, which is enthalpy driven ($\Delta\Delta H$ = 6.1 kcal/mol).

NMR melting studies

A series of 1D ^{1}H NMR spectra for the exchangeable protons were recorded at 5, 15, 25, 35, 45, and 55 ºC and are shown in Figure 3.6. The data show that the G^2 N1H imino proton resonance of the G^2:C^{11} base pair in the DDD^m was sharp and detectable only at 5 °C. The same imino peak in the DDD^{hm} began to broaden at 25 °C, and at higher temperatures this resonance disappeared. In the unmodified duplex the G^2

N1H imino proton of the G^2:C^{11} base pair was sharp up to 20 °C, above which it started to broaden. The G^{10} N1H imino proton of the 5'-neighbor base pair, C^3:G^{10} was significantly broadened at 35 °C in the DDD^m, while still sharp even at 45 °C for the DDD, and at 35 °C in DDD^{hm} this imino resonance started to broaden. The G^4 N1H imino proton of the G^4:C^9 base pair in DDD and the G^4:Z^9 proton of the DDD^{hm} duplex remained sharp at as high temperature as 45 °C, and at 55 °C started to broaden; however the same resonance in the DDD^m at 45 °C was already broadened. The N3H imino resonances of T^7 and T^8 were the sharpest in all three duplexes; they started to broaden at 55 °C in the DDD and DDD^{hm} duplexes, while in the DDD^m these peaks were not observed.

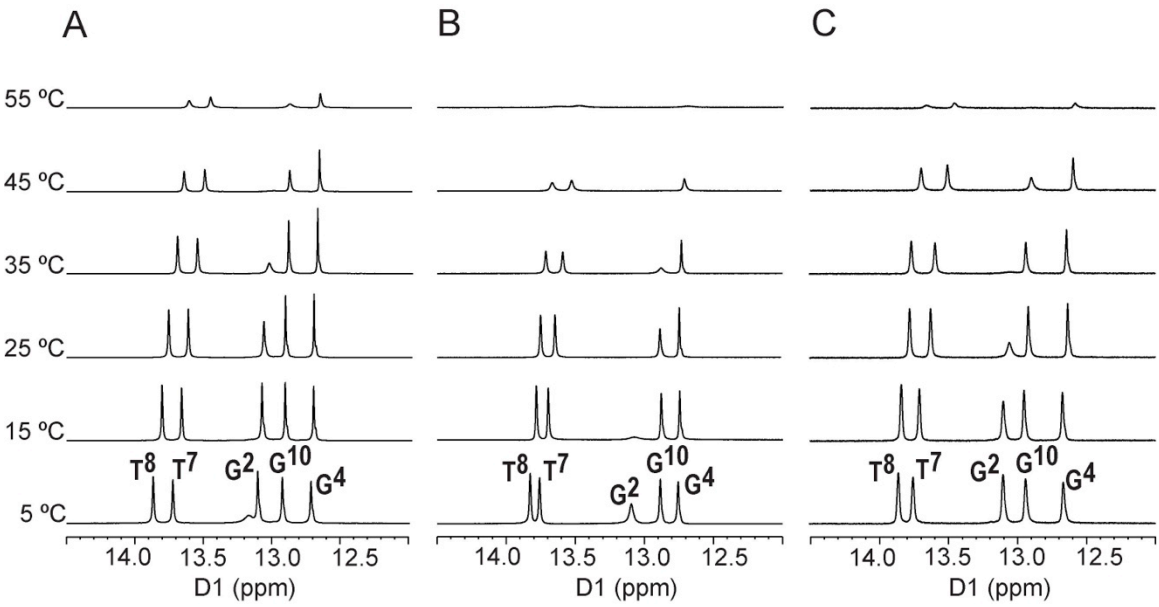


Figure 3.6. ^{1}H-NMR of imino proton resonances as a function of temperature for the unmodified DDD duplex (A), modified DDD^m (B), and modified DDD^{hm} (C) duplexes.

Exchangeable protons

Figure 3.7 shows the NOE connectivity of the purine N1H and pyrimidine N3H imino protons, for the DDD, DDD^{m}, and DDD^{hm} duplexes. The base imino protons were assigned based on their sequential connectivities in NOESY spectra, and these assignments were supported by their NOE cross-peaks to Watson-Crick base-paired amino protons (*69*). The sequential connectivities were obtained from base pairs G^{2}:C^{11}→ C^{3}:G^{10} →G^{4}:C^{9}→ A^{5}:T^{8} → A^{6}:T^{7}. For the DDD and DDD^{m} duplexes the imino-proton resonances of the terminal base pairs C^{1}:G^{12} were not observed, presumably due to fast exchange with water. In the region of the NOESY spectrum showing the NOEs between the imino and amino protons cross-peaks from modified base C^{9} H41, C^{9} H42 to the complementary base G^{4} H1 were observed, as well as interactions to neighbor bases T^{8} H3 and G^{10} H1 (Figure 3.7).

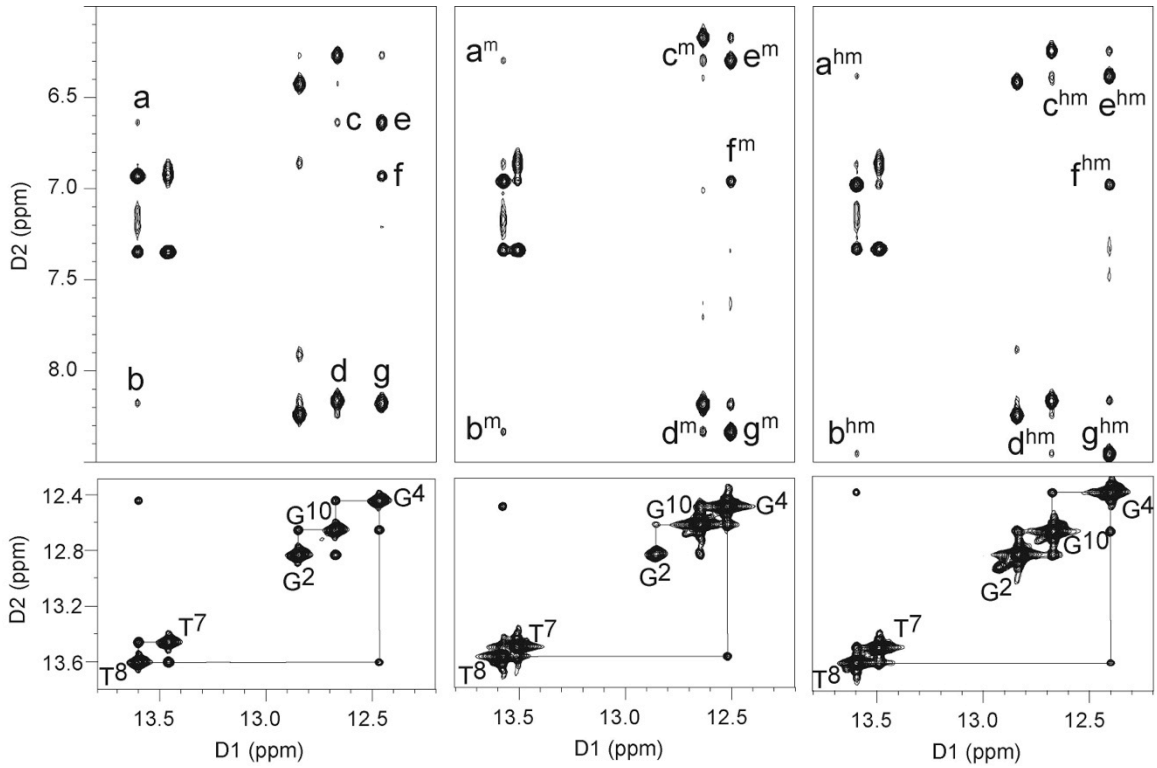


Figure 3.7. ^{1}H-^{1}H NMR NOESY spectrum showing resonances for the thymine and guanine imino protons and sequential NOE connectivity for the imino protons of the base pairs G^{2}:C^{11} to A^{6}:T^{7} for the unmodified DDD (A), modified DDD^{m} (B), and modified DDD^{hm} (C) duplexes. Expansion of the ^{1}H-^{1}H NOESY spectra for the DDD (A), DDD^{m} (B) and DDD^{hm} (C) showing the conservation of Watson-Crick base pairing and base stacking at the modification site. a, $C^{9}H41 \rightarrow T^{8}H3$; b, $C^{9}H42 \rightarrow T^{8}H3$; c, $C^{9}H41 \rightarrow G^{10}H1$; d, $C^{9}H42 \rightarrow G^{10}H1$; e, $C^{9}H41 \rightarrow G^{4}H1$; f, $A^{5}H2 \rightarrow G^{4}H1$; g, $C^{9}H42 \rightarrow G^{4}H1$. (Index ^{m} or ^{hm} refers to the base pairs in the modified duplexes, DDD^{m} and DDD^{hm}, respectively).

Nonexchangeable protons

For each of the three duplexes, the sequential assignment of non-exchangeable protons was accomplished using standard protocols (*122, 123*). In each instance, the anticipated pattern of sequential base aromatic→deoxyribose anomeric NOEs was identified from $C^{1} \rightarrow G^{12}$ (Figure 3.8). Three strong NOEs accounted for H5-H6 cross-peaks of cytosines C^{1}, C^{3}, and C^{11}.

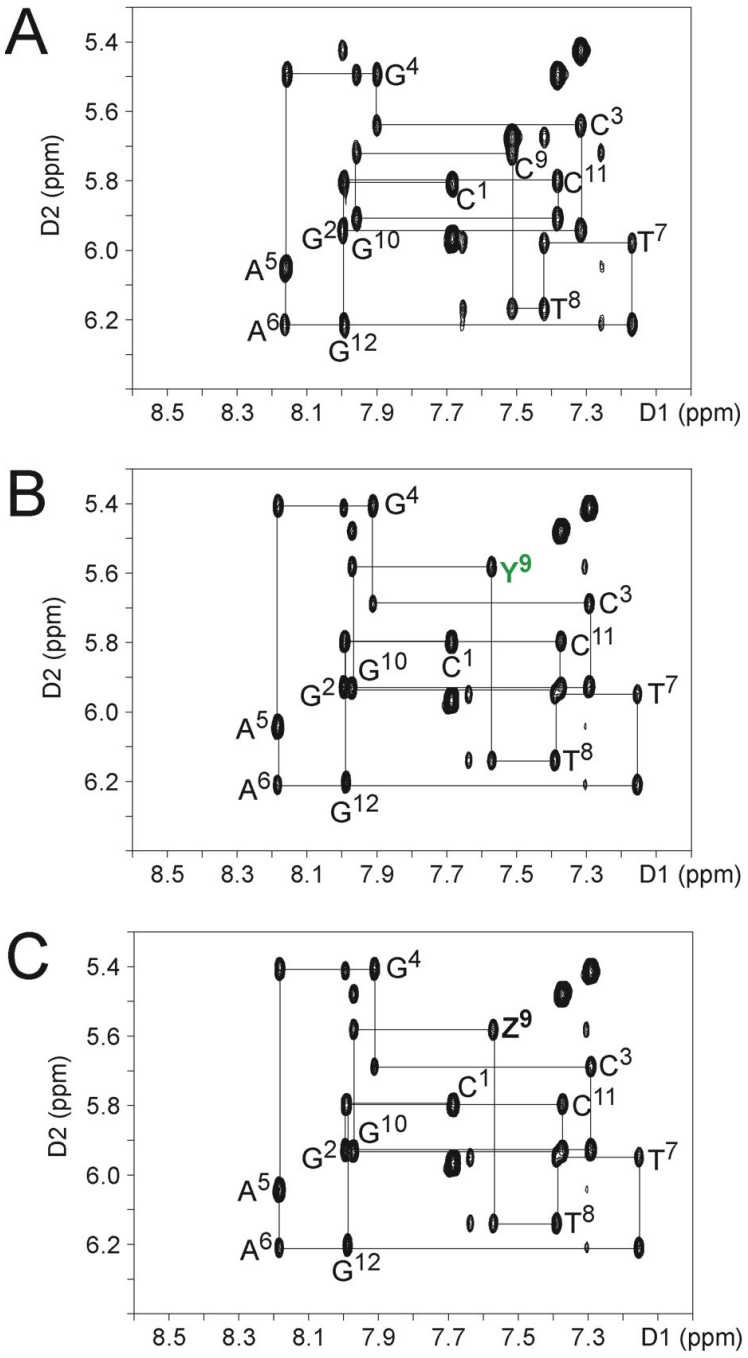


Figure 3.8. Expanded plot from the aromatic-anomeric region of the NOESY spectrum, showing sequential NOE connectivities of the unmodified DDD (A), modified DDD^{m} (B), and modified DDD^{hm} (C) duplexes.

Cytosine assignments were confirmed by a DQF-COSY spectrum recorded at identical conditions (data not shown). As compared to the DDD, in the DDD^{m} and DDD^{hm}

duplexes the cytosine H5 protons at nucleotides Y^9 or Z^9 were absent due to presence of either the methyl or hydroxymethyl substitution at the cytosine C5 position, respectively; thus, no cytosine H5-H6 cross-peaks were observed in this region of the spectrum for nucleotides Y^9 and Z^9. In all instances, for all three duplexes, the NOE cross-peaks intensities between the base protons and the deoxyribose H1' protons were of the same relative magnitudes as those between other bases in the sequence, indicating that the glycosyl bonds maintained the anticipated anti conformations associated with B-type DNA.

Dynamics of base pair opening

The imino proton exchange rates for the three duplexes were obtained by measuring magnetization transfer from water as a function of added ammonia base catalyst. The dependence of the exchange rates on the catalyst concentration is illustrated in Figure 3.9. for the C^3:G^{10}, G^4:C^9, A^5:T^8 and A^6:T^7 base pairs of the DDD, DDD^m and DDD^{hm} duplexes. This figure also shows the fits of the exchange rate as a function of increasing ammonia base concentration to equation 8.

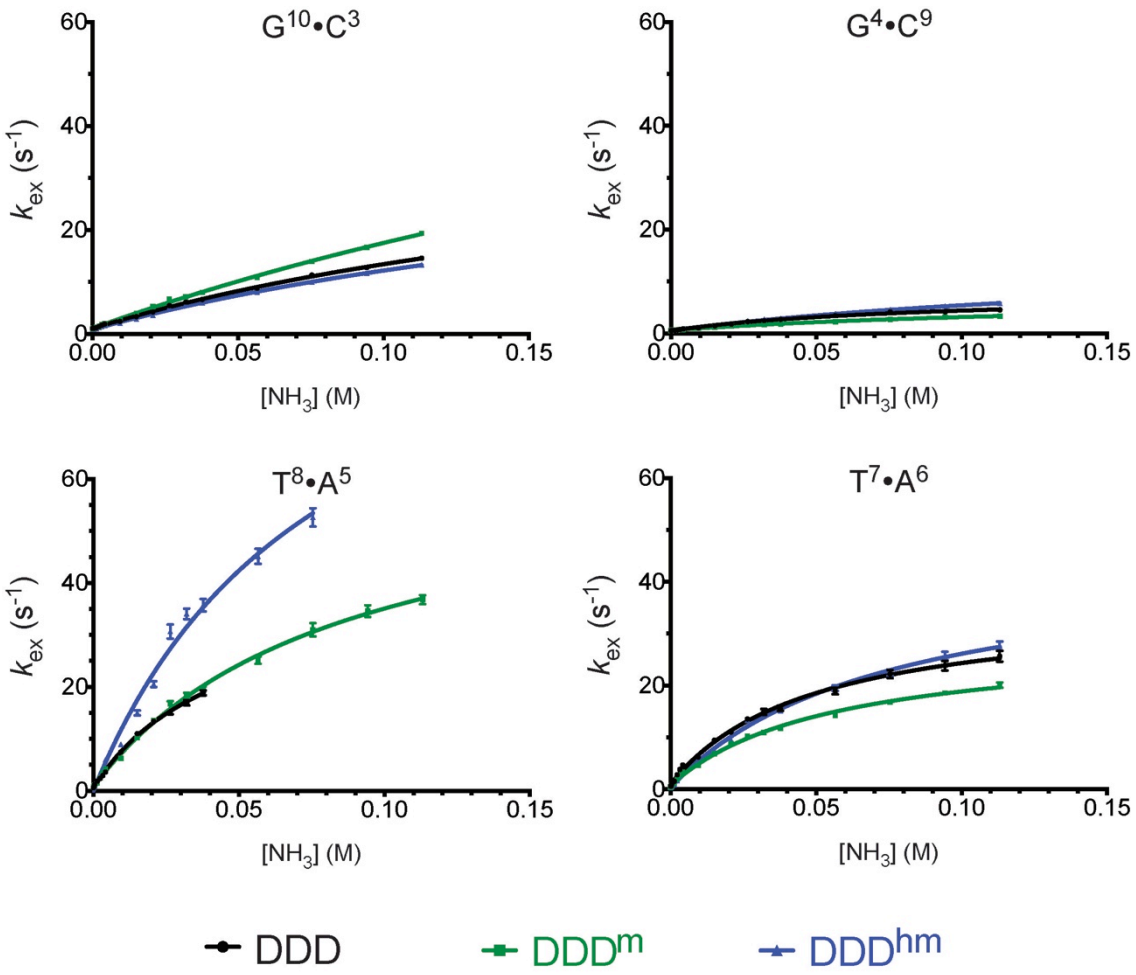


Figure 3.9. Plots showing imino proton exchange rates obtained by monitoring magnetization from water as a function of ammonia base catalyst. A. G^{10}:C^{3}. B. G^{4}:C^{9}. C. T^{8}:A^{5}. D. T^{7}:A^{6} in DDD (black), DDD^{m} (green) and DDD^{hm} (blue).

These fits were used to determine the base pair opening (k_{op}) and closing (k_{cl}) rates. To obtain the opening equilibrium constant (K_{op}) data were fitted to the eq. 9 in the lower range of catalyst concentration where the dependence between exchange rate and concentration of base catalyst is linear. The values of $k_{op,}$ k_{cl} and K_{op} in the three duplexes are summarized in Table 3.2.

	k_0 (s^{-1})[a]			K_{op} x 10^7		
	DDD	DDD^m	DDD^{hm}	DDD	DDD^m	DDD^{hm}
C^3:G^{10}	1.1 ± 0.06	1.0 ± 0.07	0.95 ± 0.05	2.9 ± 0.1	3.6 ± 0.1	2.5 ± 0.1
G^4:C^9	0.57 ± 0.03	0.56 ± 0.1	0.68 ± 0.05	1.2 ± 0.04	0.75 ± 0.04	1.2 ± 0.06
A^5:T^8	0.59 ± 0.03	0.73 ± 0.07	0.77 ± 0.03	41 ± 0.08	35 ± 2	73 ± 6
A^6:T^7	0.61 ± 0.03	0.43 ± 0.04	0.36 ± 0.01	37 ± 5	26 ± 3	29 ± 3

	k_{op} (s^{-1})			k_{cl} (× 10^{-7} s^{-1})		
	DDD	DDD^m	DDD^{hm}	DDD	DDD^m	DDD^{hm}
C^3:G^{10}	45 ± 3	90 ± 12	42 ± 5	15 ± 1	25 ± 2	17 ± 1
G^4:C^9	8 ± 0.5	7 ± 0.6	16 ± 2	6.7 ± 0.2	9.4 ± 1	13 ± 0.8
A^5:T^8	40 ± 2	65 ± 2	110 ± 13	0.97 ± 0.04	1.9 ± 0.04	1.5 ± 0.05
A^6:T^7	36 ± 1	29 ± 1	45 ± 2	0.98 ± 0.09	1.1 ± 0.08	1.6 ± 0.08

Table 3.2. Rate and Equilibrium Constants for DNA Base Pair Opening. [a]The observed exchange rate without an ammonia catalyst.

For each of the observed imino protons, in all three duplexes, as the concentration of ammonia increased, the exchange rate increased until a plateau was reached, which indicated a change for the rate-limiting step from chemical exchange to base pair opening ($k_{ex} = k_{op}$; exchange occurs at each opening event). For the G^4 N1H imino proton, the exchange rate had only a weak dependence on the concentration of ammonia. This reflected the small equilibrium constant for the opening the G^4:C^9 base pair, and resulted in low and similar K_{op} values in the DDD, DDD^m and DDD^{hm} duplexes (1.2×10^7, 0.75×10^7, and 1.2×10^7, respectively). Also, opening rates decreased more for base pair G^4:C^9 than for the other base pairs (Table 3.2), suggesting that the opening

dynamics of this base pair are lower. The opening rate for the G^4:C^9 base pair was ~2 times greater in DDD^{hm} (k_{op} = 16 s^{-1}), than in DDD and DDD^m (k_{op} = 8 and 7 s^{-1}, respectively), but the open lifetime expressed by $1/k_{cl}$ was ~2 times less in DDD^{hm} and ~1.5 in DDD^m than in DDD (Table 3.2). For the C^3:G^{10} base pair, the opening rate was greater than for the other C:G base pairs. However, it was comparable for the DDD and DDD^{hm} duplexes (42 and 45 s^{-1}, respectively). This was confirmed by the opening equilibrium constant values (Table 3.2). The greatest values for the equilibrium constants for base pair opening as well as the opening rates were observed for the T^8 N3H imino proton, which is located in the 5' neighbor base pair from the modification sites. This applied to all three duplexes. The dynamics of this base pair were different for each duplex. As compared to base pair A^5:T^8 in the DDD, in the DDD^m and DDD^{hm} duplexes this base pair opened ~1.5 and ~3 times faster, respectively. For the terminal C^1:G^{12} and penultimate G^2:C^{11} base pairs the exchange rates could not be measured because of the exchange with the solvent protons.

The atomic resolution crystal structure of modified DDD

Crystals belonged to the primitive orthorhombic space group $P2_12_12_1$ with unit cell parameters a=25.3 Å, b=40.2 Å, c=65.5 Å. The crystal structure of DDD^{hm} was refined anisotropically to 1.02 Å resolution, and is shown in Figure 3.10. The crystallographic asymmetric unit consists of two chemically equivalent self-complementary strands of the DDD^{hm} antiparallel duplex numbered (C^1 to G^{12} and C^{13} to G^{24}), one magnesium ion and 2 molecules of spermine and 221 water molecules. The electron density maps obtained were of high quality and allowed building

of a detailed model as well as modeling alternative conformations. The final model exhibited very good crystallographic statistics and was refined to a final R_{work} of 15.7 and R_{free} of 17.8 (Table 2.2).

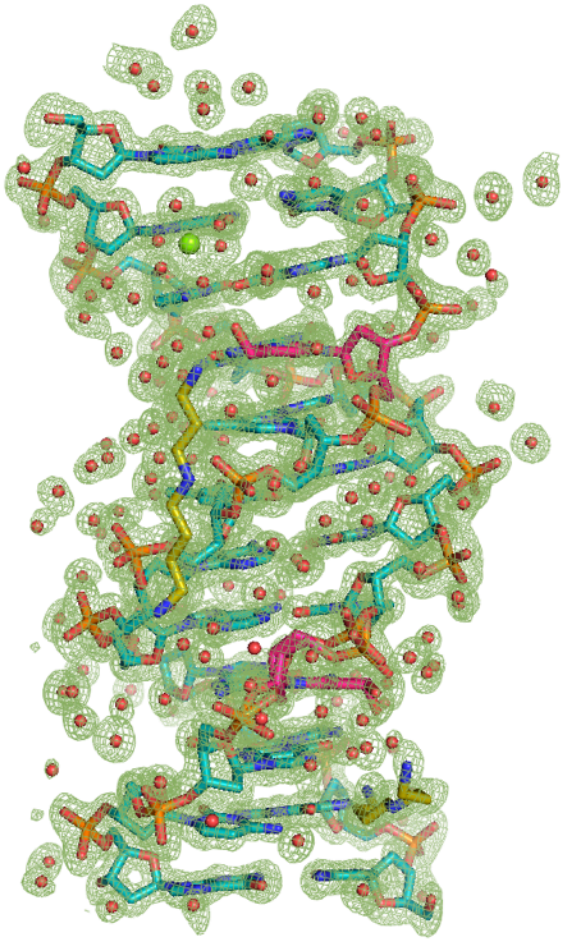


Figure 3.10. Fourier (2Fo-Fc) sum electron density contoured at the 1.0σ level (green meshwork surrounding the DDD^{hm} duplex. 5hmC modified base shown in pink.

The DDD^{hm} structure is consistent with the other DDD structures. It overlays well with the canonical structure of DDD (PDB code: 1BNA) with rmsd of 0.7 Å (least squares (LSQ) superposition using all atoms fit in COOT), and with the high resolution structure of unmodified DDD (PDB code: 355D) with rmsd of 0.7 Å. The base-stacking

pattern in the modified DDD^{hm} is very similar to the one observed in unmodified B-type DNA and helical parameters obtained by program Curves+ are comparable (Figures 3.13 and 3.14). As anticipated presence of the hydroxymethyl side chain at the cytidine nucleobase, did not significantly alter base packing between Z^{21}:G^{4}, nor did it disrupt Watson-Crick base pairing or changed the conformation of these residues (Figure 3.11 and 3.12). It also did not alter Mg^{2+} binding in the crystal structure. The single octahedrally hydrated magnesium cation is located in the major groove and it found to interacts via waters with G^{2} (chain A), *O6* of G^{22} (chain B) and with *O2P* ($P^{7'}$) and *O1P* ($P^{6'}$) of an adjacent symmetry related molecule, as it was described previously by Drew-Dickerson *et al*. (*14*), and Tereshko *et al*. (*77*). It is noteworthy that an additional hydrogen bond interaction is formed between the hydroxyl group of modified cytosine Z^{21} (chain B), and the axially coordinated water (HOH 12) (2.7 Å distance), whereby it interacts further with N^{7} of the G^{22} residue (2.8 Å distance). An additional H-bond interaction is observed between hydroxyl group on Z^{21} and *O6* of G^{22} via water molecule HOH 11 (3.0 Å distance from Z^{21} to HOH11, and 2.7 Å distance from HOH11 to *O6* of G^{22}) (Figure 3.11).

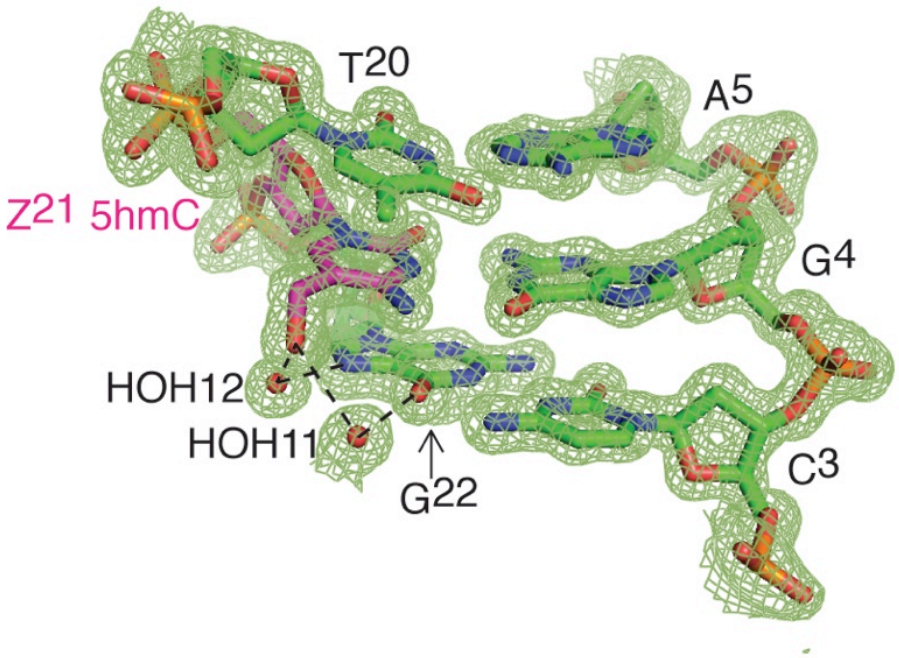


Figure 3.11. Sum electron density contoured at the 1.0σ level (green meshwork) surrounding the DDD^{hm} duplex at the modification site.

In the second modified base Z^9 (chain A) the hydroxyl moiety is observed to be in two conformations (a major conformation refined with occupancy of 0.8 and a minor with occupancy of 0.2). In the major conformation the 5-hydroxyl moiety forms a H-bond with the neighboring N^1 of spermine I residue and the *O6* of G^{10} via HOH 200 water molecule. In the minor conformation, the 5-hydroxyl moiety is turned towards backbone phosphate moiety and interacts with the neighboring waters (HOH 155). In addition, dual conformations (with occupancies 0.8/0.2) were also modeled for residues G^{10} of chain A and the complementary counterparts in chain B (the residues A^{18} and T^{19}), such that significant differences in the phosphate backbone positions are observed. Analysis of the waters forming the "spine of hydration" shows that waters involved in the minor groove hydration are conserved

and superimposable between the structures. Analysis of the alternative conformations of DDD^{hm} did not reveal significant changes in the helical parameters (Figures 3.13 and 3.14).

The differences between the two conformations primarily entail changes in torsion angles and phosphate backbone positions and might arise as a result of the association with the spermine molecules (which seem to slightly alter crystal-packing interactions. Spermine I, that is fully visible in the electron density map (all 14 non-hydrogen atoms) is located in at the end of the duplex, and extends between the bases of G^{10} and G^{14} towards the phosphate residue of the Z^{9}. The spermine was modeled in two conformations, which mainly differ in the position of the terminal ammonium groups (with occupancy 0.5 and 0.3). The interactions between spermine and DNA appear to be limited to direct interactions via N^{1}, N^{11}, and N^{14}.

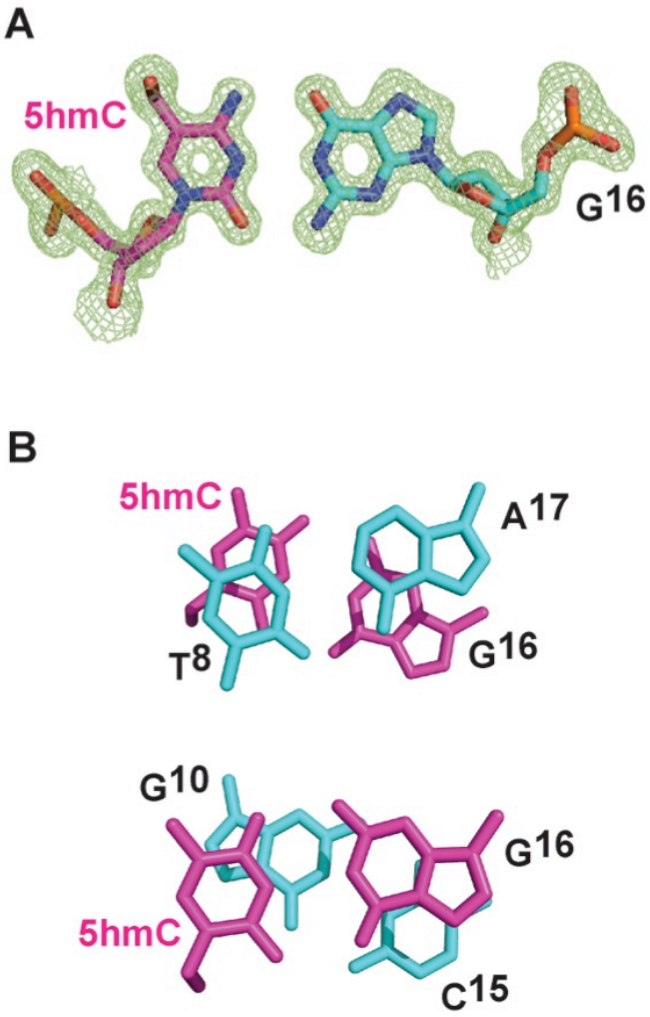


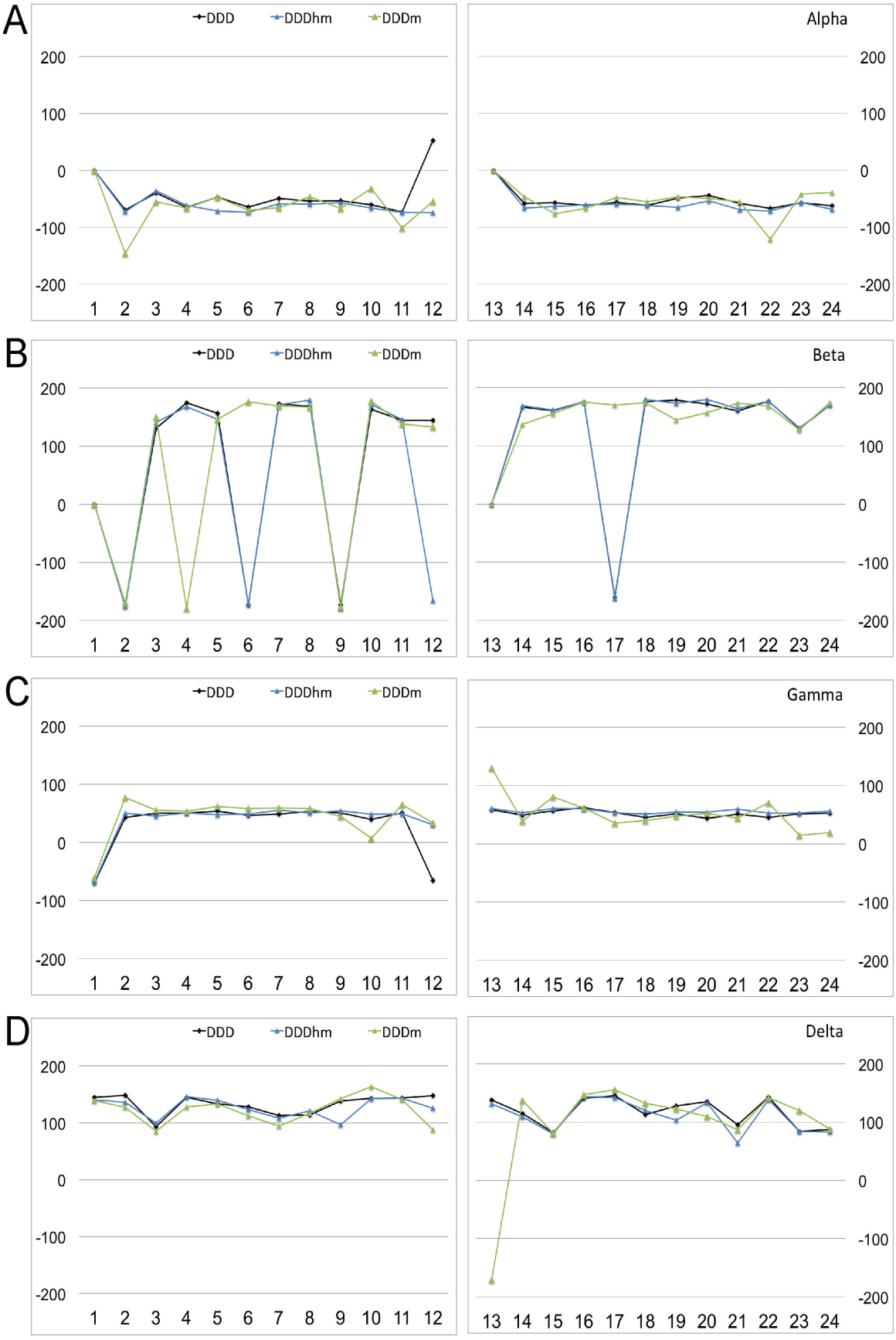
Figure 3.12. Sum electron density contoured at the 1.0σ level (green meshwork) around the 5hmC:dG modified base pair showing Watson-Crick interactions (A), and base stacking of 5hmC·dG modified base pair with T^{8}:A^{17} and G^{10}:C^{15} (B).

Spermine II is partially disordered in the density maps and only 11 residues were modeled. This molecule is positioned at the center of the duplex and is wedged between the sugar phosphate backbones of strand B (residues T^{19}, T^{20}) and the strands of the symmetry mate (residues G^{12}, chain A', and A^{18}, chain B'). The N^{5} amino group of spermine II interacts witch *O1P* (3.00 Å) and *O2P* (3.32 Å), whereas N^{10} interacts with *O1P* of T^{19} (2.91 Å) and likely triggers the dual conformation of the phosphate

moiety of the neighboring nucleotide. The phosphate moiety of T^{19} adopts two conformations such that the positions of phosphate groups differ by ~0.8 Å. We suspect that there might be a third spermine molecule, which is positioned next to the phosphates of the residues Z^{21} and G^{22}. However, it was not modeled due to the disordered density features, and instead was modeled with water molecules.

Helicoidal analysis

An analysis of the helical parameters of the crystal structure was performed in program Curves+ (*100*). The results of the analysis of torsion angles are shown in Figure 3.13. There were no significant changes in the alpha, gamma and delta torsion angles. For torsion angle beta, the DDD^{hm} and unmodified DDD are similar (*77*), however, peaks at the G^2, A^6 and A^{17} reflect changes induced by the interacting Mg^{2+} ion that is observed in both structures. The effect of ion binding is also reflected in the changes in the epsilon torsion angle. Significant changes in the zeta torsion angle between DDD^{hm} and the unmodified DDD for residue G^4 are due to the presence of the modification on the opposite base Z^9. The analysis of interbase parameters shows no substantial changes in helical rise, roll and twist, as shown in Figure 3.14.


A
DDD
DDDhm
DDDm
Alpha
B
Beta
C
Gamma
D
Delta
200
100
0
-100
-200
1 2 3 4 5 6 7 8 9 10 11 12
13 14 15 16 17 18 19 20 21 22 23 24

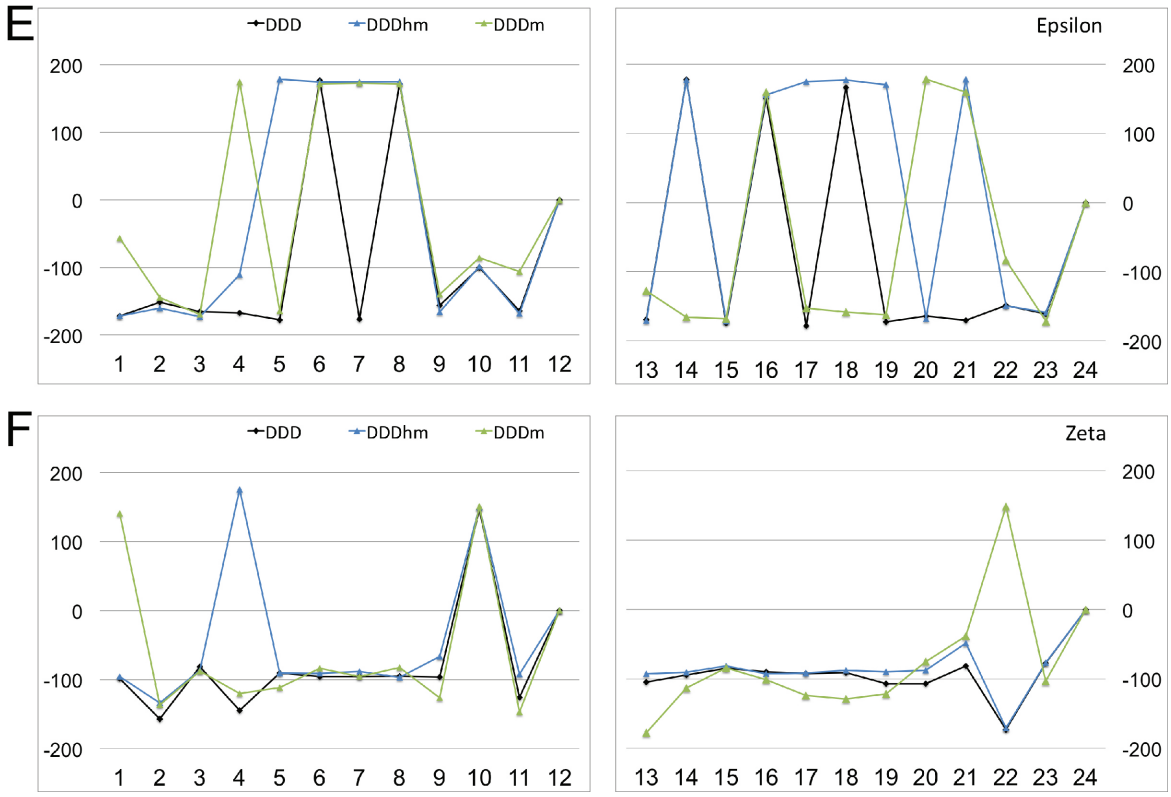


Figure 3.13. Comparison of backbone torsion angles (a) alpha, (b) beta, (c) gamma, (d) delta, (e) epsilon, and (f) zeta in the structure of the DDD^{hm} in blue, unmodified DDD (PDB entry 436D (*124*)) in black, DDD^{m} (PDB entry 265D) in green.

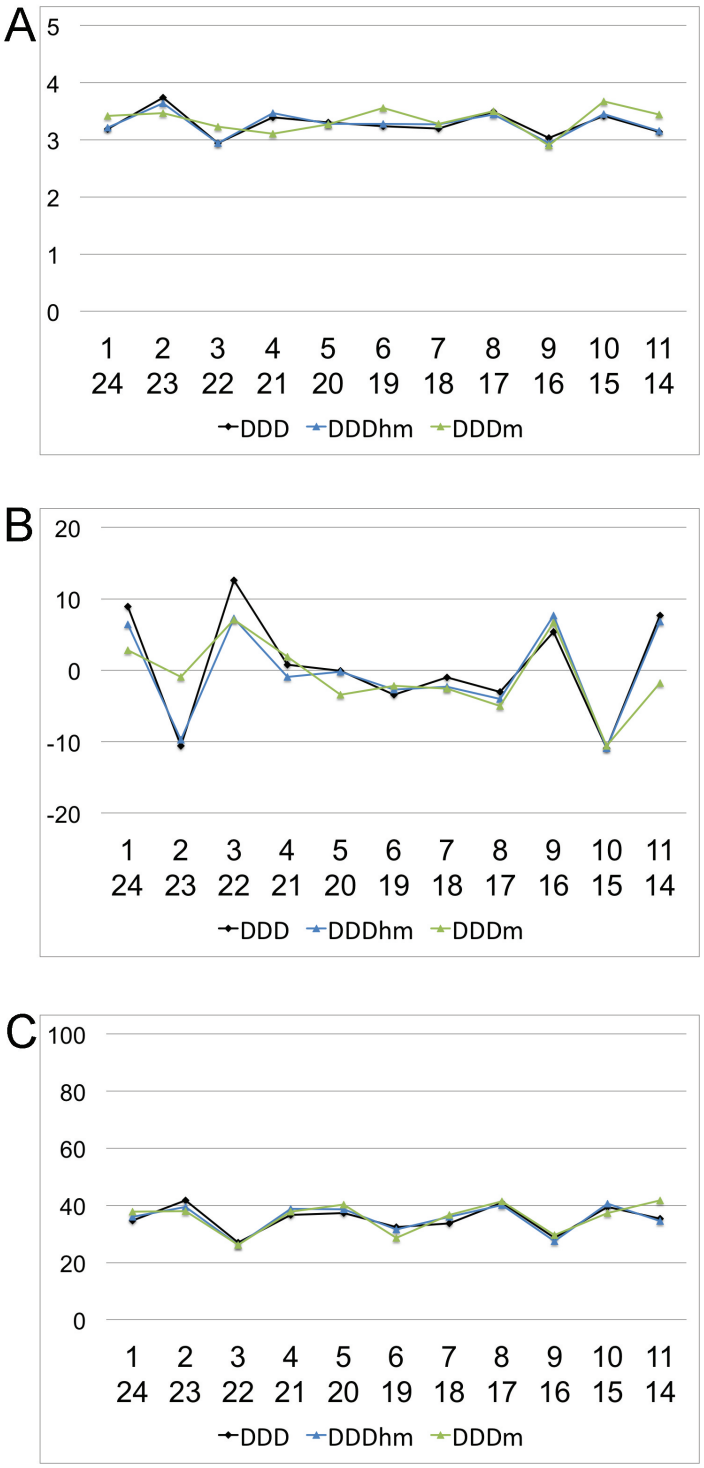


Figure 3.14. Interbase pair parameters: (a) helical rise, (b) roll, and (c) twist for the structure of DDDhm in blue, unmodified DDD (PDB entry 436D (124)) in black, DDDm (PDB entry 265D) in green.

Discussion

It has been shown that active demethylation involves oxidation by TET enzymes, which oxidize 5mC to 5hmC and further to 5fC and 5caC. These oxidation products provide substrates for BER, that would reinsert cytosines into the DNA (*104, 112, 125, 126*).

Thermodynamic Analysis of DNA with 5mC or 5hmC

DSC and UV spectroscopy experiments revealed, that with respect to the free energy of duplex formation, the presence of either 5mC or 5hmC in the DDD is similarly favorable for the DDD sequence; ΔG = -14.5 kcal/mol and -15.7 kcal/mol for DDD^{m} and DDD^{hm}, respectively; compared to ΔG = -15.7 kcal/mol in DDD (Table 3.1). All three duplexes possess similar T_M values, even at higher salt conditions (Figure 3.2 and Table 3.1). These results are in agreement with previously reported UV spectroscopic data (*116*).

Imino Proton Exchange and Base Pair Opening Kinetics

NMR spectroscopy as a function of temperature (Figure 3.6) confirms the thermodynamic results (Table 3.1). All of the imino resonances are sharp and stable at 45 °C. The G^4 N1H imino proton of G^4:Z^9, which represents the modified base pair, exhibits the sharpest resonance in the spectrum. This reflects a lower exchange rate with water as shown in the kinetic analysis. For the G^4 N1H imino proton, the exchange rate has only a weak dependence on the concentration of ammonia base (Figure 3.9).

This is reflected in the corresponding small equilibrium constant for opening of the G^4:C^9 base pair (K_{op} = 1.2 × 10^7 for DDD^{hm} and DDD, 0.75 × 10^7 for DDD^m), and reflects a short open lifetime (Table 3.2). From the 3'- site, flanking C^3:G^{22} base pair interacts via two water molecules to the hydroxymethyl group of the modified 5hmC base (Figure 3.11). This could suggest stabilization effects and lower dynamic properties. However the open lifetime expressed by $1/k_{cl}$ stays at similar levels for both base pairs (k_{cl} = 17 × 10^{-7} s^{-1} for C^3:G^{10} in DDD^{hm}, and 13 × 10^{-7} s^{-1} for G^4:Z^9 in DDD^{hm}), even though the overall opening rates for C^3:G^{10} in DDD^{hm} are somewhat higher than for the modified G^4:C^9 (k_{op} = 42 s^{-1} for C^3:G^{10} in DDD^{hm} and k_{op} = 16 s^{-1} for G^4:Z^9 in DDD^{hm}) (Table 3.2). Finally, the rates for G^4:C^9 base pair opening have similar values in DDD^{hm} and DDD, which is also reflected by comparable opening equilibrium constants. A comparison of aK_{op} for the A^5:T^8 base pair showed an increased opening equilibrium for DDD^{hm} (aK_{op} = 73 × 10^7), whereas the other imino proton resonances showed only a small or no increase. Compared with the unmodified base pair A^5:T^8 in DDD^{hm} spends twice as much time in the open state toward the solution. A key finding is that the apparent closing rate for the A^5:T^8 base pair in all three duplexes is small, while it opens 1.5 and 3 times faster in the DDD^m and DDD^{hm}, respectively, as compared with the DDD control duplex (Table 3.2). The higher exchange and base pair opening rates for thymine relative to guanine base pairs was expected (*85*). However, the 3-fold increase in opening rate and double in equilibrium constant for the 5' neighbor A^5:T^8 base pair in DDD^{hm} is significant. From the other site there is an interaction of the hydroxyl group with 3'-flanking G^{22}, which might explain the stabilizing effect of 5hmC on the overall DNA structure.

Effect of 5hmC on DNA Structure

Renciuk *et al*. (*116*) showed that the presence either 5mC or 5hmC does not influence the overall B-DNA structure of the DDD. Our solution NMR data corroborates the crystallographic data (*116*), as can be seen from a comparison of NOESY spectra of the modified and unmodified duplexes. For example, the sequential pattern of imino-imino and imino-amino NOEs is conserved in the presence of the 5hmC:G base pair, confirming that the base pairing and base stacking is maintained at the lesion sites (Figure 3.7). In addition, the pattern of sequential base aromatic→deoxyribose anomeric NOEs is also conserved (Figure 3.8), with no significant changes in chemical shifts, which indicates the presence of a normal undisrupted B-type helix. Our crystal structure confirms that the 5hmC:G base pair imparts a minimal effect on the conformation of the DDD duplex (Figures 3.10 and 3.11) and does not significantly impact Watson-Crick base pairing and base stacking (Figure 3.12) as compared to the unmodified C:G base pair. Our crystallographic data indicate that the hydroxymethyl moieties on the modified cytosines are in the major groove, oriented toward the 3' end of each strand. The hydroxyl groups interact with *O6* and N7 of the 3'-flanking G^{22} via two water molecules (Figure 3.11). One hydrogen bonding interaction was observed between the hydroxyl group at the modified cytosine Z^{21} and axially coordinated water (HOH 12) within 2.7 Å, with a further interaction to G^{22} N^7 (2.8 Å). A second hydrogen bonding interaction was observed between the Z^{21} hydroxyl group and G^{22} *O6* via water HOH 11 (3.0 Å distance from Z^{21} to HOH11, and 2.7 Å from HOH11 to G^{22} *O6*). Thus, the presence of these two water molecules might explain the low base pair opening rates for modified G^4:Z^{21} base pair.

Structure-Activity Relationships

It has been proposed that active demethylation could involve oxidation of 5mC to 5hmC, and then to 5fC and 5caC, facilitated by TET enzymes, where 5caC in the last step undergoes decarboxylation to an unmodified C (*104, 125*). Maiti and Drohat showed that TDG might be involved in demethylation when 5fC and 5caC, but not 5hmC, are removed via BER (*115*). While thymine DNA glycosylase (TDG) removes 5fC and 5caC, it is unable to remove 5hmC (*115*). They proposed that TDG activity is modulated by the electron withdrawing properties of the substituent at the 5-position of the cytosine ring, such that strongly electron withdrawing C5-substituents stabilize the developing negative charge in the base excision transition state (*127*). They concluded that the negligible electronic effect for the hydroxymethyl group ($\sigma_m = 0$) might explain why 5hmC is not excised by TDG. Similarly, 5mC ($\sigma_m = -0.07$) and cytosine ($\sigma_m = 0$) are not removed via BER (*115, 127*). Studies using NMR imino proton exchange measurements by Stivers and Song (*54*) showed that uracil DNA glycosylase (UDG) can substantially increase the equilibrium constant for opening of A:T base pairs relative to free DNA, which could provide a dynamic mode of modification identification by DNA glycosylases that require the lesion to be extrahelical. Our results show that the equilibrium constant for opening of 5hmC is low. Thus, if the Stivers model is correct, 5hmC should not be a good substrate for BER.

A role of 5hmC as an epigenetic modifier and transcriptional regulator has been proposed, with altered levels in early embryonic development, embryonic stem cell

differentiation, and tumors (*109, 113*). It was reported that the level of methylation depends on cell or tissue type and developmental stage (*128*). After fertilization, during embryonic development, many of the methylation markers are erased. This allows embryonic stem cells to differentiate into specialized cells. It is now known that 5hmC and TET deoxygenases play a crucial role in epigenetic reprogramming and regulation of tissue-specific gene expression (*129-132*). However, it was proposed that 5hmC may only be a reaction intermediate in the process of active demethylation (*39, 40*).

Acknowledgements

This work was supported by NIH grant R01 CA-55678 (M.P.S.). Funding for the NMR instruments was supplied by NIH grants S10 RR-05805, S10 RR-025677 and NSF Grant DBI 0922862, the latter funded by the American Recovery and Reinvestment Act of 2009 (Public Law 111-5). Vanderbilt University assisted with the purchase of in-house crystallographic and NMR instrumentation.

CHAPTER 4

CHARACTERIZATION OF 5-FORMYCYTOSINE IN DNA

Introduction

5-formylcytosine (5fC) is an oxidation derivative of 5-methylcytosine formed by TET enzymes in the process of active demethylation (*39, 133*). The TET family of proteins has the capacity to convert 5mC to 5fC *in vitro*. The presence of 5fC was found in the genomic DNA of mouse embryonic stem (ES) cells and organs (*39*). Genome-wide mapping in mouse ES cells showed high levels of 5fC in CpG islands of promoters and exons, which corresponded to transcriptionally active genes. Based on these studies it was proposed that 5fC plays a key role in reprogramming with specific genomic regions, that are controlled by the base excision mechanism and TDG. Excision of 5fC in ES cells is crucial for correct establishment of CpG methylation patterns during differentiation and for appropriate patterns of gene expression during development (*134*). Recently, the functional effect on transcription of 5fC in the genome has been described, for mammalian and yeast RNA polymerase II. Pol II polymerization rates and specificity constants for GTP incorporation against 5fC were reduced significantly, while there were no changes in GTP incorporation opposite C, 5mC or 5hmC templates. Thus, Pol II can read and distinguish subtle changes at the 5 position of modified cytosines and process them in different ways (*135*).

In the process of active demethylation of 5mC, 5fC is removed by thymine DNA glycosylase (TDG) via the base excision repair mechanism, resulting in the formation of unmodified cytosine (*78*). It was shown that TDG can rapidly excise 5fC from DNA in vitro, and this activity was subsequently found in mammalian cells (*136-139*). In order to understand the mechanisms of active demethylation through 5fC formation, chemical reactivity studies of the newly formed nucleobases were reported (*140*). In this study, the sensitivity of these modifications to oxidation and deamination was measured, along with C-C bond cleaving reactivity, either in the absence and presence of thiols as biologically relevant (organo) catalysts. These studies revealed that 5hmC is rapidly oxidized to 5fC, in comparison with 5mC, and the deamination reaction occurred only at a minimal level. Moreover 5fC can undergo thiol-mediated and acid-catalyzed C-C bond cleavage reactions to form unmodified dC with the release of formic acid. Thus if the DNA demethylation occurs via 5hmC oxidation to 5fC, then deformylation could take place via alternative active demethylation mechanisms (*140*). However, recent biochemical and biological studies have established that the pathway for active DNA demethylation involves BER and TDG enzyme. The latest studies of Maiti and Drohat report the investigation of the TDG mechanism excising 5fC from DNA, and the chemical properties indicating the catalytic requirements for the excision (*78, 141*). They showed that TDG can rapidly remove 5fC, with higher activity than T from G/T mispairs (*78*). TDG activity is greater for cytosine analogs with an electron-withdrawing C5-substituent ($\sigma_m>0$, σ_m - electronic substituent constant) that can stabilize negative charge developing on the excised base in the chemical transition state. As the formyl group is strongly electron-withdrawing ($\sigma_m>0.35$), it implies that TDG

could remove it. The next aspect considered by the authors was the tautomerism of 5fC under physiological conditions. It was proposed that 5fC favors an imino tautomer and adopts a Wobble-like structure, similar to the structure of G/U or G/T mispairs (Figure 4.1) (*142, 143*).

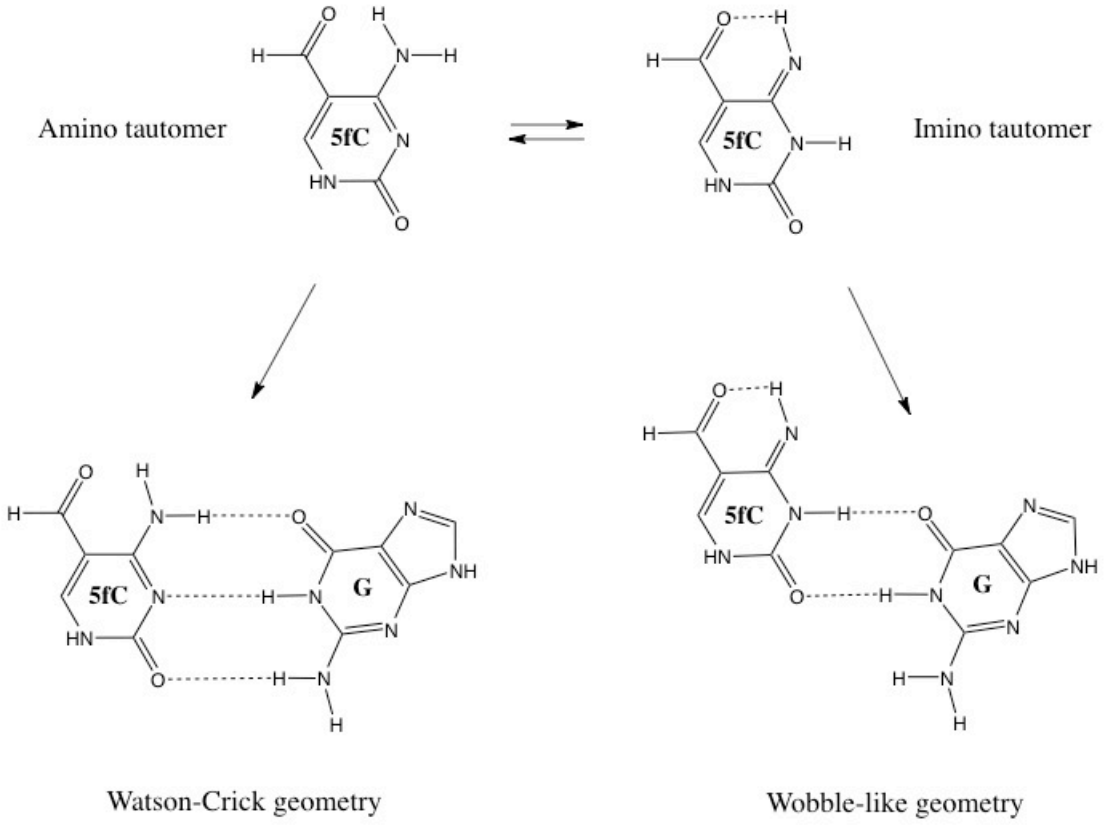


Figure 4.1. Tautomerism of 5fC modified base and base pairing schemes.

In the studies with TDG it was proposed that this unique structural geometry of a mismatched base pair is recognized by the enzyme, which results in excision of the mispaired base by TDG. The hypothesis was tested by Maiti and Drohat recently (*141*), where they calculated the relative stability of the amino and imino tautomers for 5fC in single base form. In their findings, 5fC as an anionic tautomer

is much more stable than its imino counterpart in the gas phase and in the water phase (*141*), suggesting that 5fC when paired with G forms a Watson-Crick base pair with normal geometry. In addition, NMR studies in DMSO showed the amino tautomer of 5fC was the predominant form (*144*). In the present work we will address their hypothesis and investigate duplex DNA containing 5fC modified base paired with G (DDD^f) by X-ray crystallography.

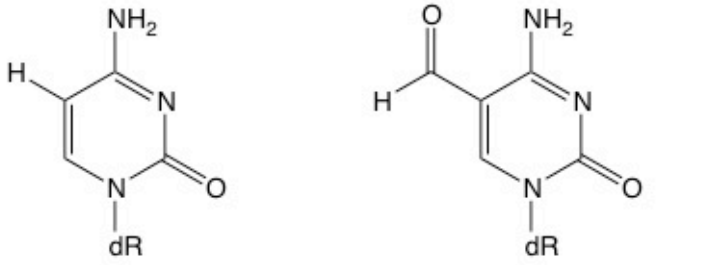
DNA glycosylases use an extrahelical base recognition mechanism, which relies on kinetically enhanced base pair opening rates for destabilized base pairs (*82*). High-resolution proton exchange NMR spectroscopy was performed to characterize the kinetics of base-pair opening in nucleic acid duplexes. The rates of exchange of imino protons with solvent protons were measured by magnetization transfer from water for each DNA duplex as a function of the concentration of exchange catalyst. In work presented here the dynamics of base pair opening in DDD^f in comparison with canonical DDD is discussed.

Results

The sequence studied in this work is well-characterized Dickerson-Drew Dodecamer, which has C/G rich termini and A/T rich core. This sequence was selected because it is self-complementary, gives well-resolved peaks in NMR spectra, and crystallizes well. The 5fC modification was incorporated into the DDD oligodeoxynucleotide at the nineth position; consequently, after annealing the duplex

were contains two modified bases. The DNA duplexes used in these studies are presented in Figure 4.2.

(a)



Cytosine; C 5-formyl-cytosine; 5fC

(b)

DDD	5′-C^2 G^2 C^3 G^4 A^5 A^6 T^7 T^8 C^9 $G^{10}C^{11}G^{12}$ -3′ 3′-$G^{12}C^{11}G^{10}C^9$ T^8 T^7 A^6 A^5 G^4 C^4 G^2 C^1 -5′
DDD^f (NMR)	5′-C^2 G^2 C^3 G^4 A^5 A^6 T^7 T^8 $\underline{X^9}$ $G^{10}C^{11}G^{12}$ -3′ 3′-$G^{12}C^{11}G^{10}\underline{X^9}$ T^8 T^7 A^6 A^5 G^4 C^4 G^2 C^1 -5′
DDD^f (Xray)	5′-C^2 G^2 C^3 G^4 A^5 A^6 T^7 T^8 $\underline{X^9}$ $G^{10}C^{11}G^{12}$ -3′ 3′-$G^{24}C^{23}G^{22}\underline{X^{21}}T^{20}T^{19}A^{18}A^{17}G^{16}C^{15}G^{14}C^{13}$ -5′

Figure 4.2. (a) Structure of dC, and 5fC. (b) Sequences and numbering of the nucleotides for unmodified DDD, and DDD^f (NMR) and DDD^f (X-ray) duplexes. In solution, the two strands of the DDD exhibit pseudo-dyad symmetry and the NMR resonances of symmetry-related nucleotides in the two strands are not individually observed. In the crystal, electron density for symmetry-related nucleotides in the two strands is observed and the nucleotides are numbered individually.

Thermal denaturation studies

The melting temperature of DDD and DDD^{f} was examined by temperature-dependent UV spectroscopy, and taking the first derivative of the resulting UC melting curves. The T_M of DDD^{f} was 46 °C, and for the unmodified DDD was 48 °C, measured at the same conditions.

NMR melting studies

A series of 1D ^{1}H NMR spectra for the exchangeable protons were recorded at 5, 15, 25, 35, 45, and 55 ºC and are shown in Figure 4.3. The data show that the N^{1}-imino proton of the G^{2}·C^{11} base pair in DDD^{f} was sharp and detectable only at 5 ºC, and broadened at 15 ºC. The same imino peak in DDD started to broaden at 35 ºC. The next base pair, G^{10}·C^{3}, which is the 5' neighbor of the modification, almost disappeared in 35 ºC in DDD^{f}, while it is still sharp even at 45 ºC for the control experiment. The N^{1}-imino proton of the X^{9}·G^{4} modified base pair in DDD^{f} remained sharp even at 35 ºC, and started to broaden at 45 ºC; but the same resonance in control DDD was still sharp and detectable at a temperature as high as 55 ºC. The N^{3}-imino resonances of T^{7} and T^{8} are sharp in DDD at 55 ºC, while these peaks are gone in DDD^{m}. There was only one observed change in the chemical shift for the G^{4} imino proton in the DDD^{f} which was upfield by 0.5 ppm when compared to G^{4} in the unmodified DDD.

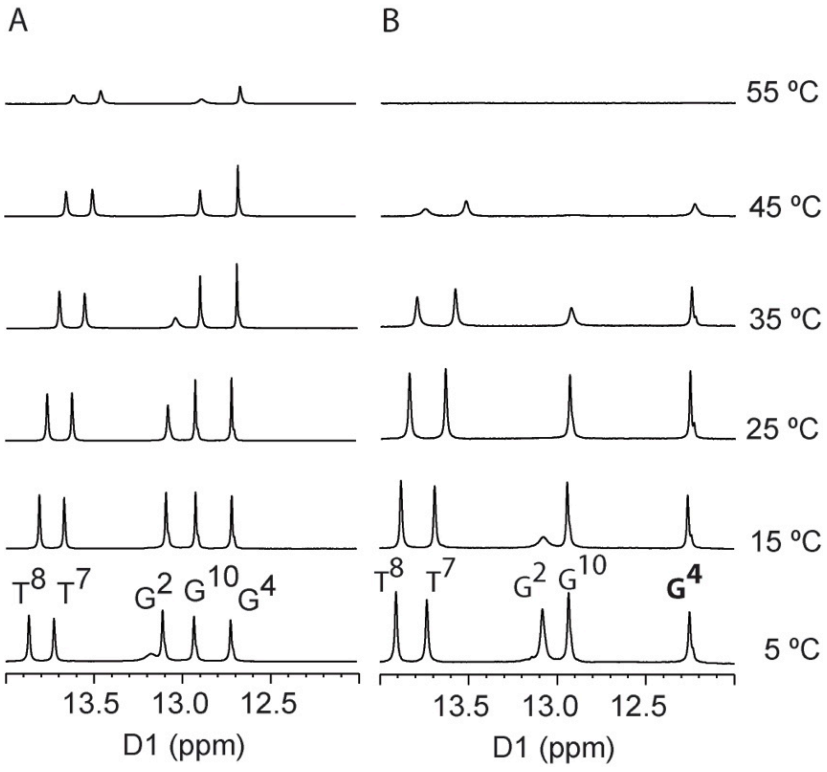


Figure 4.3. ^{1}H-NMR of imino proton resonances as a function of temperature for the unmodified duplex (A), the 5fC modified duplex (B).

Exchangeable protons.

Figure 4.4 shows the NOE connectivity of the purine N^1 and pyrimidine N^3 imino protons. The base imino protons were assigned based on their sequential connectivities in NOESY spectra, and these assignments were supported by their NOE cross-peaks to their Watson-Crick base-paired amino protons. The sequential connectivities were obtained from base pairs G^2:C^{11} → G^{10}:C^3 → G^4:C^9 → T^8:A^5 → T^7:A^6. For the DDD and DDD^f duplexes the imino-proton resonances of the terminal base pairs C^1:G^{12} are lost by fast exchange with water. The imino resonance from G^4, which is base paired with X^9 was as intense as other imino peaks in this region. However, the G^4 imino peak was shifted upfield by 0.5 ppm, reflecting the effect of base pairing opposite X^9.

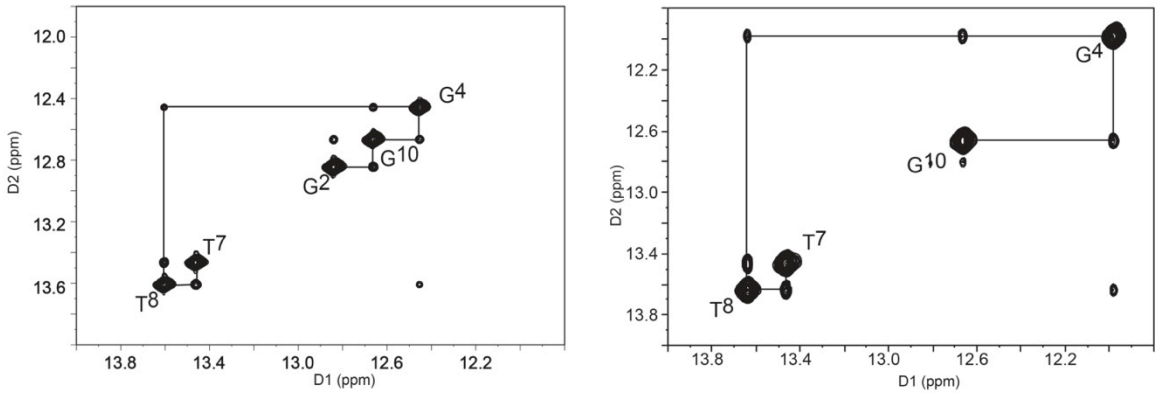


Figure 4.4. ^{1}H-^{1}H NMR NOESY spectrum showing resonances for the thymine and guanine imino protons and sequential NOE connectivity for the imino protons of the base pairs G^{2}:C^{11} to A^{6}:T^{7} for the unmodified (left), and the 5fC-modified (right) duplexes.

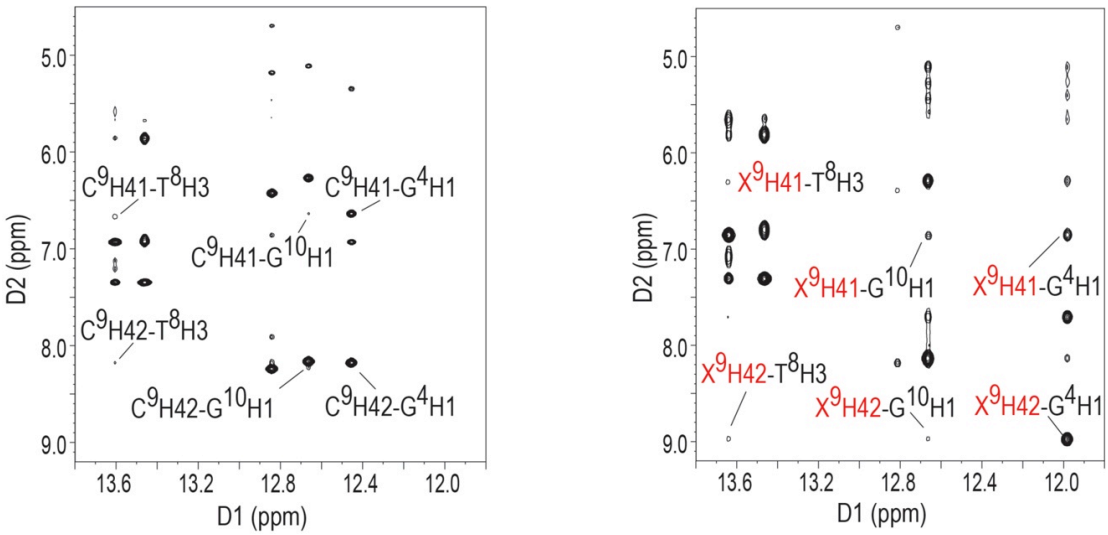


Figure 4.5. Expansion of the ^{1}H-^{1}H NOESY spectra for DDD (left), and DDD^{f} (right) showing the conservation of Watson-Crick base pairing and base stacking at the modification site.

The upfield region of the NOESY spectrum (Figure 4.5) showed the NOEs between the imino and amino protons. Cross peaks from modified base $C^{9}H41$, $C^{9}H42$ to opposite base $G^{4}H1$ were observed, as well as interactions to neighbor bases $T^{8}H3$ and $G^{10}H1$.

Nonexchangeable protons

The sequential assignment of nonexchangeable protons was accomplished using standard protocols. The unmodified duplex was used as a control for NMR assignments for DDD[f]. For modified duplexes, the anticipated pattern of sequential base aromatic → deoxyribose anomeric nuclear Overhouser enhancement (NOE) was identified from C^{1} → G^{12} (Figure 4.6).

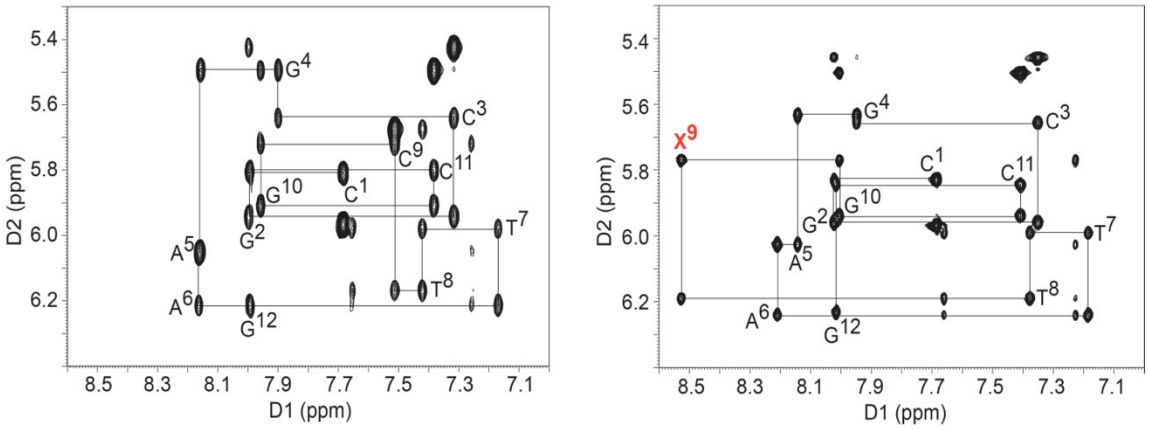


Figure 4.6. Expanded plot from the aromatic-anomeric region of the NOESY spectrum, showing sequential NOE connectivities of unmodified DDD, and modified DDD[f] duplexes.

Only one set of resonances was observed because the sequences are self-complementary. All spectra exhibited well resolved cross-peaks. Three strong NOEs accounted for H5-H6 cross-peaks of cytosine residues (C^{1}, C^{3}, and C^{11}). The H5-H6 resonance from X^{9}

is missing due to formyl substitution at 5 position in DDD^{f}. Cytosine assignments were confirmed by a DQF-COSY spectrum recorded at identical conditions (data not shown). Each base proton exhibited NOE peaks to its own and the 5'-flanking H1'-deoxyribose protons. For T^{8} and G^{10} the NOE cross-peaks intensities between the base protons and the sugar H1' of the attached deoxyribose moieties were of the same relative magnitudes as those between other bases in the sequence. The 5fC H6 resonance was observed at 8.5 ppm, shifted downfield by approximately 1 ppm with respect to that of the unmodified oligodeoxynucleotide. This was attributed to the differential electronic density for 5fC as compared to G. Proton resonances from the opposite G^{4} and A^{5} bases exhibited chemical shift changes of <0.1 ppm, compared with those bases in the unmodified DDD duplex.

Crystal structure of DDD with 5fC

Crystals belonged to the orthorhombic $P2_12_12_1$ space group with unit cell parameters a = 25.32 Å, b = 41.47 Å, c = 65.66 Å. The unmodified DDD (PDB entry 436D) was used as a search model for molecular replacement. The crystal structure of DDD^{f} was refined isotropically to a resolution of 1.74 Å. The rmsd values for bond lengths were 0.009 Å, and 1.747 deg for angles. The structure overlaid well with the canonical DDD (PDB entry: 436D (*77*)) structure with a rmsd of 0.329 Å and no structural perturbations were observed at the two modification sites (Figures 4.7). At the sites of modification, the 5fC formyl groups were in the major groove (Figure 4.8). They remained in the plane with the modified cytosine bases, thus potentiating formation of hydrogen bonds to the exocyclic amino groups at the C4 position of modified

cytosines. For both modified base pairs, the base stacking pattern in the modified DDD^{f} was similar to that observed in the unmodified DDD (Figure 4.9). The helical parameters calculated with the program Curves+ (*100*) were comparable to the unmodified DDD and are presented in the Figures 4.11 – 4.13.

Figure 4.7. Sum electron density contoured at the 1.0σ level (green meshwork) surrounding the DDD^{f} duplex at the modification site. The 5fC modified base is shown in red, and water molecules are shown in dark red.

The substitution by formyl at the C5 position of cytidine, did not disrupt Watson-Crick base pairing with G or change the conformation of these residues (Figure 4.9).

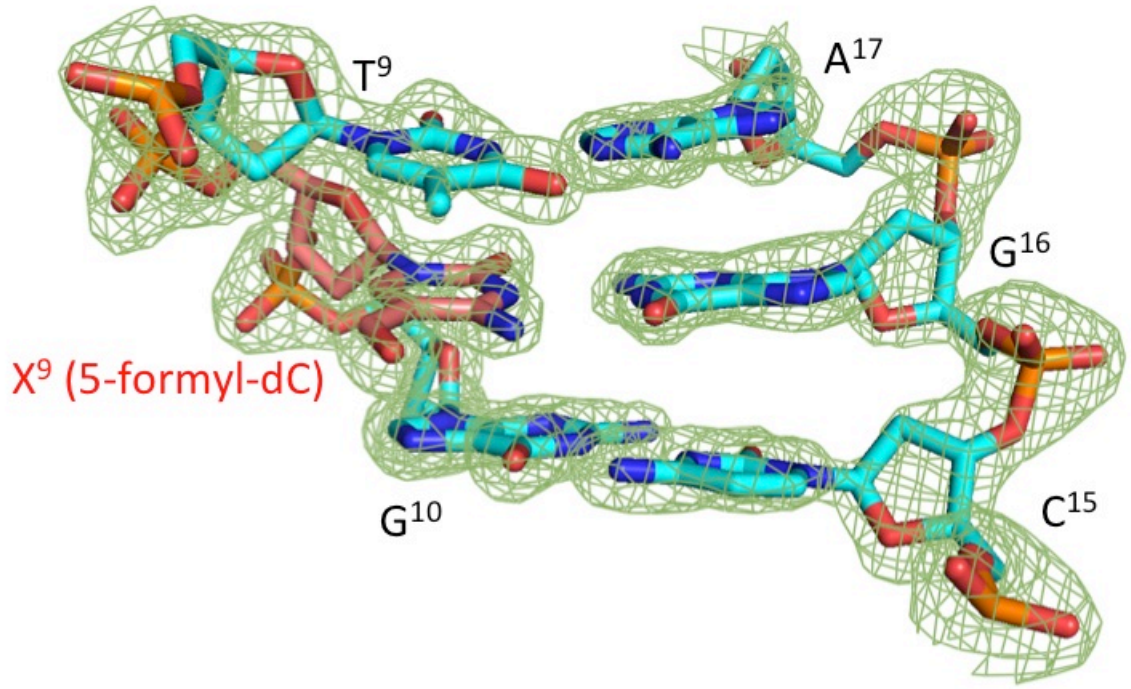


Figure 4.8. Fourier (2Fo-Fc) sum electron density contoured at the 1.0σ level (green meshwork) surrounding the DDD[f] duplex at the modification site.

Analysis of the waters forming the "spine" of hydration, shows that the waters involved in the minor groove hydration are conserved. Detailed information about data collection and refinement statistics are shown in Table 2.2.

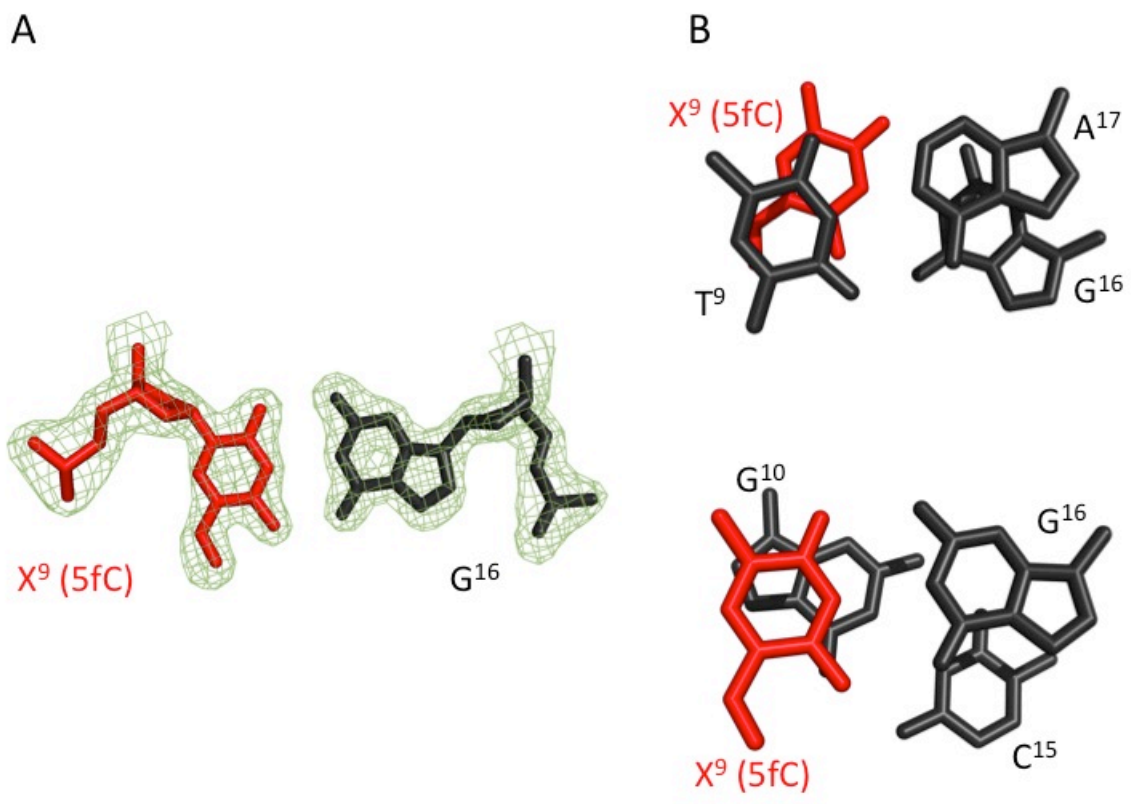


Figure 4.9. (A) Sum electron density contoured at the 1.0σ level (green meshwork) around the 5fC:dG modified base pair showing conserved Watson-Crick base pairing geometry. (B) Expanded view of the DDD[f] crystal structure, showing stacking interactions: (top) stacking of base pair T^{8}:A^{17} above base pair X^{9}:G^{16}, and (bottom) stacking of base pair X^{9}:G^{16} above base pair G^{10}:C^{15}.

Dynamics of base pair opening

To determine the dynamics of base pairs, imino proton exchange rates were measured in the presence of ammonia base catalyst, that was used because of its small size, lack of charge and high accessibility factor. Figure 4.10 shows imino proton exchange rates obtained by measuring magnetization transfer from water as a function of added ammonia base catalyst. The results of the imino proton exchange analysis are summarized in Table 4.1. For the terminal (i.e. $C^{1}\cdot G^{12}$) and penultimate

($G^2 \cdot C^{11}$) base pairs the exchange rates could not be measured because of the increased exchange with solvent protons.

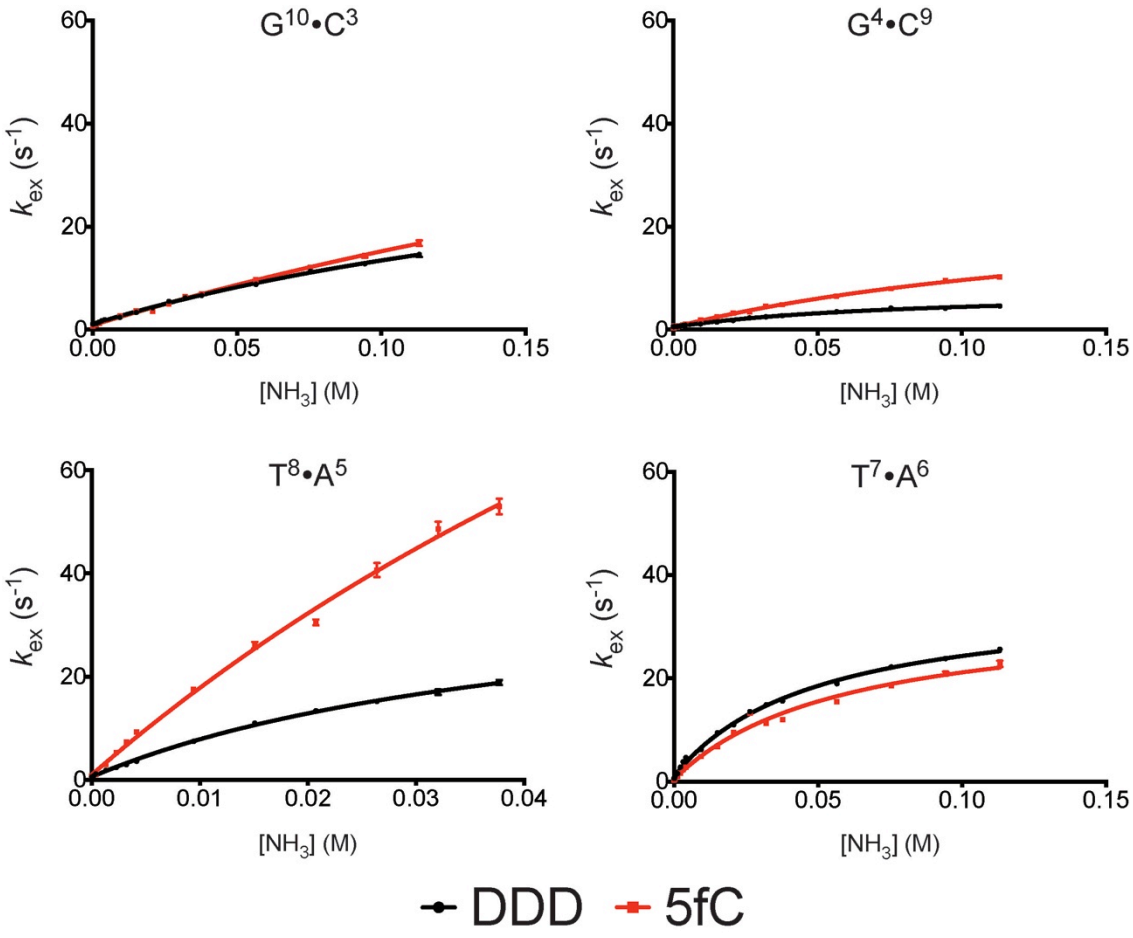


Figure 4.10. Plots showing imino proton exchange rates obtained by monitoring magnetization from water as a function of ammonia base catalyst. A. G^{10}:C^3. B. G^4:C^9. C. T^8:A^5. D. T^7:A^6 in DDD (black), and DDDf (red).

As is evident, for all of the studied imino protons in the three duplexes studied, the exchange rate reaches the EX1 regime as predicted by the eq. 9. As the concentration of ammonia increases, the exchange rates increase until a plateau is reached, which indicates a change in the rate limiting step from chemical exchange to base pair opening ($k_{ex} = k_{op}$ and exchange occurs in each opening event). For the imino proton of G^4,

the exchange rate has a weak dependence on the concentration of ammonia base. This finding reflects the small equilibrium constant for opening of the $G^4 \cdot C^9$ base pair. However, comparing G^4 from DDD^f with G^4 from DDD, indicates that the opening rate is 3 times higher (k_{op}=26 vs. k_{op}=8, for G^4 in DDD^f and DDD, respectively). The similar, low rate for base pair closing and base pair lifetime is consistent with the modified base pair spending more time exposed to solution than being embedded in a canonical base. This is also confirmed by the equilibrium constant for base pair opening, which is almost tripled for modified G^4:X^9 vs. unmodified G^4:C^9, (αK_{op} = 2.8 and 1.2, respectively).

	k_0 (s^{-1})[a]		$K_{op} \times 10^7$	
	DDD	DDD^f	DDD	DDD^f
C^3:G^{10}	1.1 ± 0.06	0.71 ± 0.03	2.9 ± 0.1	3.8 ± 0.1
G^4:C^9	0.57 ± 0.03	0.39 ± 0.04	1.2 ± 0.04	2.8 ± 0.1
A^5:T^8	0.59 ± 0.03	0.79 ± 0.04	41± 0.08	106 ± 4
A^6:T^7	0.61 ± 0.03	0.27 ± 0.03	37 ± 5	29 ±2

	k_{op} (s^{-1})		k_{cl} ($\times 10^{-7}$ s^{-1})	
	DDD	DDD^f	DDD	DDD^f
C^3:G^{10}	45 ± 3	86 ± 18	15 ± 1	21 ± 2
G^4:C^9	8 ± 0.5	26 ± 2	6.7 ± 0.2	9.3 ± 1
A^5:T^8	40 ± 2	222 ± 53	0.97 ± 0.04	2.1 ± 0.06
A^6:T^7	36 ± 1	33 ± 3	0.98 ± 0.09	1.1 ± 0.05

Table 4.1. Rate and Equilibrium Constants for DDD^f Base Pair Opening. [a]The observed exchange rate without an ammonia catalyst.

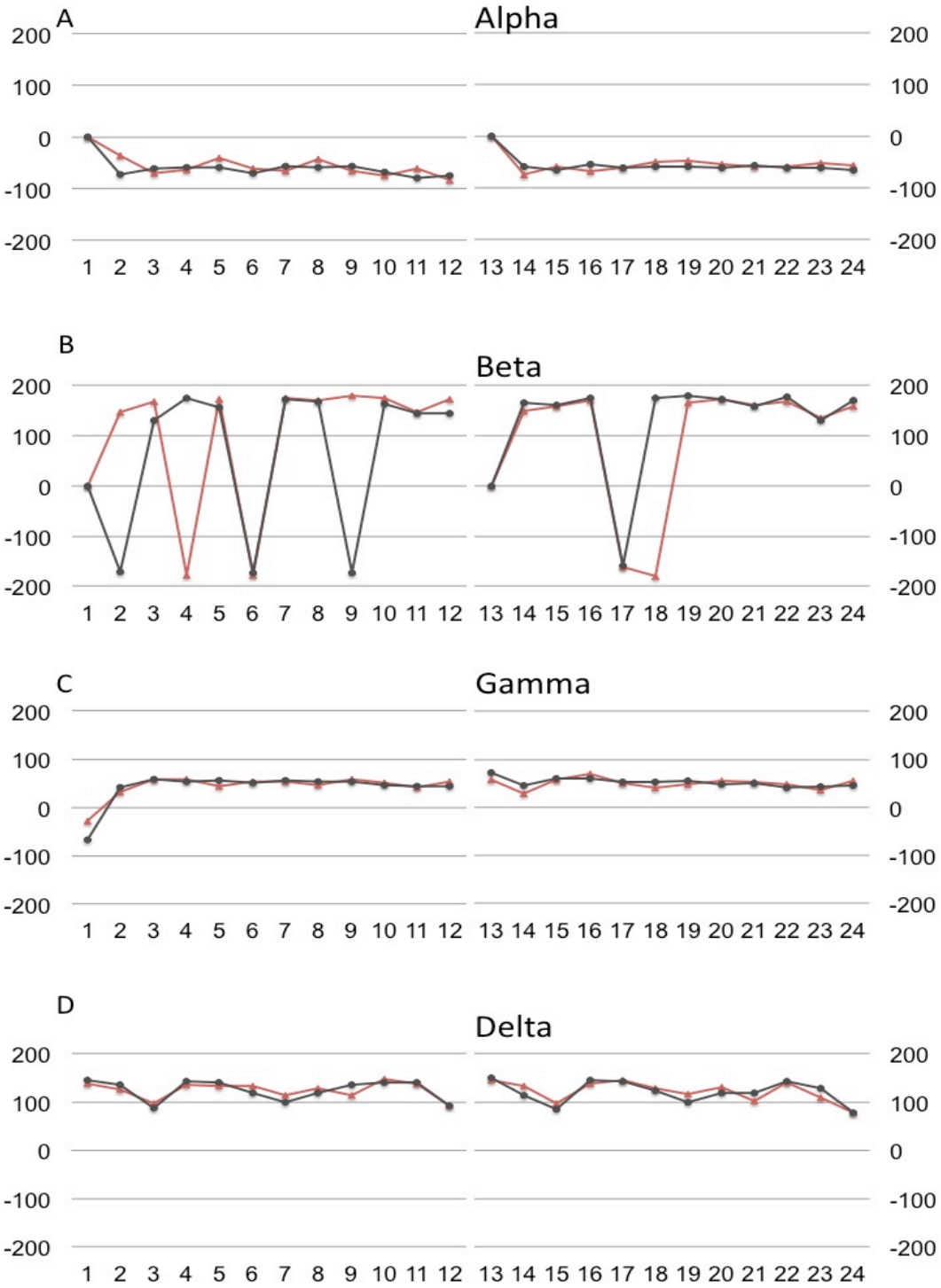
For $G^{10} \cdot C^3$ imino protons, the exchange rates are faster even at low ammonia concentrations. With an increase in the catalyst concentration, the exchange rates further

increases until the EX1 regime is reached. The equilibrium constant for base pair opening is slightly higher for the modified base pair in the DDD^f relative to the native DDD (αK_{op} = 3.8×10^7 vs. 2.9×10^7 for G^{10}:C^3). However analysis of base pair opening shows twice the rate for G^{10}:C^3 in DDD^f, relative to the same base pair in the unmodified duplex (Table 4.1). The highest value for the equilibrium constant for base pair opening is for T^8 imino proton, which is the 5'- neighbor of the modification site. This applies to both duplexes. The dynamics of this base pair is remarkably different in the DDD^f duplex. As compared to $T^8 \cdot A^5$ in the DDD, this base pair in DDD^f opens ~6-times faster, and has open lifetimes ($1/k_{cl}$) that are only 2-times shorter. A much higher rates for the equilibrium constant for base pair opening suggest that $T^8 \cdot A^5$ in DDD^f spends 2.5-fold more time in the open conformation than the corresponding base pair in the control duplex.

Helicoidal analysis

An analysis of the helical parameters of the crystal structure was performed with the program Curves+ (*100*). The analysis of the torsion angles (shown in Figure 4.11) demonstrates that there were no significant changes in the alpha, gamma, delta, epsilon and zeta angles. For the beta torsion angle, the DDD^f shows very similar conformation as the unmodified DDD (*77*). However, the peak at G^4 with the DDD^f reflects a change as a result of the presence of the modification opposite base X^9. The analysis of interbase parameters shows no substantial changes in helical rise, roll and twist, as presented in Figure 4.12. Also, shear and stretch, which are intrabase

translation parameters, are similar and comparable between bases in all studied duplexes (Figure 4.13).



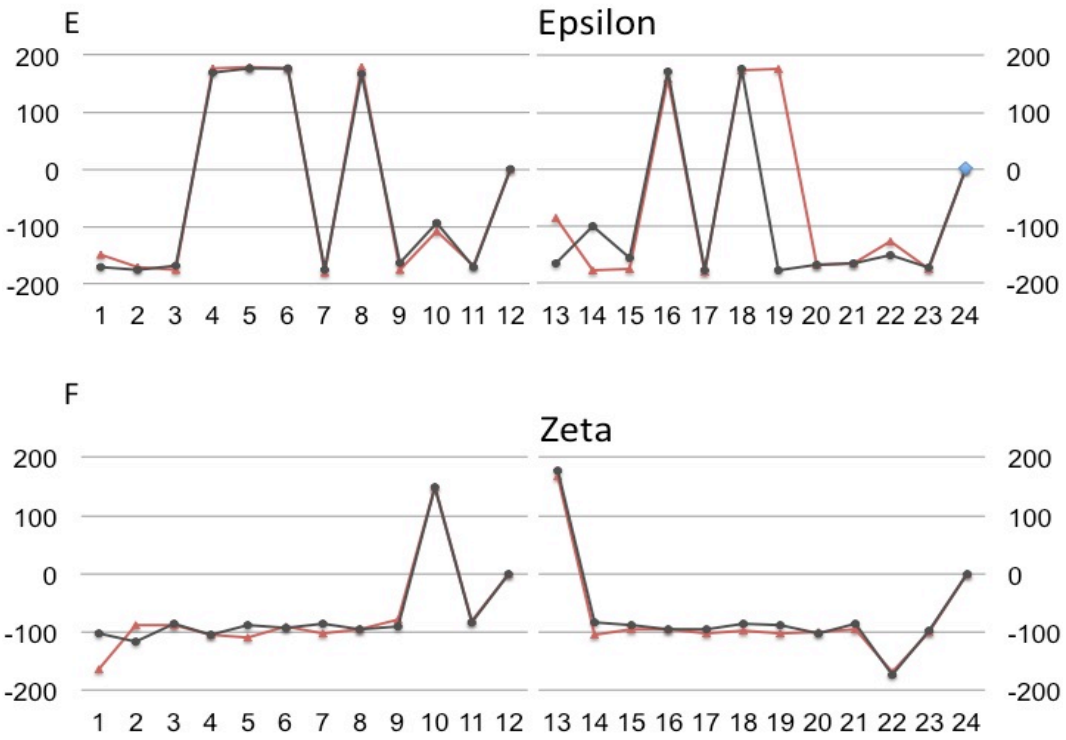


Figure 4.11. Comparison of backbone torsion angles (a) alpha, (b) beta, (c) gamma, (d) delta, (e) epsilon, and (f) zeta in the structure of DDD^{f} in red, and unmodified DDD (PDB code 436D (*77*)) in black.

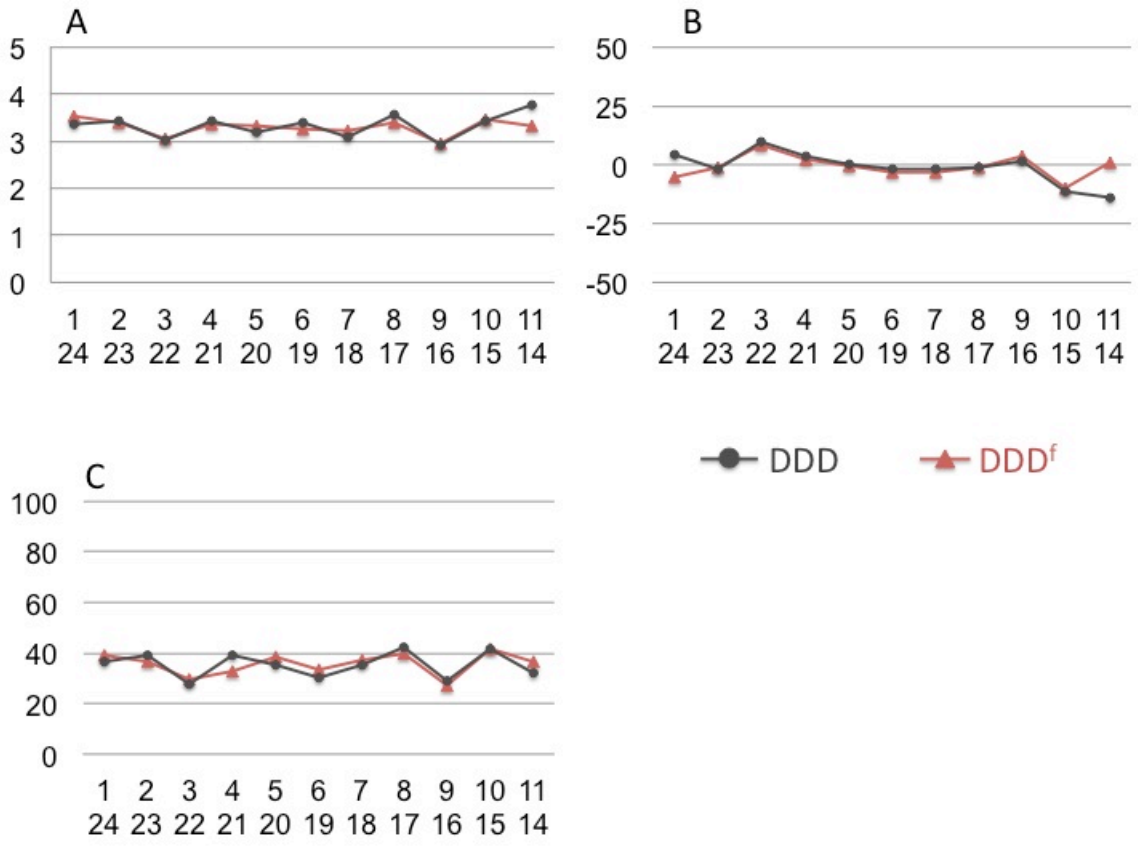


Figure 4.12. Interbase pair parameters: (a) helical rise, (b) roll, and (c) twist for the structure of DDD^f in red and unmodified DDD (PDB code 436D (*77*)) in black.

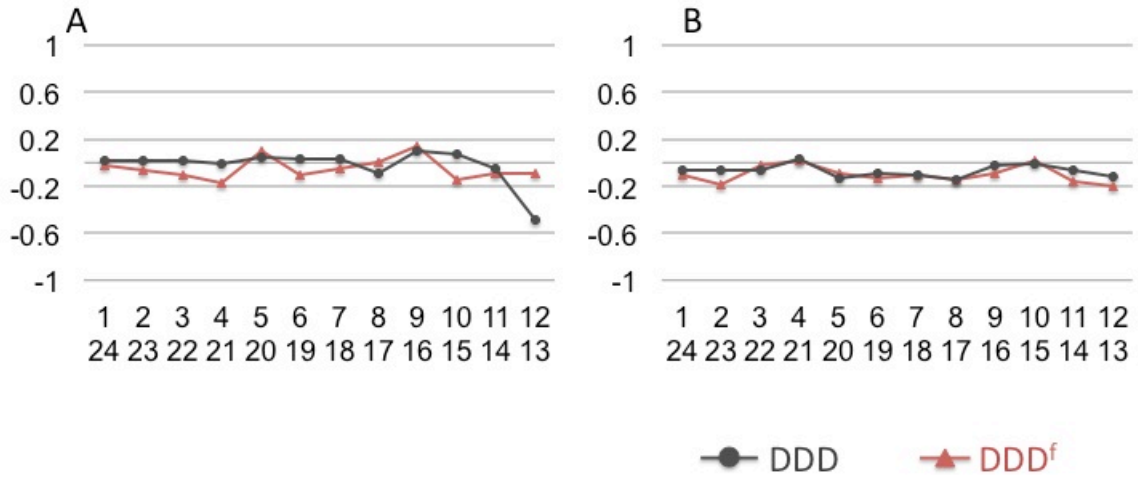


Figure 4.13. Intrabase pair parameters (a) shear, and (b) stretch for the structure of DDD^f in red, and unmodified DDD (PDB code 436D (*77*)) in black.

Discussion

The 5fC base is an active demethylation oxidation product derived from 5mC, formed post-replicatively by TET enzymes (*39*). The 5fC base has been detected in genomic DNA of mouse embryonic stem (ES) cells and organs (*39, 133*) at high levels in CpG islands of promoters and exons, which corresponded to transcriptionally active genes.

Effect of 5fC on DNA structure

TET deoxygenases act together with other DNA demethylation pathways, such as BER. In this process, 5fC is efficiently removed by thymine DNA glycosylase (TDG) resulting in the restoration of unmodified cytosine (*78, 145, 146*). Hashimoto *et al*. proposed that 5fC undergoes tautomerisation, which results in Wobble base pair formation. This unique base pair geometry was shown to be recognized by TDG in the BER pathway (*57*). In order to understand the basis by which 5fC pairs opposite guanine in DNA, the structure of 5fC:dG in DDD^{f} was determined. The structural data reveal that 5fC has a minimal effect on DNA conformation (Figure 4.8) when compared to the dC:dG base pair (*77*). The Watson-Crick base pairing geometry is conserved for the 5fC:dG base pair (Figure 4.9). There is no evidence in the electron density map for the imino tautomer of 5fC, which would result in protonation at the N3 position of the modified cytosine. This protonation would prevent the formation of a Watson-Crick base pair with the complementary guanine. In addition, the solution NMR data corroborate the crystallographic data, as can be seen from a comparison of NOESY spectra of the modified and unmodified duplexes. For example, the sequential pattern

of imino-imino and imino-amino NOEs is conserved in the presence of the 5fC:G base pair, confirming that base pairing and base stacking are maintained at the lesion sites (Figures 4.4 and 4.5). In addition, the pattern of sequential base aromatic→deoxyribose anomeric NOEs is also conserved (Figure 4.6), with no significant changes in chemical shifts, which indicates the presence of a normal undisrupted B-type helix.

Effect of 5fC on DNA stability

To investigate the thermal stability of 5fC in DNA we determined melting temperatures of the modified duplex by temperature-depended UV spectroscopy. The T_M was calculated by the first derivative of the resulting UV melting curves. The resulting T_M of DDD^f is 2 °C lower than the T_M of the unmodified duplex, suggesting lower stability of formylated DNA. The same destabilizing effect is observed in temperature-dependent NMR spectroscopy (Figure 4.3) Overall, the NMR data show that all imino protons start to broaden at 45 °C in DDD^f, while the same resonances appear as sharp peaks in the control sequence at the same temperature. The G^4 imino resonance, representing the modified base pair in DDD^f, broadened at 45 °C, but was still sharp and detectable at a temperature as high as 55 °C in the unmodified DDD.

Imino proton exchange and base pair opening kinetics

DNA glycosylase repair enzymes use an extrahelical recognition mechanism to bind and excise the damaged base from DNA (*82, 147*). In the series of studies done by Stivers *et al*., it was shown that DNA stability and dynamics play a key role in the detection of the damaged base by the repair enzyme. The hypothesis is that the base pair with lower stability will have kinetically higher opening rates (*54, 55, 148*). In order to characterize the kinetics of base pair opening in nucleic acid duplexes,

we performed high-resolution proton exchange NMR spectroscopy, where by the rates of exchange of imino protons with solvent protons were measured by magnetization transfer from water for each DNA duplex as a function of the concentration of exchange catalyst. Comparing G^4 from the DDD^f with G^4 from DDD shows a 3 times higher opening rate (k_{op}=26 vs. k_{op}=8, for G^4 in DDD^f and DDD, respectively). Together with the similar, low rate for base pair closing and base pair lifetime, that means that in fact the modified base pair spends more time exposed to the solvent than the canonical base. This is also confirmed by the equilibrium constant for base pair opening, which is 2.3-fold higher for modified G^4:X^9 vs. unmodified G^4:C^9, (αK_{op} = 2.8 × 10^7 and 1.2 × 10^7, respectively). This result implies that the possibility of 5fC adopting an extrahelical conformation for excision by thymine DNA glycosylase is high.

Summary

It was hypothesized that DNA glycosylases use an extrahelical mechanism to recognize damaged bases that may posses higher base pair opening rates and lower stability. In addition, structural characteristics recognized by TDG suggest that a Wobble base pairing is recognized geometry at the modification site. In our studies we showed that 5fC is stably stacked in DNA duplex, and that the melting temperature of the modified duplex is substantialy lower than of the canonical duplex. Moreover, the dynamics at the modification site and the opening rates are higher, which suggests

that the probability of 5fC base flipped out the DNA duplex and adopting an extrahelical conformation to be recognized by repair glycosylase is high.

Acknowledgements

This work is based upon research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines, which are supported by a grant from the National Institute of General Medical Sciences (P41 GM103403) from the National Institutes of Health. Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract No. DE-AC02-06CH11357.

This work was supported by NIH grant R01 CA-55678 (M.P.S.). Funding for the NMR instruments was provided by NIH grants S10 RR-05805, S10 RR-025677 and NSF Grant DBI 0922862, the latter funded by the American Recovery and Reinvestment Act of 2009 (Public Law 111-5). Vanderbilt University assisted with the purchase of NMR instrumentation.

CHAPTER 5

CHARACTERIZATION OF 5-CARBOXYL-CYTOSINE IN DICKERSON DREW DODECAMER

Introduction

5-carboxycytosine (5caC) is an oxidation derivative of 5mC, formed by the TET enzyme in the process of active demethylation (*39, 133*). The TET family of proteins has the capacity to convert 5mC to 5caC *in vitro*. The presence of 5caC in the genomic DNA was established in mouse embryonic stem (ES) cells (*39*). The functional effect on transcription of 5caC in the genome has been described, for mammalian and yeast RNA polymerase II. Pol II polymerization rates and specificity constants for GTP incorporation opposite 5caC were reduced significantly, although there were no changes in GTP incorporation opposite C, 5mC or 5hmC in the templates. Pol II can read and distinguish subtle changes at the 5 position of modified cytosines, and process them in different ways (*135*).

The TET-catalyzed oxidation is reminiscent of the thymine hydroxylase catalyzed conversion of thymine to iso-orotate, which was recently studied in the context of 5mC demethylation. It could be potentially achieved through a process similar to the conversion of thymine to uracil, which is achieved by conversion of thymine to iso-orotate decarboxylase (*149, 150*). However it is hypothetic pathway for DNA

demethylation, and the enzyme that is capable of decarboxylation of 5caC-containing DNA has not been identified yet. TET enzymes act together with other DNA demethylation pathways, such as BER. In this process, 5caC is removed by thymine DNA glycosylase (TDG) resulting in the formation of unmodified cytosine (*78*). It was shown, that TDG can excise 5caC from DNA in vitro, and this activity was subsequently found in mammalian cells (*136-139*). In order to understand the mechanism of active demethylation, that proceeds via 5caC formation chemical reactivity studies of the new nucleobases were reported (*140*). In this report, the sensitivity of these modifications toward the oxidation and deamination was measured, along with C-C bond cleaving reactivity, both in the absence and presence of thiols as catalysts. It turns out, that 5caC can undergo thiol-catalyzed and acid-catalyzed C-C bond cleavage reactions to form dC with the release of CO_2. This means that decarboxylation could take place as an alternative active demethylation pathway (*151*). Indeed, extensive studies by Schiesser *et al*. provided the first evidence that stem-cell nuclear extracts have the ability to decarboxylate 5caC (*152*).

However, recent biochemical and biological studies established that the pathway for active DNA demethylation involves BER and TDG enzyme. The latest studies of Maiti and Drohat report the investigation of the TDG mechanism excising 5caC from DNA, and the chemical properties indicating the catalytic requirements for the excision (*78, 141*). They showed that TDG can remove 5caC (*78*). TDG activity is greater for cytosine analogs with an electron-withdrawing substituent at the C5 position ($\sigma_m>0$, σ_m - electronic substituent constant) that can stabilize negative charge developing on the excised base in the transition state. In the case of 5caC, σ_m poorly predicts

the electronic effect for the carboxyl group. At neutral or physiological pH, the carboxyl group exists in the deprotonated form with σ_m= -0.10 (indicating electron-donating effect). However, as reported by Maiti and Drohat, the anionic carboxyl substituent lowers the pK_a of cytosine by -0.3, suggesting a stabilization effect of COO^- by NH_2 group at position C4, which would indicate an electron withdrawing effect and implying that TDG could remove it (*78, 141*). The other aspect addressed by the authors was the tautomerism of 5caC under physiological conditions. It was proposed by Hashimoto *et al*. that 5caC favors an imino tautomeric conformation leading to a Wobble-like structure of the pair with G, similar to the structure of G/U or G/T mispairs (Figure 5.1) (*142, 143*).

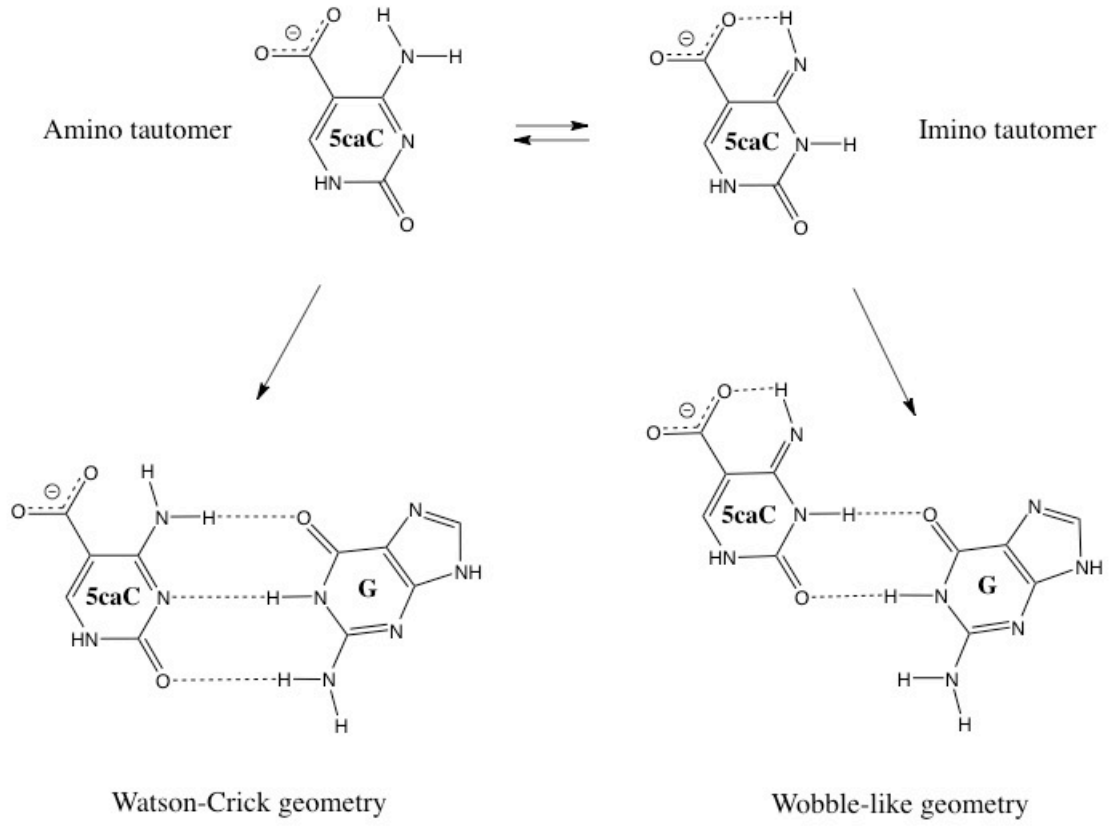


Figure 5.1. Tautomerism of 5caC modified base and base pairing schemes.

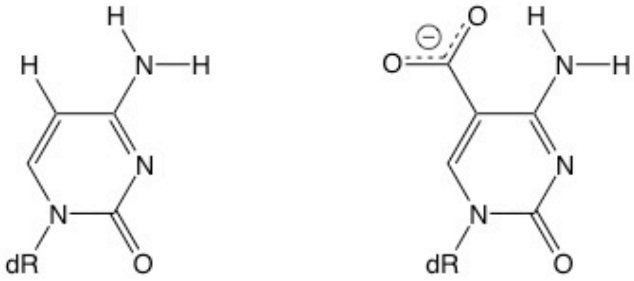
In the studies with TDG it was proposed that this unique structural geometry of the mismatch base pair is recognized by the enzyme, which results in excision of the mispaired base by TDG and repair. The hypothesis was tested by Maiti and Drohat recently, whereby they calculated the relative stability of the amino and imino tautomers for 5caC in a form of single base. Accordingly 5caC as a anionic tautomer is much more stable than its imino counterpart (Figure 5.1) in the gas phase and in the water phase (*141*), suggesting that 5caC when paired with G forms a Watson-Crick base pair with normal geometry. We addressed this hypothesis in the present work and investigated the stacking and stability of DNA containing a 5caC modified base paired with G.

DNA glycosylases use an extrahelical base recognition mechanism, which relies on kinetically enhanced base pair opening rates for destabilized base pairs (*82*). High-resolution proton exchange NMR spectroscopy was used to characterize the kinetics of base-pair opening in nucleic acid duplexes. The rates of exchange of imino protons with solvent protons were measured by magnetization transfer from water for each DNA duplex, as a function of the concentration of exchange catalyst. Here we discuss the potential of 5caC as a substrate for base excision repair.

Results

The sequence studied in this work is well-characterized Dickerson-Drew Dodecamer (*76*), which has C/G rich termini and A/T rich core. This sequence was selected because it is self-complementary and crystallized well. The 5caC modification was incorporated into the DDD oligodeoxynucleotide at the nineth position; consequently after annealing the modified duplex contains two 5caC residues. The DNA duplexes used in these studies are listed in Figure 5.2.

(a)



Cytosine; C 5-Carboxyl-cytosine; 5caC

(b)

DDD

5′-C^{2} G^{2} C^{3} G^{4} A^{5} A^{6} T^{7} T^{8} C^{9} G^{10}C^{11}G^{12} -3′
3′-G^{12}C^{11}G^{10}C^{9} T^{8} T^{7} A^{6} A^{5} G^{4} C^{4} G^{2} C^{1} -5′

DDD^{ca} (NMR)

5′-C^{2} G^{2} C^{3} G^{4} A^{5} A^{6} T^{7} T^{8} W^{9} G^{10}C^{11}G^{12} -3′
3′-G^{12}C^{11}G^{10}W^{9} T^{8} T^{7} A^{6} A^{5} G^{4} C^{4} G^{2} C^{1} -5′

DDD^{ca} (Xray)

5′-C^{2} G^{2} C^{3} G^{4} A^{5} A^{6} T^{7} T^{8} W^{9} G^{10}C^{11}G^{12} -3′
3′-G^{24}C^{23}G^{22}W^{21}T^{20}T^{19}A^{18}A^{17}G^{16}C^{15}G^{14}C^{13} -5′

Figure 5.2. (a). Structure of dC, and 5caC. (b). Sequences and numbering of the nucleotides for unmodified DDD, and DDD^{ca} (NMR) and DDD^{ca} (X-ray) duplexes. In solution, the two strands of the DDD exhibit pseudo-dyad symmetry and the NMR resonances of symmetry-related nucleotides in the two strands are not individually observed. In the crystal, electron density for symmetry-related nucleotides in the two strands is observed and the nucleotides are numbered individually.

Thermal denaturation studies

At neutral pH, the melting temperature (T_M) of the modified duplex was 54 °C as measured by UV spectroscopy in 10mM phosphate buffer at pH 7.0, 100 mM NaCl,

50 μM Na_2EDTA. The corresponding T_M for the unmodified duplex under the same conditions was 48 °C.

NMR melting studies

A series of 1D ^1H NMR spectra for the exchangeable protons were recorded in a range of 5 - 55 ºC and are shown in Figure 5.3. Overall the data indicate that all imino protons are much stable in modified DDD^{ca} than in control sample. The N^1-imino proton of the $G^2 \cdot C^{11}$ base pair in DDD was sharp and detectable up to 25 ºC, and at 35 ºC it started to broaden. The same imino peak in the DDD^{ca} duplex started to broaden at much higher temperature (45 ºC). The next base pair, $G^{10} \cdot C^3$, which is the 5'- neighbor of the modification, was sharp and quite intense in DDD^{ca}, even at a temperature as high as 55 ºC, while already broadened at the same temperature in the control experiment. The N^1-imino proton of the $W^9 \cdot G^4$ modified base pair in the DDD^{ca} is the most intense resonance among all, and remained sharp at 55 ºC. However the same resonance in DDD, even though it is the most stable, was much less intense and started to broaden. The N^3-imino resonances of T^7 and T^8 are the sharpest in both duplexes; they start to broaden at 55 ºC in the DDD and above that temperature in the DDD^{ca}, indicating that these thymine base pairs are much more stable than the same base pairs in the control sequence.

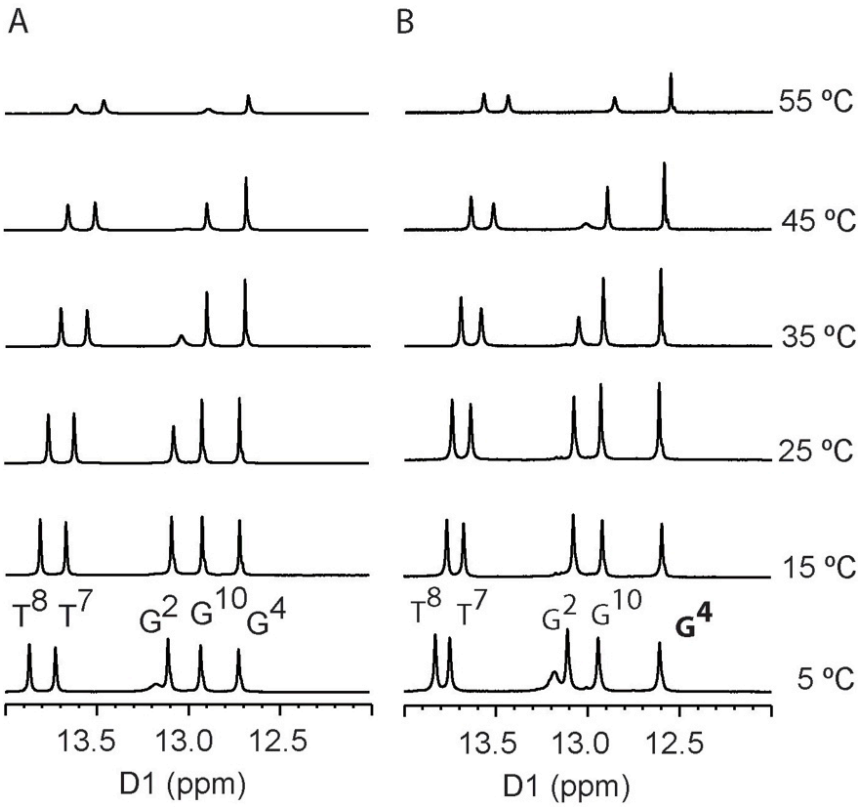


Figure 5.3. ^{1}H-NMR of imino proton resonances as a function of temperature for the unmodified DDD duplex (A), and the modified DDD^{ca} duplex (B).

Exchangeable protons.

Figure 5.4 shows the NOE connectivity of the purine N^{1} and pyrimidine N^{3} imino protons. The base imino protons were assigned based on their sequential connectivities in NOESY spectra, and these assignments were supported by their NOE cross-peaks to Watson-Crick base-paired amino protons. The sequential connectivities were obtained from base pairs G^{2}:C^{11} → G^{10}:C^{3} → G^{4}:C^{9} → T^{8}:A^{5} → T^{7}:A^{6}. For the DDD and DDD^{ca} duplexes, the imino-proton resonances of the terminal base pairs C^{1}:G^{12} are lost by fast exchange with water.

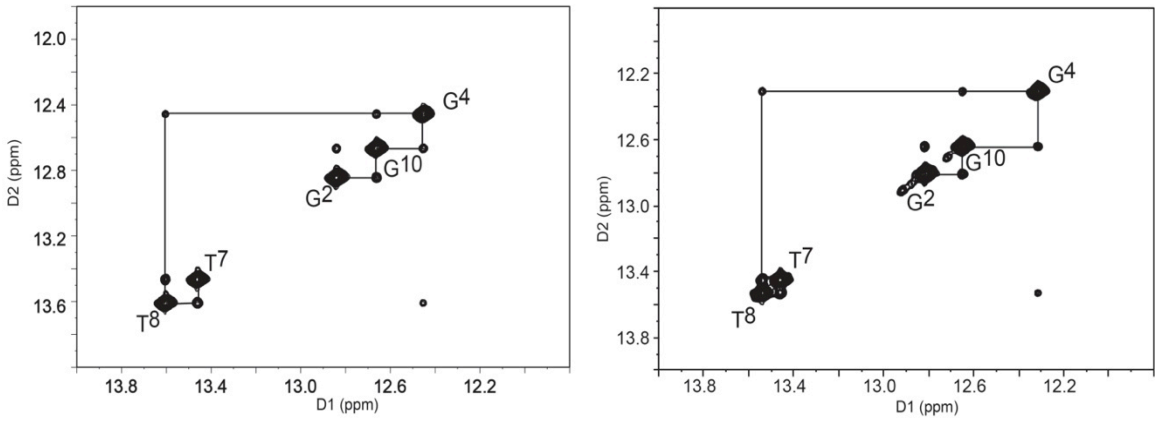


Figure 5.4. ^{1}H-^{1}H NMR NOESY spectrum showing resonances for the thymine and guanine imino protons and sequential NOE connectivity for the imino protons of the base pairs G^{2}:C^{11} to A^{6}:T^{7} for unmodified DDD (left), and modified DDD^{ca} (right) duplexes.

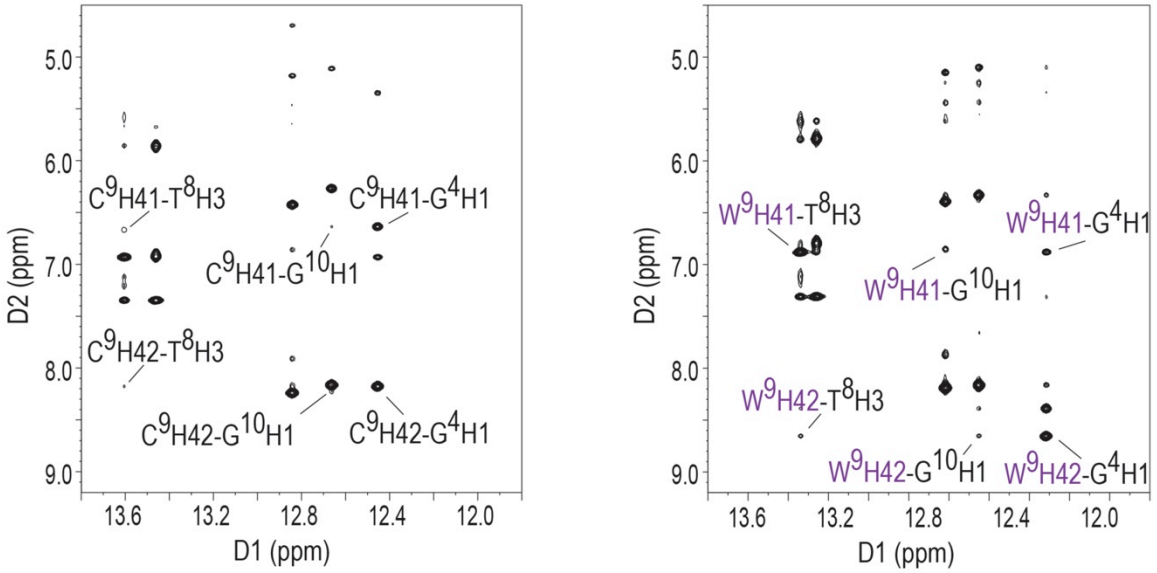


Figure 5.5. Expansion of the ^{1}H-^{1}H NOESY spectra for DDD (left), and DDD^{ca} (right) showing the conservation of Watson-Crick base pairing and base stacking at the modification site.

The upfield region of the NOESY spectrum, shown in Figure 5.5, showed the NOEs between the imino and amino protons. Cross peaks from modified base $C^{9}H41$, $C^{9}H42$ to the opposite $G^{4}H1$ base were observed, as well as interactions to neighbor $T^{8}H3$ and $G^{10}H1$ bases. No significant changes in intensities were observed in the DDD^{ca} duplex, compared with DDD.

Nonexchangeable protons

The sequential assignment of nonexchangeable protons was accomplished using standard protocols. The unmodified duplex was used as a control for NMR assignments for DDD^{ca}. For modified duplexes, the anticipated pattern of sequential base aromatic → deoxyribose anomeric nuclear Overhouser enhancement (NOE) was identified from C^{1} → G^{12} (Figure 5.6).

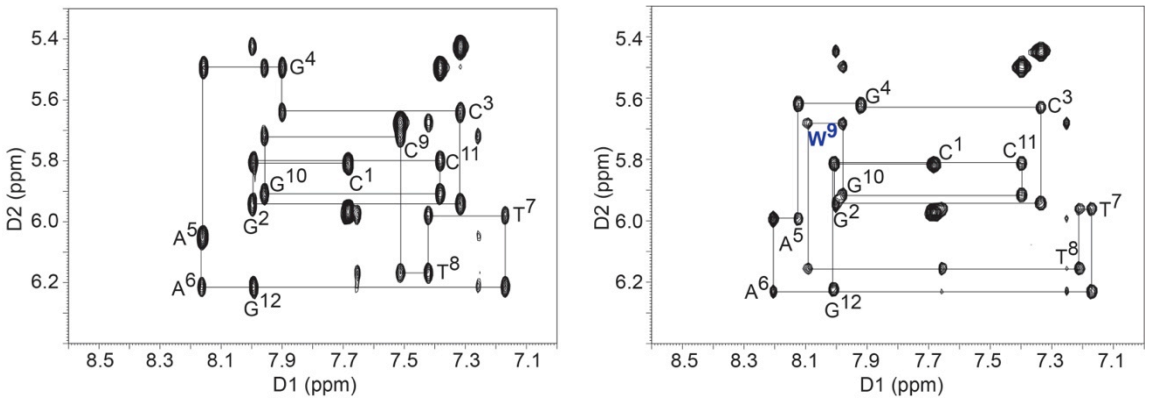


Figure 5.6. Expanded plot from the aromatic-anomeric region of the NOESY spectrum, showing sequential NOE connectivities in the unmodified DDD, and modified DDD^{ca} duplexes.

Only one set of resonances was observed because the sequences are self-complementary. The spectra exhibited well resolved cross-peaks. Three strong NOEs accounted for the H5-H6 cross-peaks of the cytosine residues (C^1, C^3, and C^{11}). The H5-H6 resonance from W^9 is missing due to carboxyl substitution at the 5-position in the DDD^{ca}. Cytosine assignments were confirmed by a DQF-COSY spectrum recorded under identical conditions (data not shown). Each base proton exhibited NOE peaks to its own and 5'-flanking H1' deoxyribose protons. For T^8, W^9 and G^{10} the NOE cross-peaks intensities between the base protons and the sugar H1' of the attached deoxyribose moieties were of the same relative magnitudes as those between other bases in the sequence. The 5caC H6 resonance was observed at 8.1 ppm, which is shifted downfield by approximately 0.6 ppm with respect to the unmodified oligodeoxynucleotide. This was attributed to the differential electronic density for 5caC as compared to G. Proton resonances from the opposite G^4 and A^5 bases exhibited chemical shift changes of <0.1 ppm compared with those bases in the unmodified DDD duplex.

Crystal structure of the DDD modified with 5caC

Crystals belonged to the orthorhombic $P2_12_12_1$ space group with unit cell parameters a = 24.25 Å, b = 41.34 Å, c = 66.41 Å. The crystal structure of DDD^{ca} was refined isotropically to a resolution of 1.95 Å. The rmsd values for bond lengths were 0.009 Å, and for angles they were 2.209 deg. The relatively high resolution achieved for the structure allowed for a detailed analysis of the geometric

and conformational changes in the duplex as a consequence of the introduced modification.

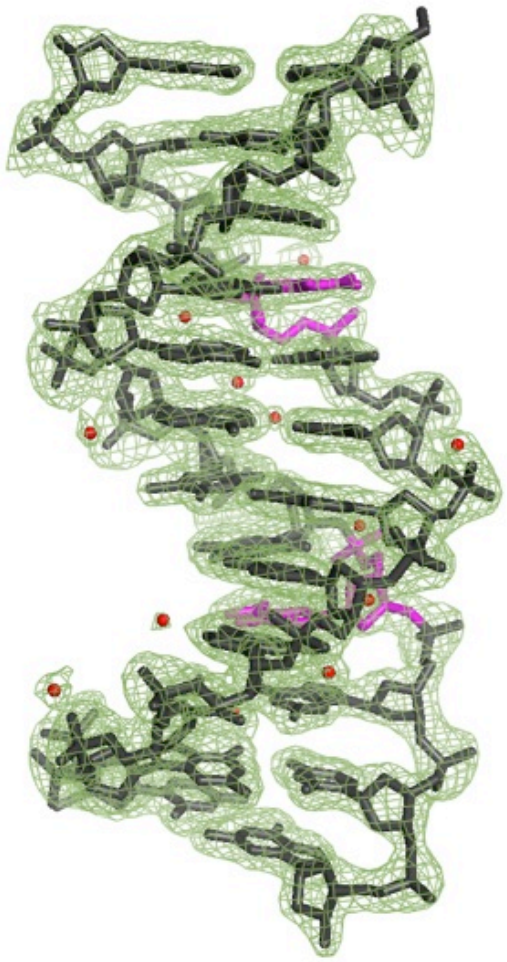


Figure 5.7. Fourier (2Fo-Fc) sum electron density contoured at the 1.0σ level (green meshwork) surrounding the DDD^{ca} duplex. The 5caC modified base is shown in magenta, and water molecules are shown in red.

Minimal perturbations of the DNA duplex were observed at the modification site (Figure 5.7). The carboxyl group was in the major groove (Figure 5.8), the oxygen from the COO^- group was within hydrogen bonding distance of the amino group at the C4 position of the modified cytosine.

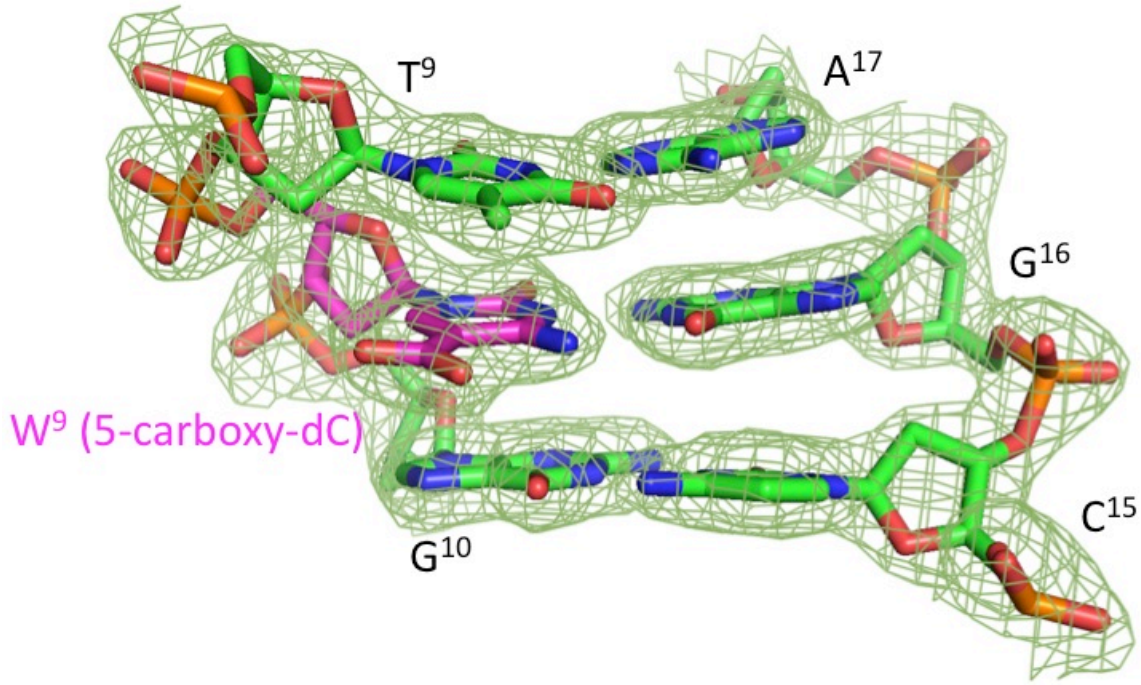


Figure 5.8. Fourier (2Fo-Fc) sum electron density contoured at the 1.0σ level (green meshwork) surrounding the DDD^{ca} duplex at one of the modification sites.

The DDD^{ca} structure is very similar to the other DDD structures. It overlays well with the canonical structure of DDD (PDB code: 436D) with an rmsd of 0.44 Å. The base stacking pattern in the modified DDD^{ca} is very similar to the one observed in unmodified B-type DNA (Figure 4.9), and helical parameters obtained with the program Curves+ (*100*) are comparable (Figures 4.11 – 4.13). The substitution of the cytidine nucleobase at the C5 position with the carboxyl group did not disrupt Watson-Crick base pairing or change the conformations of these residues (Figure 5.9). Analysis of the waters forming the "spine" of hydration, shows that the waters involved in the minor groove

hydration are conserved. Detailed information about data collection and refinement statistics are shown in Table 2.2.

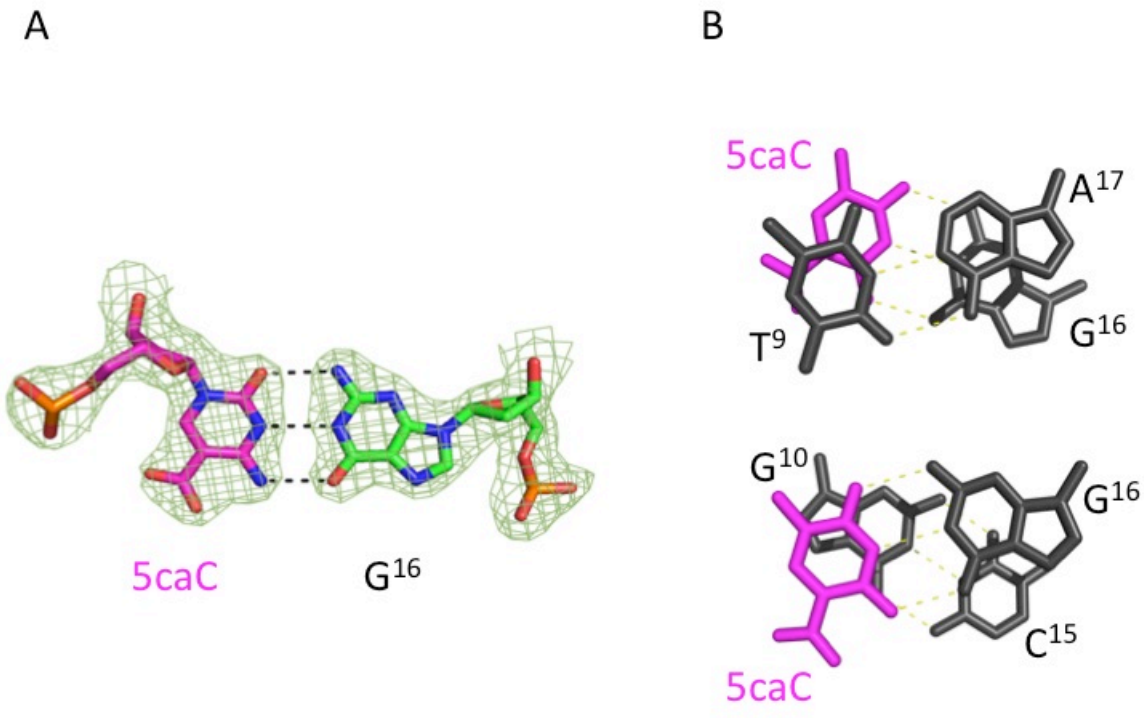


Figure 5.9. (A) Fourier (2Fo-Fc) sum electron density contoured at the 1.0σ level (green meshwork) around the 5caC:dG modified base pair showing conserved Watson-Crick base pairing geometry. (B) Stacking interactions in the DDD^{ca} duplex: (top) stacking of base pair T^{8}:A^{17} above base pair W^{9}:G^{16}, and (bottom) stacking of base pair W^{9}:G^{16} above base pair G^{10}:C^{15}.

Dynamics of base pair opening

To determine the dynamics of the base pairs, imino exchange rates were measured in the presence of base catalyst. In this work, ammonia base as the proton acceptor in imino proton exchange was used, due to its small size, lack of charge and high accessibility factor. Figure 5.10 shows imino proton exchange rates obtained by measuring magnetization transfer from water as a function of added ammonia base

catalyst. The results of the imino proton exchange analysis are summarized in Table 5.1. For the terminal (i.e. $C^{1}\cdot G^{12}$) and penultimate ($G^{2}\cdot C^{11}$) base pairs the exchange rates could not be measured because of the increased exchange with the solvent protons.

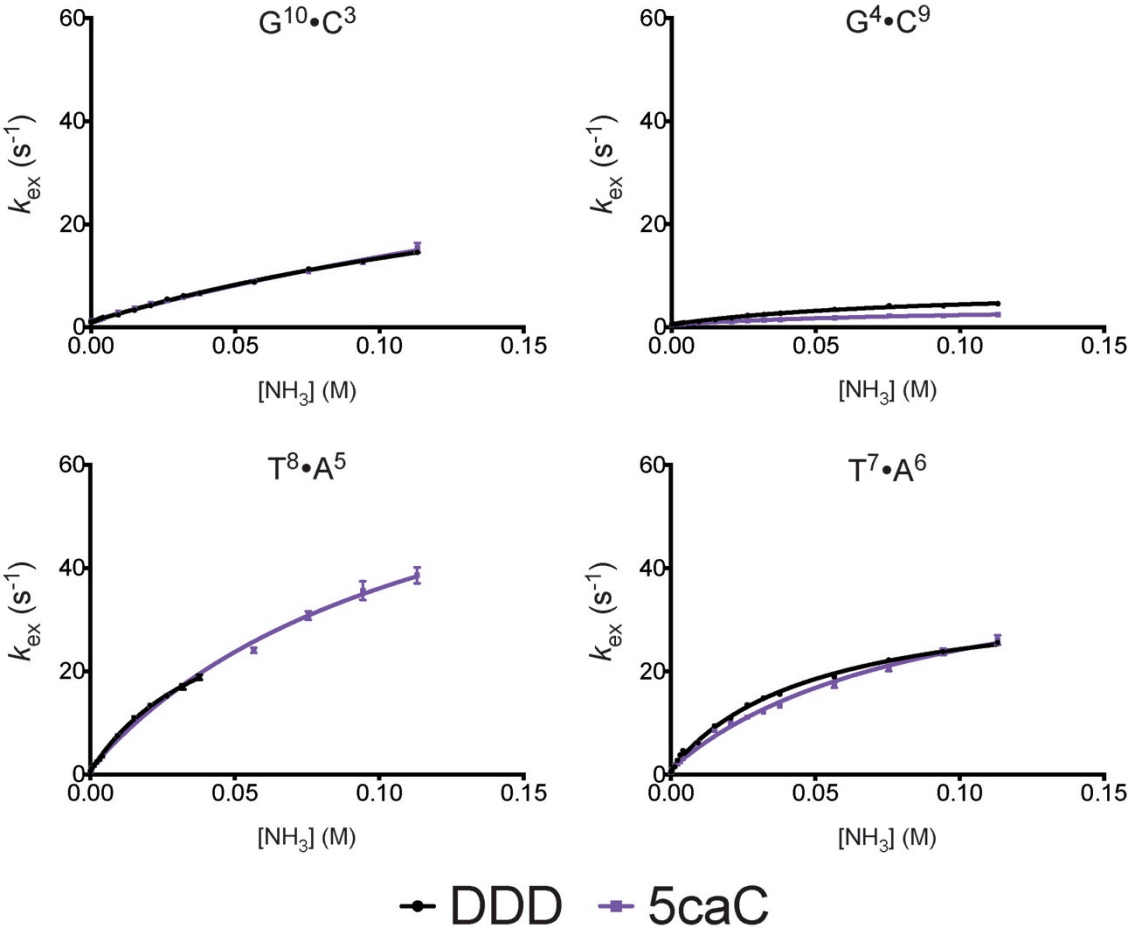


Figure 5.10. Plots showing imino proton exchange rates obtained by monitoring magnetization from water as a function of ammonia base catalyst. A. G^{10}:C^{3}. B. G^{4}:C^{9}. C. T^{8}:A^{5}. D. T^{7}:A^{6} in DDD (black), and DDD^{ca} (purple).

For all studied imino protons, in both duplexes, the exchange rate reaches the EX1 regime, as predicted by eq. 8. As the concentration of base ammonia catalyst increases, the values of the exchange rate increase until a plateau is reached, indicating a change in the rate limiting step from chemical exchange to base pair opening ($k_{ex} = k_{op}$

and exchange occuring at each opening event). For the imino proton of G^4, the exchange rate has a weak dependence on the concentration of ammonia base. This finding reflects the small equilibrium constant for opening of the $G^4 \cdot C^9$ base pair; in DDD^{ca} it is ~2 times smaller than in the DDD ($\alpha K_{op} = 0.6 \times 10^7$ vs. 1.2×10^7 for $G^{10}:C^3$ in DDD^{ca} and DDD, respectively). The same trend is observed with the opening rate, whereby $G^4 \cdot W^9$ base pair opens with a rate of 4 s^{-1}, while unmodified $G^4 \cdot C^9$ base pair has a rate of 8 s^{-1}. With almost the same an opening rate that is as the closing rate, the $G^4 \cdot W^9$ modified base pair is the least dynamic base pair in the analyzed duplexes.

	k_0 (s^{-1})[a]		$K_{op} \times 10^7$	
	DDD	DDD^{ca}	DDD	DDD^{ca}
$C^3:G^{10}$	1.1 ± 0.06	1.1 ± 0.06	2.9 ± 0.1	3.2 ± 0.08
$G^4:C^9$	0.57 ± 0.03	0.49 ± 0.04	1.2 ± 0.04	0.6 ± 0.04
$A^5:T^8$	0.59 ± 0.03	0.59 ± 0.03	41± 0.08	34 ± 1
$A^6:T^7$	0.61 ± 0.03	0.55 ± 0.03	37 ± 5	34 ±0.9

	k_{op} (s^{-1})		k_{cl} ($\times 10^{-7}$ s^{-1})	
	DDD	DDD^{ca}	DDD	DDD^{ca}
$C^3:G^{10}$	45 ± 3	64 ± 12	15 ± 1	20 ± 1
$G^4:C^9$	8 ± 0.5	4 ± 0.3	6.7 ± 0.2	6.6 ± 0.1
$A^5:T^8$	40 ± 2	78 ± 5	0.97 ± 0.04	2.3 ± 0.05
$A^6:T^7$	36 ± 1	45 ± 3	0.98 ± 0.09	1.3 ± 0.01

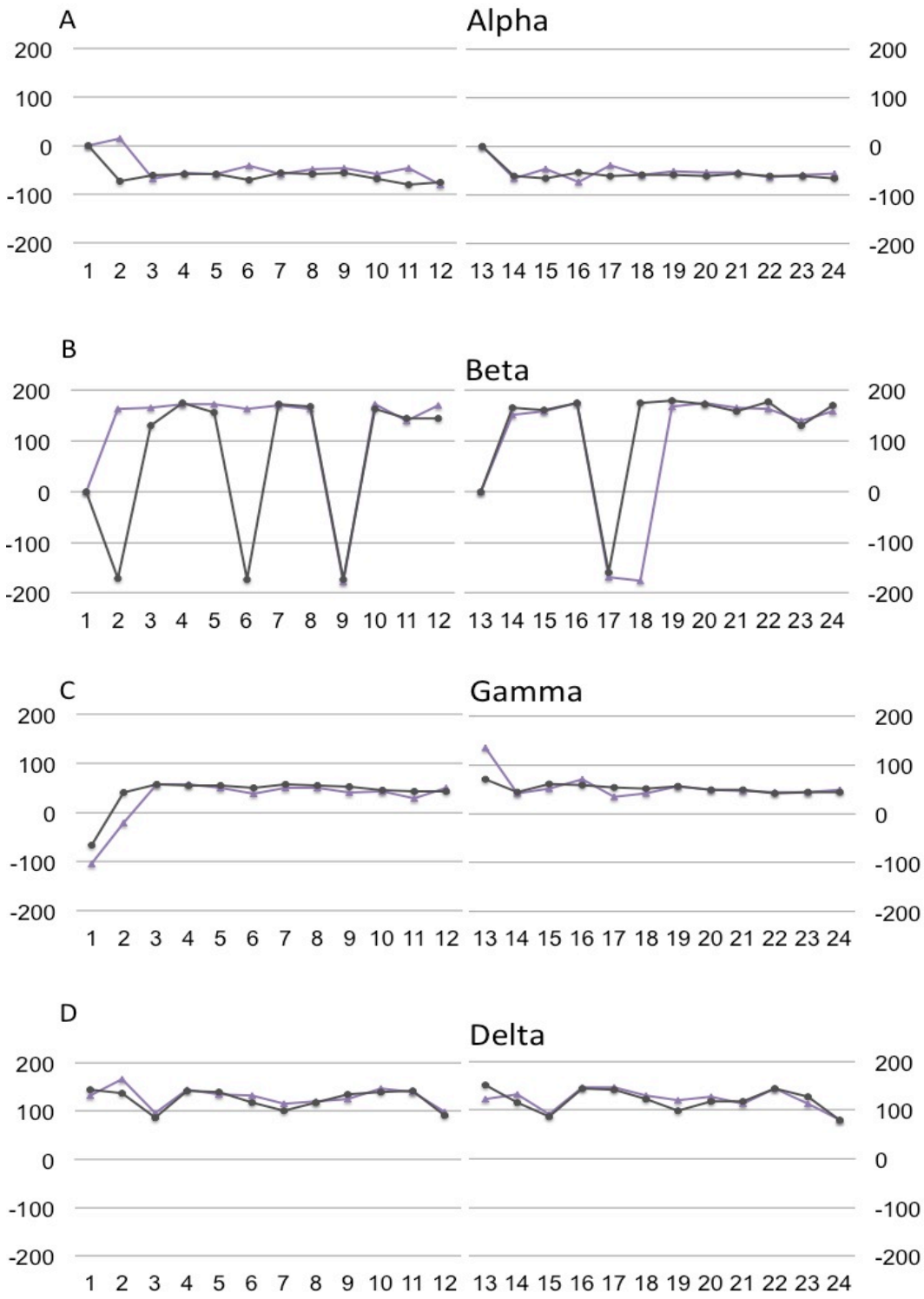
Table 5.1. Rate and Equilibrium Constants for DNA Base Pair Opening. [a]The observed exchange rate without an ammonia catalyst.

For $G^{10} \cdot C^3$ imino protons, the exchange rates and all kinetic parameters are at the same or similarly low level for both the DDD^{ca} and the DDD duplex.

The highest value for the equilibrium constant of base pair opening is for the T^8 imino proton, which is the 5' neighbor of the modification site. This applies to all three duplexes. The dynamics behavior of this base pair is remarkably different in the studied duplexes. With an opening rates that are almost doubled in DDD^{ca}, and doubled closing rate (Table 5.1), it appears that the frequency for base pair opening and closing are the same.

Helical analysis

An analysis of the helical parameters of the DDD^{ca} was performed with the program Curves+ (*100*). The analysis of the backbone torsion angles (Figure 5.11) showed no significant changes in alpha, gamma, delta, and zeta angles. For the beta torsion angle, the DDD^{ca} shows similar conformations as the unmodified DDD (*77*). However, peaks at the A^6 and A^{17} reflect changes induced by the interacting Mg^{2+} ion observed in the unmodified DDD structure. Similarly the peak at T^{20} residue for DDD^f relative to DDD for torsion angle epsilon reflects the change as the result of the presence of the modification at the flanking X^{21} residue. The analysis of interbase parameters shows no substantial changes in helical rise, roll and twist, as shown in Figure 5.12. Moreover shear and stretch, that characterize intrabase translations are comparable between bases in all studied duplexes (Figure 5.13).


A
Alpha
200
100
0
-100
-200
1 2 3 4 5 6 7 8 9 10 11 12
13 14 15 16 17 18 19 20 21 22 23 24
B
Beta
C
Gamma
D
Delta

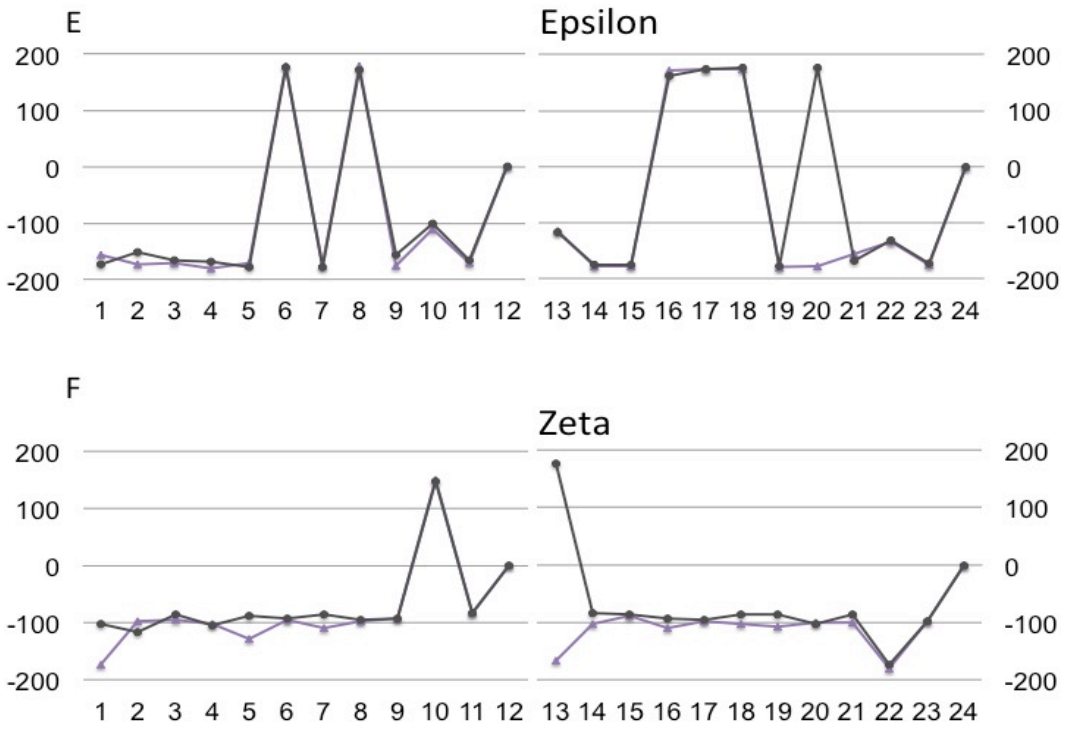


Figure 5.11. Comparison of backbone torsion angles (a) alpha, (b) beta, (c) gamma, (d) delta, (e) epsilon, and (f) zeta in the structure of the DDD^{ca} (in purple), and in the unmodified DDD (PDB entry 436D (124)) (in black).

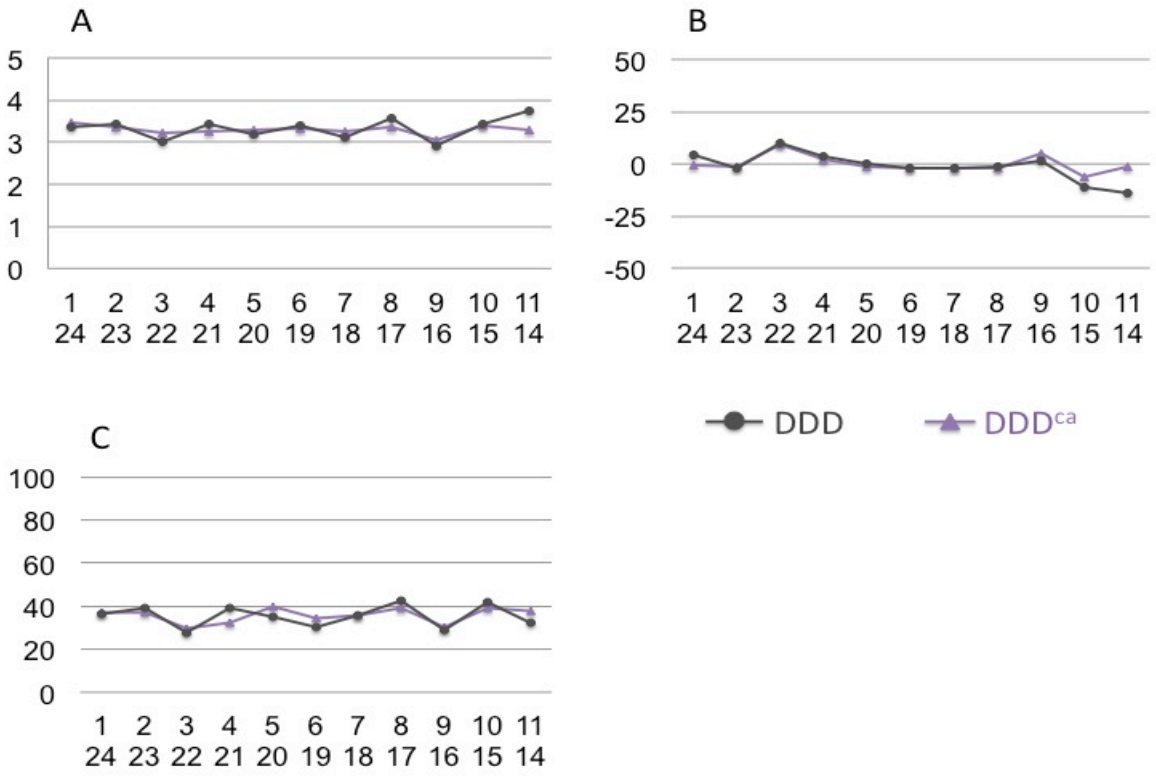


Figure 5.12. Interbase pair parameters: (a) helical rise, (b) roll, and (c) twist for the structure of DDD^{ca} in purple, and unmodified DDD (PDB entry 436D (124)) in black.

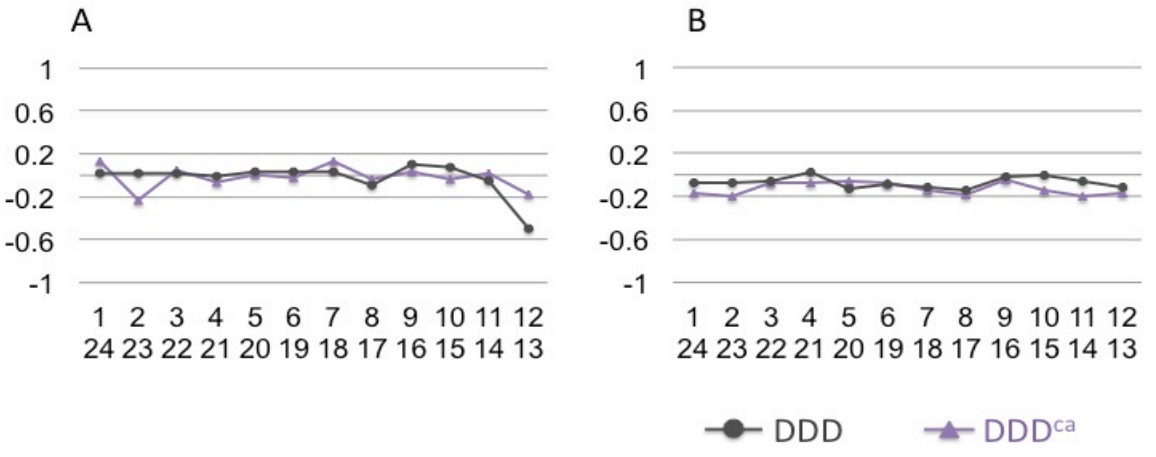


Figure 5.13. Intrabase pair parameters: (a) shear, (b) stretch in the structure of the DDD^{ca} (in purple), and the unmodified DDD (PDB entry 436D (124)) (in black).

Discussion

The 5caC base is an active demethylation oxidation product derived from 5mC, formed post-replicatively by TET enzymes (*39*). The 5caC base has been detected I genomic DNA of mouse embryonic stem (ES) cells and was found to be at level of 3 5caC in every 10^6 C (*39*).

Effect of 5caC on DNA stability.

The T_m of the DDD^{ca} duplex is 6 °C higher than the T_m of the unmodified duplex, demonstrating a stabilizing effect of carboxylated DNA. The same effect is observed in temperature-dependent NMR spectroscopy (Figure 5.3). Overall, the NMR data show that all imino protons appear as sharp peaks at a temperature as high as 55 °C. Only the penultimate base pair in DDD^{ca} is broadened at 55 °C, but resonances coming from other base pairs are still observed. The G^4 imino resonance, representing the modified base pair G^4:W^9 appears as the sharpest and most intense peak in the spectrum. The imino protons of the unmodified sequence are broadening at lower temperature. Previous studies on DNA electrostatics and stability showed that the presence of positive charge on the DNA stabilizes the duplex (*153*). It would appear by analogy that additional negative charge in DNA would result in destabilization. In fact, Maiti and Drohat (*141*) showed that 5caC ionizes with pK_a values of 4.28 at the N3 position, and 2.45 at the carboxyl, which suggests that 5caC exists as a monoanion at physiological pH. Substitution with 5caC introduces a negatively charged carboxyl group, which affects the electrostatics of the nucleobase. Double bond character on the oxygen from the carboxyl group serves as a proton acceptor and is capable to form a hydrogen bond. The amino group at C4 is in close proximity and donates a hydrogen. This additional hydrogen bond is formed within the 5caC:G base pair, which enlarges the modified

cytosine by an additional ring. In this fashion it mimics the two-ring purine nucleic base. This results in increased stacking and increased DNA stability (Figure 5.9). Base stacking interactions are of importance in stabilizing nucleic acid duplex.

Imino Proton Exchange and Base Pair Opening Kinetics.

DNA glycosylase repair enzymes are believed to employ an extrahelical recognition mechanism to bind and excise damaged bases from DNA (*147, 154-158*). Stivers and co-workers (*54, 148, 159*) showed that DNA stability and dynamics play a crucial role in detection of damaged base by repair enzyme. The assumption is that base pairs with lower stabilities will have greater opening rates. For the G^4 imino proton, which represents the modified base pair, our data show that the exchange rate has a weak dependence on the concentration of the ammonia base, resulting in a small equilibrium constant for opening the G^4:W^9 base pair. In the DDD^{ca}, αK_{op} is small and $\alpha K_{op} = 0.6 \times 10^7$ for the modified base pair, while in control DDD $\alpha K_{op} = 1.2 \times 10^7$. The same trend is observed for the opening rate, whereby $G^4 \cdot W^9$ base pair opens 4 times per second, while the unmodified $G^4 \cdot C^9$ base pair opens with a rate of 8 s^{-1}. With similar values for opening and closing rates, the $G^4 \cdot W^9$ modified base pair represents the least dynamic base pair among all pairs in the studied duplexes. This result implies that the possibility of 5caC adopting an extrahelical conformation for excision by thymine DNA glycosylase is unlikely.

Effect of 5caC on DNA Structure and Structure-Activity Relationships.

In order to understand the mechanism of active demethylation through 5caC formation, chemical reactivity studies on the modified nucleobases were reported (*140*). The sensitivity of these modifications toward oxidation and deamination was measured,

along with C-C bond cleaving reactivity, in the presence and absence of thiols as catalysts. 5caC can undergo thiol-catalyzed and acid-catalyzed C-C bond cleavage reactions to form dC with the release of CO_2. Thus, decarboxylation could occur as an alternative active demethylation pathway (*151*). It was proposed that in the process of active DNA demethylation, in the last step, 5caC could decarboxylate to dC (*160*). In that case it is believed that DNA decarboxylase might share some structural, sequential and mechanistic similarities with isoorotate decarboxylase (IDCase), which catalyzes decarboxylation of 5caU to U in fungi (*160*). Indeed, Xu *et al*. (*160*) reported biochemical studies with fungi IDCase catalyzing decarboxylation of 5caC to dC, albeit with weak activity. Nonetheless this was the first *in vitro* evidence for direct decarboxylation of 5caC to dC by the enzyme. However, with respect to DNA demethylation in mammals, no enzyme that is capable of decarboxylation of 5caC-containing DNA has been identified. TET enzymes act together with other DNA demethylation pathways such as BER. In this process 5caC is removed by thymine DNA glycosylase (TDG) resulting in the restoration of unmodified cytosine (*43, 78, 161, 162*). TDG can excise 5caC from DNA in vitro, and this activity was confirmed in mammalian cells (*136-139*).

It was proposed by Hashimoto *et al*. (*142, 143*) that 5caC undergoes tautomerization caused by the presence of the negatively charged carboxyl group at C5 on modified cytosine, and forms a Wobble base pair with the opposite guanine, similar to the structure of G:U or G:T mispairs (Figure 5.1) (*142, 143*). This mismatch base pair geometry is recognized by TDG, and results in the excision of the mispaired base and repair. In order to determine the basis by which 5caC pairs with the opposite guanine in DNA, the structure of 5caC placed opposite dG (DDD^{ca}) was determined. The crystals of DDD^{ca} diffracted at a resolution of 1.95 Å. The structure revealed that 5caC has a minimal effect on duplex conformation (Figures 5.7 and 5.8), and base pair geometry (Figure 5.9), as compared to canonical dC:dG base pair. The Watson-Crick

geometry was conserved for the 5caC:G base pair, indicating the amino tautomer for the modified base (Figure 5.9). In addition, the solution NMR data corroborate the crystallographic data as can be seen from a comparison of NOESY spectra of the modified and unmodified duplexes. The sequential pattern of imino-imino and imino-amino NOEs is conserved in the presence of the modified base pair, confirming that the base pairing and base stacking is maintained at the lesion sites (Figures 5.4 and 5.5). In addition, the pattern of sequential base aromatic→deoxyribose anomeric NOEs is also conserved (Figure 5.6), with no significant changes in chemical shifts, which indicates the presence of a normal undisrupted B-type helix. It was confirmed at the nucleobase level as well, by Maiti and Drohat (*78, 141*), who showed higher relative stabilities of the amino tautomers for 5caC. These authors also reported an investigation of the TDG mechanism excising 5caC from DNA, and the chemical properties indicating the catalytic requirements for the excision. They showed that TDG activity is greater for cytosine analogs with an electron-withdrawing substituent at the C5 position, which can stabilize negative charge developing on the excised base in the transition state. In the case of 5caC, σ_m does not predict the electronic effect for the carboxyl group well. At physiological pH, the carboxyl group is in the anionic form, and $\sigma_m = -0.10$, which indicated an electron-donating effect. However, Drohat *et al*. (*78, 141*) calculated that the anionic carboxyl substituent lowers the pK_a of cytosine by -0.3. This suggests a stabilizing effect of COO^- by NH_2 group at position C4, which would indicate an electron withdrawing effect. It implies that TDG could remove 5caC from DNA (*78, 141*).

Summary

The 5caC base is stably stacked in the DNA duplex, and the melting temperature of the DDD^{ca} is higher than that of the canonical duplex. The dynamics at the modification site and the opening rates are low, which suggests a low probability of 5caC being flipped out of the duplex and adopting an extrahelical conformation to be recognized by thymine DNA glycosylase is low.

Acknowledgements

This work is based upon research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines, which are supported by a grant from the National Institute of General Medical Sciences (P41 GM103403) from the National Institutes of Health. Use of the Advanced Photon Source, an Office of Science User Facility operated for the U.S. Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the U.S. DOE under Contract No. DE-AC02-06CH11357.

This work was supported by NIH grant R01 CA-55678 (M.P.S.). Funding for NMR was supplied by NIH grants S10 RR-05805, S10 RR-025677 and NSF Grant DBI 0922862, the latter funded by the American Recovery and Reinvestment Act of 2009 (Public Law 111-5). Vanderbilt University assisted with the purchase of NMR instrumentation.

CHAPTER 6

SUMMARY, CONCLUSION AND FUTURE DIRECTIONS

This dissertation describes studies toward the characterization of oxidation derivatives of 5-methylcytosine (5mC): 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC), which were site-specifically incorporated into the Dickerson-Drew Dodecamer (DDD) DNA sequence. The structures of modified DNA in the crystalline state were determined by X-ray crystallography, and the characterization of modified DNA in solution was performed using high-resolution NMR spectroscopy. In order to characterize the kinetics of base-pair opening in the native and modified nucleic acid duplexes, high-resolution proton exchange NMR spectroscopy was used. The rates of exchange of imino protons with solvent protons were measured by magnetization transfer from water for each DNA duplex as a function of the concentration of exchange base catalyst.

The methylation of cytosine in DNA is very important in epigenetics, as cells use methylation cells to switch genes on and off on their way to specialization (*23-25, 163*). While it is clear how the methylation occurs, the question of demethylation pathways is of high interest and many different studies have addressed this issue recently. The search for active demethylation pathways led to the discovery of 5hmC in the mammalian genome. As it turns out 5hmC is oxidized by TET enzymes from 5mC (*33, 34, 163*). Since 5hmC was found to be tissue specific and high levels of 5hmC were found in the central nervous system, it was proposed that 5hmC is an epigenetic marker

similar to 5mC (*33, 34, 163*). However, at approximately at the same time it was found that TET enzymes can further oxidize 5hmC to 5fC and 5caC, and the latter could be decarboxylated to unmodified cytosine. Therefore it was proposed that active demethylation of 5mC proceeds through oxidation to 5hmC, 5fC, and 5caC by TET enzymes (*39*). However, TET enzymes act together with other DNA demethylation pathways such as base excision repair (BER). Maiti and Drohat showed that TDG might be involved in demethylation when 5fC and 5caC, but not 5hmC, are removed via BER (*115*). While thymine DNA glycosylase (TDG) removes 5fC and 5caC, it is unable to remove 5hmC (*115*).

The initial and most important step in BER involves finding the damaged base by the repair enzyme. There are many different hypotheses about the recognition of the lesion by the enzyme, and some of them were addressed in this dissertation. In one of the hypothesis, the structural characteristics recognized by TDG involve Wobble base pairing geometry at the modification site (*57, 141-143, 146, 164*). 5hmC was found not to be a substrate for TDG and BER, but rather an epigenetic marker (*78, 165*). High resolution NMR spectroscopy and the X-ray crystal structure of the DDD substituted with 5hmC at 1 Å resolution showed no perturbations in the DNA duplex conformation as the structure was of normal B-type geometry. At the modification site the 5hmC base was stably stacked and paired with opposite G according to Watson-Crick geometry. The hydroxymethyl group of the modified base was oriented toward the 3'- end of the duplex and interacted via two structural water molecules with O6 and N7 on flanking 3'- G^{10} base. That suggested a stabilizing effect as a result of the oxidized modified base. The further oxidation products 5fC and 5caC were proposed to undergo tautomerization,

which would lead to protonation at N3 of the modified cytosine and affect pairing with guanine (*142, 143*). Thus, in order to avoid two potentially repulsive hydrogens in the base pair, both nucleobases would have to relocate into a different conformation to form, for instance, Wobble-like base pair geometry, as it was observed in case of G/T or G/U mispairs in DNA, which were shown to be recognized and excised by TDG (*57*). However there is no evidence for dual or any different conformations for modified 5fC:G and 5caC:G base pair in the studied DDD^{f} and DDD^{ca} structures. Both modifications form normal Watson-Crick geometry in the crystalline state, and there was no evidence for imino tautomers present in solution as shown by NMR high-resolution spectroscopy.

It was hypothesized that DNA glycosylases use an extrahelical mechanism to recognize the modified base possessing higher opening rates and lower stability. When 5hmC analog was introduced to DDD sequence, no change the thermal stability was observed relative to unmodified DDD. The dynamics at the modification site and the opening rates are low, which suggests that the probability of 5hmC to flip out the DNA duplex and adopting an extrahelical conformation to be recognized by repair glycosylase is low. These findings are consistent with current studies done by other labs in that 5hmC is not a substrate for TDG glycosylase (*78*). The presence of the 5fC in a DNA duplex decreases its stability and the kinetics are remarkably different from the control sequence. The opening rate for the modified base pair is 3-times higher than that for the canonical sample. In addition, the dynamics for base pair opening for the flanking 5'-neighboring base pair is also five times higher. The high frequency for opening of bases at the modification site indicates a highly dynamic “behavior” that could be captured by the repair DNA glycosylase and attract the enzyme

to the damaged base. In the case of 5caC, the melting temperature is higher relative to the canonical duplex. The dynamics at the modification site and the opening rates are low, suggesting a low probability of 5caC to flip out the duplex and adopt an extrahelical conformation that can be recognized by thymine DNA glycosylase. Since the 5caC modification was found to be excised from DNA by TDG it raises the question about the recognition of the lesion by the enzyme. There was no structural indication of unusual base pairing geometry in the crystal structure. Moreover, the low dynamic properties and high stability of the base pair does not support this hypothesis. That means that TDG may use a different recognition mechanism to process the damaged base.

The work on many different glycosylases demonstrated that there is not one universal mechanism used by all the enzymes to locate the lesion (*147, 166-169*). In the studies done by Plum and Breslauer the relationship between DNA lesions and DNA repair involving a "thermodynamic signature" was proposed (*170*). This led to the hypothesis that repair enzymes can sense the damaged nucleotide based on its local thermodynamic effect, even when the structures of the DNA with and without the lesion are indistinguishable by structural analysis using either crystallography or NMR. Indeed, this scenario was tested in studies with 5-hydroxy-2'-deoxycytidine (5ohC) embedded in the DDD sequence (*171*). This oxidative lesion is capable of Watson-Crick pairing with guanine, although it is highly destabilizing when compared to unmodified DDD, due to the reduced enthalpic term as a result of disrupted base stacking. Temperature-dependent NMR studies revealed local destabilization of the modified base pair, whereas it is the most stable base pair in native sequence. The crystal structure with the 5ohC

modification is indistinguishable from the canonical sequence (*172*). Thus it was proposed, that DNA glycosylase uses the thermodynamic signature mechanism to find the lesion based on local thermodynamic destabilization (*171*). The detailed analysis of the 5hmC modification in the DDD duplex revealed no thermodynamic changes compared to the native DDD. 5hmC is a stable base in DNA and with no enzyme discovered yet that can excise 5hmC from the genome, it is possible that it serves as an epigenetic marker. For the other oxidation products such as 5fC and 5caC, the complete thermodynamic analysis is still to be performed. In particular, as far as 5caC is concerned, additional information about the effects of this modification on the DNA stability could provide the specifics, which could be sensed by the repair enzyme.

APPENDIX

PDB COORDINATE FILES

File A-1: Crystal structure 5-hydroxymethyl-2'-deoxycytidine in B-type DNA, DDD^{hm} (PDB code 4I9V).

```
HEADER    DNA                                     05-DEC-12   4I9V
TITLE     THE ATOMIC STRUCTURE OF 5-HYDROXYMETHYL 2'-DEOXYCITIDINE BASE PAIRED
TITLE    2 WITH 2'-DEOXYGUANOSINE IN DICKERSON DREW DODECAMER
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: DNA (5'-D(*CP*GP*CP*GP*AP*AP*TP*TP*(5HC)P*GP*CP*G)-3');
COMPND   3 CHAIN: A, B;
COMPND   4 ENGINEERED: YES
SOURCE    MOL_ID: 1;
SOURCE   2 SYNTHETIC: YES;
SOURCE   3 ORGANISM_SCIENTIFIC: SYNTHETIC DNA;
SOURCE   4 ORGANISM_TAXID: 32630
KEYWDS    5-HYDROXYMETHYL 2' DEOXYCITIDINE, 5-HYDROXYMETHYL-DC ADDUCT, MODIFIED
KEYWDS   2 DDD, DODECAMER OF B-DNA, DNA
EXPDTA    X-RAY DIFFRACTION
AUTHOR    M.W.SZULIK, B.NOCEK, A.JOACHIMIAK,M.P.STONE
REVDAT   1   20-NOV-13 4I9V    0
JRNL        AUTH   M.W.SZULIK,B.NOCEK,A.JOACHIMIAK,M.P.STONE
JRNL        TITL   THE ATOMIC STRUCTURE OF 5-HYDROXYMETHYL 2`-DEOXYCITIDINE
JRNL        TITL 2 BASE PAIRED WITH 2`-DEOXYGUANOSINE IN DICKERSON DREW
JRNL        TITL 3 DODECAMER
JRNL        REF    TO BE PUBLISHED
JRNL        REFN
REMARK   2
REMARK   2 RESOLUTION.    1.02 ANGSTROMS.
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : REFMAC 5.7.0029
REMARK   3   AUTHORS     : MURSHUDOV,VAGIN,DODSON
REMARK   3
REMARK   3    REFINEMENT TARGET : MAXIMUM LIKELIHOOD
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) : 1.02
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) : 14.88
REMARK   3   DATA CUTOFF            (SIGMA(F)) : NULL
REMARK   3   COMPLETENESS FOR RANGE        (%) : 95.0
REMARK   3   NUMBER OF REFLECTIONS             : 32113
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT.
REMARK   3   CROSS-VALIDATION METHOD          : THROUGHOUT
REMARK   3   FREE R VALUE TEST SET SELECTION  : RANDOM
REMARK   3   R VALUE     (WORKING + TEST SET) : 0.161
REMARK   3   R VALUE           (WORKING SET) : 0.159
REMARK   3   FREE R VALUE                     : 0.182
```

```
REMARK   3   FREE R VALUE TEST SET SIZE   (%) : 5.000
REMARK   3   FREE R VALUE TEST SET COUNT      : 1700
REMARK   3
REMARK   3  FIT IN THE HIGHEST RESOLUTION BIN.
REMARK   3   TOTAL NUMBER OF BINS USED           : 20
REMARK   3   BIN RESOLUTION RANGE HIGH       (A) : 1.02
REMARK   3   BIN RESOLUTION RANGE LOW        (A) : 1.05
REMARK   3   REFLECTION IN BIN     (WORKING SET) : 1374
REMARK   3   BIN COMPLETENESS (WORKING+TEST) (%) : 56.96
REMARK   3   BIN R VALUE           (WORKING SET) : 0.3350
REMARK   3   BIN FREE R VALUE SET COUNT          : 82
REMARK   3   BIN FREE R VALUE                    : 0.3020
REMARK   3
REMARK   3  NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK   3   PROTEIN ATOMS            : 0
REMARK   3   NUCLEIC ACID ATOMS       : 490
REMARK   3   HETEROGEN ATOMS          : 26
REMARK   3   SOLVENT ATOMS            : 187
REMARK   3
REMARK   3  B VALUES.
REMARK   3   FROM WILSON PLOT           (A**2) : NULL
REMARK   3   MEAN B VALUE      (OVERALL, A**2) : 22.68
REMARK   3   OVERALL ANISOTROPIC B VALUE.
REMARK   3    B11 (A**2) : 0.24000
REMARK   3    B22 (A**2) : -1.49000
REMARK   3    B33 (A**2) : 1.26000
REMARK   3    B12 (A**2) : 0.00000
REMARK   3    B13 (A**2) : 0.00000
REMARK   3    B23 (A**2) : 0.00000
REMARK   3
REMARK   3  ESTIMATED OVERALL COORDINATE ERROR.
REMARK   3   ESU BASED ON R VALUE                            (A): 0.029
REMARK   3   ESU BASED ON FREE R VALUE                       (A): 0.030
REMARK   3   ESU BASED ON MAXIMUM LIKELIHOOD                 (A): 0.023
REMARK   3   ESU FOR B VALUES BASED ON MAXIMUM LIKELIHOOD (A**2): 0.990
REMARK   3
REMARK   3 CORRELATION COEFFICIENTS.
REMARK   3   CORRELATION COEFFICIENT FO-FC      : 0.979
REMARK   3   CORRELATION COEFFICIENT FO-FC FREE : 0.974
REMARK   3
REMARK   3  RMS DEVIATIONS FROM IDEAL VALUES        COUNT    RMS    WEIGHT
REMARK   3   BOND LENGTHS REFINED ATOMS        (A):   629 ; 0.020 ; 0.012
REMARK   3   BOND LENGTHS OTHERS               (A):   375 ; 0.006 ; 0.020
REMARK   3   BOND ANGLES REFINED ATOMS   (DEGREES):   937 ; 2.488 ; 1.354
REMARK   3   BOND ANGLES OTHERS          (DEGREES):   888 ; 2.562 ; 3.000
REMARK   3   TORSION ANGLES, PERIOD 1    (DEGREES):  NULL ;  NULL ;  NULL
REMARK   3   TORSION ANGLES, PERIOD 2    (DEGREES):  NULL ;  NULL ;  NULL
REMARK   3   TORSION ANGLES, PERIOD 3    (DEGREES):  NULL ;  NULL ;  NULL
REMARK   3   TORSION ANGLES, PERIOD 4    (DEGREES):  NULL ;  NULL ;  NULL
REMARK   3   CHIRAL-CENTER RESTRAINTS       (A**3):    78 ; 0.406 ; 0.200
REMARK   3   GENERAL PLANES REFINED ATOMS      (A):   337 ; 0.041 ; 0.020
REMARK   3   GENERAL PLANES OTHERS             (A):   137 ; 0.008 ; 0.020
REMARK   3   NON-BONDED CONTACTS REFINED ATOMS (A):  NULL ;  NULL ;  NULL
REMARK   3   NON-BONDED CONTACTS OTHERS        (A):  NULL ;  NULL ;  NULL
REMARK   3   NON-BONDED TORSION REFINED ATOMS  (A):  NULL ;  NULL ;  NULL
REMARK   3   NON-BONDED TORSION OTHERS         (A):  NULL ;  NULL ;  NULL
```

```
REMARK   3   H-BOND (X...Y) REFINED ATOMS      (A):  NULL ;  NULL ;  NULL
REMARK   3   H-BOND (X...Y) OTHERS             (A):  NULL ;  NULL ;  NULL
REMARK   3   POTENTIAL METAL-ION REFINED ATOMS (A):  NULL ;  NULL ;  NULL
REMARK   3   POTENTIAL METAL-ION OTHERS        (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY VDW REFINED ATOMS        (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY VDW OTHERS               (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY H-BOND REFINED ATOMS     (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY H-BOND OTHERS            (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY METAL-ION REFINED ATOMS  (A):  NULL ;  NULL ;  NULL
REMARK   3   SYMMETRY METAL-ION OTHERS         (A):  NULL ;  NULL ;  NULL
REMARK   3
REMARK   3  ISOTROPIC THERMAL FACTOR RESTRAINTS.     COUNT   RMS    WEIGHT
REMARK   3   MAIN-CHAIN BOND REFINED ATOMS  (A**2):  NULL ;  NULL ;  NULL
REMARK   3   MAIN-CHAIN BOND OTHER ATOMS    (A**2):  NULL ;  NULL ;  NULL
REMARK   3   MAIN-CHAIN ANGLE REFINED ATOMS (A**2):  NULL ;  NULL ;  NULL
REMARK   3   MAIN-CHAIN ANGLE OTHER ATOMS   (A**2):  NULL ;  NULL ;  NULL
REMARK   3   SIDE-CHAIN BOND REFINED ATOMS  (A**2):  NULL ;  NULL ;  NULL
REMARK   3   SIDE-CHAIN BOND OTHER ATOMS    (A**2):  NULL ;  NULL ;  NULL
REMARK   3   SIDE-CHAIN ANGLE REFINED ATOMS (A**2):  NULL ;  NULL ;  NULL
REMARK   3   SIDE-CHAIN ANGLE OTHER ATOMS   (A**2):  NULL ;  NULL ;  NULL
REMARK   3   LONG RANGE B REFINED ATOMS     (A**2):  NULL ;  NULL ;  NULL
REMARK   3   LONG RANGE B OTHER ATOMS       (A**2):  NULL ;  NULL ;  NULL
REMARK   3
REMARK   3 ANISOTROPIC THERMAL FACTOR RESTRAINTS.    COUNT   RMS   WEIGHT
REMARK   3   RIGID-BOND RESTRAINTS          (A**2):  1002 ; 4.594 ; 3.000
REMARK   3   SPHERICITY; FREE ATOMS         (A**2):     2 ; 8.840 ; 5.000
REMARK   3   SPHERICITY; BONDED ATOMS       (A**2):  1134 ;10.669 ; 5.000
REMARK   3
REMARK   3  NCS RESTRAINTS STATISTICS
REMARK   3   NUMBER OF DIFFERENT NCS GROUPS : NULL
REMARK   3
REMARK   3  TLS DETAILS
REMARK   3   NUMBER OF TLS GROUPS  : NULL
REMARK   3
REMARK   3  BULK SOLVENT MODELLING.
REMARK   3   METHOD USED : MASK
REMARK   3   PARAMETERS FOR MASK CALCULATION
REMARK   3   VDW PROBE RADIUS   : 1.20
REMARK   3   ION PROBE RADIUS   : 0.80
REMARK   3   SHRINKAGE RADIUS   : 0.80
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS: HYDROGENS HAVE BEEN ADDED IN THE RIDING
REMARK   3  POSITIONS
REMARK   4
REMARK   4 4I9V COMPLIES WITH FORMAT V. 3.30, 13-JUL-11
REMARK 100
REMARK 100 THIS ENTRY HAS BEEN PROCESSED BY RCSB ON 18-JAN-13.
REMARK 100 THE RCSB ID CODE IS RCSB076486.
REMARK 200
REMARK 200 EXPERIMENTAL DETAILS
REMARK 200  EXPERIMENT TYPE                : X-RAY DIFFRACTION
REMARK 200  DATE OF DATA COLLECTION        : 14-JUN-12
REMARK 200  TEMPERATURE           (KELVIN) : 100
REMARK 200  PH                             : 7.0
REMARK 200  NUMBER OF CRYSTALS USED        : 1
REMARK 200
```

```
REMARK 200  SYNCHROTRON              (Y/N) : Y
REMARK 200  RADIATION SOURCE               : APS
REMARK 200  BEAMLINE                       : 19-ID
REMARK 200  X-RAY GENERATOR MODEL          : NULL
REMARK 200  MONOCHROMATIC OR LAUE    (M/L) : M
REMARK 200  WAVELENGTH OR RANGE        (A) : 0.979
REMARK 200  MONOCHROMATOR                  : DOUBLE CRYSTAL
REMARK 200  OPTICS                         : MIRRORS
REMARK 200
REMARK 200  DETECTOR TYPE                  : CCD
REMARK 200  DETECTOR MANUFACTURER          : ADSC QUANTUM 315R
REMARK 200  INTENSITY-INTEGRATION SOFTWARE : HKL3000
REMARK 200  DATA SCALING SOFTWARE          : HKL3000
REMARK 200
REMARK 200  NUMBER OF UNIQUE REFLECTIONS   : 35974
REMARK 200  RESOLUTION RANGE HIGH      (A) : 1.020
REMARK 200  RESOLUTION RANGE LOW       (A) : 14.900
REMARK 200  REJECTION CRITERIA  (SIGMA(I)) : 2.000
REMARK 200
REMARK 200 OVERALL.
REMARK 200  COMPLETENESS FOR RANGE     (%) : 98.0
REMARK 200  DATA REDUNDANCY                : 8.800
REMARK 200  R MERGE                    (I) : 0.04300
REMARK 200  R SYM                      (I) : NULL
REMARK 200  <I/SIGMA(I)> FOR THE DATA SET  : 44.6700
REMARK 200
REMARK 200 IN THE HIGHEST RESOLUTION SHELL.
REMARK 200  HIGHEST RESOLUTION SHELL, RANGE HIGH (A) : 1.02
REMARK 200  HIGHEST RESOLUTION SHELL, RANGE LOW  (A) : 1.03
REMARK 200  COMPLETENESS FOR SHELL     (%) : 99.9
REMARK 200  DATA REDUNDANCY IN SHELL       : 6.50
REMARK 200  R MERGE FOR SHELL          (I) : 0.78000
REMARK 200  R SYM FOR SHELL            (I) : NULL
REMARK 200  <I/SIGMA(I)> FOR SHELL         : NULL
REMARK 200
REMARK 200 DIFFRACTION PROTOCOL: SINGLE WAVELENGTH
REMARK 200 METHOD USED TO DETERMINE THE STRUCTURE: MOLECULAR REPLACEMENT
REMARK 200 SOFTWARE USED: MOLREP
REMARK 200 STARTING MODEL: PDB ENTRY 436D
REMARK 200
REMARK 200 REMARK: NULL
REMARK 280
REMARK 280 CRYSTAL
REMARK 280 SOLVENT CONTENT, VS   (%): 46.63
REMARK 280 MATTHEWS COEFFICIENT, VM (ANGSTROMS**3/DA): 2.30
REMARK 280
REMARK 280 CRYSTALLIZATION CONDITIONS: 10% MPD,40MM CACODYLATE, 80MM NACL,12
REMARK 280  MM SPERMINE TETRACHLORIDE, 20MM MGCL2, PH 7.0, VAPOR DIFFUSION,
REMARK 280  HANGING DROP, TEMPERATURE 291K
REMARK 290
REMARK 290 CRYSTALLOGRAPHIC SYMMETRY
REMARK 290 SYMMETRY OPERATORS FOR SPACE GROUP: P 21 21 21
REMARK 290
REMARK 290      SYMOP   SYMMETRY
REMARK 290     NNNMMM   OPERATOR
REMARK 290       1555   X,Y,Z
```

```
REMARK 290       2555   -X+1/2,-Y,Z+1/2
REMARK 290       3555   -X,Y+1/2,-Z+1/2
REMARK 290       4555   X+1/2,-Y+1/2,-Z
REMARK 290
REMARK 290     WHERE NNN -> OPERATOR NUMBER
REMARK 290           MMM -> TRANSLATION VECTOR
REMARK 290
REMARK 290 CRYSTALLOGRAPHIC SYMMETRY TRANSFORMATIONS
REMARK 290 THE FOLLOWING TRANSFORMATIONS OPERATE ON THE ATOM/HETATM
REMARK 290 RECORDS IN THIS ENTRY TO PRODUCE CRYSTALLOGRAPHICALLY
REMARK 290 RELATED MOLECULES.
REMARK 290   SMTRY1   1  1.000000  0.000000  0.000000        0.00000
REMARK 290   SMTRY2   1  0.000000  1.000000  0.000000        0.00000
REMARK 290   SMTRY3   1  0.000000  0.000000  1.000000        0.00000
REMARK 290   SMTRY1   2 -1.000000  0.000000  0.000000       12.80300
REMARK 290   SMTRY2   2  0.000000 -1.000000  0.000000        0.00000
REMARK 290   SMTRY3   2  0.000000  0.000000  1.000000       32.15950
REMARK 290   SMTRY1   3 -1.000000  0.000000  0.000000        0.00000
REMARK 290   SMTRY2   3  0.000000  1.000000  0.000000       20.67200
REMARK 290   SMTRY3   3  0.000000  0.000000 -1.000000       32.15950
REMARK 290   SMTRY1   4  1.000000  0.000000  0.000000       12.80300
REMARK 290   SMTRY2   4  0.000000 -1.000000  0.000000       20.67200
REMARK 290   SMTRY3   4  0.000000  0.000000 -1.000000        0.00000
REMARK 290
REMARK 290 REMARK: NULL
REMARK 300
REMARK 300 BIOMOLECULE: 1
REMARK 300 SEE REMARK 350 FOR THE AUTHOR PROVIDED AND/OR PROGRAM
REMARK 300 GENERATED ASSEMBLY INFORMATION FOR THE STRUCTURE IN
REMARK 300 THIS ENTRY. THE REMARK MAY ALSO PROVIDE INFORMATION ON
REMARK 300 BURIED SURFACE AREA.
REMARK 350
REMARK 350 COORDINATES FOR A COMPLETE MULTIMER REPRESENTING THE KNOWN
REMARK 350 BIOLOGICALLY SIGNIFICANT OLIGOMERIZATION STATE OF THE
REMARK 350 MOLECULE CAN BE GENERATED BY APPLYING BIOMT TRANSFORMATIONS
REMARK 350 GIVEN BELOW.  BOTH NON-CRYSTALLOGRAPHIC AND
REMARK 350 CRYSTALLOGRAPHIC OPERATIONS ARE GIVEN.
REMARK 350
REMARK 350 BIOMOLECULE: 1
REMARK 350 AUTHOR DETERMINED BIOLOGICAL UNIT: DIMERIC
REMARK 350 SOFTWARE DETERMINED QUATERNARY STRUCTURE: DIMERIC
REMARK 350 SOFTWARE USED: PISA
REMARK 350 TOTAL BURIED SURFACE AREA: 1470 ANGSTROM**2
REMARK 350 SURFACE AREA OF THE COMPLEX: 5150 ANGSTROM**2
REMARK 350 CHANGE IN SOLVENT FREE ENERGY: -11.0 KCAL/MOL
REMARK 350 APPLY THE FOLLOWING TO CHAINS: A, B
REMARK 350   BIOMT1   1  1.000000  0.000000  0.000000        0.00000
REMARK 350   BIOMT2   1  0.000000  1.000000  0.000000        0.00000
REMARK 350   BIOMT3   1  0.000000  0.000000  1.000000        0.00000
REMARK 500
REMARK 500 GEOMETRY AND STEREOCHEMISTRY
REMARK 500 SUBTOPIC: COVALENT BOND ANGLES
REMARK 500
REMARK 500 THE STEREOCHEMICAL PARAMETERS OF THE FOLLOWING RESIDUES
REMARK 500 HAVE VALUES WHICH DEVIATE FROM EXPECTED VALUES BY MORE
REMARK 500 THAN 6*RMSD (M=MODEL NUMBER; RES=RESIDUE NAME; C=CHAIN
```

```
REMARK 500 IDENTIFIER; SSEQ=SEQUENCE NUMBER; I=INSERTION CODE).
REMARK 500
REMARK 500 STANDARD TABLE:
REMARK 500 FORMAT: (10X,I3,1X,A3,1X,A1,I4,A1,3(1X,A4,2X),12X,F5.1)
REMARK 500
REMARK 500 EXPECTED VALUES PROTEIN: ENGH AND HUBER, 1999
REMARK 500 EXPECTED VALUES NUCLEIC ACID: CLOWNEY ET AL 1996
REMARK 500
REMARK 500  M RES CSSEQI ATM1   ATM2   ATM3
REMARK 500     DG A  12   O5' -  P   -  OP2 ANGL. DEV. = -16.2 DEGREES
REMARK 500
REMARK 500 REMARK: NULL
REMARK 525
REMARK 525 SOLVENT
REMARK 525
REMARK 525 THE SOLVENT MOLECULES HAVE CHAIN IDENTIFIERS THAT
REMARK 525 INDICATE THE POLYMER CHAIN WITH WHICH THEY ARE MOST
REMARK 525 CLOSELY ASSOCIATED. THE REMARK LISTS ALL THE SOLVENT
REMARK 525 MOLECULES WHICH ARE MORE THAN 5A AWAY FROM THE
REMARK 525 NEAREST POLYMER CHAIN (M = MODEL NUMBER;
REMARK 525 RES=RESIDUE NAME; C=CHAIN IDENTIFIER; SSEQ=SEQUENCE
REMARK 525 NUMBER; I=INSERTION CODE):
REMARK 525
REMARK 525  M RES CSSEQI
REMARK 525    HOH B 283        DISTANCE =  6.25 ANGSTROMS
REMARK 525    HOH B 292        DISTANCE =  5.42 ANGSTROMS
REMARK 525    HOH B 293        DISTANCE =  6.88 ANGSTROMS
REMARK 610
REMARK 610 MISSING HETEROATOM
REMARK 610 THE FOLLOWING RESIDUES HAVE MISSING ATOMS (M=MODEL NUMBER;
REMARK 610 RES=RESIDUE NAME; C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER;
REMARK 610 I=INSERTION CODE):
REMARK 610   M RES C SSEQI
REMARK 610     SPK B  101
REMARK 615
REMARK 615 ZERO OCCUPANCY ATOM
REMARK 615 THE FOLLOWING RESIDUES HAVE ATOMS MODELED WITH ZERO
REMARK 615 OCCUPANCY. THE LOCATION AND PROPERTIES OF THESE ATOMS
REMARK 615 MAY NOT BE RELIABLE. (M=MODEL NUMBER; RES=RESIDUE NAME;
REMARK 615 C=CHAIN IDENTIFIER; SSEQ=SEQUENCE NUMBER; I=INSERTION CODE):
REMARK 615   M RES C  SSEQI
REMARK 615     HOH B   285
REMARK 620
REMARK 620 METAL COORDINATION
REMARK 620 (M=MODEL NUMBER; RES=RESIDUE NAME; C=CHAIN IDENTIFIER;
REMARK 620 SSEQ=SEQUENCE NUMBER; I=INSERTION CODE):
REMARK 620
REMARK 620 COORDINATION ANGLES FOR:  M RES CSSEQI METAL
REMARK 620                              MG A 101  MG
REMARK 620 N RES CSSEQI ATOM
REMARK 620 1 HOH B 270   O
REMARK 620 2 HOH A 208   O    93.2
REMARK 620 3 HOH A 249   O    89.9  93.8
REMARK 620 4 HOH B 206   O   179.8  86.7  90.0
REMARK 620 5 HOH B 261   O    88.8  86.2 178.7  91.4
REMARK 620 6 HOH B 262   O    92.3 173.7  89.2  87.9  90.9
```

```
REMARK 620 N                    1     2     3     4     5
REMARK 800
REMARK 800 SITE
REMARK 800 SITE_IDENTIFIER: AC1
REMARK 800 EVIDENCE_CODE: SOFTWARE
REMARK 800 SITE_DESCRIPTION: BINDING SITE FOR RESIDUE MG A 101
REMARK 800
REMARK 800 SITE_IDENTIFIER: AC2
REMARK 800 EVIDENCE_CODE: SOFTWARE
REMARK 800 SITE_DESCRIPTION: BINDING SITE FOR RESIDUE SPK B 101
REMARK 800
REMARK 800 SITE_IDENTIFIER: AC3
REMARK 800 EVIDENCE_CODE: SOFTWARE
REMARK 800 SITE_DESCRIPTION: BINDING SITE FOR RESIDUE SPK B 102
REMARK 900
REMARK 900 RELATED ENTRIES
REMARK 900 RELATED ID: 355D   RELATED DB: PDB
DBREF  4I9V A    1    12  PDB    4I9V     4I9V             1     12
DBREF  4I9V B   13    24  PDB    4I9V     4I9V            13     24
SEQRES   1 A   12   DC  DG  DC  DG  DA  DA  DT  DT 5HC  DG  DC  DG
SEQRES   1 B   12   DC  DG  DC  DG  DA  DA  DT  DT 5HC  DG  DC  DG
MODRES 4I9V 5HC A    9   DC
MODRES 4I9V 5HC B   21   DC
HET    5HC  A   9      23
HET    5HC  B  21      21
HET     MG  A 101       1
HET    SPK  B 101      11
HET    SPK  B 102      28
HETNAM     5HC 2'-DEOXY-5-(HYDROXYMETHYL)CYTIDINE 5'-(DIHYDROGEN
HETNAM   2 5HC  PHOSPHATE)
HETNAM      MG MAGNESIUM ION
HETNAM     SPK SPERMINE (FULLY PROTONATED FORM)
FORMUL   1  5HC    2(C10 H16 N3 O8 P)
FORMUL   3   MG    MG 2+
FORMUL   4  SPK    2(C10 H30 N4 4+)
FORMUL   6  HOH   *187(H2 O)
LINK         O3'  DT A   8                 P   5HC A   9     1555   1555  1.59
LINK         O3' 5HC A   9                 P    DG A  10     1555   1555  1.53
LINK         O3'  DT B  20                 P   5HC B  21     1555   1555  1.57
LINK         O3' 5HC B  21                 P    DG B  22     1555   1555  1.54
LINK        MG    MG A 101                 O   HOH B 270     1555   1555  2.04
LINK        MG    MG A 101                 O   HOH A 208     1555   1555  2.04
LINK        MG    MG A 101                 O   HOH A 249     1555   1555  2.08
LINK        MG    MG A 101                 O   HOH B 206     1555   1555  2.08
LINK        MG    MG A 101                 O   HOH B 261     1555   1555  2.09
LINK        MG    MG A 101                 O   HOH B 262     1555   1555  2.09
SITE     1 AC1  6 HOH A 208  HOH A 249  HOH B 206  HOH B 261
SITE     2 AC1  6 HOH B 262  HOH B 270
SITE     1 AC2  7  DG A  12  HOH A 264  HOH A 271   DA B  18
SITE     2 AC2  7  DT B  19   DT B  20  HOH B 250
SITE     1 AC3  8 5HC A   9   DG A  10  HOH A 231   DG B  14
SITE     2 AC3  8  DC B  15   DC B  23  HOH B 219  HOH B 257
CRYST1   25.606   41.344   64.319  90.00  90.00  90.00 P 21 21 21    8
ORIGX1      1.000000  0.000000  0.000000        0.00000
ORIGX2      0.000000  1.000000  0.000000        0.00000
ORIGX3      0.000000  0.000000  1.000000        0.00000
```

```
SCALE1      0.039053  0.000000  0.000000        0.00000
SCALE2      0.000000  0.024187  0.000000        0.00000
SCALE3      0.000000  0.000000  0.015548        0.00000
ATOM      1  O5'  DC A   1      -8.013  15.825 -22.967  1.00 23.56           O
ANISOU    1  O5'  DC A   1     2844   1616   4488   -240   -258    343       O
ATOM      2  C5'  DC A   1      -7.285  14.638 -23.310  1.00 20.62           C
ANISOU    2  C5'  DC A   1     2631   1722   3481   -370   -244    141       C
ATOM      3  C4'  DC A   1      -8.198  13.451 -23.135  1.00 20.18           C
ANISOU    3  C4'  DC A   1     2716   1507   3442   -207   -150     -8       C
ATOM      4  O4'  DC A   1      -7.529  12.266 -23.624  1.00 20.51           O
ANISOU    4  O4'  DC A   1     3049   1590   3154   -131   -204    173       O
ATOM      5  C3'  DC A   1      -8.576  13.131 -21.695  1.00 21.54           C
ANISOU    5  C3'  DC A   1     2907   1690   3586    -95     10     73       C
ATOM      6  O3'  DC A   1      -9.964  12.764 -21.710  1.00 23.86           O
ANISOU    6  O3'  DC A   1     2883   2083   4096    -18     88    201       O
ATOM      7  C2'  DC A   1      -7.599  12.032 -21.314  1.00 20.54           C
ANISOU    7  C2'  DC A   1     3130   1586   3086    -78    -26    -42       C
ATOM      8  C1'  DC A   1      -7.444  11.277 -22.613  1.00 19.55           C
ANISOU    8  C1'  DC A   1     2929   1498   3000    -76     -4     48       C
ATOM      9  N1   DC A   1      -6.181  10.566 -22.816  1.00 18.68           N
ANISOU    9  N1   DC A   1     2789   1439   2867   -198   -219     14       N
ATOM     10  C2   DC A   1      -6.211   9.314 -23.439  1.00 18.66           C
ANISOU   10  C2   DC A   1     2736   1438   2916   -397   -105      7       C
ATOM     11  O2   DC A   1      -7.306   8.809 -23.716  1.00 20.32           O
ANISOU   11  O2   DC A   1     2650   1535   3535   -261     29   -236       O
ATOM     12  N3   DC A   1      -5.046   8.678 -23.701  1.00 18.09           N
ANISOU   12  N3   DC A   1     2592   1496   2782   -369    -83      4       N
ATOM     13  C4   DC A   1      -3.887   9.264 -23.398  1.00 18.74           C
ANISOU   13  C4   DC A   1     2478   1618   3024   -353   -272    166       C
ATOM     14  N4   DC A   1      -2.762   8.606 -23.681  1.00 19.80           N
ANISOU   14  N4   DC A   1     2437   1794   3291   -381   -281     67       N
ATOM     15  C5   DC A   1      -3.828  10.553 -22.795  1.00 19.77           C
ANISOU   15  C5   DC A   1     2549   1772   3188   -365   -303    -11       C
ATOM     16  C6   DC A   1      -4.987  11.167 -22.537  1.00 19.55           C
ANISOU   16  C6   DC A   1     2729   1588   3109   -203   -303     35       C
ATOM     17  P    DG A   2     -10.765  12.548 -20.383  1.00 26.04           P
ANISOU   17  P    DG A   2     3323   2376   4194     58    480    291       P
ATOM     18  OP1  DG A   2     -12.130  12.923 -20.616  1.00 29.73           O
ANISOU   18  OP1  DG A   2     3423   2872   5002    582    819    579       O
ATOM     19  OP2  DG A   2     -10.071  13.067 -19.272  1.00 29.00           O
ANISOU   19  OP2  DG A   2     3850   3431   3737   -619    532     18       O
ATOM     20  O5'  DG A   2     -10.673  10.977 -20.163  1.00 23.68           O
ANISOU   20  O5'  DG A   2     2834   2190   3972   -249    301    204       O
ATOM     21  C5'  DG A   2     -11.428  10.127 -21.004  1.00 22.65           C
ANISOU   21  C5'  DG A   2     2579   2588   3437     95    246     80       C
ATOM     22  C4'  DG A   2     -11.239   8.706 -20.550  1.00 21.99           C
ANISOU   22  C4'  DG A   2     2253   2490   3611    -64    224    175       C
ATOM     23  O4'  DG A   2      -9.897   8.270 -20.844  1.00 24.02           O
ANISOU   23  O4'  DG A   2     2351   3330   3445    116    365    -66       O
ATOM     24  C3'  DG A   2     -11.452   8.467 -19.060  1.00 21.91           C
ANISOU   24  C3'  DG A   2     2809   1787   3725    124    597    386       C
ATOM     25  O3'  DG A   2     -12.228   7.270 -19.048  1.00 29.79           O
ANISOU   25  O3'  DG A   2     4090   2395   4831   -635   1827   -166       O
ATOM     26  C2'  DG A   2     -10.039   8.294 -18.532  1.00 23.04           C
ANISOU   26  C2'  DG A   2     2865   2403   3484    717    674    412       C
ATOM     27  C1'  DG A   2      -9.333   7.669 -19.716  1.00 21.83           C
```

```
ANISOU   27  C1'  DG A   2     2651   2196   3444     64    670     29       C
ATOM     28  N9   DG A   2      -7.886   7.824 -19.805  1.00 19.47           N
ANISOU   28  N9   DG A   2     2494   1812   3088    481    299    129       N
ATOM     29  C8   DG A   2      -7.132   8.909 -19.442  1.00 19.61           C
ANISOU   29  C8   DG A   2     2708   1725   3017    309    262    187       C
ATOM     30  N7   DG A   2      -5.865   8.760 -19.703  1.00 18.90           N
ANISOU   30  N7   DG A   2     2713   1537   2929    313    193    178       N
ATOM     31  C5   DG A   2      -5.778   7.502 -20.278  1.00 17.60           C
ANISOU   31  C5   DG A   2     2289   1572   2825    209     -6    156       C
ATOM     32  C6   DG A   2      -4.651   6.782 -20.753  1.00 16.44           C
ANISOU   32  C6   DG A   2     2246   1491   2507    216    -71     89       C
ATOM     33  O6   DG A   2      -3.481   7.125 -20.773  1.00 17.75           O
ANISOU   33  O6   DG A   2     2223   1701   2818     63    -49      1       O
ATOM     34  N1   DG A   2      -5.016   5.544 -21.263  1.00 16.93           N
ANISOU   34  N1   DG A   2     2072   1788   2571    233    -82     76       N
ATOM     35  C2   DG A   2      -6.291   5.053 -21.295  1.00 16.68           C
ANISOU   35  C2   DG A   2     2139   1642   2554    257    -73    100       C
ATOM     36  N2   DG A   2      -6.438   3.840 -21.820  1.00 17.23           N
ANISOU   36  N2   DG A   2     1966   1849   2730     55    -53     -1       N
ATOM     37  N3   DG A   2      -7.348   5.714 -20.863  1.00 17.69           N
ANISOU   37  N3   DG A   2     2120   1851   2748    288     57    226       N
ATOM     38  C4   DG A   2      -7.020   6.923 -20.371  1.00 17.73           C
ANISOU   38  C4   DG A   2     2405   1558   2772    248    179    140       C
ATOM     39  P    DC A   3     -13.104   6.799 -17.821  1.00 25.26           P
ANISOU   39  P    DC A   3     3187   2125   4283    299   1239    301       P
ATOM     40  OP1  DC A   3     -14.465   6.749 -18.257  1.00 35.26           O
ANISOU   40  OP1  DC A   3     3301   4648   5446   1453   1135   1252       O
ATOM     41  OP2  DC A   3     -12.786   7.522 -16.667  1.00 32.75           O
ANISOU   41  OP2  DC A   3     5338   2468   4638    760   1702   -146       O
ATOM     42  O5'  DC A   3     -12.717   5.272 -17.713  1.00 22.98           O
ANISOU   42  O5'  DC A   3     2926   2212   3594    546    276     59       O
ATOM     43  C5'  DC A   3     -12.462   4.488 -18.858  1.00 19.40           C
ANISOU   43  C5'  DC A   3     2142   1922   3306    128   -157    139       C
ATOM     44  C4'  DC A   3     -11.304   3.565 -18.568  1.00 18.42           C
ANISOU   44  C4'  DC A   3     2240   1849   2907    129    -95     32       C
ATOM     45  O4'  DC A   3     -10.083   4.304 -18.569  1.00 18.58           O
ANISOU   45  O4'  DC A   3     1974   2115   2967    138   -104    163       O
ATOM     46  C3'  DC A   3     -11.339   2.846 -17.219  1.00 19.10           C
ANISOU   46  C3'  DC A   3     2263   2096   2897     76     14     26       C
ATOM     47  O3'  DC A   3     -11.831   1.540 -17.514  1.00 20.43           O
ANISOU   47  O3'  DC A   3     2189   2118   3455    -43   -164    266       O
ATOM     48  C2'  DC A   3      -9.894   2.888 -16.733  1.00 19.14           C
ANISOU   48  C2'  DC A   3     2240   2094   2937    -36     -5    240       C
ATOM     49  C1'  DC A   3      -9.117   3.539 -17.867  1.00 18.51           C
ANISOU   49  C1'  DC A   3     2096   2023   2913    104    -74     90       C
ATOM     50  N1   DC A   3      -8.047   4.456 -17.482  1.00 17.85           N
ANISOU   50  N1   DC A   3     2124   1906   2752    129   -129     61       N
ATOM     51  C2   DC A   3      -6.730   4.200 -17.880  1.00 17.27           C
ANISOU   51  C2   DC A   3     2255   1674   2630    257    -26    152       C
ATOM     52  O2   DC A   3      -6.472   3.143 -18.456  1.00 17.85           O
ANISOU   52  O2   DC A   3     2160   1840   2781    228   -168      0       O
ATOM     53  N3   DC A   3      -5.771   5.111 -17.611  1.00 17.19           N
ANISOU   53  N3   DC A   3     2113   1740   2679    197    -35    155       N
ATOM     54  C4   DC A   3      -6.087   6.243 -16.986  1.00 17.32           C
ANISOU   54  C4   DC A   3     2242   1700   2637    227    -43    161       C
ATOM     55  N4   DC A   3      -5.113   7.114 -16.743  1.00 17.71           N
```

```
ANISOU   55  N4   DC A   3     2282   1741   2703    151    -64    178       N
ATOM     56  C5   DC A   3      -7.419   6.533 -16.581  1.00 18.69           C
ANISOU   56  C5   DC A   3     2217   2088   2795    249    -27    -70       C
ATOM     57  C6   DC A   3      -8.360   5.628 -16.858  1.00 18.34           C
ANISOU   57  C6   DC A   3     2227   1977   2763    334   -122     14       C
ATOM     58  P    DG A   4     -12.189   0.490 -16.381  1.00 21.70           P
ANISOU   58  P    DG A   4     2141   2288   3814     62     78    389       P
ATOM     59  OP1  DG A   4     -13.100  -0.475 -16.951  1.00 25.45           O
ANISOU   59  OP1  DG A   4     2312   2407   4948   -268   -163    466       O
ATOM     60  OP2  DG A   4     -12.537   1.163 -15.199  1.00 23.67           O
ANISOU   60  OP2  DG A   4     2544   2838   3612    325    472    565       O
ATOM     61  O5'  DG A   4     -10.814  -0.234 -16.073  1.00 19.80           O
ANISOU   61  O5'  DG A   4     2078   2312   3132     26      0    292       O
ATOM     62  C5'  DG A   4     -10.138  -0.970 -17.073  1.00 19.65           C
ANISOU   62  C5'  DG A   4     2355   2139   2972   -104    -51    176       C
ATOM     63  C4'  DG A   4      -8.743  -1.279 -16.592  1.00 19.72           C
ANISOU   63  C4'  DG A   4     2463   1904   3125      9   -163     58       C
ATOM     64  O4'  DG A   4      -7.975  -0.060 -16.507  1.00 18.58           O
ANISOU   64  O4'  DG A   4     2216   1911   2930     16   -128     77       O
ATOM     65  C3'  DG A   4      -8.683  -1.903 -15.197  1.00 19.70           C
ANISOU   65  C3'  DG A   4     2537   1747   3202     46    -21    177       C
ATOM     66  O3'  DG A   4      -7.591  -2.822 -15.160  1.00 21.01           O
ANISOU   66  O3'  DG A   4     2506   1870   3606     86   -189    230       O
ATOM     67  C2'  DG A   4      -8.301  -0.724 -14.327  1.00 18.78           C
ANISOU   67  C2'  DG A   4     2183   1875   3075    -43   -138    205       C
ATOM     68  C1'  DG A   4      -7.337  -0.055 -15.264  1.00 18.13           C
ANISOU   68  C1'  DG A   4     2208   1817   2863     18    -13     59       C
ATOM     69  N9   DG A   4      -6.944   1.307 -14.942  1.00 17.98           N
ANISOU   69  N9   DG A   4     2249   1803   2780    117   -130    127       N
ATOM     70  C8   DG A   4      -7.671   2.279 -14.307  1.00 18.66           C
ANISOU   70  C8   DG A   4     2408   1868   2814    259   -120    109       C
ATOM     71  N7   DG A   4      -6.999   3.380 -14.123  1.00 18.42           N
ANISOU   71  N7   DG A   4     2272   1880   2843    269    -35    154       N
ATOM     72  C5   DG A   4      -5.763   3.128 -14.699  1.00 17.78           C
ANISOU   72  C5   DG A   4     2275   1757   2723    205   -135     90       C
ATOM     73  C6   DG A   4      -4.601   3.937 -14.783  1.00 17.31           C
ANISOU   73  C6   DG A   4     2226   1716   2633    176   -164    139       C
ATOM     74  O6   DG A   4      -4.447   5.099 -14.406  1.00 18.39           O
ANISOU   74  O6   DG A   4     2392   1575   3020    102   -159    111       O
ATOM     75  N1   DG A   4      -3.554   3.265 -15.404  1.00 17.25           N
ANISOU   75  N1   DG A   4     2251   1687   2617    101    -88    233       N
ATOM     76  C2   DG A   4      -3.613   1.979 -15.874  1.00 17.22           C
ANISOU   76  C2   DG A   4     2072   1835   2636    109   -150    156       C
ATOM     77  N2   DG A   4      -2.506   1.506 -16.443  1.00 17.77           N
ANISOU   77  N2   DG A   4     2308   1842   2600    260   -115     73       N
ATOM     78  N3   DG A   4      -4.684   1.216 -15.793  1.00 17.40           N
ANISOU   78  N3   DG A   4     2238   1683   2688    180   -108    117       N
ATOM     79  C4   DG A   4      -5.708   1.843 -15.185  1.00 17.27           C
ANISOU   79  C4   DG A   4     2265   1828   2467     82    -86    123       C
ATOM     80  P    DA A   5      -7.823  -4.354 -15.073  1.00 26.07           P
ANISOU   80  P    DA A   5     2700   1805   5400   -215   -281    202       P
ATOM     81  OP1  DA A   5      -8.585  -4.719 -16.242  1.00 34.24           O
ANISOU   81  OP1  DA A   5     3082   2918   7009   -320   -912  -1342       O
ATOM     82  OP2  DA A   5      -8.290  -4.567 -13.720  1.00 31.70           O
ANISOU   82  OP2  DA A   5     3493   3237   5314    714     88   1091       O
ATOM     83  O5'  DA A   5      -6.355  -4.926 -15.183  1.00 25.02           O
```

```
ANISOU   83  O5'  DA A   5     2898   1758   4850    161   -844     76       O
ATOM     84  C5'  DA A   5      -5.666  -4.872 -16.400  1.00 25.58           C
ANISOU   84  C5'  DA A   5     2828   2432   4458    326   -868   -541       C
ATOM     85  C4'  DA A   5      -4.189  -4.656 -16.171  1.00 21.18           C
ANISOU   85  C4'  DA A   5     2686   1952   3407    429   -622   -666       C
ATOM     86  O4'  DA A   5      -3.928  -3.295 -15.744  1.00 21.12           O
ANISOU   86  O4'  DA A   5     2938   1786   3298    383   -703   -353       O
ATOM     87  C3'  DA A   5      -3.541  -5.550 -15.118  1.00 22.28           C
ANISOU   87  C3'  DA A   5     2746   1758   3960    554   -845   -378       C
ATOM     88  O3'  DA A   5      -2.273  -5.885 -15.698  1.00 22.49           O
ANISOU   88  O3'  DA A   5     2905   1798   3842    404   -716   -432       O
ATOM     89  C2'  DA A   5      -3.487  -4.650 -13.895  1.00 19.79           C
ANISOU   89  C2'  DA A   5     2573   1456   3490     45   -505   -107       C
ATOM     90  C1'  DA A   5      -3.224  -3.293 -14.513  1.00 18.90           C
ANISOU   90  C1'  DA A   5     2494   1453   3232    214   -533   -253       C
ATOM     91  N9   DA A   5      -3.680  -2.136 -13.746  1.00 18.25           N
ANISOU   91  N9   DA A   5     2540   1495   2899     49   -403    -76       N
ATOM     92  C8   DA A   5      -4.920  -1.933 -13.205  1.00 19.12           C
ANISOU   92  C8   DA A   5     2614   1682   2967    -35   -321    -63       C
ATOM     93  N7   DA A   5      -5.068  -0.760 -12.643  1.00 18.43           N
ANISOU   93  N7   DA A   5     2467   1622   2911     17   -253     76       N
ATOM     94  C5   DA A   5      -3.838  -0.149 -12.821  1.00 17.03           C
ANISOU   94  C5   DA A   5     2363   1552   2554     82   -225     99       C
ATOM     95  C6   DA A   5      -3.355   1.123 -12.469  1.00 16.52           C
ANISOU   95  C6   DA A   5     2301   1472   2502    192   -241     59       C
ATOM     96  N6   DA A   5      -4.076   2.035 -11.812  1.00 16.72           N
ANISOU   96  N6   DA A   5     2257   1521   2574    104    -84    103       N
ATOM     97  N1   DA A   5      -2.078   1.413 -12.785  1.00 16.93           N
ANISOU   97  N1   DA A   5     2375   1477   2578     85   -115    112       N
ATOM     98  C2   DA A   5      -1.362   0.510 -13.458  1.00 17.09           C
ANISOU   98  C2   DA A   5     2346   1448   2698    161   -215     42       C
ATOM     99  N3   DA A   5      -1.709  -0.709 -13.859  1.00 17.55           N
ANISOU   99  N3   DA A   5     2484   1467   2715    217   -246     19       N
ATOM    100  C4   DA A   5      -2.975  -0.981 -13.508  1.00 17.26           C
ANISOU  100  C4   DA A   5     2446   1501   2611    121   -322     68       C
ATOM    101  P    DA A   6      -1.195  -6.801 -14.946  1.00 21.29           P
ANISOU  101  P    DA A   6     2864   1740   3485    289   -432    -76       P
ATOM    102  OP1  DA A   6      -0.352  -7.448 -15.980  1.00 22.29           O
ANISOU  102  OP1  DA A   6     2900   1976   3590    386   -379   -195       O
ATOM    103  OP2  DA A   6      -1.891  -7.542 -13.910  1.00 19.81           O
ANISOU  103  OP2  DA A   6     2433   1867   3227    -54   -414     44       O
ATOM    104  O5'  DA A   6      -0.314  -5.761 -14.155  1.00 20.69           O
ANISOU  104  O5'  DA A   6     2622   1862   3375    307   -645   -251       O
ATOM    105  C5'  DA A   6       0.604  -4.951 -14.874  1.00 21.58           C
ANISOU  105  C5'  DA A   6     2857   1900   3440    213   -357   -276       C
ATOM    106  C4'  DA A   6       1.423  -4.227 -13.844  1.00 19.89           C
ANISOU  106  C4'  DA A   6     2763   1941   2853    603   -317   -260       C
ATOM    107  O4'  DA A   6       0.570  -3.264 -13.199  1.00 20.52           O
ANISOU  107  O4'  DA A   6     2989   1949   2859    851   -198   -259       O
ATOM    108  C3'  DA A   6       1.994  -5.115 -12.732  1.00 19.27           C
ANISOU  108  C3'  DA A   6     2601   1821   2897    389   -337   -250       C
ATOM    109  O3'  DA A   6       3.395  -4.886 -12.801  1.00 19.32           O
ANISOU  109  O3'  DA A   6     2550   1974   2815    427   -148   -182       O
ATOM    110  C2'  DA A   6       1.321  -4.617 -11.461  1.00 17.99           C
ANISOU  110  C2'  DA A   6     2221   1647   2967    284   -245   -143       C
ATOM    111  C1'  DA A   6       0.888  -3.212 -11.827  1.00 17.77           C
```

```
ANISOU  111  C1'  DA A   6     2361   1681   2710    358   -182   -160       C
ATOM    112  N9   DA A   6      -0.299  -2.724 -11.152  1.00 16.84           N
ANISOU  112  N9   DA A   6     2298   1458   2640    197   -101    -30       N
ATOM    113  C8   DA A   6      -1.503  -3.367 -11.041  1.00 17.41           C
ANISOU  113  C8   DA A   6     2179   1689   2746    148   -267    -97       C
ATOM    114  N7   DA A   6      -2.446  -2.635 -10.504  1.00 16.93           N
ANISOU  114  N7   DA A   6     2321   1397   2713    131   -201    -48       N
ATOM    115  C5   DA A   6      -1.816  -1.428 -10.230  1.00 16.24           C
ANISOU  115  C5   DA A   6     2328   1373   2469     65    -90     63       C
ATOM    116  C6   DA A   6      -2.277  -0.221  -9.670  1.00 15.50           C
ANISOU  116  C6   DA A   6     2180   1324   2381    107   -181    125       C
ATOM    117  N6   DA A   6      -3.524  -0.031  -9.236  1.00 16.31           N
ANISOU  117  N6   DA A   6     2151   1475   2569    123    -91      6       N
ATOM    118  N1   DA A   6      -1.388   0.784  -9.536  1.00 16.10           N
ANISOU  118  N1   DA A   6     2207   1375   2535    106    -89    108       N
ATOM    119  C2   DA A   6      -0.138   0.603  -9.981  1.00 16.39           C
ANISOU  119  C2   DA A   6     2171   1451   2606    173   -144     98       C
ATOM    120  N3   DA A   6       0.407  -0.473 -10.542  1.00 16.61           N
ANISOU  120  N3   DA A   6     2295   1376   2638    182    -35     86       N
ATOM    121  C4   DA A   6      -0.498  -1.462 -10.647  1.00 16.35           C
ANISOU  121  C4   DA A   6     2337   1277   2595    142   -161     63       C
ATOM    122  P    DT A   7       4.430  -5.497 -11.750  1.00 18.20           P
ANISOU  122  P    DT A   7     2379   1638   2895    363   -152   -214       P
ATOM    123  OP1  DT A   7       5.718  -5.600 -12.454  1.00 20.55           O
ANISOU  123  OP1  DT A   7     2628   1954   3224    279    -51   -356       O
ATOM    124  OP2  DT A   7       3.886  -6.716 -11.089  1.00 18.55           O
ANISOU  124  OP2  DT A   7     2459   1509   3081    190   -349    -31       O
ATOM    125  O5'  DT A   7       4.517  -4.365 -10.642  1.00 17.72           O
ANISOU  125  O5'  DT A   7     2370   1501   2861    222    -90   -204       O
ATOM    126  C5'  DT A   7       4.940  -3.056 -11.019  1.00 17.94           C
ANISOU  126  C5'  DT A   7     2331   1632   2852    146    -12   -143       C
ATOM    127  C4'  DT A   7       4.733  -2.109  -9.863  1.00 18.23           C
ANISOU  127  C4'  DT A   7     2338   1624   2962    218     34   -278       C
ATOM    128  O4'  DT A   7       3.329  -1.995  -9.571  1.00 17.48           O
ANISOU  128  O4'  DT A   7     2323   1632   2685    250     13    -20       O
ATOM    129  C3'  DT A   7       5.410  -2.537  -8.557  1.00 18.60           C
ANISOU  129  C3'  DT A   7     2443   1673   2948    156     32   -264       C
ATOM    130  O3'  DT A   7       6.509  -1.634  -8.378  1.00 20.96           O
ANISOU  130  O3'  DT A   7     2371   2129   3463    185     17   -560       O
ATOM    131  C2'  DT A   7       4.302  -2.460  -7.519  1.00 18.04           C
ANISOU  131  C2'  DT A   7     2277   1750   2825     92   -144   -127       C
ATOM    132  C1'  DT A   7       3.236  -1.629  -8.208  1.00 17.24           C
ANISOU  132  C1'  DT A   7     2385   1470   2693    231      5     -9       C
ATOM    133  N1   DT A   7       1.850  -1.858  -7.782  1.00 16.44           N
ANISOU  133  N1   DT A   7     2416   1232   2596    137    -25    -24       N
ATOM    134  C2   DT A   7       1.145  -0.804  -7.243  1.00 16.52           C
ANISOU  134  C2   DT A   7     2376   1324   2576    257    -38    103       C
ATOM    135  O2   DT A   7       1.637   0.291  -7.027  1.00 17.09           O
ANISOU  135  O2   DT A   7     2339   1356   2798    195     31     -9       O
ATOM    136  N3   DT A   7      -0.160  -1.084  -6.938  1.00 16.15           N
ANISOU  136  N3   DT A   7     2280   1299   2556    246    -52     61       N
ATOM    137  C4   DT A   7      -0.814  -2.289  -7.100  1.00 16.39           C
ANISOU  137  C4   DT A   7     2333   1345   2549    103    -54    202       C
ATOM    138  O4   DT A   7      -2.004  -2.377  -6.811  1.00 17.18           O
ANISOU  138  O4   DT A   7     2312   1492   2722    126    -20    127       O
ATOM    139  C5   DT A   7      -0.008  -3.366  -7.641  1.00 16.66           C
```

```
ANISOU  139  C5   DT A   7     2439   1295   2593     83   -151    145       C
ATOM    140  C7   DT A   7      -0.626  -4.717  -7.832  1.00 17.97           C
ANISOU  140  C7   DT A   7     2539   1425   2864     12    -33    -64       C
ATOM    141  C6   DT A   7       1.263  -3.096  -7.963  1.00 16.58           C
ANISOU  141  C6   DT A   7     2381   1343   2573    144   -230     23       C
ATOM    142  P    DT A   8       7.404  -1.655  -7.058  1.00 23.59           P
ANISOU  142  P    DT A   8     2520   2276   4165    600   -618   -704       P
ATOM    143  OP1  DT A   8       8.716  -1.057  -7.396  1.00 28.71           O
ANISOU  143  OP1  DT A   8     2039   3643   5225    714   -261  -1370       O
ATOM    144  OP2  DT A   8       7.350  -2.994  -6.446  1.00 27.86           O
ANISOU  144  OP2  DT A   8     3104   2502   4980    856  -1331   -168       O
ATOM    145  O5'  DT A   8       6.646  -0.681  -6.072  1.00 20.37           O
ANISOU  145  O5'  DT A   8     2633   1887   3216    -80   -275   -232       O
ATOM    146  C5'  DT A   8       6.439   0.676  -6.459  1.00 19.53           C
ANISOU  146  C5'  DT A   8     2383   1973   3063      6    152   -204       C
ATOM    147  C4'  DT A   8       5.680   1.380  -5.362  1.00 20.08           C
ANISOU  147  C4'  DT A   8     2200   2310   3117    -54    180   -294       C
ATOM    148  O4'  DT A   8       4.354   0.825  -5.306  1.00 20.46           O
ANISOU  148  O4'  DT A   8     2191   2642   2941   -131    -43    -96       O
ATOM    149  C3'  DT A   8       6.267   1.220  -3.954  1.00 22.98           C
ANISOU  149  C3'  DT A   8     2364   3222   3145     50    -43   -773       C
ATOM    150  O3'  DT A   8       6.556   2.525  -3.484  1.00 26.53           O
ANISOU  150  O3'  DT A   8     2665   3750   3666  -1024    321  -1190       O
ATOM    151  C2'  DT A   8       5.164   0.558  -3.145  1.00 20.72           C
ANISOU  151  C2'  DT A   8     2509   2420   2944    469   -192   -281       C
ATOM    152  C1'  DT A   8       3.924   0.828  -3.960  1.00 19.02           C
ANISOU  152  C1'  DT A   8     2187   2171   2868    131     -5    -23       C
ATOM    153  N1   DT A   8       2.862  -0.168  -3.853  1.00 17.78           N
ANISOU  153  N1   DT A   8     2332   1678   2743    206     70     88       N
ATOM    154  C2   DT A   8       1.612   0.244  -3.454  1.00 17.15           C
ANISOU  154  C2   DT A   8     2260   1589   2665    154     57     45       C
ATOM    155  O2   DT A   8       1.362   1.387  -3.105  1.00 17.74           O
ANISOU  155  O2   DT A   8     2250   1559   2930    123     58     21       O
ATOM    156  N3   DT A   8       0.659  -0.738  -3.465  1.00 17.07           N
ANISOU  156  N3   DT A   8     2376   1457   2652    135     66      1       N
ATOM    157  C4   DT A   8       0.827  -2.062  -3.822  1.00 17.23           C
ANISOU  157  C4   DT A   8     2525   1496   2523    244    -42    104       C
ATOM    158  O4   DT A   8      -0.132  -2.822  -3.797  1.00 18.34           O
ANISOU  158  O4   DT A   8     2732   1476   2758    119     90      3       O
ATOM    159  C5   DT A   8       2.169  -2.433  -4.217  1.00 17.88           C
ANISOU  159  C5   DT A   8     2594   1615   2582    406    -27    187       C
ATOM    160  C7   DT A   8       2.452  -3.851  -4.602  1.00 19.45           C
ANISOU  160  C7   DT A   8     2818   1653   2917    457    127    104       C
ATOM    161  C6   DT A   8       3.106  -1.476  -4.220  1.00 18.15           C
ANISOU  161  C6   DT A   8     2474   1809   2611    411    -39    -16       C
HETATM  162  P   5HC A   9       7.082   2.882  -2.025  1.00 26.54           P
ANISOU  162  P   5HC A   9     2766   3801   3515   -219    -15   -949       P
HETATM  163  OP1 5HC A   9       7.981   4.304  -2.363  1.00 32.38           O
ANISOU  163  OP1 5HC A   9     2654   5035   4612   -898   -413  -1226       O
HETATM  164  OP2 5HC A   9       7.671   1.403  -1.322  1.00 29.03           O
ANISOU  164  OP2 5HC A   9     3330   4006   3691   -151   -424   -596       O
HETATM  165  O5' 5HC A   9       5.662   3.284  -1.320  1.00 24.30           O
ANISOU  165  O5' 5HC A   9     2482   3409   3340    -32   -331   -740       O
HETATM  166  C5' 5HC A   9       4.856   4.263  -1.813  1.00 21.88           C
ANISOU  166  C5' 5HC A   9     2601   2590   3121   -394     -8   -336       C
HETATM  167  C4' 5HC A   9       3.652   4.368  -0.897  1.00 19.77           C
```

```
ANISOU  167  C4' 5HC A   9     2579   1836   3097   -465    152     63       C
HETATM  168  O4' 5HC A   9       2.865   3.190  -0.969  1.00 19.06           O
ANISOU  168  O4' 5HC A   9     2488   1801   2953   -409     59     -1       O
HETATM  169  C3' 5HC A   9       4.027   4.561   0.573  1.00 23.06           C
ANISOU  169  C3' 5HC A   9     3570   2079   3112   -273    253   -347       C
HETATM  170  O3' 5HC A   9       3.638   5.840   1.005  1.00 26.29           O
ANISOU  170  O3' 5HC A   9     4596   1791   3601    126   1136   -150       O
HETATM  171  C2' 5HC A   9       3.441   3.440   1.258  1.00 20.08           C
ANISOU  171  C2' 5HC A   9     2447   2275   2905   -301    149    -72       C
HETATM  172  C1' 5HC A   9       2.396   2.869   0.356  1.00 19.02           C
ANISOU  172  C1' 5HC A   9     2534   1745   2949   -209    198     -8       C
HETATM  173  N1  5HC A   9       2.070   1.425   0.298  1.00 17.21           N
ANISOU  173  N1  5HC A   9     2206   1705   2626    -75    -87     94       N
HETATM  174  C2  5HC A   9       0.816   0.996   0.402  1.00 16.83           C
ANISOU  174  C2  5HC A   9     2188   1612   2593    -24     57      2       C
HETATM  175  O2  5HC A   9      -0.097   1.753   0.704  1.00 18.21           O
ANISOU  175  O2  5HC A   9     2278   1768   2872    -82    120    -19       O
HETATM  176  N3  5HC A   9       0.543  -0.304   0.249  1.00 17.07           N
ANISOU  176  N3  5HC A   9     2332   1658   2495    -49     85     82       N
HETATM  177  C4  5HC A   9       1.508  -1.157  -0.060  1.00 16.89           C
ANISOU  177  C4  5HC A   9     2355   1619   2443     80    -49    -41       C
HETATM  178  N4  5HC A   9       1.139  -2.459  -0.257  1.00 18.16           N
ANISOU  178  N4  5HC A   9     2606   1448   2845    201     96     41       N
HETATM  179  C5  5HC A   9       2.826  -0.737  -0.188  1.00 18.02           C
ANISOU  179  C5  5HC A   9     2302   1939   2604    176   -291    -16       C
HETATM  180  C5MA5HC A   9       3.976  -1.675  -0.518  0.80 19.74           C
ANISOU  180  C5MA5HC A   9     2375   2030   3096    171   -346   -119       C
HETATM  181  C5MB5HC A   9       3.840  -1.760  -0.548  0.20 19.99           C
ANISOU  181  C5MB5HC A   9     2473   2323   2798    343    -17    111       C
HETATM  182  O5 A5HC A   9       4.206  -2.565   0.569  0.80 23.25           O
ANISOU  182  O5 A5HC A   9     3083   2171   3576    415   -645   -123       O
HETATM  183  O5 B5HC A   9       5.041  -1.307  -0.089  0.20 19.97           O
ANISOU  183  O5 B5HC A   9     2202   2585   2798    567     40    208       O
HETATM  184  C6  5HC A   9       3.048   0.595  -0.017  0.80 18.28           C
ANISOU  184  C6  5HC A   9     2446   1898   2601    114   -193     56       C
ATOM    185  P    DG A  10       3.993   6.501   2.333  0.80 31.69           P
ANISOU  185  P    DG A  10     6116   1836   4088  -1307   1879   -671       P
ATOM    186  OP1  DG A  10       3.727   7.932   2.185  1.00 41.30           O
ANISOU  186  OP1  DG A  10     7630   3204   4857    436   2837    155       O
ATOM    187  OP2  DG A  10       5.266   6.076   2.754  1.00 42.79           O
ANISOU  187  OP2  DG A  10     6028   4987   5240   -432   1473    242       O
ATOM    188  O5'  DG A  10       2.943   5.870   3.345  1.00 24.57           O
ANISOU  188  O5'  DG A  10     3823   1829   3682   -814    939   -367       O
ATOM    189  C5'  DG A  10       1.546   6.114   3.298  1.00 23.54           C
ANISOU  189  C5'  DG A  10     4063   1602   3280   -248    393    -30       C
ATOM    190  C4'  DG A  10       0.867   5.179   4.273  1.00 20.58           C
ANISOU  190  C4'  DG A  10     2825   1678   3314    -63    377   -174       C
ATOM    191  O4'  DG A  10       0.960   3.829   3.765  1.00 19.55           O
ANISOU  191  O4'  DG A  10     2631   1635   3158    -41    116   -220       O
ATOM    192  C3'  DG A  10       1.499   5.138   5.669  1.00 21.57           C
ANISOU  192  C3'  DG A  10     3088   1887   3219   -463    429   -151       C
ATOM    193  O3'  DG A  10       0.495   5.018   6.673  1.00 21.79           O
ANISOU  193  O3'  DG A  10     3075   1857   3345   -387    447   -322       O
ATOM    194  C2'  DG A  10       2.279   3.840   5.658  1.00 21.13           C
ANISOU  194  C2'  DG A  10     2768   2093   3164   -405     15   -263       C
ATOM    195  C1'  DG A  10       1.322   3.003   4.850  1.00 19.40           C
```

```
ANISOU  195  C1'  DG A  10     2396   1821   3153   -116     45   -215       C
ATOM    196  N9   DG A  10       1.779   1.714   4.331  1.00 18.52           N
ANISOU  196  N9   DG A  10     2351   1743   2940    -69    -65    -24       N
ATOM    197  C8   DG A  10       3.046   1.308   3.996  1.00 18.90           C
ANISOU  197  C8   DG A  10     2247   1971   2963   -163    -68    -83       C
ATOM    198  N7   DG A  10       3.090   0.083   3.552  1.00 18.74           N
ANISOU  198  N7   DG A  10     2197   1965   2957    -51    -68    -46       N
ATOM    199  C5   DG A  10       1.767  -0.339   3.588  1.00 17.85           C
ANISOU  199  C5   DG A  10     2187   1851   2741    -33    -74    172       C
ATOM    200  C6   DG A  10       1.183  -1.591   3.244  1.00 17.40           C
ANISOU  200  C6   DG A  10     2394   1588   2627    111    -46     55       C
ATOM    201  O6   DG A  10       1.728  -2.595   2.789  1.00 18.91           O
ANISOU  201  O6   DG A  10     2504   1719   2959    176     62     36       O
ATOM    202  N1   DG A  10      -0.189  -1.589   3.463  1.00 17.26           N
ANISOU  202  N1   DG A  10     2312   1545   2701     49    -48     33       N
ATOM    203  C2   DG A  10      -0.906  -0.536   3.965  1.00 17.06           C
ANISOU  203  C2   DG A  10     2255   1514   2712    -10    105     67       C
ATOM    204  N2   DG A  10      -2.214  -0.720   4.110  1.00 17.62           N
ANISOU  204  N2   DG A  10     2307   1538   2850    -86     71     17       N
ATOM    205  N3   DG A  10      -0.377   0.634   4.268  1.00 17.60           N
ANISOU  205  N3   DG A  10     2219   1625   2843     31    -14     31       N
ATOM    206  C4   DG A  10       0.950   0.660   4.060  1.00 18.03           C
ANISOU  206  C4   DG A  10     2280   1761   2808    -85     10    186       C
ATOM    207  P    DC A  11      -0.029   6.285   7.476  1.00 22.76           P
ANISOU  207  P    DC A  11     3301   1860   3487   -318    438   -338       P
ATOM    208  OP1  DC A  11       0.033   7.472   6.653  1.00 25.52           O
ANISOU  208  OP1  DC A  11     3766   1740   4188   -541    454    -97       O
ATOM    209  OP2  DC A  11       0.628   6.239   8.711  1.00 26.17           O
ANISOU  209  OP2  DC A  11     3624   2717   3601   -445    471   -400       O
ATOM    210  O5'  DC A  11      -1.524   5.867   7.785  1.00 22.04           O
ANISOU  210  O5'  DC A  11     3192   1774   3407     24    231   -324       O
ATOM    211  C5'  DC A  11      -2.505   5.873   6.774  1.00 22.16           C
ANISOU  211  C5'  DC A  11     3134   1800   3485     12    398      9       C
ATOM    212  C4'  DC A  11      -3.490   4.750   6.985  1.00 20.24           C
ANISOU  212  C4'  DC A  11     2880   1639   3171    210    297     72       C
ATOM    213  O4'  DC A  11      -2.857   3.481   6.709  1.00 20.19           O
ANISOU  213  O4'  DC A  11     3123   1546   3000    172    262     99       O
ATOM    214  C3'  DC A  11      -4.075   4.612   8.381  1.00 20.69           C
ANISOU  214  C3'  DC A  11     3085   1601   3174    120    279     61       C
ATOM    215  O3'  DC A  11      -5.423   4.210   8.168  1.00 23.34           O
ANISOU  215  O3'  DC A  11     3019   2382   3465    156    306    306       O
ATOM    216  C2'  DC A  11      -3.234   3.505   8.997  1.00 20.20           C
ANISOU  216  C2'  DC A  11     2991   1629   3054     94    257     61       C
ATOM    217  C1'  DC A  11      -3.018   2.594   7.809  1.00 19.06           C
ANISOU  217  C1'  DC A  11     2567   1737   2936     70     47    126       C
ATOM    218  N1   DC A  11      -1.842   1.729   7.823  1.00 17.91           N
ANISOU  218  N1   DC A  11     2391   1590   2823      2     -2     62       N
ATOM    219  C2   DC A  11      -1.981   0.394   7.430  1.00 17.91           C
ANISOU  219  C2   DC A  11     2404   1585   2816   -105     35     62       C
ATOM    220  O2   DC A  11      -3.108  -0.047   7.223  1.00 18.39           O
ANISOU  220  O2   DC A  11     2409   1571   3005     55    -58     33       O
ATOM    221  N3   DC A  11      -0.883  -0.375   7.317  1.00 17.84           N
ANISOU  221  N3   DC A  11     2386   1632   2758   -136   -137     37       N
ATOM    222  C4   DC A  11       0.322   0.145   7.544  1.00 18.55           C
ANISOU  222  C4   DC A  11     2443   1730   2874   -127   -164    -45       C
ATOM    223  N4   DC A  11       1.380  -0.634   7.367  1.00 18.95           N
```

```
ANISOU  223  N4   DC A  11     2374   1795   3030   -121   -173     86       N
ATOM    224  C5   DC A  11       0.495   1.510   7.915  1.00 19.97           C
ANISOU  224  C5   DC A  11     2655   1736   3195   -333    -45   -121       C
ATOM    225  C6   DC A  11      -0.603   2.259   8.035  1.00 19.96           C
ANISOU  225  C6   DC A  11     2550   1925   3108   -252      1    -54       C
ATOM    226  P    DG A  12      -6.471   4.222   9.346  1.00 27.72           P
ANISOU  226  P    DG A  12     3890   2384   4257    449   1216    243       P
ATOM    227  OP1  DG A  12      -7.712   4.652   8.853  1.00 27.56           O
ANISOU  227  OP1  DG A  12     4070   1611   4787    430   1126    204       O
ATOM    228  OP2  DG A  12      -5.880   4.726  10.519  1.00 32.88           O
ANISOU  228  OP2  DG A  12     3826   4655   4011    873    933    779       O
ATOM    229  O5'  DG A  12      -6.416   2.793   9.948  1.00 33.13           O
ANISOU  229  O5'  DG A  12     4367   3386   4833    413    740   1459       O
ATOM    230  C5'  DG A  12      -6.983   1.731   9.279  1.00 26.12           C
ANISOU  230  C5'  DG A  12     4018   2511   3394  -1214     -8   -405       C
ATOM    231  C4'  DG A  12      -7.091   0.575  10.253  1.00 19.74           C
ANISOU  231  C4'  DG A  12     2340   2197   2962      2    -51    154       C
ATOM    232  O4'  DG A  12      -5.967  -0.320  10.200  1.00 26.91           O
ANISOU  232  O4'  DG A  12     2590   4225   3408   1099   -236   -570       O
ATOM    233  C3'  DG A  12      -7.276   0.962  11.716  1.00 19.40           C
ANISOU  233  C3'  DG A  12     2370   1821   3180    -52    -15    -17       C
ATOM    234  O3'  DG A  12      -8.437   0.297  12.187  1.00 19.83           O
ANISOU  234  O3'  DG A  12     2304   1934   3295    151    109    151       O
ATOM    235  C2'  DG A  12      -6.016   0.476  12.412  1.00 20.33           C
ANISOU  235  C2'  DG A  12     2408   2328   2987    -67    -80   -149       C
ATOM    236  C1'  DG A  12      -5.400  -0.543  11.466  1.00 21.19           C
ANISOU  236  C1'  DG A  12     2293   2660   3098    303   -126    -67       C
ATOM    237  N9   DG A  12      -3.955  -0.487  11.294  1.00 19.72           N
ANISOU  237  N9   DG A  12     2266   2133   3091     -6    -95     82       N
ATOM    238  C8   DG A  12      -3.141   0.613  11.373  1.00 19.31           C
ANISOU  238  C8   DG A  12     2328   2080   2928     17    -97    287       C
ATOM    239  N7   DG A  12      -1.898   0.350  11.081  1.00 19.76           N
ANISOU  239  N7   DG A  12     2268   2160   3079   -209    -74    138       N
ATOM    240  C5   DG A  12      -1.897  -1.000  10.763  1.00 18.91           C
ANISOU  240  C5   DG A  12     2280   1971   2934    -80   -157    284       C
ATOM    241  C6   DG A  12      -0.832  -1.858  10.375  1.00 18.84           C
ANISOU  241  C6   DG A  12     2165   2137   2855    -25     66    194       C
ATOM    242  O6   DG A  12       0.355  -1.585  10.213  1.00 20.63           O
ANISOU  242  O6   DG A  12     2267   2267   3303   -172    -45    277       O
ATOM    243  N1   DG A  12      -1.277  -3.156  10.156  1.00 19.50           N
ANISOU  243  N1   DG A  12     2201   2162   3043    -22    -75    190       N
ATOM    244  C2   DG A  12      -2.566  -3.583  10.311  1.00 18.89           C
ANISOU  244  C2   DG A  12     2027   2360   2788    -74     57   -101       C
ATOM    245  N2   DG A  12      -2.790  -4.876  10.081  1.00 19.80           N
ANISOU  245  N2   DG A  12     2201   2206   3113    204     41   -134       N
ATOM    246  N3   DG A  12      -3.562  -2.797  10.675  1.00 18.80           N
ANISOU  246  N3   DG A  12     2023   2068   3052    -52    -19    -13       N
ATOM    247  C4   DG A  12      -3.160  -1.528  10.878  1.00 18.99           C
ANISOU  247  C4   DG A  12     2088   2251   2875     30    -17    175       C
TER     248       DG A  12
ATOM    249  O5'  DC B  13       4.517  -9.654   9.217  1.00 35.24           O
ANISOU  249  O5'  DC B  13     3239   4591   5560    696     60    -31       O
ATOM    250  C5'  DC B  13       3.797 -10.590  10.029  1.00 29.09           C
ANISOU  250  C5'  DC B  13     2877   3536   4637    630   -305   -548       C
ATOM    251  C4'  DC B  13       2.328 -10.515   9.691  1.00 25.51           C
ANISOU  251  C4'  DC B  13     2689   2485   4518    346   -199   -305       C
```

```
ATOM    252  O4'  DC B  13       1.738  -9.234   9.994  1.00 23.51           O
ANISOU  252  O4'  DC B  13     2730   2316   3886    324   -133   -160       O
ATOM    253  C3'  DC B  13       1.960 -10.856   8.245  1.00 25.96           C
ANISOU  253  C3'  DC B  13     3069   2481   4314    175     70    -67       C
ATOM    254  O3'  DC B  13       0.875 -11.787   8.202  1.00 26.95           O
ANISOU  254  O3'  DC B  13     3325   2033   4881    190   -436   -427       O
ATOM    255  C2'  DC B  13       1.552  -9.526   7.641  1.00 25.51           C
ANISOU  255  C2'  DC B  13     2903   2837   3952    399    152    115       C
ATOM    256  C1'  DC B  13       1.112  -8.680   8.842  1.00 24.13           C
ANISOU  256  C1'  DC B  13     2889   2273   4003    665    250    230       C
ATOM    257  N1   DC B  13       1.468  -7.247   8.815  1.00 22.20           N
ANISOU  257  N1   DC B  13     2316   2634   3481     76     76    -41       N
ATOM    258  C2   DC B  13       0.501  -6.289   9.162  1.00 21.44           C
ANISOU  258  C2   DC B  13     2363   2222   3561     96   -246    151       C
ATOM    259  O2   DC B  13      -0.664  -6.654   9.353  1.00 21.81           O
ANISOU  259  O2   DC B  13     2494   2316   3475    307     43    109       O
ATOM    260  N3   DC B  13       0.855  -4.985   9.223  1.00 20.55           N
ANISOU  260  N3   DC B  13     2163   2487   3155     24    -98      6       N
ATOM    261  C4   DC B  13       2.116  -4.626   8.981  1.00 20.39           C
ANISOU  261  C4   DC B  13     2172   2394   3179   -127    -35    -28       C
ATOM    262  N4   DC B  13       2.423  -3.331   9.071  1.00 21.23           N
ANISOU  262  N4   DC B  13     2287   2420   3359   -139   -191    -34       N
ATOM    263  C5   DC B  13       3.130  -5.583   8.679  1.00 22.11           C
ANISOU  263  C5   DC B  13     2214   2522   3664    -65    211    -37       C
ATOM    264  C6   DC B  13       2.769  -6.871   8.623  1.00 23.23           C
ANISOU  264  C6   DC B  13     2309   2561   3957    -85    -98    -33       C
ATOM    265  P    DG B  14       0.431 -12.476   6.825  1.00 29.20           P
ANISOU  265  P    DG B  14     3962   2238   4896    349   -420   -442       P
ATOM    266  OP1  DG B  14      -0.166 -13.796   7.166  1.00 34.49           O
ANISOU  266  OP1  DG B  14     4505   1899   6698    514   -722   -331       O
ATOM    267  OP2  DG B  14       1.568 -12.436   5.862  1.00 36.05           O
ANISOU  267  OP2  DG B  14     3738   3314   6642    912    399  -1545       O
ATOM    268  O5'  DG B  14      -0.698 -11.466   6.343  1.00 24.63           O
ANISOU  268  O5'  DG B  14     2953   2188   4215    394   -145   -598       O
ATOM    269  C5'  DG B  14      -1.863 -11.411   7.143  1.00 22.11           C
ANISOU  269  C5'  DG B  14     3129   1915   3354    112   -296    -36       C
ATOM    270  C4'  DG B  14      -2.697 -10.233   6.720  1.00 19.91           C
ANISOU  270  C4'  DG B  14     2493   1700   3370   -177    128   -262       C
ATOM    271  O4'  DG B  14      -1.969  -9.030   6.977  1.00 21.17           O
ANISOU  271  O4'  DG B  14     2653   1536   3855    -37     80   -252       O
ATOM    272  C3'  DG B  14      -2.997 -10.225   5.224  1.00 21.45           C
ANISOU  272  C3'  DG B  14     2602   1974   3571     17   -115    107       C
ATOM    273  O3'  DG B  14      -4.371 -10.525   5.087  1.00 20.62           O
ANISOU  273  O3'  DG B  14     2691   1520   3622   -268     19    -14       O
ATOM    274  C2'  DG B  14      -2.672  -8.817   4.768  1.00 21.93           C
ANISOU  274  C2'  DG B  14     2879   2072   3381   -240    -33      2       C
ATOM    275  C1'  DG B  14      -2.439  -8.063   6.061  1.00 21.04           C
ANISOU  275  C1'  DG B  14     2781   1410   3802    -16   -110   -200       C
ATOM    276  N9   DG B  14      -1.411  -7.044   5.948  1.00 20.97           N
ANISOU  276  N9   DG B  14     2391   1492   4082     48    -77   -164       N
ATOM    277  C8   DG B  14      -0.116  -7.261   5.545  1.00 21.40           C
ANISOU  277  C8   DG B  14     2366   1770   3993     -6    -51      1       C
ATOM    278  N7   DG B  14       0.610  -6.179   5.555  1.00 20.65           N
ANISOU  278  N7   DG B  14     2433   1745   3665    -49    -12   -171       N
ATOM    279  C5   DG B  14      -0.254  -5.189   6.000  1.00 18.90           C
ANISOU  279  C5   DG B  14     2265   1654   3258     18   -146     31       C
```

```
ATOM    280  C6   DG B  14      -0.032  -3.809   6.197  1.00 17.21           C
ANISOU  280  C6   DG B  14     2102   1591   2844    -86   -125    130       C
ATOM    281  O6   DG B  14       1.018  -3.177   6.059  1.00 18.11           O
ANISOU  281  O6   DG B  14     2246   1643   2988   -120   -102    115       O
ATOM    282  N1   DG B  14      -1.176  -3.164   6.656  1.00 17.18           N
ANISOU  282  N1   DG B  14     2187   1596   2743   -130   -103     48       N
ATOM    283  C2   DG B  14      -2.390  -3.768   6.874  1.00 17.31           C
ANISOU  283  C2   DG B  14     2208   1500   2868    -86   -104     65       C
ATOM    284  N2   DG B  14      -3.381  -2.964   7.296  1.00 17.84           N
ANISOU  284  N2   DG B  14     2293   1547   2937   -129    -21    151       N
ATOM    285  N3   DG B  14      -2.611  -5.064   6.695  1.00 18.31           N
ANISOU  285  N3   DG B  14     2186   1454   3316    -84    -61     39       N
ATOM    286  C4   DG B  14      -1.504  -5.708   6.260  1.00 18.67           C
ANISOU  286  C4   DG B  14     2256   1414   3420     15    -82     84       C
ATOM    287  P    DC B  15      -5.063 -10.543   3.652  1.00 21.03           P
ANISOU  287  P    DC B  15     3190   1621   3179   -138    135   -121       P
ATOM    288  OP1  DC B  15      -6.189 -11.484   3.750  1.00 22.84           O
ANISOU  288  OP1  DC B  15     3755   1536   3385   -565    -10    -97       O
ATOM    289  OP2  DC B  15      -4.025 -10.756   2.613  1.00 24.83           O
ANISOU  289  OP2  DC B  15     3865   2288   3281    412    393   -217       O
ATOM    290  O5'  DC B  15      -5.581  -9.055   3.465  1.00 20.14           O
ANISOU  290  O5'  DC B  15     2887   1744   3021   -155    160      4       O
ATOM    291  C5'  DC B  15      -6.578  -8.504   4.328  1.00 18.89           C
ANISOU  291  C5'  DC B  15     2582   1743   2850   -313    119     45       C
ATOM    292  C4'  DC B  15      -6.552  -6.999   4.236  1.00 19.01           C
ANISOU  292  C4'  DC B  15     2726   1712   2784   -285     10    -81       C
ATOM    293  O4'  DC B  15      -5.232  -6.528   4.539  1.00 20.19           O
ANISOU  293  O4'  DC B  15     2992   1572   3108   -371   -280    114       O
ATOM    294  C3'  DC B  15      -6.829  -6.413   2.861  1.00 18.81           C
ANISOU  294  C3'  DC B  15     2590   1767   2788   -312    163    -65       C
ATOM    295  O3'  DC B  15      -8.231  -6.410   2.647  1.00 19.90           O
ANISOU  295  O3'  DC B  15     2690   1898   2971   -407     70     66       O
ATOM    296  C2'  DC B  15      -6.216  -5.031   2.975  1.00 20.48           C
ANISOU  296  C2'  DC B  15     2816   1765   3197   -333   -105    162       C
ATOM    297  C1'  DC B  15      -5.060  -5.248   3.921  1.00 19.84           C
ANISOU  297  C1'  DC B  15     2742   1594   3203   -157    -39     96       C
ATOM    298  N1   DC B  15      -3.727  -5.223   3.329  1.00 19.11           N
ANISOU  298  N1   DC B  15     2686   1538   3033    -62    -94      9       N
ATOM    299  C2   DC B  15      -2.994  -4.043   3.428  1.00 17.83           C
ANISOU  299  C2   DC B  15     2512   1420   2842    -48     94     56       C
ATOM    300  O2   DC B  15      -3.542  -3.037   3.890  1.00 18.22           O
ANISOU  300  O2   DC B  15     2483   1386   3052     33     31      1       O
ATOM    301  N3   DC B  15      -1.714  -4.022   3.005  1.00 18.25           N
ANISOU  301  N3   DC B  15     2587   1499   2845    150     93    -20       N
ATOM    302  C4   DC B  15      -1.186  -5.105   2.435  1.00 20.21           C
ANISOU  302  C4   DC B  15     2825   1732   3120    178     40   -246       C
ATOM    303  N4   DC B  15       0.077  -5.036   2.018  1.00 21.23           N
ANISOU  303  N4   DC B  15     2944   1752   3368    267    128   -226       N
ATOM    304  C5   DC B  15      -1.932  -6.305   2.260  1.00 22.52           C
ANISOU  304  C5   DC B  15     3269   1589   3695    113    110   -150       C
ATOM    305  C6   DC B  15      -3.189  -6.321   2.719  1.00 21.63           C
ANISOU  305  C6   DC B  15     3028   1720   3470     54    -94   -319       C
ATOM    306  P    DG B  16      -8.845  -6.049   1.228  1.00 21.07           P
ANISOU  306  P    DG B  16     2850   1989   3167   -448   -202      5       P
ATOM    307  OP1  DG B  16     -10.248  -6.523   1.248  1.00 24.23           O
ANISOU  307  OP1  DG B  16     3275   2365   3566   -930   -400     84       O
```

```
ATOM    308  OP2  DG B  16      -7.938  -6.522   0.147  1.00 22.93           O
ANISOU  308  OP2  DG B  16     3515   2169   3028   -241   -230   -197       O
ATOM    309  O5'  DG B  16      -8.836  -4.461   1.193  1.00 19.96           O
ANISOU  309  O5'  DG B  16     2604   2012   2965   -340     47     67       O
ATOM    310  C5'  DG B  16      -9.564  -3.684   2.157  1.00 19.81           C
ANISOU  310  C5'  DG B  16     2461   2100   2964    -53    146    226       C
ATOM    311  C4'  DG B  16      -9.280  -2.221   1.926  1.00 20.01           C
ANISOU  311  C4'  DG B  16     2526   2200   2874   -418    -75    198       C
ATOM    312  O4'  DG B  16      -7.856  -1.984   2.051  1.00 18.88           O
ANISOU  312  O4'  DG B  16     2411   1926   2834   -218     -1    135       O
ATOM    313  C3'  DG B  16      -9.675  -1.681   0.548  1.00 20.05           C
ANISOU  313  C3'  DG B  16     2263   2214   3139      5   -121    302       C
ATOM    314  O3'  DG B  16     -10.152  -0.343   0.773  1.00 20.65           O
ANISOU  314  O3'  DG B  16     2440   2246   3159      9     40    392       O
ATOM    315  C2'  DG B  16      -8.390  -1.789  -0.242  1.00 20.01           C
ANISOU  315  C2'  DG B  16     2515   2128   2961   -309    -23     99       C
ATOM    316  C1'  DG B  16      -7.330  -1.541   0.813  1.00 18.79           C
ANISOU  316  C1'  DG B  16     2347   1878   2913   -138    -45    201       C
ATOM    317  N9   DG B  16      -6.079  -2.241   0.534  1.00 17.72           N
ANISOU  317  N9   DG B  16     2405   1644   2683   -173     71     52       N
ATOM    318  C8   DG B  16      -5.915  -3.477  -0.043  1.00 18.68           C
ANISOU  318  C8   DG B  16     2563   1660   2871   -166     61     -2       C
ATOM    319  N7   DG B  16      -4.664  -3.778  -0.263  1.00 19.08           N
ANISOU  319  N7   DG B  16     2610   1786   2851   -191      3   -132       N
ATOM    320  C5   DG B  16      -3.963  -2.665   0.182  1.00 17.51           C
ANISOU  320  C5   DG B  16     2480   1544   2627    -65     16     85       C
ATOM    321  C6   DG B  16      -2.573  -2.375   0.145  1.00 17.43           C
ANISOU  321  C6   DG B  16     2487   1461   2674    -46     60     23       C
ATOM    322  O6   DG B  16      -1.648  -3.088  -0.261  1.00 18.30           O
ANISOU  322  O6   DG B  16     2679   1565   2710     27    216     14       O
ATOM    323  N1   DG B  16      -2.301  -1.099   0.634  1.00 17.05           N
ANISOU  323  N1   DG B  16     2271   1546   2661    -57     66      7       N
ATOM    324  C2   DG B  16      -3.240  -0.227   1.122  1.00 16.74           C
ANISOU  324  C2   DG B  16     2259   1502   2598    -36    103    159       C
ATOM    325  N2   DG B  16      -2.782   0.963   1.547  1.00 17.18           N
ANISOU  325  N2   DG B  16     2339   1447   2740    -20     59     66       N
ATOM    326  N3   DG B  16      -4.533  -0.492   1.175  1.00 16.99           N
ANISOU  326  N3   DG B  16     2311   1440   2705    -52      6    116       N
ATOM    327  C4   DG B  16      -4.824  -1.707   0.669  1.00 16.89           C
ANISOU  327  C4   DG B  16     2318   1517   2580   -104     78    139       C
ATOM    328  P    DA B  17     -10.173   0.833  -0.331  1.00 21.05           P
ANISOU  328  P    DA B  17     2404   2346   3245     68      6    434       P
ATOM    329  OP1  DA B  17     -11.279   1.754   0.053  1.00 22.46           O
ANISOU  329  OP1  DA B  17     2387   2667   3480    198     90    506       O
ATOM    330  OP2  DA B  17     -10.137   0.270  -1.695  1.00 23.21           O
ANISOU  330  OP2  DA B  17     2612   2778   3429     56   -122    473       O
ATOM    331  O5'  DA B  17      -8.787   1.572  -0.094  1.00 19.99           O
ANISOU  331  O5'  DA B  17     2363   2225   3005     -3    158    298       O
ATOM    332  C5'  DA B  17      -8.470   2.156   1.180  1.00 19.29           C
ANISOU  332  C5'  DA B  17     2412   2049   2865    113    265    191       C
ATOM    333  C4'  DA B  17      -7.340   3.135   0.999  1.00 18.87           C
ANISOU  333  C4'  DA B  17     2286   1933   2949    238    249    267       C
ATOM    334  O4'  DA B  17      -6.147   2.432   0.593  1.00 18.85           O
ANISOU  334  O4'  DA B  17     2337   1937   2886    223    254    194       O
ATOM    335  C3'  DA B  17      -7.584   4.188  -0.080  1.00 19.37           C
ANISOU  335  C3'  DA B  17     2380   1815   3163    238    380    272       C
```

```
ATOM    336  O3'  DA B  17      -7.017   5.418   0.430  1.00 20.63           O
ANISOU  336  O3'  DA B  17     2684   1723   3429    200    596    140       O
ATOM    337  C2'  DA B  17      -6.885   3.614  -1.307  1.00 18.90           C
ANISOU  337  C2'  DA B  17     2333   1771   3077    208    203    212       C
ATOM    338  C1'  DA B  17      -5.712   2.872  -0.702  1.00 17.85           C
ANISOU  338  C1'  DA B  17     2243   1683   2856    163    172    130       C
ATOM    339  N9   DA B  17      -5.309   1.684  -1.439  1.00 17.40           N
ANISOU  339  N9   DA B  17     2284   1614   2712      0    186    206       N
ATOM    340  C8   DA B  17      -6.124   0.642  -1.798  1.00 18.01           C
ANISOU  340  C8   DA B  17     2212   1866   2764     27    130    153       C
ATOM    341  N7   DA B  17      -5.486  -0.359  -2.353  1.00 17.59           N
ANISOU  341  N7   DA B  17     2314   1681   2687    -19    140    159       N
ATOM    342  C5   DA B  17      -4.156   0.031  -2.318  1.00 16.97           C
ANISOU  342  C5   DA B  17     2251   1542   2652    -97     65    146       C
ATOM    343  C6   DA B  17      -2.974  -0.599  -2.743  1.00 16.69           C
ANISOU  343  C6   DA B  17     2336   1550   2455     64     92    114       C
ATOM    344  N6   DA B  17      -2.932  -1.823  -3.290  1.00 17.33           N
ANISOU  344  N6   DA B  17     2408   1541   2635      1     30     40       N
ATOM    345  N1   DA B  17      -1.814   0.077  -2.582  1.00 16.97           N
ANISOU  345  N1   DA B  17     2457   1347   2642     -7    153    155       N
ATOM    346  C2   DA B  17      -1.848   1.292  -2.020  1.00 16.80           C
ANISOU  346  C2   DA B  17     2291   1470   2621    102    214     99       C
ATOM    347  N3   DA B  17      -2.893   1.980  -1.567  1.00 16.78           N
ANISOU  347  N3   DA B  17     2203   1492   2678     98    197    192       N
ATOM    348  C4   DA B  17      -4.031   1.288  -1.754  1.00 16.78           C
ANISOU  348  C4   DA B  17     2203   1602   2568     30    168     57       C
ATOM    349  P  A DA B  18      -6.973   6.716  -0.455  0.80 21.71           P
ANISOU  349  P  A DA B  18     2701   1716   3831    375    459    196       P
ATOM    350  P  B DA B  18      -7.014   6.862  -0.413  0.20 23.21           P
ANISOU  350  P  B DA B  18     2871   1940   4007     93    383    304       P
ATOM    351  OP1A DA B  18      -6.961   7.867   0.460  0.80 24.47           O
ANISOU  351  OP1A DA B  18     3353   1705   4239    532    835   -131       O
ATOM    352  OP1B DA B  18      -6.985   7.995   0.560  0.20 26.26           O
ANISOU  352  OP1B DA B  18     3379   2319   4280     66    427    -23       O
ATOM    353  OP2A DA B  18      -7.993   6.634  -1.528  0.80 23.59           O
ANISOU  353  OP2A DA B  18     2990   1986   3987    300    341    348       O
ATOM    354  OP2B DA B  18      -8.071   6.823  -1.458  0.20 23.62           O
ANISOU  354  OP2B DA B  18     3151   1777   4045    595    148    686       O
ATOM    355  O5'A DA B  18      -5.571   6.655  -1.191  0.80 20.14           O
ANISOU  355  O5'A DA B  18     2782   1589   3280    137    397    100       O
ATOM    356  O5'B DA B  18      -5.608   6.872  -1.164  0.20 21.93           O
ANISOU  356  O5'B DA B  18     2847   1990   3496    128    285    167       O
ATOM    357  C5'A DA B  18      -4.360   6.670  -0.449  0.80 19.88           C
ANISOU  357  C5'A DA B  18     2659   1697   3194    -35    432     14       C
ATOM    358  C5'B DA B  18      -4.362   6.853  -0.449  0.20 21.44           C
ANISOU  358  C5'B DA B  18     2815   2083   3246    246    378     10       C
ATOM    359  C4'A DA B  18      -3.211   6.593  -1.418  0.80 20.12           C
ANISOU  359  C4'A DA B  18     3160   1160   3325     -1    716     27       C
ATOM    360  C4'B DA B  18      -3.225   6.832  -1.441  0.20 21.41           C
ANISOU  360  C4'B DA B  18     3152   1772   3210    418    522    -87       C
ATOM    361  O4'A DA B  18      -3.190   5.295  -2.035  0.80 19.35           O
ANISOU  361  O4'A DA B  18     3122   1202   3026    269    504      0       O
ATOM    362  O4'B DA B  18      -3.116   5.515  -2.021  0.20 20.69           O
ANISOU  362  O4'B DA B  18     3120   1687   3053    336    437     17       O
ATOM    363  C3'A DA B  18      -3.249   7.615  -2.562  0.80 27.46           C
ANISOU  363  C3'A DA B  18     5158   1849   3425   1195   1741    245       C
```

```
ATOM    364  C3'B DA B  18      -3.398   7.804  -2.611  0.20 28.34           C
ANISOU  364  C3'B DA B  18     5179   2142   3446   1490    841    113       C
ATOM    365  O3'A DA B  18      -2.021   8.360  -2.363  0.80 37.27           O
ANISOU  365  O3'A DA B  18     6728    898   6535  -1500   3792    -47       O
ATOM    366  O3'B DA B  18      -2.264   8.673  -2.648  0.20 38.21           O
ANISOU  366  O3'B DA B  18     8218   1664   4633   -241   1874   -505       O
ATOM    367  C2'A DA B  18      -3.380   6.757  -3.824  0.80 21.78           C
ANISOU  367  C2'A DA B  18     3057   1706   3509    547    726    464       C
ATOM    368  C2'B DA B  18      -3.442   6.921  -3.850  0.20 22.17           C
ANISOU  368  C2'B DA B  18     3039   1976   3408    356    509    385       C
ATOM    369  C1'A DA B  18      -2.791   5.432  -3.386  0.80 19.05           C
ANISOU  369  C1'A DA B  18     2845   1314   3079    327    611    143       C
ATOM    370  C1'B DA B  18      -2.777   5.642  -3.392  0.20 20.34           C
ANISOU  370  C1'B DA B  18     2798   1855   3072    301    476     89       C
ATOM    371  N9 A DA B  18      -3.263   4.249  -4.083  0.80 18.34           N
ANISOU  371  N9 A DA B  18     2757   1442   2767    499    353    155       N
ATOM    372  N9 B DA B  18      -3.253   4.440  -4.058  0.20 18.85           N
ANISOU  372  N9 B DA B  18     2638   1717   2805    478    288    254       N
ATOM    373  C8 A DA B  18      -4.568   3.846  -4.177  0.80 18.68           C
ANISOU  373  C8 A DA B  18     2584   1807   2707    580    326    207       C
ATOM    374  C8 B DA B  18      -4.558   4.032  -4.156  0.20 18.56           C
ANISOU  374  C8 B DA B  18     2659   1711   2679    386    214    316       C
ATOM    375  N7 A DA B  18      -4.720   2.684  -4.763  0.80 17.95           N
ANISOU  375  N7 A DA B  18     2321   1846   2651    397    238    219       N
ATOM    376  N7 B DA B  18      -4.705   2.869  -4.739  0.20 17.94           N
ANISOU  376  N7 B DA B  18     2299   1825   2691    213    238    236       N
ATOM    377  C5 A DA B  18      -3.424   2.278  -5.043  0.80 16.94           C
ANISOU  377  C5 A DA B  18     2274   1492   2669    266    162    330       C
ATOM    378  C5 B DA B  18      -3.405   2.469  -5.014  0.20 17.59           C
ANISOU  378  C5 B DA B  18     2323   1606   2752    267    145    411       C
ATOM    379  C6 A DA B  18      -2.904   1.111  -5.634  0.80 16.44           C
ANISOU  379  C6 A DA B  18     2303   1465   2478    151      3    188       C
ATOM    380  C6 B DA B  18      -2.881   1.308  -5.605  0.20 17.73           C
ANISOU  380  C6 B DA B  18     2344   1721   2669    127     96    203       C
ATOM    381  N6 A DA B  18      -3.660   0.097  -6.066  0.80 17.27           N
ANISOU  381  N6 A DA B  18     2289   1677   2596    125    170    272       N
ATOM    382  N6 B DA B  18      -3.633   0.299  -6.050  0.20 18.91           N
ANISOU  382  N6 B DA B  18     2327   2094   2763     51     98     56       N
ATOM    383  N1 A DA B  18      -1.566   1.028  -5.784  0.80 16.64           N
ANISOU  383  N1 A DA B  18     2430   1348   2542    219     19    130       N
ATOM    384  N1 B DA B  18      -1.540   1.225  -5.745  0.20 17.76           N
ANISOU  384  N1 B DA B  18     2361   1671   2715    246     11    204       N
ATOM    385  C2 A DA B  18      -0.806   2.028  -5.322  0.80 16.47           C
ANISOU  385  C2 A DA B  18     2366   1270   2620    204     59    130       C
ATOM    386  C2 B DA B  18      -0.785   2.229  -5.283  0.20 17.56           C
ANISOU  386  C2 B DA B  18     2317   1585   2770    203    112    251       C
ATOM    387  N3 A DA B  18      -1.176   3.165  -4.733  0.80 17.37           N
ANISOU  387  N3 A DA B  18     2583   1322   2693    213    198     52       N
ATOM    388  N3 B DA B  18      -1.161   3.365  -4.698  0.20 18.59           N
ANISOU  388  N3 B DA B  18     2577   1624   2862    269    187    235       N
ATOM    389  C4 A DA B  18      -2.515   3.230  -4.623  0.80 17.61           C
ANISOU  389  C4 A DA B  18     2590   1355   2744    340    167     98       C
ATOM    390  C4 B DA B  18      -2.500   3.423  -4.593  0.20 18.03           C
ANISOU  390  C4 B DA B  18     2561   1487   2800    299    214    227       C
ATOM    391  P  A DT B  19      -1.541   9.518  -3.335  0.80 37.49           P
ANISOU  391  P  A DT B  19     6209   1119   6913    431   1620     37       P
```

```
ATOM    392  P  B DT B  19      -2.057   9.603  -3.908  0.20 30.04           P
ANISOU  392  P  B DT B  19     4515   2253   4645    534   1472    -77       P
ATOM    393  OP1A DT B  19      -0.693  10.471  -2.562  0.80 66.07           O
ANISOU  393  OP1A DT B  19    10410   3780  10911   -889   1278  -2410       O
ATOM    394  OP1B DT B  19      -1.464  10.851  -3.390  0.20 21.48           O
ANISOU  394  OP1B DT B  19     3829   1516   2815   1641    594    689       O
ATOM    395  OP2A DT B  19      -2.712   9.937  -4.105  0.80 59.66           O
ANISOU  395  OP2A DT B  19     7409   6697   8561    902    391   -805       O
ATOM    396  OP2B DT B  19      -3.326   9.671  -4.683  0.20 21.09           O
ANISOU  396  OP2B DT B  19     4167    351   3495   -593   2162   -457       O
ATOM    397  O5'A DT B  19      -0.728   8.748  -4.443  0.80 25.76           O
ANISOU  397  O5'A DT B  19     3778   2013   3996    379   1258    685       O
ATOM    398  O5'B DT B  19      -0.999   8.800  -4.783  0.20 21.19           O
ANISOU  398  O5'B DT B  19     3355   1503   3191    602    460    187       O
ATOM    399  C5'A DT B  19       0.469   8.079  -4.106  0.80 22.11           C
ANISOU  399  C5'A DT B  19     3527   1598   3275    162    579    -20       C
ATOM    400  C5'B DT B  19       0.142   8.166  -4.191  0.20 17.66           C
ANISOU  400  C5'B DT B  19     2742   1410   2556     67    311   -115       C
ATOM    401  C4'A DT B  19       0.918   7.292  -5.307  0.80 19.18           C
ANISOU  401  C4'A DT B  19     2697   1620   2970   -112     73    -50       C
ATOM    402  C4'B DT B  19       0.827   7.367  -5.271  0.20 15.03           C
ANISOU  402  C4'B DT B  19     2073   1059   2578     31     90   -121       C
ATOM    403  O4'A DT B  19      -0.042   6.275  -5.637  0.80 19.64           O
ANISOU  403  O4'A DT B  19     2907   1446   3108   -203   -128     81       O
ATOM    404  O4'B DT B  19      -0.025   6.262  -5.629  0.20 14.40           O
ANISOU  404  O4'B DT B  19     2034    951   2486      8     60     36       O
ATOM    405  C3'A DT B  19       1.056   8.140  -6.570  0.80 18.48           C
ANISOU  405  C3'A DT B  19     2584   1483   2953   -186    145    -67       C
ATOM    406  C3'B DT B  19       1.048   8.162  -6.561  0.20 17.12           C
ANISOU  406  C3'B DT B  19     2395   1566   2543   -150    114    -29       C
ATOM    407  O3'A DT B  19       2.450   8.345  -6.734  0.80 20.95           O
ANISOU  407  O3'A DT B  19     2720   2181   3056   -496    269   -283       O
ATOM    408  O3'B DT B  19       2.450   8.340  -6.735  0.20 19.41           O
ANISOU  408  O3'B DT B  19     2484   1823   3067   -275    210   -256       O
ATOM    409  C2'A DT B  19       0.376   7.328  -7.666  0.80 18.81           C
ANISOU  409  C2'A DT B  19     2700   1634   2810    -89     73      6       C
ATOM    410  C2'B DT B  19       0.355   7.348  -7.647  0.20 15.70           C
ANISOU  410  C2'B DT B  19     2185   1361   2419     -2    176     38       C
ATOM    411  C1'A DT B  19       0.158   5.982  -7.009  0.80 18.13           C
ANISOU  411  C1'A DT B  19     2535   1382   2969     43   -153    -59       C
ATOM    412  C1'B DT B  19       0.163   5.993  -7.006  0.20 14.89           C
ANISOU  412  C1'B DT B  19     2003   1242   2409    103    109     -6       C
ATOM    413  N1 A DT B  19      -0.993   5.204  -7.463  0.80 17.76           N
ANISOU  413  N1 A DT B  19     2520   1414   2813     96   -179     73       N
ATOM    414  N1 B DT B  19      -0.992   5.212  -7.460  0.20 14.11           N
ANISOU  414  N1 B DT B  19     1885   1104   2370    266    116     18       N
ATOM    415  C2 A DT B  19      -0.758   3.973  -8.032  0.80 17.56           C
ANISOU  415  C2 A DT B  19     2494   1295   2883    -21   -171     91       C
ATOM    416  C2 B DT B  19      -0.760   3.979  -8.023  0.20 13.80           C
ANISOU  416  C2 B DT B  19     1744   1080   2416    328    199     60       C
ATOM    417  O2 A DT B  19       0.357   3.542  -8.258  0.80 18.19           O
ANISOU  417  O2 A DT B  19     2525   1394   2993    106    -87   -135       O
ATOM    418  O2 B DT B  19       0.358   3.548  -8.245  0.20 14.13           O
ANISOU  418  O2 B DT B  19     1673   1136   2560    257    232    129       O
ATOM    419  N3 A DT B  19      -1.886   3.273  -8.358  0.80 17.05           N
ANISOU  419  N3 A DT B  19     2429   1269   2780     -9   -233     19       N
```

```
ATOM    420  N3 B DT B  19      -1.892   3.281  -8.357  0.20 14.08           N
ANISOU  420  N3 B DT B  19     1773   1223   2353    208    185    206       N
ATOM    421  C4 A DT B  19      -3.195   3.664  -8.174  0.80 17.88           C
ANISOU  421  C4 A DT B  19     2415   1437   2941     93   -253    166       C
ATOM    422  C4 B DT B  19      -3.202   3.678  -8.173  0.20 13.56           C
ANISOU  422  C4 B DT B  19     1745   1129   2275    161    148    209       C
ATOM    423  O4 A DT B  19      -4.103   2.908  -8.501  0.80 18.22           O
ANISOU  423  O4 A DT B  19     2218   1578   3124    101   -246     35       O
ATOM    424  O4 B DT B  19      -4.117   2.932  -8.510  0.20 14.05           O
ANISOU  424  O4 B DT B  19     1911   1091   2334     50    207    137       O
ATOM    425  C5 A DT B  19      -3.373   4.983  -7.599  0.80 18.56           C
ANISOU  425  C5 A DT B  19     2458   1561   3032    178    -49     37       C
ATOM    426  C5 B DT B  19      -3.373   4.987  -7.580  0.20 14.19           C
ANISOU  426  C5 B DT B  19     1827   1234   2327    350     98    112       C
ATOM    427  C7 A DT B  19      -4.760   5.496  -7.370  0.80 18.90           C
ANISOU  427  C7 A DT B  19     2532   1734   2914    287    -61   -138       C
ATOM    428  C7 B DT B  19      -4.758   5.503  -7.346  0.20 14.48           C
ANISOU  428  C7 B DT B  19     1789   1526   2186    307     91   -105       C
ATOM    429  C6 A DT B  19      -2.275   5.670  -7.260  0.80 18.38           C
ANISOU  429  C6 A DT B  19     2634   1425   2924    234    -76     55       C
ATOM    430  C6 B DT B  19      -2.273   5.673  -7.247  0.20 14.57           C
ANISOU  430  C6 B DT B  19     1882   1274   2377    338     73    154       C
ATOM    431  P    DT B  20       3.036   9.005  -8.063  1.00 23.35           P
ANISOU  431  P    DT B  20     3632   1527   3712  -1010    506   -547       P
ATOM    432  OP1  DT B  20       4.325   9.581  -7.694  1.00 30.87           O
ANISOU  432  OP1  DT B  20     4382   3080   4264   -787    510   -330       O
ATOM    433  OP2  DT B  20       2.084   9.773  -8.765  1.00 25.54           O
ANISOU  433  OP2  DT B  20     4327   1244   4133     49    409    269       O
ATOM    434  O5'  DT B  20       3.318   7.740  -8.990  1.00 19.35           O
ANISOU  434  O5'  DT B  20     2986   1006   3359   -369    206   -141       O
ATOM    435  C5'  DT B  20       4.091   6.648  -8.524  1.00 20.44           C
ANISOU  435  C5'  DT B  20     2626   2094   3044   -341     37    -59       C
ATOM    436  C4'  DT B  20       4.193   5.564  -9.574  1.00 18.85           C
ANISOU  436  C4'  DT B  20     2653   1342   3164   -356    -31   -122       C
ATOM    437  O4'  DT B  20       2.904   4.964  -9.816  1.00 18.38           O
ANISOU  437  O4'  DT B  20     2628   1107   3248    -44     43    152       O
ATOM    438  C3'  DT B  20       4.721   6.022 -10.937  1.00 20.07           C
ANISOU  438  C3'  DT B  20     2601   1725   3298   -164     99   -202       C
ATOM    439  O3'  DT B  20       5.718   5.067 -11.262  1.00 22.30           O
ANISOU  439  O3'  DT B  20     2626   2342   3502     13   -152   -245       O
ATOM    440  C2'  DT B  20       3.495   5.960 -11.834  1.00 17.95           C
ANISOU  440  C2'  DT B  20     2782    801   3236    -11    127     36       C
ATOM    441  C1'  DT B  20       2.687   4.834 -11.202  1.00 17.77           C
ANISOU  441  C1'  DT B  20     2752    794   3205    -93    -25     88       C
ATOM    442  N1   DT B  20       1.232   4.831 -11.406  1.00 16.82           N
ANISOU  442  N1   DT B  20     2693    467   3229     66    105    132       N
ATOM    443  C2   DT B  20       0.606   3.663 -11.790  1.00 16.13           C
ANISOU  443  C2   DT B  20     2608    531   2986   -151    -18    120       C
ATOM    444  O2   DT B  20       1.211   2.658 -12.120  1.00 17.23           O
ANISOU  444  O2   DT B  20     2600    586   3359    205    -47     58       O
ATOM    445  N3   DT B  20      -0.764   3.735 -11.823  1.00 16.14           N
ANISOU  445  N3   DT B  20     2550    496   3084    -60    -71    103       N
ATOM    446  C4   DT B  20      -1.557   4.817 -11.474  1.00 16.52           C
ANISOU  446  C4   DT B  20     2600    548   3129    298    -76    166       C
ATOM    447  O4   DT B  20      -2.770   4.714 -11.503  1.00 16.72           O
ANISOU  447  O4   DT B  20     2573    646   3132    243   -113     61       O
```

```
ATOM    448  C5   DT B  20      -0.831   5.998 -11.052  1.00 16.26           C
ANISOU  448  C5   DT B  20     2713    262   3200     67   -124    144       C
ATOM    449  C7   DT B  20      -1.597   7.230 -10.678  1.00 17.35           C
ANISOU  449  C7   DT B  20     2748    663   3179     60    -54      0       C
ATOM    450  C6   DT B  20       0.505   5.939 -11.029  1.00 16.49           C
ANISOU  450  C6   DT B  20     2647    355   3263    110   -117    168       C
HETATM  451  P   5HC B  21       6.694   5.225 -12.477  1.00 24.00           P
ANISOU  451  P   5HC B  21     2655   2626   3838   -398     66   -527       P
HETATM  452  OP1 5HC B  21       7.842   4.371 -12.214  1.00 27.67           O
ANISOU  452  OP1 5HC B  21     2702   3594   4215   -121    -69   -954       O
HETATM  453  OP2 5HC B  21       6.834   6.610 -12.784  1.00 27.44           O
ANISOU  453  OP2 5HC B  21     3388   3242   3795  -1060    507     38       O
HETATM  454  O5' 5HC B  21       5.884   4.644 -13.707  1.00 20.89           O
ANISOU  454  O5' 5HC B  21     2638   1852   3447   -201    121   -290       O
HETATM  455  C5' 5HC B  21       5.620   3.252 -13.748  1.00 20.57           C
ANISOU  455  C5' 5HC B  21     2560   1628   3627     21    107   -257       C
HETATM  456  C4' 5HC B  21       4.578   2.959 -14.798  1.00 21.24           C
ANISOU  456  C4' 5HC B  21     2510   1946   3614    175      6   -280       C
HETATM  457  O4' 5HC B  21       3.317   3.532 -14.439  1.00 19.62           O
ANISOU  457  O4' 5HC B  21     2431   1633   3390     91      0     28       O
HETATM  458  C3' 5HC B  21       4.875   3.524 -16.176  1.00 21.71           C
ANISOU  458  C3' 5HC B  21     2647   1874   3725    211     35   -435       C
HETATM  459  O3' 5HC B  21       5.438   2.419 -16.864  1.00 25.37           O
ANISOU  459  O3' 5HC B  21     2658   3105   3873    754   -337   -623       O
HETATM  460  C2' 5HC B  21       3.502   3.901 -16.723  1.00 21.90           C
ANISOU  460  C2' 5HC B  21     2669   2326   3325    389    -88   -374       C
HETATM  461  C1' 5HC B  21       2.548   3.507 -15.618  1.00 20.59           C
ANISOU  461  C1' 5HC B  21     2455   1877   3490    234   -191   -117       C
HETATM  462  N1  5HC B  21       1.381   4.370 -15.400  1.00 18.70           N
ANISOU  462  N1  5HC B  21     2480   1304   3321    295    -93     52       N
HETATM  463  C2  5HC B  21       0.108   3.820 -15.537  1.00 18.13           C
ANISOU  463  C2  5HC B  21     2495   1254   3137      3    -77   -234       C
HETATM  464  O2  5HC B  21      -0.001   2.656 -15.940  1.00 18.29           O
ANISOU  464  O2  5HC B  21     2452   1078   3420     69    -79   -187       O
HETATM  465  N3  5HC B  21      -0.969   4.572 -15.238  1.00 16.88           N
ANISOU  465  N3  5HC B  21     2322   1020   3070    -13    -47    109       N
HETATM  466  C4  5HC B  21      -0.811   5.825 -14.815  1.00 16.90           C
ANISOU  466  C4  5HC B  21     2513    921   2987     78     -8     58       C
HETATM  467  N4  5HC B  21      -1.903   6.521 -14.514  1.00 17.96           N
ANISOU  467  N4  5HC B  21     2508   1105   3210     17    -38     95       N
HETATM  468  C5  5HC B  21       0.477   6.421 -14.689  1.00 17.49           C
ANISOU  468  C5  5HC B  21     2560    805   3280    -84    -23    360       C
HETATM  469  C5M 5HC B  21       0.690   7.876 -14.210  1.00 18.79           C
ANISOU  469  C5M 5HC B  21     2462   1763   2912    -81    -28    148       C
HETATM  470  O5  5HC B  21       0.201   8.787 -15.220  1.00 20.00           O
ANISOU  470  O5  5HC B  21     2760   1691   3148    121    -80    174       O
HETATM  471  C6  5HC B  21       1.535   5.658 -14.968  1.00 17.74           C
ANISOU  471  C6  5HC B  21     2518   1004   3219    245    -10     61       C
ATOM    472  P    DG B  22       5.957   2.505 -18.307  1.00 29.70           P
ANISOU  472  P    DG B  22     2495   4802   3987    179    -24   -916       P
ATOM    473  OP1  DG B  22       6.885   1.386 -18.491  1.00 33.39           O
ANISOU  473  OP1  DG B  22     2666   5472   4546    950   -410  -1860       O
ATOM    474  OP2  DG B  22       6.319   3.881 -18.636  1.00 29.57           O
ANISOU  474  OP2  DG B  22     2597   5036   3601   -368    417   -212       O
ATOM    475  O5'  DG B  22       4.684   2.285 -19.230  1.00 27.62           O
ANISOU  475  O5'  DG B  22     2380   4235   3879    238   -180   -964       O
```

```
ATOM    476  C5'  DG B  22       4.041   1.033 -19.269  1.00 25.38           C
ANISOU  476  C5'  DG B  22     2478   3560   3604    547   -109   -728       C
ATOM    477  C4'  DG B  22       2.800   1.148 -20.118  1.00 21.58           C
ANISOU  477  C4'  DG B  22     2557   2108   3534    415   -318   -873       C
ATOM    478  O4'  DG B  22       1.813   1.979 -19.462  1.00 19.97           O
ANISOU  478  O4'  DG B  22     2448   1682   3458    449   -123   -254       O
ATOM    479  C3'  DG B  22       3.025   1.782 -21.488  1.00 20.21           C
ANISOU  479  C3'  DG B  22     2191   1971   3515    214    -50   -683       C
ATOM    480  O3'  DG B  22       2.262   1.031 -22.433  1.00 20.61           O
ANISOU  480  O3'  DG B  22     2269   1932   3627    133   -105   -442       O
ATOM    481  C2'  DG B  22       2.396   3.156 -21.352  1.00 19.75           C
ANISOU  481  C2'  DG B  22     2472   1664   3366     -7    -58   -382       C
ATOM    482  C1'  DG B  22       1.252   2.830 -20.434  1.00 18.52           C
ANISOU  482  C1'  DG B  22     2424   1213   3400    325    -28   -185       C
ATOM    483  N9   DG B  22       0.613   3.950 -19.753  1.00 17.92           N
ANISOU  483  N9   DG B  22     2318   1303   3187    186   -168   -128       N
ATOM    484  C8   DG B  22       1.194   5.119 -19.330  1.00 20.22           C
ANISOU  484  C8   DG B  22     2494   1881   3308    127   -124   -372       C
ATOM    485  N7   DG B  22       0.358   5.917 -18.732  1.00 18.58           N
ANISOU  485  N7   DG B  22     2286   1570   3202    186    -63   -149       N
ATOM    486  C5   DG B  22      -0.851   5.237 -18.766  1.00 16.85           C
ANISOU  486  C5   DG B  22     2461    764   3175    -59   -124    -61       C
ATOM    487  C6   DG B  22      -2.133   5.615 -18.294  1.00 17.29           C
ANISOU  487  C6   DG B  22     2341   1160   3065    -16   -116     61       C
ATOM    488  O6   DG B  22      -2.456   6.638 -17.702  1.00 17.63           O
ANISOU  488  O6   DG B  22     2471    957   3267    212    -83    -42       O
ATOM    489  N1   DG B  22      -3.083   4.639 -18.551  1.00 16.41           N
ANISOU  489  N1   DG B  22     2246    966   3020    211    -17    -41       N
ATOM    490  C2   DG B  22      -2.835   3.452 -19.185  1.00 16.45           C
ANISOU  490  C2   DG B  22     2246    992   3009    204    -88    171       C
ATOM    491  N2   DG B  22      -3.881   2.646 -19.351  1.00 16.72           N
ANISOU  491  N2   DG B  22     2306   1027   3016    167    -56    126       N
ATOM    492  N3   DG B  22      -1.646   3.086 -19.631  1.00 17.04           N
ANISOU  492  N3   DG B  22     2258   1154   3062     90    -47     89       N
ATOM    493  C4   DG B  22      -0.709   4.023 -19.396  1.00 17.26           C
ANISOU  493  C4   DG B  22     2328   1080   3149    -15   -168     48       C
ATOM    494  P    DC B  23       2.756   0.808 -23.886  1.00 19.80           P
ANISOU  494  P    DC B  23     1990   2396   3136     52     48   -425       P
ATOM    495  OP1  DC B  23       3.873  -0.136 -23.892  1.00 23.05           O
ANISOU  495  OP1  DC B  23     2205   2822   3730    302     28   -727       O
ATOM    496  OP2  DC B  23       2.907   2.108 -24.494  1.00 21.38           O
ANISOU  496  OP2  DC B  23     2474   2333   3317   -312     90   -202       O
ATOM    497  O5'  DC B  23       1.485   0.108 -24.533  1.00 19.03           O
ANISOU  497  O5'  DC B  23     2148   2068   3012     83    -46   -327       O
ATOM    498  C5'  DC B  23       0.972  -1.125 -24.013  1.00 19.03           C
ANISOU  498  C5'  DC B  23     2107   1944   3178    140   -176   -298       C
ATOM    499  C4'  DC B  23      -0.520  -1.031 -23.792  1.00 18.25           C
ANISOU  499  C4'  DC B  23     2121   1919   2892    203    -92    -92       C
ATOM    500  O4'  DC B  23      -0.820  -0.072 -22.761  1.00 18.05           O
ANISOU  500  O4'  DC B  23     2136   1869   2851    212    -37   -139       O
ATOM    501  C3'  DC B  23      -1.349  -0.568 -24.982  1.00 18.05           C
ANISOU  501  C3'  DC B  23     2269   1705   2882    107   -179   -203       C
ATOM    502  O3'  DC B  23      -1.636  -1.658 -25.842  1.00 19.30           O
ANISOU  502  O3'  DC B  23     2450   1634   3248    254   -200   -323       O
ATOM    503  C2'  DC B  23      -2.620  -0.106 -24.307  1.00 17.85           C
ANISOU  503  C2'  DC B  23     2150   1754   2876    100   -106    -92       C
```

```
ATOM    504  C1'  DC B  23      -2.087   0.529 -23.050  1.00 17.71           C
ANISOU  504  C1'  DC B  23     2004   1751   2973    132    -84    -50       C
ATOM    505  N1   DC B  23      -1.915   1.990 -23.046  1.00 17.09           N
ANISOU  505  N1   DC B  23     2003   1687   2801     82   -101    -86       N
ATOM    506  C2   DC B  23      -2.998   2.782 -22.663  1.00 16.37           C
ANISOU  506  C2   DC B  23     2005   1694   2520    124    -95     16       C
ATOM    507  O2   DC B  23      -4.094   2.246 -22.477  1.00 17.16           O
ANISOU  507  O2   DC B  23     1978   1718   2824    172    -73    -68       O
ATOM    508  N3   DC B  23      -2.815   4.112 -22.494  1.00 17.01           N
ANISOU  508  N3   DC B  23     2025   1739   2698     59   -102    -23       N
ATOM    509  C4   DC B  23      -1.621   4.656 -22.735  1.00 17.14           C
ANISOU  509  C4   DC B  23     2004   1755   2752     53      3    -79       C
ATOM    510  N4   DC B  23      -1.475   5.963 -22.518  1.00 17.90           N
ANISOU  510  N4   DC B  23     2006   1811   2983     74    -20   -144       N
ATOM    511  C5   DC B  23      -0.515   3.877 -23.176  1.00 17.98           C
ANISOU  511  C5   DC B  23     2070   1873   2888     45     80   -238       C
ATOM    512  C6   DC B  23      -0.700   2.561 -23.303  1.00 17.79           C
ANISOU  512  C6   DC B  23     1996   1873   2888     64      0    -93       C
ATOM    513  P    DG B  24      -2.027  -1.440 -27.343  1.00 19.52           P
ANISOU  513  P    DG B  24     2302   1945   3169    198   -184   -386       P
ATOM    514  OP1  DG B  24      -1.938  -2.767 -27.981  1.00 22.14           O
ANISOU  514  OP1  DG B  24     2561   2042   3808    445   -462   -698       O
ATOM    515  OP2  DG B  24      -1.251  -0.302 -27.899  1.00 20.89           O
ANISOU  515  OP2  DG B  24     2455   2431   3050     53     -3   -418       O
ATOM    516  O5'  DG B  24      -3.539  -0.949 -27.293  1.00 18.90           O
ANISOU  516  O5'  DG B  24     2301   1915   2961     82   -157   -173       O
ATOM    517  C5'  DG B  24      -4.611  -1.815 -26.891  1.00 17.95           C
ANISOU  517  C5'  DG B  24     2247   1630   2942     53   -263   -103       C
ATOM    518  C4'  DG B  24      -5.870  -1.000 -26.711  1.00 17.69           C
ANISOU  518  C4'  DG B  24     2377   1411   2932     -2   -283    -91       C
ATOM    519  O4'  DG B  24      -5.665  -0.029 -25.664  1.00 17.05           O
ANISOU  519  O4'  DG B  24     2224   1476   2777    -22   -143    -86       O
ATOM    520  C3'  DG B  24      -6.310  -0.178 -27.917  1.00 17.37           C
ANISOU  520  C3'  DG B  24     2233   1538   2827    -46   -220   -155       C
ATOM    521  O3'  DG B  24      -7.058  -1.008 -28.808  1.00 18.53           O
ANISOU  521  O3'  DG B  24     2439   1531   3068    -43   -312   -224       O
ATOM    522  C2'  DG B  24      -7.108   0.944 -27.279  1.00 17.50           C
ANISOU  522  C2'  DG B  24     2248   1517   2883    -36   -100    -54       C
ATOM    523  C1'  DG B  24      -6.357   1.186 -25.989  1.00 17.02           C
ANISOU  523  C1'  DG B  24     2250   1418   2796   -106    -56     14       C
ATOM    524  N9   DG B  24      -5.359   2.251 -26.043  1.00 16.84           N
ANISOU  524  N9   DG B  24     2168   1485   2746    -98   -147    -72       N
ATOM    525  C8   DG B  24      -4.063   2.145 -26.488  1.00 17.48           C
ANISOU  525  C8   DG B  24     2292   1665   2683   -142    -62    -34       C
ATOM    526  N7   DG B  24      -3.370   3.239 -26.316  1.00 17.33           N
ANISOU  526  N7   DG B  24     2205   1656   2721   -102    -52    -82       N
ATOM    527  C5   DG B  24      -4.269   4.128 -25.742  1.00 16.61           C
ANISOU  527  C5   DG B  24     1996   1670   2642    -48   -140     -1       C
ATOM    528  C6   DG B  24      -4.088   5.467 -25.307  1.00 16.94           C
ANISOU  528  C6   DG B  24     2173   1637   2625   -256   -124     36       C
ATOM    529  O6   DG B  24      -3.066   6.163 -25.358  1.00 18.23           O
ANISOU  529  O6   DG B  24     2273   1692   2960   -330    -79    -51       O
ATOM    530  N1   DG B  24      -5.262   5.999 -24.782  1.00 17.20           N
ANISOU  530  N1   DG B  24     2265   1581   2689   -266     27    -72       N
ATOM    531  C2   DG B  24      -6.458   5.334 -24.694  1.00 16.88           C
ANISOU  531  C2   DG B  24     2297   1476   2639   -204     31   -108       C
```

```
ATOM    532  N2   DG B  24      -7.485   6.014 -24.173  1.00 17.90           N
ANISOU  532  N2   DG B  24     2389   1516   2894   -171    175   -165       N
ATOM    533  N3   DG B  24      -6.632   4.076 -25.066  1.00 16.72           N
ANISOU  533  N3   DG B  24     2135   1535   2681   -170    -42   -143       N
ATOM    534  C4   DG B  24      -5.501   3.535 -25.570  1.00 16.50           C
ANISOU  534  C4   DG B  24     2214   1528   2527   -210    -93    -59       C
TER     535       DG B  24
HETATM  536 MG    MG A 101      -2.090  10.717 -18.618  1.00 19.21          MG
ANISOU  536 MG    MG A 101     2532   2214   2553     82   -100     -7      MG
HETATM  537  N1  SPK B 101       3.058  13.484 -12.303  1.00 38.88           N
ANISOU  537  N1  SPK B 101     4846   4645   5280   -362   -350   -636       N
HETATM  538  C2  SPK B 101       2.338  13.371 -11.029  1.00 35.81           C
ANISOU  538  C2  SPK B 101     4971   4257   4376    -16   -185   -373       C
HETATM  539  C3  SPK B 101       3.293  12.989  -9.913  1.00 30.62           C
ANISOU  539  C3  SPK B 101     4522   3285   3826    364    193   -399       C
HETATM  540  C4  SPK B 101       2.535  12.987  -8.612  1.00 33.05           C
ANISOU  540  C4  SPK B 101     4292   3869   4397   -391   -165    331       C
HETATM  541  N5  SPK B 101       3.342  12.640  -7.440  1.00 32.00           N
ANISOU  541  N5  SPK B 101     4490   3570   4098   -253    238    134       N
HETATM  542  C6  SPK B 101       2.473  12.579  -6.256  1.00 37.35           C
ANISOU  542  C6  SPK B 101     5311   5164   3712    233    381    414       C
HETATM  543  C7  SPK B 101       3.267  12.078  -5.063  1.00 32.55           C
ANISOU  543  C7  SPK B 101     4172   4341   3852      6     33     -2       C
HETATM  544  C8  SPK B 101       2.320  11.846  -3.906  1.00 33.83           C
ANISOU  544  C8  SPK B 101     4600   4791   3461   -732     35    -96       C
HETATM  545  C9  SPK B 101       3.011  11.338  -2.676  1.00 37.83           C
ANISOU  545  C9  SPK B 101     4599   5260   4513   -792    939   1191       C
HETATM  546  N10 SPK B 101       2.252  11.856  -1.535  1.00 38.86           N
ANISOU  546  N10 SPK B 101     5459   3604   5703   -727    755    383       N
HETATM  547  C11 SPK B 101       2.948  11.432  -0.327  1.00 70.65           C
ANISOU  547  C11 SPK B 101    14443  11216  14481  -2296  -5727   2814       C
HETATM  548  N1 ASPK B 102       3.224  -6.111   4.543  0.50 30.43           N
ANISOU  548  N1 ASPK B 102     3750   3385   4425    211    -11    -18       N
HETATM  549  N1 BSPK B 102       3.498  -4.976   2.823  0.30 33.20           N
ANISOU  549  N1 BSPK B 102     4193   3925   4495    227    -95     60       N
HETATM  550  C2 ASPK B 102       3.134  -4.633   4.562  0.50 27.48           C
ANISOU  550  C2 ASPK B 102     3553   3273   3613    141     79     -3       C
HETATM  551  C2 BSPK B 102       3.170  -4.714   4.292  0.30 29.88           C
ANISOU  551  C2 BSPK B 102     3538   3557   4256    627   -479    386       C
HETATM  552  C3 ASPK B 102       4.449  -4.054   5.015  0.50 32.21           C
ANISOU  552  C3 ASPK B 102     3951   4325   3961   -149   -196   -275       C
HETATM  553  C3 BSPK B 102       4.374  -4.182   5.043  0.30 33.74           C
ANISOU  553  C3 BSPK B 102     4024   4514   4280     55   -283   -179       C
HETATM  554  C4 ASPK B 102       4.318  -2.673   5.567  0.50 29.32           C
ANISOU  554  C4 ASPK B 102     3401   3734   4003   -148    123    320       C
HETATM  555  C4 BSPK B 102       4.280  -2.739   5.455  0.30 31.94           C
ANISOU  555  C4 BSPK B 102     3467   4429   4240    167    -29      3       C
HETATM  556  N5 ASPK B 102       5.570  -1.933   5.760  0.50 31.92           N
ANISOU  556  N5 ASPK B 102     3612   4300   4216   -311   -138    -33       N
HETATM  557  N5 BSPK B 102       5.530  -2.012   5.745  0.30 31.37           N
ANISOU  557  N5 BSPK B 102     3674   4276   3968     69     39    -51       N
HETATM  558  C6 ASPK B 102       6.567  -2.158   4.741  0.50 32.61           C
ANISOU  558  C6 ASPK B 102     3649   4545   4196    643   -457    171       C
HETATM  559  C6 BSPK B 102       6.553  -2.175   4.736  0.30 25.06           C
ANISOU  559  C6 BSPK B 102     2603   3407   3511    254   -586    485       C
HETATM  560  C7 ASPK B 102       7.575  -1.044   4.737  0.50 36.82           C
```

```
ANISOU  560  C7 ASPK B 102     4652   4911   4424    138   -342    515       C
HETATM  561  C7 BSPK B 102       7.544  -1.046   4.752  0.30 29.77           C
ANISOU  561  C7 BSPK B 102     3304   4107   3898   -332   -215    360       C
HETATM  562  C8 ASPK B 102       8.835  -1.685   4.254  0.50 38.26           C
ANISOU  562  C8 ASPK B 102     4654   5107   4775   -127   -146    186       C
HETATM  563  C8 BSPK B 102       8.820  -1.673   4.286  0.30 33.44           C
ANISOU  563  C8 BSPK B 102     3414   4944   4348    147   -380     94       C
HETATM  564  C9 ASPK B 102       9.757  -0.809   3.453  0.50 37.48           C
ANISOU  564  C9 ASPK B 102     4887   4384   4966   -107   -354    144       C
HETATM  565  C9 BSPK B 102       9.731  -0.809   3.459  0.30 38.30           C
ANISOU  565  C9 BSPK B 102     4718   4826   5005    -13    -52    353       C
HETATM  566  N10ASPK B 102      10.279   0.219   4.320  0.50 33.65           N
ANISOU  566  N10ASPK B 102     3823   4509   4450   -848   -242    449       N
HETATM  567  N10BSPK B 102      10.276   0.226   4.304  0.30 38.11           N
ANISOU  567  N10BSPK B 102     4656   5177   4646    155   -116    339       N
HETATM  568  C11ASPK B 102       9.899   1.569   3.898  0.50 38.40           C
ANISOU  568  C11ASPK B 102     5004   4488   5096   -395    -44    343       C
HETATM  569  C11BSPK B 102       9.883   1.567   3.867  0.30 38.57           C
ANISOU  569  C11BSPK B 102     4897   4860   4897   -260    322    395       C
HETATM  570  C12ASPK B 102      10.427   1.801   2.514  0.50 37.63           C
ANISOU  570  C12ASPK B 102     4482   4950   4866   -177     32     51       C
HETATM  571  C12BSPK B 102      10.415   1.785   2.482  0.30 36.34           C
ANISOU  571  C12BSPK B 102     4199   4787   4822   -131    255     89       C
HETATM  572  C13ASPK B 102      10.114   3.126   1.874  0.50 39.84           C
ANISOU  572  C13ASPK B 102     5072   4906   5157   -194     93    137       C
HETATM  573  C13BSPK B 102       9.903   2.979   1.718  0.30 38.21           C
ANISOU  573  C13BSPK B 102     4740   4714   5064   -137    192     45       C
HETATM  574  N14ASPK B 102      10.183   2.948   0.411  0.50 38.49           N
ANISOU  574  N14ASPK B 102     4750   4733   5140   -373     99     89       N
HETATM  575  N14BSPK B 102      11.050   3.887   1.584  0.30 36.30           N
ANISOU  575  N14BSPK B 102     4259   4545   4988    236    510    -56       N
HETATM  576  O   HOH A 201       5.846  -7.853 -14.025  1.00 21.22           O
ANISOU  576  O   HOH A 201     3182   1741   3137    -44    332   -111       O
HETATM  577  O   HOH A 202       2.577  -4.929  -0.769  1.00 26.57           O
ANISOU  577  O   HOH A 202     4030   2357   3708    887   -159   -105       O
HETATM  578  O   HOH A 203       2.952   2.709  -7.560  1.00 18.18           O
ANISOU  578  O   HOH A 203     2248   1663   2997     45     10     86       O
HETATM  579  O   HOH A 204      -1.508   3.996  -0.035  1.00 18.25           O
ANISOU  579  O   HOH A 204     2450   1511   2970    -27    246     21       O
HETATM  580  O   HOH A 205     -10.798   2.045 -13.263  1.00 31.40           O
ANISOU  580  O   HOH A 205     3285   4050   4592   -510    933   -673       O
HETATM  581  O   HOH A 206      -2.267   2.679   4.092  1.00 20.23           O
ANISOU  581  O   HOH A 206     2738   1943   3005    172     27   -121       O
HETATM  582  O   HOH A 207      -6.922   0.503 -10.931  1.00 27.30           O
ANISOU  582  O   HOH A 207     3088   3739   3546    162     39   -501       O
HETATM  583  O   HOH A 208      -1.957   9.193 -19.974  1.00 19.11           O
ANISOU  583  O   HOH A 208     2565   1570   3125    -17   -190    101       O
HETATM  584  O   HOH A 209       2.769  -6.156  -8.572  1.00 22.56           O
ANISOU  584  O   HOH A 209     2926   1821   3823    672    445    512       O
HETATM  585  O   HOH A 210       4.099   0.826   7.490  1.00 23.58           O
ANISOU  585  O   HOH A 210     2554   2720   3683   -272   -259     47       O
HETATM  586  O   HOH A 211       6.435   0.997   1.218  1.00 36.42           O
ANISOU  586  O   HOH A 211     4374   3612   5850    940   -721   -261       O
HETATM  587  O   HOH A 212      -4.084  -0.970 -17.389  1.00 21.97           O
ANISOU  587  O   HOH A 212     3127   1964   3255    174    -16    -15       O
HETATM  588  O   HOH A 213      -0.767  -5.435  -4.004  1.00 30.33           O
```

```
ANISOU  588  O   HOH A 213     5424   1742   4357   -406    844   -244       O
HETATM  589  O   HOH A 214      -1.733  -6.733 -11.114  1.00 24.67           O
ANISOU  589  O   HOH A 214     3046   1672   4655    140   -115   -151       O
HETATM  590  O   HOH A 215      -2.796   4.676   2.374  1.00 22.82           O
ANISOU  590  O   HOH A 215     3053   1953   3664    310    366     -9       O
HETATM  591  O   HOH A 216      -8.374   5.451 -12.923  1.00 25.51           O
ANISOU  591  O   HOH A 216     3009   2705   3979    503     21   -505       O
HETATM  592  O   HOH A 217      -0.165  -1.774 -15.968  1.00 31.40           O
ANISOU  592  O   HOH A 217     5557   2131   4241    482    996      8       O
HETATM  593  O   HOH A 218      -1.321   8.773   4.715  1.00 35.18           O
ANISOU  593  O   HOH A 218     5814   2277   5276   -569   -874    216       O
HETATM  594  O   HOH A 219      -3.958  -4.276  -6.741  1.00 28.51           O
ANISOU  594  O   HOH A 219     3443   2564   4822   -628    870   -729       O
HETATM  595  O   HOH A 220      -4.851  -3.816  -9.974  1.00 36.03           O
ANISOU  595  O   HOH A 220     2850   3864   6973   -556     84   1053       O
HETATM  596  O   HOH A 221     -15.519   4.872 -20.124  1.00 36.79           O
ANISOU  596  O   HOH A 221     2585   5781   5611    380   -358    254       O
HETATM  597  O   HOH A 222      -0.280  -5.917 -18.463  1.00 30.59           O
ANISOU  597  O   HOH A 222     5292   2197   4135    418   -184     23       O
HETATM  598  O   HOH A 223      -8.216  11.712 -17.811  1.00 26.78           O
ANISOU  598  O   HOH A 223     3520   2790   3863   -502    712   -212       O
HETATM  599  O   HOH A 224      -6.103  -1.516  -8.798  1.00 28.55           O
ANISOU  599  O   HOH A 224     3426   3361   4060   -613    371   -337       O
HETATM  600  O   HOH A 225      -4.251   3.582  12.794  1.00 27.78           O
ANISOU  600  O   HOH A 225     3656   2479   4418    399    315    706       O
HETATM  601  O   HOH A 226      -5.654  11.954 -25.879  1.00 24.74           O
ANISOU  601  O   HOH A 226     3458   2284   3657   -788    -10    210       O
HETATM  602  O   HOH A 227       6.060  -4.783  -5.268  1.00 36.77           O
ANISOU  602  O   HOH A 227     4834   3463   5673   1354    892    588       O
HETATM  603  O   HOH A 228      -5.742   9.863 -15.940  1.00 32.89           O
ANISOU  603  O   HOH A 228     5975   2598   3923    863   -154   -265       O
HETATM  604  O   HOH A 229      -0.067   2.401  11.646  1.00 35.80           O
ANISOU  604  O   HOH A 229     4458   4340   4803  -1673    215  -1165       O
HETATM  605  O   HOH A 230       6.061  -2.940  -3.236  1.00 34.36           O
ANISOU  605  O   HOH A 230     4231   3400   5424   1650   -789   -536       O
HETATM  606  O   HOH A 231       5.432  -0.922   2.438  1.00 25.28           O
ANISOU  606  O   HOH A 231     2882   3171   3552    587   -381    -97       O
HETATM  607  O   HOH A 232      -5.381   7.425 -13.438  1.00 28.30           O
ANISOU  607  O   HOH A 232     4590   2609   3554   1349   -203   -148       O
HETATM  608  O   HOH A 233      -8.708  10.080 -15.776  1.00 43.06           O
ANISOU  608  O   HOH A 233     4784   5110   6465   1451  -1040  -1009       O
HETATM  609  O   HOH A 234      -4.313  14.386 -26.084  1.00 41.33           O
ANISOU  609  O   HOH A 234     7568   4570   3564  -4477   1126   -539       O
HETATM  610  O   HOH A 235      -0.954  -3.255 -18.162  1.00 32.72           O
ANISOU  610  O   HOH A 235     5048   3045   4339    118   -588   -609       O
HETATM  611  O   HOH A 236      -0.305   9.442 -22.178  1.00 26.37           O
ANISOU  611  O   HOH A 236     3183   2874   3961   -643    195   -394       O
HETATM  612  O   HOH A 237      -1.199  12.650 -24.605  1.00 28.72           O
ANISOU  612  O   HOH A 237     3654   2552   4707   -297   -115   -503       O
HETATM  613  O   HOH A 238      -7.144   5.616   4.399  1.00 31.43           O
ANISOU  613  O   HOH A 238     4570   2735   4637    571    633   -223       O
HETATM  614  O   HOH A 239      -2.967   7.314   3.066  1.00 30.50           O
ANISOU  614  O   HOH A 239     4921   2268   4399     15    165   -517       O
HETATM  615  O   HOH A 240      -7.778   8.335 -13.432  1.00 38.02           O
ANISOU  615  O   HOH A 240     4855   4692   4898   1303   -488   -294       O
HETATM  616  O   HOH A 241       0.290  -8.273  -9.973  1.00 27.17           O
```

```
ANISOU  616  O   HOH A 241     3516   3113   3691   1182    313    277       O
HETATM  617  O   HOH A 242       2.347   9.016   6.032  1.00 43.86           O
ANISOU  617  O   HOH A 242     6442   3649   6572  -1681    628   -142       O
HETATM  618  O   HOH A 243       5.830   3.305   3.984  1.00 39.16           O
ANISOU  618  O   HOH A 243     3858   4989   6031   -895  -1710    952       O
HETATM  619  O   HOH A 244       7.765   5.493   1.254  1.00 64.61           O
ANISOU  619  O   HOH A 244     5470   5662  13417   1054    577  -2395       O
HETATM  620  O   HOH A 245       7.623   6.811  -1.183  1.00 55.02           O
ANISOU  620  O   HOH A 245     6938   6040   7925   -586   3284   -890       O
HETATM  621  O   HOH A 246       8.861   0.126  -9.839  1.00 45.23           O
ANISOU  621  O   HOH A 246     4136   6350   6699    560   1369   2188       O
HETATM  622  O   HOH A 247       9.514   3.058  -6.160  1.00 44.59           O
ANISOU  622  O   HOH A 247     6123   4693   6125   -204   -665  -1094       O
HETATM  623  O   HOH A 248       9.951   0.432  -5.269  1.00 45.44           O
ANISOU  623  O   HOH A 248     3065   7107   7093    -96   -576  -3114       O
HETATM  624  O   HOH A 249      -4.168  10.730 -18.627  1.00 20.60           O
ANISOU  624  O   HOH A 249     2855   1718   3252    221   -173     14       O
HETATM  625  O   HOH A 250      -3.180  11.072 -26.087  1.00 30.86           O
ANISOU  625  O   HOH A 250     4985   2517   4222   -560    408    258       O
HETATM  626  O   HOH A 251      -2.466  13.105 -28.138  1.00 46.90           O
ANISOU  626  O   HOH A 251     8693   3807   5319   1082    994   1116       O
HETATM  627  O   HOH A 252     -10.487   8.382 -15.000  0.50 32.50           O
ANISOU  627  O   HOH A 252     3781   4716   3848    552    604   1337       O
HETATM  628  O   HOH A 253     -13.058  -3.117 -17.601  1.00 40.37           O
ANISOU  628  O   HOH A 253     4272   3803   7261   -656   -151  -1232       O
HETATM  629  O   HOH A 254     -12.401  -3.541 -13.825  1.00 55.40           O
ANISOU  629  O   HOH A 254     9652   4972   6425   -855     53   1702       O
HETATM  630  O   HOH A 255      -9.737   0.565 -11.326  1.00 36.24           O
ANISOU  630  O   HOH A 255     3346   5962   4462    495    479     80       O
HETATM  631  O   HOH A 256      -9.996  -2.331 -11.645  1.00 51.95           O
ANISOU  631  O   HOH A 256     8425   6025   5287  -1779   1536    844       O
HETATM  632  O   HOH A 257     -10.303   3.992   9.471  1.00 33.34           O
ANISOU  632  O   HOH A 257     3606   2601   6458    -72   -910   -165       O
HETATM  633  O   HOH A 258     -10.265   2.496  11.682  1.00 36.71           O
ANISOU  633  O   HOH A 258     4112   3207   6627    945    -19    404       O
HETATM  634  O   HOH A 259       5.075  -5.162  -2.051  1.00 44.14           O
ANISOU  634  O   HOH A 259     5345   6083   5341   2915   1061   1118       O
HETATM  635  O   HOH A 260       7.799  -1.435  -1.628  1.00 44.95           O
ANISOU  635  O   HOH A 260     5505   6608   4964   2344  -1518   -806       O
HETATM  636  O   HOH A 261      -8.208  16.400 -20.331  1.00 36.90           O
ANISOU  636  O   HOH A 261     4113   4015   5892   -812    731  -1124       O
HETATM  637  O   HOH A 262     -10.191  15.734 -18.403  0.50 34.87           O
ANISOU  637  O   HOH A 262     3782   3441   6023    169    377   -418       O
HETATM  638  O   HOH A 263     -15.642   8.143 -19.912  1.00 43.10           O
ANISOU  638  O   HOH A 263     4192   6048   6136   -994     64   1385       O
HETATM  639  O   HOH A 264      -8.234   4.649  12.473  1.00113.09           O
ANISOU  639  O   HOH A 264    26290   7910   8768   6416   7413    670       O
HETATM  640  O   HOH A 265       7.201   4.585  -7.653  1.00 45.43           O
ANISOU  640  O   HOH A 265     4940   8013   4305  -2529    266   -571       O
HETATM  641  O   HOH A 266     -11.054  -4.604 -17.070  0.50 33.63           O
ANISOU  641  O   HOH A 266     4674   2648   5454   -509   -889     -4       O
HETATM  642  O   HOH A 267       1.471  -6.433 -20.456  0.50 35.63           O
ANISOU  642  O   HOH A 267     5505   3538   4494    249    279   -259       O
HETATM  643  O   HOH A 268       6.072  -4.148 -14.926  1.00 41.40           O
ANISOU  643  O   HOH A 268     6838   3456   5436   1281   2206    562       O
HETATM  644  O   HOH A 269       5.869   8.208  -1.107  0.50 45.75           O
```

```
ANISOU  644  O   HOH A 269     6524   4822   6037   -935   -927   -276       O
HETATM  645  O   HOH A 270      -0.028   9.926   9.373  0.50 51.17           O
ANISOU  645  O   HOH A 270     3836   9268   6339  -2288   1767  -3587       O
HETATM  646  O   HOH A 271      -7.384   6.482   6.969  1.00 34.53           O
ANISOU  646  O   HOH A 271     4442   2938   5740    766    766   -253       O
HETATM  647  O  AHOH A 272      -6.243  -7.856 -18.372  0.50 37.97           O
ANISOU  647  O  AHOH A 272     6359   3629   4438    175   -138   -807       O
HETATM  648  O  BHOH A 272      -5.609  -8.461 -16.705  0.50 38.45           O
ANISOU  648  O  BHOH A 272     5311   3226   6070   -317    -40    747       O
HETATM  649  O   HOH A 273       2.287   3.921   9.668  0.50 49.09           O
ANISOU  649  O   HOH A 273     8460   4255   5936     79   -391   2175       O
HETATM  650  O   HOH A 274      -1.140   6.737  11.085  0.50 30.77           O
ANISOU  650  O   HOH A 274     4300   3575   3814    209    497     20       O
HETATM  651  O   HOH A 275      -0.510  -7.490  -5.475  1.00 52.18           O
ANISOU  651  O   HOH A 275    10433   2578   6814   -760   2553  -1512       O
HETATM  652  O   HOH A 276     -12.897   5.720 -14.135  0.50 27.97           O
ANISOU  652  O   HOH A 276     3568   3527   3532    178    167   -309       O
HETATM  653  O   HOH A 277       3.719   8.141  -2.619  0.50 32.92           O
ANISOU  653  O   HOH A 277     4463   2763   5280   -940   -672   -522       O
HETATM  654  O   HOH A 278       4.551   7.369   7.001  0.50 36.36           O
ANISOU  654  O   HOH A 278     4808   2798   6206   -774    402    247       O
HETATM  655  O   HOH A 279       1.787   8.779  10.502  0.50 37.48           O
ANISOU  655  O   HOH A 279     4443   4398   5400    -87    -94    102       O
HETATM  656  O   HOH A 280     -11.605  17.221 -21.811  0.50 40.39           O
ANISOU  656  O   HOH A 280     4953   3736   6658   -693   1537   -354       O
HETATM  657  O   HOH A 281      -1.539   8.821   1.241  1.00 35.87           O
ANISOU  657  O   HOH A 281     5334   3512   4783     30    484     42       O
HETATM  658  O   HOH A 282       6.541  -4.764   1.452  1.00 53.26           O
ANISOU  658  O   HOH A 282     5674   7651   6909   2279  -1107   -355       O
HETATM  659  O   HOH A 283       0.942   8.099  -0.326  1.00 30.53           O
ANISOU  659  O   HOH A 283     4699   1918   4980   -345   -645   -305       O
HETATM  660  O   HOH B 201       1.208   4.167  -3.482  1.00 18.99           O
ANISOU  660  O   HOH B 201     2613   1473   3128     62    103     90       O
HETATM  661  O   HOH B 202       1.032  11.287 -14.460  1.00 29.99           O
ANISOU  661  O   HOH B 202     4798   2030   4563    -38  -1182    302       O
HETATM  662  O   HOH B 203      -4.929   6.390 -10.991  1.00 23.33           O
ANISOU  662  O   HOH B 203     2754   2153   3955    503      5    457       O
HETATM  663  O   HOH B 204      -2.808   9.187 -13.976  1.00 30.03           O
ANISOU  663  O   HOH B 204     3322   1883   6202    333    368    511       O
HETATM  664  O   HOH B 205       0.236   5.707  -1.401  1.00 21.69           O
ANISOU  664  O   HOH B 205     3031   1549   3661    -64    722    -43       O
HETATM  665  O   HOH B 206      -2.107   9.244 -17.147  1.00 20.68           O
ANISOU  665  O   HOH B 206     2824   1832   3198     94   -131     -1       O
HETATM  666  O   HOH B 207       1.120   8.388 -17.725  1.00 19.98           O
ANISOU  666  O   HOH B 207     2582   1816   3190     33   -140    -12       O
HETATM  667  O   HOH B 208      -2.765   9.025  -7.569  1.00 23.32           O
ANISOU  667  O   HOH B 208     3262   1631   3967    425   -139     13       O
HETATM  668  O   HOH B 209       3.323   4.503  -5.373  1.00 21.06           O
ANISOU  668  O   HOH B 209     2437   2347   3219   -208     69   -114       O
HETATM  669  O   HOH B 210       5.514  -2.950   9.131  1.00 27.08           O
ANISOU  669  O   HOH B 210     2599   3651   4036   -650   -139    360       O
HETATM  670  O   HOH B 211       1.549   9.196 -11.289  1.00 24.52           O
ANISOU  670  O   HOH B 211     3663   1671   3980   -455     96    -98       O
HETATM  671  O   HOH B 212      -7.533  -3.787 -28.440  1.00 21.94           O
ANISOU  671  O   HOH B 212     2876   1826   3633   -267    -79    -89       O
HETATM  672  O   HOH B 213      -6.950   6.216  -4.271  1.00 25.45           O
```

```
ANISOU  672  O   HOH B 213     3068   2790   3809    850     68    233       O
HETATM  673  O   HOH B 214      -0.843   2.370 -27.422  1.00 22.00           O
ANISOU  673  O   HOH B 214     2526   2356   3475     55     83   -121       O
HETATM  674  O   HOH B 215       2.859   0.560 -11.571  1.00 19.88           O
ANISOU  674  O   HOH B 215     2375   1753   3422    176    234    146       O
HETATM  675  O   HOH B 216      -7.216   5.223 -10.481  1.00 27.23           O
ANISOU  675  O   HOH B 216     3226   3055   4064    429   -298    191       O
HETATM  676  O   HOH B 217       1.364   3.522 -26.298  1.00 24.69           O
ANISOU  676  O   HOH B 217     2968   2684   3728   -333    506   -134       O
HETATM  677  O   HOH B 218      -7.253   1.815  -5.266  1.00 26.95           O
ANISOU  677  O   HOH B 218     2879   4109   3249    264     87    145       O
HETATM  678  O   HOH B 219     -12.908   0.487   1.932  1.00 27.48           O
ANISOU  678  O   HOH B 219     2510   4243   3687   -108    -44    820       O
HETATM  679  O   HOH B 220      -4.881   8.664  -9.367  1.00 26.12           O
ANISOU  679  O   HOH B 220     3294   2689   3939    444   -173   -253       O
HETATM  680  O   HOH B 221      -6.855  -2.260  -3.880  1.00 30.59           O
ANISOU  680  O   HOH B 221     3920   3587   4115   -877    342  -1077       O
HETATM  681  O   HOH B 222      -0.764  10.780  -8.470  1.00 28.92           O
ANISOU  681  O   HOH B 222     3894   2097   4994   -204    493    367       O
HETATM  682  O   HOH B 223       4.567   1.638  -9.587  1.00 21.53           O
ANISOU  682  O   HOH B 223     2373   2332   3473     92    154     -3       O
HETATM  683  O   HOH B 224      -9.427   2.142  -3.659  1.00 29.63           O
ANISOU  683  O   HOH B 224     3099   4067   4092    329    416    983       O
HETATM  684  O   HOH B 225       1.318   6.994 -22.868  1.00 27.93           O
ANISOU  684  O   HOH B 225     3055   3135   4419   -955    436   -283       O
HETATM  685  O   HOH B 226       4.322  -2.546 -24.998  1.00 33.62           O
ANISOU  685  O   HOH B 226     3418   3595   5760    841   -383  -1419       O
HETATM  686  O   HOH B 227      -0.568  10.709 -12.365  1.00 38.66           O
ANISOU  686  O   HOH B 227     4687   2774   7227    -69  -1041    -59       O
HETATM  687  O   HOH B 228      -1.492  -5.561  -1.404  1.00 91.33           O
ANISOU  687  O   HOH B 228    20165   4937   9597   1784   2806  -1276       O
HETATM  688  O   HOH B 229      -6.143   8.785  -5.168  1.00 31.08           O
ANISOU  688  O   HOH B 229     3850   3117   4841    334     42   1170       O
HETATM  689  O   HOH B 230      -6.494  -0.685  -6.185  1.00 27.45           O
ANISOU  689  O   HOH B 230     2883   3521   4022    -25     85   -557       O
HETATM  690  O   HOH B 231      -0.963   0.258 -19.741  1.00 24.36           O
ANISOU  690  O   HOH B 231     2966   2255   4032    325    132    -24       O
HETATM  691  O   HOH B 232     -11.485   4.360  -0.265  1.00 31.52           O
ANISOU  691  O   HOH B 232     3198   2712   6066    628    409    297       O
HETATM  692  O   HOH B 233       0.761   0.345 -17.322  1.00 28.54           O
ANISOU  692  O   HOH B 233     4644   2172   4025    128    625    257       O
HETATM  693  O   HOH B 234      -6.582   2.748  -9.500  1.00 25.77           O
ANISOU  693  O   HOH B 234     2958   2659   4174     40   -537    285       O
HETATM  694  O   HOH B 235      -2.468 -14.775   5.778  1.00 60.04           O
ANISOU  694  O   HOH B 235     5572   6312  10927    687    624  -1562       O
HETATM  695  O   HOH B 236      -9.975   5.017  -2.626  1.00 32.96           O
ANISOU  695  O   HOH B 236     4050   4037   4436     72    297    713       O
HETATM  696  O   HOH B 237      -3.802  -8.987   0.610  1.00 40.56           O
ANISOU  696  O   HOH B 237     6819   3316   5276   -551   1407    285       O
HETATM  697  O   HOH B 238       5.576   8.867  -5.530  1.00 39.43           O
ANISOU  697  O   HOH B 238     5197   4878   4904  -1770   -985     27       O
HETATM  698  O   HOH B 239       0.480  -2.203 -20.508  1.00 37.54           O
ANISOU  698  O   HOH B 239     4418   4737   5107   1193    -91   -468       O
HETATM  699  O   HOH B 240     -12.279  -6.259   3.031  1.00 39.02           O
ANISOU  699  O   HOH B 240     4242   5382   5200  -1741    673   -493       O
HETATM  700  O   HOH B 241      -8.420   2.908  -7.461  1.00 34.05           O
```

```
ANISOU  700  O   HOH B 241     4404   4454   4077   1230     77    297       O
HETATM  701  O   HOH B 242     -10.273  -9.187   2.528  1.00 41.62           O
ANISOU  701  O   HOH B 242     5966   4686   5158  -2102      5   1009       O
HETATM  702  O   HOH B 243       4.413   5.754 -19.474  1.00 35.34           O
ANISOU  702  O   HOH B 243     3352   5036   5039  -1119    347   -457       O
HETATM  703  O   HOH B 244      -0.523   5.710 -26.410  1.00 28.37           O
ANISOU  703  O   HOH B 244     3209   3201   4366   -328    509   -110       O
HETATM  704  O   HOH B 245     -13.296  -3.589   2.389  1.00 43.39           O
ANISOU  704  O   HOH B 245     4011   6231   6243    177   -701  -1001       O
HETATM  705  O   HOH B 246       3.910   6.115 -22.191  1.00 31.12           O
ANISOU  705  O   HOH B 246     3318   3722   4785   -477   -369    474       O
HETATM  706  O   HOH B 247       5.326   6.311  -5.026  1.00 82.42           O
ANISOU  706  O   HOH B 247    11574   8469  11271  -3133  -5399   6007       O
HETATM  707  O   HOH B 248       4.443   7.607 -14.904  1.00 27.67           O
ANISOU  707  O   HOH B 248     2871   2857   4785   -351   -181    707       O
HETATM  708  O   HOH B 249       3.874   9.504 -12.859  1.00 33.03           O
ANISOU  708  O   HOH B 249     3890   3827   4831   -797    290    648       O
HETATM  709  O   HOH B 250     -10.648   6.203   1.297  0.50 30.90           O
ANISOU  709  O   HOH B 250     4705   2511   4523    274    255    360       O
HETATM  710  O   HOH B 251      -4.093  13.457 -15.884  1.00 33.85           O
ANISOU  710  O   HOH B 251     3413   4675   4770    720   -301  -1439       O
HETATM  711  O   HOH B 252      -8.480   5.016  -6.150  1.00 37.19           O
ANISOU  711  O   HOH B 252     4094   5567   4468   -609    198   -404       O
HETATM  712  O   HOH B 253      -8.275   6.732  -8.510  1.00 39.66           O
ANISOU  712  O   HOH B 253     4432   6236   4399   1298   -426  -1019       O
HETATM  713  O   HOH B 254      -7.029   8.894  -7.746  1.00 32.37           O
ANISOU  713  O   HOH B 254     4189   3701   4410   1087   -101    506       O
HETATM  714  O   HOH B 255      -5.872   7.775   3.047  1.00 32.68           O
ANISOU  714  O   HOH B 255     5065   2364   4988    669    500     15       O
HETATM  715  O   HOH B 256       0.102 -14.670   9.631  1.00 37.03           O
ANISOU  715  O   HOH B 256     4927   2523   6617    253    181     48       O
HETATM  716  O   HOH B 257       7.620   2.741   2.779  0.50 34.87           O
ANISOU  716  O   HOH B 257     2704   5122   5421    665    158   -897       O
HETATM  717  O   HOH B 258       7.157   2.257 -10.065  1.00 35.13           O
ANISOU  717  O   HOH B 258     2444   5144   5758    222    122   1879       O
HETATM  718  O   HOH B 259       8.052   7.180  -8.396  1.00 54.77           O
ANISOU  718  O   HOH B 259     6130   9365   5315  -1718   -676   -801       O
HETATM  719  O   HOH B 260       7.020  -2.654 -25.999  1.00 35.32           O
ANISOU  719  O   HOH B 260     3206   3170   7042     78   1084   -807       O
HETATM  720  O   HOH B 261      -0.005  10.737 -18.644  1.00 20.32           O
ANISOU  720  O   HOH B 261     2671   1772   3275     76   -208    -33       O
HETATM  721  O   HOH B 262      -2.116  12.137 -17.086  1.00 20.92           O
ANISOU  721  O   HOH B 262     2913   1813   3221    -15   -193   -147       O
HETATM  722  O   HOH B 263       3.917   8.126 -17.425  1.00 38.73           O
ANISOU  722  O   HOH B 263     2599   7142   4975   -894     45   -562       O
HETATM  723  O   HOH B 264       4.746  -1.496 -21.640  1.00 47.14           O
ANISOU  723  O   HOH B 264     7382   5315   5214   3120  -1025  -1646       O
HETATM  724  O   HOH B 265       4.739   4.057 -24.046  1.00 35.72           O
ANISOU  724  O   HOH B 265     4053   3999   5519  -1522   -737    421       O
HETATM  725  O   HOH B 266       4.040  11.534 -17.525  1.00 38.94           O
ANISOU  725  O   HOH B 266     3784   2887   8124    134     18   -790       O
HETATM  726  O   HOH B 267       6.503   9.438  -9.282  1.00107.68           O
ANISOU  726  O   HOH B 267     7985  15830  17098  -4205   5000  -5534       O
HETATM  727  O   HOH B 268       7.885  12.176  -7.962  1.00 46.26           O
ANISOU  727  O   HOH B 268     4531   5048   7998  -1489   -141   1016       O
HETATM  728  O   HOH B 269      -2.797   8.488 -26.981  1.00 32.64           O
```

```
ANISOU  728  O   HOH B 269     4521   3306   4572   -339   -454    501       O
HETATM  729  O   HOH B 270      -2.080  12.157 -20.062  1.00 20.75           O
ANISOU  729  O   HOH B 270     2844   1634   3405     -4   -353    -21       O
HETATM  730  O   HOH B 271       1.392 -10.091   4.267  1.00 48.78           O
ANISOU  730  O   HOH B 271     6494   4383   7655   -109   2158  -1768       O
HETATM  731  O   HOH B 272      -1.220 -11.002   2.292  1.00 44.13           O
ANISOU  731  O   HOH B 272     5183   6601   4982   1182   1028    338       O
HETATM  732  O   HOH B 273      -4.077  -6.118  -1.618  1.00 91.16           O
ANISOU  732  O   HOH B 273    20103   2930  11603   3457   2426  -4366       O
HETATM  733  O   HOH B 274      -4.825  -4.059  -4.000  1.00 35.08           O
ANISOU  733  O   HOH B 274     4110   3385   5830  -1047    849  -1504       O
HETATM  734  O   HOH B 275      -9.477   4.959   3.159  1.00 30.52           O
ANISOU  734  O   HOH B 275     4038   2270   5289    583   1174    800       O
HETATM  735  O   HOH B 276      -2.091  -8.116  -0.892  1.00115.08           O
ANISOU  735  O   HOH B 276    27773   3331  12621   -473   3351   1236       O
HETATM  736  O   HOH B 277       0.788  -7.488   0.563  1.00100.48           O
ANISOU  736  O   HOH B 277    13724   6460  17994   1204   2824  -1127       O
HETATM  737  O   HOH B 278     -11.763   2.704  -5.070  0.50 33.48           O
ANISOU  737  O   HOH B 278     3318   4599   4803   1064    -89    890       O
HETATM  738  O   HOH B 279      -4.548  12.687  -5.905  1.00 50.61           O
ANISOU  738  O   HOH B 279     6072   4619   8539    433   -295  -2169       O
HETATM  739  O   HOH B 280      -3.699  14.543  -4.043  0.50 37.16           O
ANISOU  739  O   HOH B 280     5562   2876   5679    222   1514    -13       O
HETATM  740  O   HOH B 281       7.368   0.558 -21.270  0.50116.86           O
ANISOU  740  O   HOH B 281     6066  29679   8654   5224  -1497  -5527       O
HETATM  741  O   HOH B 282       7.459   4.907 -20.980  1.00 31.27           O
ANISOU  741  O   HOH B 282     2612   5169   4099    141     82   -184       O
HETATM  742  O   HOH B 283     -11.007   4.398  -9.441  0.50 44.87           O
ANISOU  742  O   HOH B 283     4336   8746   3963  -2176    670  -1173       O
HETATM  743  O   HOH B 284      -8.065  -5.150  -2.303  0.50 33.22           O
ANISOU  743  O   HOH B 284     4126   4788   3706   -404   -737   -199       O
HETATM  744  O  AHOH B 285       3.281  -0.970 -14.016  0.00 30.46           O
ANISOU  744  O  AHOH B 285     4205   3413   3957    722    447   -238       O
HETATM  745  O  BHOH B 285       4.696  -0.403 -13.308  0.50 30.55           O
ANISOU  745  O  BHOH B 285     5017   2847   3741   1776   1611   1055       O
HETATM  746  O   HOH B 286      -7.390  10.652  -2.905  0.50 51.22           O
ANISOU  746  O   HOH B 286     4287   7663   7511   1005   -360    721       O
HETATM  747  O  AHOH B 287      -0.115  -4.964 -23.684  0.50 40.01           O
ANISOU  747  O  AHOH B 287     3313   4142   7745    325    950    479       O
HETATM  748  O  BHOH B 287      -1.617  -4.775 -25.825  0.50 33.51           O
ANISOU  748  O  BHOH B 287     5764   1417   5550    600   -762    224       O
HETATM  749  O   HOH B 288       2.883   4.930 -28.258  0.50 29.95           O
ANISOU  749  O   HOH B 288     3243   3275   4860   -485    244    727       O
HETATM  750  O   HOH B 289       1.349   7.221 -25.391  0.50 36.90           O
ANISOU  750  O   HOH B 289     3463   6711   3845  -1508    566   -358       O
HETATM  751  O   HOH B 290      -5.602  10.344  -0.148  1.00 88.23           O
ANISOU  751  O   HOH B 290     5543  10851  17129    703   2732  -1796       O
HETATM  752  O  AHOH B 291     -11.204  -1.946  -3.074  0.50 32.47           O
ANISOU  752  O  AHOH B 291     3794   4167   4374   -140   -287    197       O
HETATM  753  O  BHOH B 291      -9.727  -1.753  -3.519  0.50 28.64           O
ANISOU  753  O  BHOH B 291     3080   4214   3586   -226   -109   -367       O
HETATM  754  O   HOH B 292       3.035  10.214 -20.664  1.00 45.28           O
ANISOU  754  O   HOH B 292     5553   3829   7820   -603    292   -443       O
HETATM  755  O   HOH B 293       3.919  10.217 -22.849  1.00 51.99           O
ANISOU  755  O   HOH B 293     6933   4557   8264  -1564   1697  -1140       O
HETATM  756  O   HOH B 294       2.777  -4.135 -26.264  1.00 62.82           O
```

```
ANISOU  756  O   HOH B 294     6572   6106  11190    383    454  -1432       O
HETATM  757  O   HOH B 295      10.997   0.183 -19.547  1.00 43.78           O
ANISOU  757  O   HOH B 295     5694   6380   4558  -2444   -910    306       O
HETATM  758  O   HOH B 296       9.623   3.051 -18.066  1.00 54.33           O
ANISOU  758  O   HOH B 296     3240   9812   7591     66  -1597   2395       O
HETATM  759  O   HOH B 297       8.538   5.519 -17.276  1.00 36.67           O
ANISOU  759  O   HOH B 297     3983   5411   4538  -1536    738   -460       O
HETATM  760  O   HOH B 298      11.494   3.365 -14.553  1.00 47.74           O
ANISOU  760  O   HOH B 298     5135   3548   9454   -334   2302   -663       O
HETATM  761  O   HOH B 299       9.634   4.034 -15.804  1.00 58.14           O
ANISOU  761  O   HOH B 299     4550   9345   8195   -989   1336   1572       O
HETATM  762  O   HOH B 300       9.242   2.528 -13.567  1.00 39.07           O
ANISOU  762  O   HOH B 300     4961   3977   5907   1141   1444    516       O
HETATM  763  O   HOH B 301       8.060   7.240 -15.865  0.50 36.86           O
ANISOU  763  O   HOH B 301     4912   4149   4942    379   -324   1125       O
HETATM  764  O   HOH B 302       7.362   7.443 -21.063  0.50 32.97           O
ANISOU  764  O   HOH B 302     3879   3990   4658   -884    565   -264       O
HETATM  765  O   HOH B 303       7.048  -0.623 -17.035  1.00 41.84           O
ANISOU  765  O   HOH B 303     5338   5546   5011    518    332    340       O
HETATM  766  O   HOH B 304       7.023  10.304  -3.981  1.00 44.49           O
ANISOU  766  O   HOH B 304     6078   5021   5804  -1367   -356   -626       O
CONECT  150  162
CONECT  162  150  163  164  165
CONECT  163  162
CONECT  164  162
CONECT  165  162  166
CONECT  166  165  167
CONECT  167  166  168  169
CONECT  168  167  172
CONECT  169  167  170  171
CONECT  170  169  185
CONECT  171  169  172
CONECT  172  168  171  173
CONECT  173  172  174  184
CONECT  174  173  175  176
CONECT  175  174
CONECT  176  174  177
CONECT  177  176  178  179
CONECT  178  177
CONECT  179  177  180  181  184
CONECT  180  179  182
CONECT  181  179  183
CONECT  182  180
CONECT  183  181
CONECT  184  173  179
CONECT  185  170
CONECT  439  451
CONECT  451  439  452  453  454
CONECT  452  451
CONECT  453  451
CONECT  454  451  455
CONECT  455  454  456
CONECT  456  455  457  458
CONECT  457  456  461
CONECT  458  456  459  460
CONECT  459  458  472
```

```
CONECT  460  458  461
CONECT  461  457  460  462
CONECT  462  461  463  471
CONECT  463  462  464  465
CONECT  464  463
CONECT  465  463  466
CONECT  466  465  467  468
CONECT  467  466
CONECT  468  466  469  471
CONECT  469  468  470
CONECT  470  469
CONECT  471  462  468
CONECT  472  459
CONECT  536  583  624  665  720
CONECT  536  721  729
CONECT  537  538
CONECT  538  537  539
CONECT  539  538  540
CONECT  540  539  541
CONECT  541  540  542
CONECT  542  541  543
CONECT  543  542  544
CONECT  544  543  545
CONECT  545  544  546
CONECT  546  545  547
CONECT  547  546
CONECT  548  550
CONECT  549  551
CONECT  550  548  552
CONECT  551  549  553
CONECT  552  550  554
CONECT  553  551  555
CONECT  554  552  556
CONECT  555  553  557
CONECT  556  554  558
CONECT  557  555  559
CONECT  558  556  560
CONECT  559  557  561
CONECT  560  558  562
CONECT  561  559  563
CONECT  562  560  564
CONECT  563  561  565
CONECT  564  562  566
CONECT  565  563  567
CONECT  566  564  568
CONECT  567  565  569
CONECT  568  566  570
CONECT  569  567  571
CONECT  570  568  572
CONECT  571  569  573
CONECT  572  570  574
CONECT  573  571  575
CONECT  574  572
CONECT  575  573
CONECT  583  536
CONECT  624  536
```

```
CONECT  665  536
CONECT  720  536
CONECT  721  536
CONECT  729  536
MASTER      319    0    5    0    0    0    6    6  703    2   95    2
END
```

File A-3: PDB coordinates for the crystal structure of 5-formyl-2'-deoxycytidine in B-type DNA, DDD[f].

```
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : REFMAC 5.8.0049
REMARK   3   AUTHORS     : MURSHUDOV,SKUBAK,LEBEDEV,PANNU,
REMARK   3                 STEINER,NICHOLLS,WINN,LONG,VAGIN
REMARK   3
REMARK   3    REFINEMENT TARGET : MAXIMUM LIKELIHOOD
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) :   1.74
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) :  35.06
REMARK   3   DATA CUTOFF            (SIGMA(F)) : NONE
REMARK   3   COMPLETENESS FOR RANGE        (%) :  98.98
REMARK   3   NUMBER OF REFLECTIONS             :    7121
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT.
REMARK   3   CROSS-VALIDATION METHOD          : THROUGHOUT
REMARK   3   FREE R VALUE TEST SET SELECTION  : RANDOM
REMARK   3   R VALUE     (WORKING + TEST SET) : 0.23571
REMARK   3   R VALUE           (WORKING SET) :  0.23315
REMARK   3   FREE R VALUE                     :  0.29312
REMARK   3   FREE R VALUE TEST SET SIZE   (%) :  4.6
REMARK   3   FREE R VALUE TEST SET COUNT      :   344
REMARK   3
REMARK   3  FIT IN THE HIGHEST RESOLUTION BIN.
REMARK   3   TOTAL NUMBER OF BINS USED           :      20
REMARK   3   BIN RESOLUTION RANGE HIGH           :    1.740
REMARK   3   BIN RESOLUTION RANGE LOW            :    1.785
REMARK   3   REFLECTION IN BIN     (WORKING SET) :      521
REMARK   3   BIN COMPLETENESS (WORKING+TEST) (%) :    99.28
REMARK   3   BIN R VALUE           (WORKING SET) :    0.300
REMARK   3   BIN FREE R VALUE SET COUNT          :       27
REMARK   3   BIN FREE R VALUE                    :    0.358
REMARK   3
REMARK   3  NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK   3   ALL ATOMS                :      547
REMARK   3
REMARK   3  B VALUES.
REMARK   3   FROM WILSON PLOT           (A**2) : NULL
REMARK   3   MEAN B VALUE      (OVERALL, A**2) :  35.181
REMARK   3   OVERALL ANISOTROPIC B VALUE.
REMARK   3    B11 (A**2) :     2.03
REMARK   3    B22 (A**2) :     0.80
REMARK   3    B33 (A**2) :    -2.82
REMARK   3    B12 (A**2) :     0.00
REMARK   3    B13 (A**2) :    -0.00
REMARK   3    B23 (A**2) :    -0.00
```

```
REMARK   3
REMARK   3  ESTIMATED OVERALL COORDINATE ERROR.
REMARK   3   ESU BASED ON R VALUE                            (A):   0.138
REMARK   3   ESU BASED ON FREE R VALUE                       (A):   0.144
REMARK   3   ESU BASED ON MAXIMUM LIKELIHOOD                 (A):   0.117
REMARK   3   ESU FOR B VALUES BASED ON MAXIMUM LIKELIHOOD (A**2):   3.908
REMARK   3
REMARK   3 CORRELATION COEFFICIENTS.
REMARK   3   CORRELATION COEFFICIENT FO-FC      :   0.959
REMARK   3   CORRELATION COEFFICIENT FO-FC FREE :   0.932
REMARK   3
REMARK   3  RMS DEVIATIONS FROM IDEAL VALUES        COUNT    RMS    WEIGHT
REMARK   3   BOND LENGTHS REFINED ATOMS        (A):   552 ; 0.009 ; 0.011
REMARK   3   BOND LENGTHS OTHERS               (A):   268 ; 0.002 ; 0.020
REMARK   3   BOND ANGLES REFINED ATOMS   (DEGREES):   850 ; 1.747 ; 1.182
REMARK   3   BOND ANGLES OTHERS          (DEGREES):   634 ; 1.611 ; 3.000
REMARK   3   CHIRAL-CENTER RESTRAINTS       (A**3):    72 ; 0.071 ; 0.200
REMARK   3   GENERAL PLANES REFINED ATOMS      (A):   290 ; 0.023 ; 0.020
REMARK   3   GENERAL PLANES OTHERS             (A):   122 ; 0.003 ; 0.020
REMARK   3
REMARK   3  ISOTROPIC THERMAL FACTOR RESTRAINTS.     COUNT   RMS    WEIGHT
REMARK   3   SIDE-CHAIN BOND REFINED ATOMS  (A**2):   552 ; 3.873 ; 3.639
REMARK   3   SIDE-CHAIN BOND OTHER ATOMS  (A**2)  :   551 ; 3.874 ; 3.640
REMARK   3   SIDE-CHAIN ANGLE OTHER ATOMS (A**2)  :   850 ; 5.591 ; 5.470
REMARK   3   LONG RANGE B REFINED ATOMS (A**2)    :   819 ; 5.904 ;34.255
REMARK   3   LONG RANGE B OTHER ATOMS (A**2)      :   804 ; 5.858 ;34.257
REMARK   3
REMARK   3  NCS RESTRAINTS STATISTICS
REMARK   3   NUMBER OF NCS GROUPS : NULL
REMARK   3
REMARK   3  TWIN DETAILS
REMARK   3   NUMBER OF TWIN DOMAINS  : NULL
REMARK   3
REMARK   3
REMARK   3  TLS DETAILS
REMARK   3   NUMBER OF TLS GROUPS  : NULL
REMARK   3
REMARK   3
REMARK   3  BULK SOLVENT MODELLING.
REMARK   3   METHOD USED :  MASK
REMARK   3   PARAMETERS FOR MASK CALCULATION
REMARK   3   VDW PROBE RADIUS   :   1.20
REMARK   3   ION PROBE RADIUS   :   0.80
REMARK   3   SHRINKAGE RADIUS   :   0.80
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS:
REMARK   3  HYDROGENS HAVE BEEN ADDED IN THE RIDING POSITIONS
REMARK   3  U VALUES      : REFINED INDIVIDUALLY
REMARK   3
CRYST1   25.316   41.470   65.669  90.00  90.00  90.00 P 21 21 21
SCALE1      0.039501  0.000000  0.000000        0.00000
SCALE2     -0.000000  0.024114  0.000000        0.00000
SCALE3      0.000000 -0.000000  0.015228        0.00000
ATOM      1  O5'  DC A   1       7.063  14.256  22.460  1.00 53.02      A    O
ATOM      2  C5'  DC A   1       6.906  13.495  23.675  1.00 44.30      A    C
ATOM      3  C4'  DC A   1       7.388  12.081  23.461  1.00 41.26      A    C
```

```
ATOM      4  C3'  DC A   1       7.300  11.584  22.015  1.00 47.01      A    C
ATOM      5  O3'  DC A   1       8.504  10.848  21.790  1.00 49.71      A    O
ATOM      6  C2'  DC A   1       6.101  10.652  22.051  1.00 45.11      A    C
ATOM      7  C1'  DC A   1       6.300  10.047  23.426  1.00 38.25      A    C
ATOM      8  O4'  DC A   1       6.583  11.159  24.234  1.00 38.74      A    O
ATOM      9  N1   DC A   1       5.148   9.354  23.997  1.00 34.51      A    N
ATOM     10  C6   DC A   1       3.915   9.933  23.917  1.00 38.58      A    C
ATOM     11  C5   DC A   1       2.834   9.319  24.412  1.00 44.80      A    C
ATOM     12  C4   DC A   1       3.020   8.032  25.005  1.00 38.41      A    C
ATOM     13  N4   DC A   1       1.987   7.382  25.525  1.00 41.79      A    N
ATOM     14  N3   DC A   1       4.223   7.469  25.108  1.00 30.93      A    N
ATOM     15  C2   DC A   1       5.300   8.084  24.591  1.00 31.81      A    C
ATOM     16  O2   DC A   1       6.422   7.562  24.639  1.00 33.15      A    O
ATOM     17  P    DG A   2       9.192  10.792  20.363  1.00 59.87      A    P
ATOM     18  OP1  DG A   2      10.029  12.012  20.229  1.00 60.91      A    O
ATOM     19  O5'  DG A   2      10.177   9.547  20.515  1.00 49.54      A    O
ATOM     20  C5'  DG A   2      10.843   9.268  21.759  1.00 51.72      A    C
ATOM     21  C4'  DG A   2      11.008   7.777  21.926  1.00 45.38      A    C
ATOM     22  C3'  DG A   2      11.207   6.999  20.620  1.00 43.40      A    C
ATOM     23  O3'  DG A   2      12.385   6.204  20.655  1.00 44.27      A    O
ATOM     24  C2'  DG A   2      10.037   6.046  20.535  1.00 42.59      A    C
ATOM     25  C1'  DG A   2       9.387   6.058  21.902  1.00 42.86      A    C
ATOM     26  O4'  DG A   2       9.833   7.232  22.564  1.00 45.70      A    O
ATOM     27  N9   DG A   2       7.920   6.111  21.844  1.00 36.66      A    N
ATOM     28  C4   DG A   2       7.057   5.129  22.254  1.00 30.40      A    C
ATOM     29  C5   DG A   2       5.798   5.599  21.951  1.00 31.88      A    C
ATOM     30  N7   DG A   2       5.866   6.879  21.423  1.00 34.20      A    N
ATOM     31  C8   DG A   2       7.143   7.133  21.356  1.00 34.88      A    C
ATOM     32  N3   DG A   2       7.408   3.922  22.745  1.00 32.05      A    N
ATOM     33  C2   DG A   2       6.341   3.189  23.054  1.00 28.25      A    C
ATOM     34  N2   DG A   2       6.500   1.973  23.572  1.00 30.09      A    N
ATOM     35  N1   DG A   2       5.050   3.604  22.863  1.00 28.69      A    N
ATOM     36  C6   DG A   2       4.676   4.821  22.311  1.00 28.46      A    C
ATOM     37  O6   DG A   2       3.480   5.099  22.193  1.00 32.17      A    O
ATOM     38  OP2  DG A   2       8.161  10.441  19.332  1.00 47.41      A    O
ATOM     39  P    DC A   3      12.899   5.496  19.330  1.00 51.42      A    P
ATOM     40  OP1  DC A   3      14.356   5.350  19.468  1.00 57.83      A    O
ATOM     41  O5'  DC A   3      12.171   4.087  19.303  1.00 43.82      A    O
ATOM     42  C5'  DC A   3      12.528   3.078  20.259  1.00 39.21      A    C
ATOM     43  C4'  DC A   3      11.504   1.979  20.180  1.00 36.20      A    C
ATOM     44  C3'  DC A   3      11.370   1.330  18.811  1.00 33.94      A    C
ATOM     45  O3'  DC A   3      12.156   0.142  18.796  1.00 40.30      A    O
ATOM     46  C2'  DC A   3       9.912   0.925  18.746  1.00 38.63      A    C
ATOM     47  C1'  DC A   3       9.243   1.687  19.852  1.00 37.19      A    C
ATOM     48  O4'  DC A   3      10.216   2.552  20.428  1.00 32.62      A    O
ATOM     49  N1   DC A   3       8.152   2.530  19.404  1.00 32.54      A    N
ATOM     50  C6   DC A   3       8.401   3.731  18.804  1.00 33.17      A    C
ATOM     51  C5   DC A   3       7.404   4.560  18.487  1.00 37.56      A    C
ATOM     52  C4   DC A   3       6.081   4.146  18.818  1.00 32.41      A    C
ATOM     53  N4   DC A   3       5.040   4.918  18.504  1.00 33.84      A    N
ATOM     54  N3   DC A   3       5.834   2.962  19.372  1.00 30.95      A    N
ATOM     55  C2   DC A   3       6.856   2.170  19.747  1.00 27.59      A    C
ATOM     56  O2   DC A   3       6.670   1.072  20.272  1.00 30.70      A    O
ATOM     57  OP2  DC A   3      12.277   6.183  18.156  1.00 49.20      A    O
ATOM     58  P    DG A   4      12.291  -0.677  17.462  1.00 41.60      A    P
ATOM     59  OP1  DG A   4      13.447  -1.554  17.603  1.00 49.14      A    O
```

```
ATOM     60  O5'  DG A   4      11.045  -1.655  17.536  1.00 38.14      A    O
ATOM     61  C5'  DG A   4      10.945  -2.646  18.574  1.00 39.91      A    C
ATOM     62  C4'  DG A   4       9.681  -3.453  18.370  1.00 47.85      A    C
ATOM     63  C3'  DG A   4       9.611  -4.190  17.030  1.00 47.02      A    C
ATOM     64  O3'  DG A   4       9.131  -5.521  17.175  1.00 57.79      A    O
ATOM     65  C2'  DG A   4       8.662  -3.365  16.192  1.00 39.78      A    C
ATOM     66  C1'  DG A   4       7.789  -2.663  17.216  1.00 40.09      A    C
ATOM     67  O4'  DG A   4       8.521  -2.593  18.424  1.00 42.15      A    O
ATOM     68  N9   DG A   4       7.374  -1.309  16.870  1.00 36.73      A    N
ATOM     69  C4   DG A   4       6.105  -0.797  17.026  1.00 34.90      A    C
ATOM     70  C5   DG A   4       6.163   0.478  16.531  1.00 30.05      A    C
ATOM     71  N7   DG A   4       7.443   0.774  16.090  1.00 28.84      A    N
ATOM     72  C8   DG A   4       8.112  -0.330  16.268  1.00 29.03      A    C
ATOM     73  N3   DG A   4       5.045  -1.453  17.545  1.00 36.92      A    N
ATOM     74  C2   DG A   4       3.945  -0.719  17.515  1.00 30.57      A    C
ATOM     75  N2   DG A   4       2.801  -1.233  17.985  1.00 36.63      A    N
ATOM     76  N1   DG A   4       3.895   0.572  17.062  1.00 31.16      A    N
ATOM     77  C6   DG A   4       4.976   1.275  16.543  1.00 32.62      A    C
ATOM     78  O6   DG A   4       4.833   2.456  16.205  1.00 30.36      A    O
ATOM     79  OP2  DG A   4      12.137   0.267  16.340  1.00 44.37      A    O
ATOM     80  P    DA A   5       8.920  -6.401  15.858  1.00 64.69      A    P
ATOM     81  OP1  DA A   5       9.442  -7.754  16.152  1.00 68.70      A    O
ATOM     82  OP2  DA A   5       9.398  -5.627  14.674  1.00 58.00      A    O
ATOM     83  O5'  DA A   5       7.342  -6.370  15.661  1.00 48.79      A    O
ATOM     84  C5'  DA A   5       6.480  -6.477  16.794  1.00 48.18      A    C
ATOM     85  C4'  DA A   5       5.063  -6.253  16.341  1.00 44.71      A    C
ATOM     86  C3'  DA A   5       4.685  -6.980  15.047  1.00 45.71      A    C
ATOM     87  O3'  DA A   5       3.415  -7.585  15.263  1.00 48.22      A    O
ATOM     88  C2'  DA A   5       4.628  -5.870  14.009  1.00 43.09      A    C
ATOM     89  C1'  DA A   5       4.231  -4.677  14.844  1.00 34.40      A    C
ATOM     90  O4'  DA A   5       4.888  -4.847  16.078  1.00 40.05      A    O
ATOM     91  N9   DA A   5       4.663  -3.396  14.313  1.00 28.61      A    N
ATOM     92  C4   DA A   5       3.851  -2.294  14.288  1.00 27.73      A    C
ATOM     93  C5   DA A   5       4.613  -1.274  13.753  1.00 27.16      A    C
ATOM     94  N7   DA A   5       5.901  -1.711  13.488  1.00 26.97      A    N
ATOM     95  C8   DA A   5       5.878  -2.978  13.838  1.00 26.54      A    C
ATOM     96  N3   DA A   5       2.556  -2.238  14.635  1.00 28.73      A    N
ATOM     97  C2   DA A   5       2.064  -1.022  14.403  1.00 28.64      A    C
ATOM     98  N1   DA A   5       2.682   0.062  13.934  1.00 28.07      A    N
ATOM     99  C6   DA A   5       3.990  -0.024  13.612  1.00 26.54      A    C
ATOM    100  N6   DA A   5       4.589   1.046  13.097  1.00 25.01      A    N
ATOM    101  P    DA A   6       2.703  -8.413  14.078  1.00 49.41      A    P
ATOM    102  OP1  DA A   6       2.130  -9.628  14.710  1.00 47.72      A    O
ATOM    103  OP2  DA A   6       3.629  -8.506  12.958  1.00 42.25      A    O
ATOM    104  O5'  DA A   6       1.509  -7.450  13.707  1.00 42.99      A    O
ATOM    105  C5'  DA A   6       0.561  -7.126  14.713  1.00 38.95      A    C
ATOM    106  C4'  DA A   6      -0.488  -6.219  14.131  1.00 38.25      A    C
ATOM    107  C3'  DA A   6      -1.151  -6.805  12.893  1.00 36.76      A    C
ATOM    108  O3'  DA A   6      -2.548  -6.616  13.142  1.00 40.26      A    O
ATOM    109  C2'  DA A   6      -0.481  -6.053  11.753  1.00 35.95      A    C
ATOM    110  C1'  DA A   6      -0.202  -4.684  12.375  1.00 35.61      A    C
ATOM    111  O4'  DA A   6       0.076  -4.941  13.740  1.00 39.54      A    O
ATOM    112  N9   DA A   6       0.951  -3.969  11.827  1.00 30.39      A    N
ATOM    113  C4   DA A   6       1.024  -2.630  11.523  1.00 27.08      A    C
ATOM    114  C5   DA A   6       2.314  -2.419  11.080  1.00 25.98      A    C
ATOM    115  N7   DA A   6       3.060  -3.588  11.161  1.00 27.17      A    N
```

```
ATOM    116  C8   DA A   6       2.223  -4.459  11.665  1.00 27.03      A    C
ATOM    117  N3   DA A   6       0.012  -1.741  11.508  1.00 29.47      A    N
ATOM    118  C2   DA A   6       0.416  -0.573  11.015  1.00 30.01      A    C
ATOM    119  N1   DA A   6       1.641  -0.202  10.614  1.00 27.91      A    N
ATOM    120  C6   DA A   6       2.633  -1.124  10.640  1.00 26.25      A    C
ATOM    121  N6   DA A   6       3.846  -0.764  10.235  1.00 26.47      A    N
ATOM    122  P    DT A   7      -3.675  -7.057  12.036  1.00 46.76      A    P
ATOM    123  OP1  DT A   7      -4.956  -7.270  12.751  1.00 48.13      A    O
ATOM    124  O5'  DT A   7      -3.824  -5.738  11.162  1.00 37.89      A    O
ATOM    125  C5'  DT A   7      -4.333  -4.544  11.722  1.00 35.27      A    C
ATOM    126  C4'  DT A   7      -4.242  -3.447  10.693  1.00 31.30      A    C
ATOM    127  C3'  DT A   7      -4.928  -3.792   9.369  1.00 32.43      A    C
ATOM    128  O3'  DT A   7      -6.028  -2.904   9.254  1.00 32.90      A    O
ATOM    129  C2'  DT A   7      -3.857  -3.555   8.302  1.00 35.76      A    C
ATOM    130  C1'  DT A   7      -2.834  -2.713   9.044  1.00 33.09      A    C
ATOM    131  O4'  DT A   7      -2.867  -3.181  10.361  1.00 30.26      A    O
ATOM    132  N1   DT A   7      -1.449  -2.799   8.590  1.00 29.02      A    N
ATOM    133  C6   DT A   7      -0.777  -3.999   8.592  1.00 31.24      A    C
ATOM    134  C5   DT A   7       0.496  -4.132   8.208  1.00 29.65      A    C
ATOM    135  C7   DT A   7       1.213  -5.444   8.221  1.00 34.02      A    C
ATOM    136  C4   DT A   7       1.199  -2.951   7.767  1.00 29.11      A    C
ATOM    137  N3   DT A   7       0.477  -1.778   7.830  1.00 27.76      A    N
ATOM    138  C2   DT A   7      -0.842  -1.639   8.188  1.00 29.00      A    C
ATOM    139  O2   DT A   7      -1.419  -0.565   8.159  1.00 26.02      A    O
ATOM    140  O4   DT A   7       2.364  -2.925   7.416  1.00 27.37      A    O
ATOM    141  OP2  DT A   7      -3.073  -8.059  11.105  1.00 43.54      A    O
ATOM    142  P    DT A   8      -6.995  -3.006   7.972  1.00 40.98      A    P
ATOM    143  OP1  DT A   8      -8.370  -2.749   8.444  1.00 46.45      A    O
ATOM    144  O5'  DT A   8      -6.500  -1.771   7.107  1.00 34.55      A    O
ATOM    145  C5'  DT A   8      -6.228  -0.532   7.741  1.00 32.48      A    C
ATOM    146  C4'  DT A   8      -5.556   0.445   6.809  1.00 31.80      A    C
ATOM    147  C3'  DT A   8      -6.173   0.561   5.413  1.00 35.67      A    C
ATOM    148  O3'  DT A   8      -6.408   1.935   5.163  1.00 37.98      A    O
ATOM    149  C2'  DT A   8      -5.120  -0.020   4.479  1.00 34.16      A    C
ATOM    150  C1'  DT A   8      -3.826   0.249   5.254  1.00 32.91      A    C
ATOM    151  O4'  DT A   8      -4.173   0.068   6.620  1.00 30.98      A    O
ATOM    152  N1   DT A   8      -2.716  -0.648   4.999  1.00 28.55      A    N
ATOM    153  C6   DT A   8      -2.911  -1.998   5.183  1.00 26.93      A    C
ATOM    154  C5   DT A   8      -1.958  -2.912   5.029  1.00 28.09      A    C
ATOM    155  C7   DT A   8      -2.209  -4.375   5.222  1.00 34.22      A    C
ATOM    156  C4   DT A   8      -0.631  -2.456   4.640  1.00 28.62      A    C
ATOM    157  N3   DT A   8      -0.501  -1.091   4.494  1.00 25.97      A    N
ATOM    158  C2   DT A   8      -1.482  -0.143   4.636  1.00 28.11      A    C
ATOM    159  O2   DT A   8      -1.279   1.042   4.419  1.00 28.16      A    O
ATOM    160  O4   DT A   8       0.355  -3.169   4.523  1.00 27.94      A    O
ATOM    161  OP2  DT A   8      -6.664  -4.231   7.217  1.00 36.27      A    O
HETATM  162  P   5fC A   9      -7.040   2.426   3.783  1.00 46.65      A    P
HETATM  163  OP1 5fC A   9      -7.703   3.698   4.094  1.00 48.80      A    O
HETATM  164  O5' 5fC A   9      -5.760   2.817   2.928  1.00 41.00      A    O
HETATM  165  C5' 5fC A   9      -4.954   3.937   3.361  1.00 37.44      A    C
HETATM  166  C4' 5fC A   9      -3.802   4.108   2.403  1.00 37.18      A    C
HETATM  167  C3' 5fC A   9      -4.231   4.327   0.951  1.00 38.78      A    C
HETATM  168  O3' 5fC A   9      -3.781   5.658   0.760  1.00 44.04      A    O
HETATM  169  C2' 5fC A   9      -3.539   3.216   0.159  1.00 35.05      A    C
HETATM  170  C1' 5fC A   9      -2.482   2.683   1.108  1.00 32.92      A    C
HETATM  171  O4' 5fC A   9      -2.988   2.924   2.415  1.00 34.45      A    O
```

```
HETATM  172  N1  5fC A   9      -2.122   1.261   1.061  1.00 32.89      A    N
HETATM  173  C6  5fC A   9      -3.042   0.288   1.336  1.00 33.23      A    C
HETATM  174  C5A 5fC A   9      -3.729  -1.976   1.678  1.00 34.45      A    C
HETATM  175  O5A 5fC A   9      -3.537  -3.228   1.731  1.00 40.65      A    O
HETATM  176  C5  5fC A   9      -2.686  -1.004   1.402  1.00 30.16      A    C
HETATM  177  C4  5fC A   9      -1.312  -1.319   1.187  1.00 27.02      A    C
HETATM  178  N4  5fC A   9      -0.893  -2.587   1.246  1.00 29.47      A    N
HETATM  179  N3  5fC A   9      -0.406  -0.376   0.912  1.00 27.85      A    N
HETATM  180  C2  5fC A   9      -0.779   0.926   0.866  1.00 29.08      A    C
HETATM  181  O2  5fC A   9       0.017   1.823   0.584  1.00 29.99      A    O
HETATM  182  OP2 5fC A   9      -7.746   1.286   3.154  1.00 46.31      A    O
ATOM    183  P    DG A  10      -4.099   6.435  -0.551  1.00 53.68      A    P
ATOM    184  OP1  DG A  10      -3.680   7.859  -0.329  1.00 56.56      A    O
ATOM    185  O5'  DG A  10      -3.052   5.863  -1.605  1.00 48.51      A    O
ATOM    186  C5'  DG A  10      -1.681   6.252  -1.555  1.00 44.58      A    C
ATOM    187  C4'  DG A  10      -0.908   5.465  -2.590  1.00 37.26      A    C
ATOM    188  C3'  DG A  10      -1.520   5.524  -3.990  1.00 35.18      A    C
ATOM    189  O3'  DG A  10      -0.450   5.472  -4.930  1.00 41.27      A    O
ATOM    190  C2'  DG A  10      -2.327   4.254  -4.033  1.00 34.05      A    C
ATOM    191  C1'  DG A  10      -1.350   3.345  -3.394  1.00 30.38      A    C
ATOM    192  O4'  DG A  10      -0.939   4.066  -2.223  1.00 33.85      A    O
ATOM    193  N9   DG A  10      -1.724   1.985  -3.009  1.00 31.02      A    N
ATOM    194  C4   DG A  10      -0.823   0.958  -2.854  1.00 27.56      A    C
ATOM    195  C5   DG A  10      -1.572  -0.135  -2.483  1.00 27.59      A    C
ATOM    196  N7   DG A  10      -2.910   0.217  -2.326  1.00 28.63      A    N
ATOM    197  C8   DG A  10      -2.956   1.476  -2.673  1.00 29.82      A    C
ATOM    198  N3   DG A  10       0.501   1.036  -3.089  1.00 26.34      A    N
ATOM    199  C2   DG A  10       1.115  -0.110  -2.867  1.00 25.50      A    C
ATOM    200  N2   DG A  10       2.430  -0.189  -3.066  1.00 25.55      A    N
ATOM    201  N1   DG A  10       0.468  -1.266  -2.480  1.00 28.74      A    N
ATOM    202  C6   DG A  10      -0.908  -1.374  -2.260  1.00 26.64      A    C
ATOM    203  O6   DG A  10      -1.389  -2.454  -1.892  1.00 27.40      A    O
ATOM    204  OP2  DG A  10      -5.417   6.007  -1.059  1.00 46.61      A    O
ATOM    205  P    DC A  11      -0.005   6.780  -5.749  1.00 47.60      A    P
ATOM    206  OP1  DC A  11      -0.140   7.939  -4.826  1.00 49.58      A    O
ATOM    207  O5'  DC A  11       1.542   6.552  -5.988  1.00 39.08      A    O
ATOM    208  C5'  DC A  11       2.461   6.429  -4.899  1.00 40.60      A    C
ATOM    209  C4'  DC A  11       3.584   5.492  -5.279  1.00 42.97      A    C
ATOM    210  C3'  DC A  11       4.142   5.651  -6.691  1.00 38.22      A    C
ATOM    211  O3'  DC A  11       5.571   5.550  -6.612  1.00 43.05      A    O
ATOM    212  C2'  DC A  11       3.512   4.507  -7.460  1.00 36.70      A    C
ATOM    213  C1'  DC A  11       3.342   3.429  -6.403  1.00 36.16      A    C
ATOM    214  O4'  DC A  11       3.145   4.122  -5.179  1.00 35.36      A    O
ATOM    215  N1   DC A  11       2.191   2.543  -6.538  1.00 34.57      A    N
ATOM    216  C6   DC A  11       0.933   3.061  -6.656  1.00 34.02      A    C
ATOM    217  C5   DC A  11      -0.141   2.269  -6.624  1.00 32.64      A    C
ATOM    218  C4   DC A  11       0.071   0.885  -6.369  1.00 32.57      A    C
ATOM    219  N4   DC A  11      -0.963   0.051  -6.294  1.00 31.31      A    N
ATOM    220  N3   DC A  11       1.294   0.374  -6.228  1.00 29.10      A    N
ATOM    221  C2   DC A  11       2.366   1.194  -6.226  1.00 30.81      A    C
ATOM    222  O2   DC A  11       3.510   0.769  -6.041  1.00 29.32      A    O
ATOM    223  OP2  DC A  11      -0.685   6.733  -7.085  1.00 47.16      A    O
ATOM    224  P    DG A  12       6.494   5.956  -7.860  1.00 46.38      A    P
ATOM    225  OP1  DG A  12       7.893   6.126  -7.365  1.00 48.60      A    O
ATOM    226  O5'  DG A  12       6.420   4.678  -8.818  1.00 46.73      A    O
ATOM    227  C5'  DG A  12       7.268   3.517  -8.601  1.00 44.27      A    C
```

```
ATOM    228  C4'  DG A  12       6.852   2.399  -9.527  1.00 39.03      A    C
ATOM    229  C3'  DG A  12       6.779   2.780 -11.005  1.00 40.02      A    C
ATOM    230  O3'  DG A  12       8.019   2.512 -11.636  1.00 39.96      A    O
ATOM    231  C2'  DG A  12       5.662   1.914 -11.559  1.00 40.49      A    C
ATOM    232  C1'  DG A  12       5.043   1.263 -10.317  1.00 39.93      A    C
ATOM    233  O4'  DG A  12       5.523   1.998  -9.198  1.00 38.45      A    O
ATOM    234  N9   DG A  12       3.590   1.272 -10.246  1.00 33.33      A    N
ATOM    235  C4   DG A  12       2.805   0.223  -9.849  1.00 29.93      A    C
ATOM    236  C5   DG A  12       1.513   0.708  -9.856  1.00 34.72      A    C
ATOM    237  N7   DG A  12       1.493   2.055 -10.195  1.00 35.39      A    N
ATOM    238  C8   DG A  12       2.747   2.354 -10.397  1.00 38.74      A    C
ATOM    239  N3   DG A  12       3.237  -1.024  -9.558  1.00 27.18      A    N
ATOM    240  C2   DG A  12       2.230  -1.844  -9.290  1.00 29.13      A    C
ATOM    241  N2   DG A  12       2.475  -3.139  -9.034  1.00 30.35      A    N
ATOM    242  N1   DG A  12       0.910  -1.461  -9.283  1.00 28.64      A    N
ATOM    243  C6   DG A  12       0.447  -0.178  -9.546  1.00 30.47      A    C
ATOM    244  O6   DG A  12      -0.763   0.071  -9.464  1.00 33.53      A    O
ATOM    245  OP2  DG A  12       5.803   7.035  -8.586  1.00 50.37      A    O
TER     246       DG A  12
ATOM    246  O5'  DC B  13      -3.942  -8.932  -9.675  1.00 60.68      B    O
ATOM    247  C5'  DC B  13      -2.729  -9.569 -10.124  1.00 54.83      B    C
ATOM    248  C4'  DC B  13      -1.569  -9.016  -9.328  1.00 51.71      B    C
ATOM    249  C3'  DC B  13      -1.699  -9.163  -7.811  1.00 50.71      B    C
ATOM    250  O3'  DC B  13      -0.454  -9.373  -7.174  1.00 55.06      B    O
ATOM    251  C2'  DC B  13      -2.033  -7.765  -7.348  1.00 52.73      B    C
ATOM    252  C1'  DC B  13      -1.231  -6.959  -8.340  1.00 45.42      B    C
ATOM    253  O4'  DC B  13      -1.468  -7.595  -9.576  1.00 47.77      B    O
ATOM    254  N1   DC B  13      -1.657  -5.569  -8.454  1.00 40.50      B    N
ATOM    255  C6   DC B  13      -2.976  -5.243  -8.328  1.00 44.08      B    C
ATOM    256  C5   DC B  13      -3.381  -3.973  -8.398  1.00 47.50      B    C
ATOM    257  C4   DC B  13      -2.383  -2.977  -8.582  1.00 42.21      B    C
ATOM    258  N4   DC B  13      -2.733  -1.699  -8.683  1.00 43.64      B    N
ATOM    259  N3   DC B  13      -1.094  -3.289  -8.735  1.00 34.79      B    N
ATOM    260  C2   DC B  13      -0.698  -4.570  -8.645  1.00 31.99      B    C
ATOM    261  O2   DC B  13       0.497  -4.891  -8.719  1.00 35.45      B    O
ATOM    262  P    DG B  14       0.158 -10.839  -7.027  1.00 61.99      B    P
ATOM    263  OP1  DG B  14       0.415 -11.402  -8.368  1.00 51.57      B    O
ATOM    264  O5'  DG B  14       1.630 -10.517  -6.538  1.00 53.67      B    O
ATOM    265  C5'  DG B  14       2.532  -9.990  -7.520  1.00 48.43      B    C
ATOM    266  C4'  DG B  14       3.541  -9.084  -6.868  1.00 43.78      B    C
ATOM    267  C3'  DG B  14       3.853  -9.428  -5.416  1.00 40.04      B    C
ATOM    268  O3'  DG B  14       5.259  -9.425  -5.278  1.00 43.30      B    O
ATOM    269  C2'  DG B  14       3.206  -8.320  -4.610  1.00 41.74      B    C
ATOM    270  C1'  DG B  14       3.194  -7.165  -5.582  1.00 43.03      B    C
ATOM    271  O4'  DG B  14       3.021  -7.745  -6.864  1.00 41.00      B    O
ATOM    272  N9   DG B  14       2.078  -6.259  -5.376  1.00 35.89      B    N
ATOM    273  C4   DG B  14       2.091  -4.899  -5.570  1.00 33.28      B    C
ATOM    274  C5   DG B  14       0.802  -4.485  -5.326  1.00 33.87      B    C
ATOM    275  N7   DG B  14      -0.008  -5.563  -5.000  1.00 32.49      B    N
ATOM    276  C8   DG B  14       0.783  -6.596  -5.076  1.00 35.17      B    C
ATOM    277  N3   DG B  14       3.174  -4.153  -5.879  1.00 32.12      B    N
ATOM    278  C2   DG B  14       2.882  -2.863  -5.927  1.00 30.96      B    C
ATOM    279  N2   DG B  14       3.837  -1.973  -6.211  1.00 32.19      B    N
ATOM    280  N1   DG B  14       1.634  -2.354  -5.698  1.00 25.52      B    N
ATOM    281  C6   DG B  14       0.509  -3.099  -5.373  1.00 30.41      B    C
ATOM    282  O6   DG B  14      -0.587  -2.529  -5.208  1.00 31.80      B    O
```

```
ATOM    283  OP2  DG B  14      -0.608 -11.574  -5.988  1.00 65.24      B    O
ATOM    284  P    DC B  15       5.895  -9.786  -3.878  1.00 47.66      B    P
ATOM    285  OP1  DC B  15       7.083 -10.582  -4.175  1.00 45.77      B    O
ATOM    286  O5'  DC B  15       6.352  -8.357  -3.345  1.00 38.19      B    O
ATOM    287  C5'  DC B  15       7.265  -7.568  -4.121  1.00 36.87      B    C
ATOM    288  C4'  DC B  15       7.129  -6.128  -3.706  1.00 32.10      B    C
ATOM    289  C3'  DC B  15       7.394  -5.931  -2.225  1.00 32.83      B    C
ATOM    290  O3'  DC B  15       8.727  -5.527  -2.135  1.00 33.30      B    O
ATOM    291  C2'  DC B  15       6.503  -4.770  -1.844  1.00 36.15      B    C
ATOM    292  C1'  DC B  15       5.503  -4.652  -2.974  1.00 33.11      B    C
ATOM    293  O4'  DC B  15       5.773  -5.703  -3.895  1.00 34.52      B    O
ATOM    294  N1   DC B  15       4.118  -4.776  -2.546  1.00 28.61      B    N
ATOM    295  C6   DC B  15       3.615  -5.979  -2.141  1.00 32.82      B    C
ATOM    296  C5   DC B  15       2.333  -6.105  -1.783  1.00 32.74      B    C
ATOM    297  C4   DC B  15       1.535  -4.923  -1.781  1.00 31.98      B    C
ATOM    298  N4   DC B  15       0.251  -4.983  -1.426  1.00 33.10      B    N
ATOM    299  N3   DC B  15       2.033  -3.736  -2.138  1.00 27.49      B    N
ATOM    300  C2   DC B  15       3.329  -3.626  -2.486  1.00 25.72      B    C
ATOM    301  O2   DC B  15       3.814  -2.552  -2.866  1.00 26.37      B    O
ATOM    302  OP2  DC B  15       4.825 -10.254  -2.978  1.00 46.52      B    O
ATOM    303  P    DG B  16       9.438  -5.339  -0.751  1.00 35.28      B    P
ATOM    304  OP1  DG B  16      10.835  -5.682  -0.951  1.00 43.92      B    O
ATOM    305  O5'  DG B  16       9.265  -3.779  -0.498  1.00 33.04      B    O
ATOM    306  C5'  DG B  16       9.931  -2.839  -1.302  1.00 30.23      B    C
ATOM    307  C4'  DG B  16       9.400  -1.467  -0.991  1.00 30.82      B    C
ATOM    308  C3'  DG B  16       9.796  -1.009   0.403  1.00 30.57      B    C
ATOM    309  O3'  DG B  16      10.181   0.363   0.256  1.00 33.66      B    O
ATOM    310  C2'  DG B  16       8.583  -1.370   1.223  1.00 30.16      B    C
ATOM    311  C1'  DG B  16       7.452  -1.147   0.247  1.00 27.65      B    C
ATOM    312  O4'  DG B  16       7.943  -1.501  -1.015  1.00 28.73      B    O
ATOM    313  N9   DG B  16       6.243  -1.902   0.502  1.00 25.37      B    N
ATOM    314  C4   DG B  16       4.970  -1.405   0.395  1.00 26.60      B    C
ATOM    315  C5   DG B  16       4.144  -2.429   0.790  1.00 26.09      B    C
ATOM    316  N7   DG B  16       4.881  -3.550   1.145  1.00 26.78      B    N
ATOM    317  C8   DG B  16       6.117  -3.203   0.931  1.00 24.67      B    C
ATOM    318  N3   DG B  16       4.637  -0.163  -0.016  1.00 29.64      B    N
ATOM    319  C2   DG B  16       3.339   0.046   0.065  1.00 25.65      B    C
ATOM    320  N2   DG B  16       2.841   1.218  -0.295  1.00 26.47      B    N
ATOM    321  N1   DG B  16       2.428  -0.910   0.436  1.00 26.79      B    N
ATOM    322  C6   DG B  16       2.751  -2.197   0.840  1.00 27.95      B    C
ATOM    323  O6   DG B  16       1.866  -2.969   1.164  1.00 27.91      B    O
ATOM    324  OP2  DG B  16       8.669  -6.002   0.297  1.00 39.64      B    O
ATOM    325  P    DA B  17      10.234   1.424   1.446  1.00 38.13      B    P
ATOM    326  OP1  DA B  17      11.183   2.496   1.064  1.00 37.61      B    O
ATOM    327  OP2  DA B  17      10.339   0.727   2.754  1.00 33.58      B    O
ATOM    328  O5'  DA B  17       8.805   2.135   1.373  1.00 34.32      B    O
ATOM    329  C5'  DA B  17       8.384   2.861   0.191  1.00 33.02      B    C
ATOM    330  C4'  DA B  17       7.231   3.758   0.549  1.00 31.13      B    C
ATOM    331  C3'  DA B  17       7.468   4.649   1.766  1.00 35.18      B    C
ATOM    332  O3'  DA B  17       6.770   5.856   1.499  1.00 39.66      B    O
ATOM    333  C2'  DA B  17       6.786   3.905   2.901  1.00 33.80      B    C
ATOM    334  C1'  DA B  17       5.651   3.181   2.213  1.00 30.96      B    C
ATOM    335  O4'  DA B  17       6.077   2.951   0.876  1.00 31.53      B    O
ATOM    336  N9   DA B  17       5.324   1.875   2.773  1.00 28.32      B    N
ATOM    337  C4   DA B  17       4.072   1.380   3.037  1.00 28.30      B    C
ATOM    338  C5   DA B  17       4.266   0.085   3.485  1.00 27.61      B    C
```

```
ATOM    339  N7   DA B  17       5.612  -0.249   3.450  1.00 26.12      B    N
ATOM    340  C8   DA B  17       6.195   0.847   3.028  1.00 26.53      B    C
ATOM    341  N3   DA B  17       2.902   2.032   2.941  1.00 27.16      B    N
ATOM    342  C2   DA B  17       1.899   1.254   3.342  1.00 27.23      B    C
ATOM    343  N1   DA B  17       1.934  -0.019   3.779  1.00 27.91      B    N
ATOM    344  C6   DA B  17       3.125  -0.651   3.825  1.00 24.10      B    C
ATOM    345  N6   DA B  17       3.166  -1.906   4.269  1.00 27.66      B    N
ATOM    346  P    DA B  18       6.778   7.058   2.537  1.00 38.20      B    P
ATOM    347  OP1  DA B  18       6.645   8.308   1.721  1.00 43.32      B    O
ATOM    348  OP2  DA B  18       7.855   6.823   3.538  1.00 41.19      B    O
ATOM    349  O5'  DA B  18       5.373   6.914   3.234  1.00 32.27      B    O
ATOM    350  C5'  DA B  18       4.210   6.789   2.412  1.00 36.14      B    C
ATOM    351  C4'  DA B  18       3.009   6.675   3.311  1.00 34.86      B    C
ATOM    352  C3'  DA B  18       3.032   7.587   4.538  1.00 36.89      B    C
ATOM    353  O3'  DA B  18       1.779   8.296   4.506  1.00 42.86      B    O
ATOM    354  C2'  DA B  18       3.231   6.625   5.702  1.00 32.51      B    C
ATOM    355  C1'  DA B  18       2.673   5.311   5.175  1.00 32.89      B    C
ATOM    356  O4'  DA B  18       2.963   5.323   3.807  1.00 32.09      B    O
ATOM    357  N9   DA B  18       3.244   4.077   5.700  1.00 29.18      B    N
ATOM    358  C4   DA B  18       2.559   2.937   6.064  1.00 30.18      B    C
ATOM    359  C5   DA B  18       3.526   2.020   6.425  1.00 28.20      B    C
ATOM    360  N7   DA B  18       4.795   2.534   6.210  1.00 27.86      B    N
ATOM    361  C8   DA B  18       4.573   3.748   5.767  1.00 29.38      B    C
ATOM    362  N3   DA B  18       1.224   2.772   6.150  1.00 27.11      B    N
ATOM    363  C2   DA B  18       0.926   1.555   6.607  1.00 28.91      B    C
ATOM    364  N1   DA B  18       1.752   0.571   6.992  1.00 27.71      B    N
ATOM    365  C6   DA B  18       3.082   0.759   6.863  1.00 27.13      B    C
ATOM    366  N6   DA B  18       3.906  -0.231   7.198  1.00 25.58      B    N
ATOM    367  P    DT B  19       1.354   9.281   5.750  1.00 46.96      B    P
ATOM    368  OP1  DT B  19       0.586  10.374   5.187  1.00 59.97      B    O
ATOM    369  O5'  DT B  19       0.397   8.373   6.659  1.00 42.24      B    O
ATOM    370  C5'  DT B  19      -0.683   7.564   6.172  1.00 36.68      B    C
ATOM    371  C4'  DT B  19      -1.136   6.605   7.254  1.00 36.07      B    C
ATOM    372  C3'  DT B  19      -1.392   7.236   8.631  1.00 34.72      B    C
ATOM    373  O3'  DT B  19      -2.781   7.079   8.891  1.00 38.41      B    O
ATOM    374  C2'  DT B  19      -0.522   6.437   9.588  1.00 35.10      B    C
ATOM    375  C1'  DT B  19      -0.285   5.138   8.810  1.00 32.49      B    C
ATOM    376  O4'  DT B  19      -0.138   5.588   7.478  1.00 32.49      B    O
ATOM    377  N1   DT B  19       0.929   4.370   9.164  1.00 30.17      B    N
ATOM    378  C6   DT B  19       2.157   4.978   9.025  1.00 29.58      B    C
ATOM    379  C5   DT B  19       3.323   4.363   9.254  1.00 29.57      B    C
ATOM    380  C7   DT B  19       4.646   5.036   9.074  1.00 33.09      B    C
ATOM    381  C4   DT B  19       3.293   2.978   9.664  1.00 30.05      B    C
ATOM    382  N3   DT B  19       2.022   2.430   9.805  1.00 27.48      B    N
ATOM    383  C2   DT B  19       0.823   3.045   9.537  1.00 28.01      B    C
ATOM    384  O2   DT B  19      -0.245   2.475   9.666  1.00 30.76      B    O
ATOM    385  O4   DT B  19       4.288   2.318   9.958  1.00 27.38      B    O
ATOM    386  OP2  DT B  19       2.562   9.524   6.601  1.00 46.52      B    O
ATOM    387  P    DT B  20      -3.425   7.545  10.310  1.00 43.75      B    P
ATOM    388  OP1  DT B  20      -4.785   8.037   9.992  1.00 56.62      B    O
ATOM    389  O5'  DT B  20      -3.569   6.173  11.090  1.00 37.85      B    O
ATOM    390  C5'  DT B  20      -4.267   5.115  10.469  1.00 34.81      B    C
ATOM    391  C4'  DT B  20      -4.106   3.895  11.331  1.00 38.13      B    C
ATOM    392  C3'  DT B  20      -4.581   4.176  12.750  1.00 36.83      B    C
ATOM    393  O3'  DT B  20      -5.444   3.092  13.028  1.00 40.63      B    O
ATOM    394  C2'  DT B  20      -3.301   4.222  13.568  1.00 32.79      B    C
```

```
ATOM    395  C1'  DT B  20      -2.440   3.233  12.801  1.00 35.72      B    C
ATOM    396  O4'  DT B  20      -2.710   3.522  11.436  1.00 34.54      B    O
ATOM    397  N1   DT B  20      -0.990   3.317  12.967  1.00 31.05      B    N
ATOM    398  C6   DT B  20      -0.373   4.523  12.722  1.00 30.10      B    C
ATOM    399  C5   DT B  20       0.952   4.690  12.736  1.00 32.95      B    C
ATOM    400  C7   DT B  20       1.607   6.008  12.473  1.00 33.96      B    C
ATOM    401  C4   DT B  20       1.781   3.523  12.983  1.00 29.71      B    C
ATOM    402  N3   DT B  20       1.093   2.349  13.210  1.00 29.85      B    N
ATOM    403  C2   DT B  20      -0.266   2.174  13.209  1.00 28.53      B    C
ATOM    404  O2   DT B  20      -0.782   1.098  13.450  1.00 30.71      B    O
ATOM    405  O4   DT B  20       3.003   3.527  13.003  1.00 29.18      B    O
ATOM    406  OP2  DT B  20      -2.417   8.367  11.028  1.00 42.86      B    O
HETATM  407  P   5fC B  21      -6.397   3.156  14.316  1.00 45.81      B    P
HETATM  408  OP1 5fC B  21      -7.641   2.441  13.949  1.00 53.55      B    O
HETATM  409  O5' 5fC B  21      -5.621   2.260  15.383  1.00 38.52      B    O
HETATM  410  C5' 5fC B  21      -5.305   0.887  15.106  1.00 39.92      B    C
HETATM  411  C4' 5fC B  21      -4.219   0.396  16.028  1.00 38.57      B    C
HETATM  412  C3' 5fC B  21      -4.499   0.629  17.518  1.00 37.82      B    C
HETATM  413  O3' 5fC B  21      -4.904  -0.636  18.058  1.00 42.45      B    O
HETATM  414  C2' 5fC B  21      -3.181   1.147  18.091  1.00 41.97      B    C
HETATM  415  C1' 5fC B  21      -2.196   0.881  16.975  1.00 39.77      B    C
HETATM  416  O4' 5fC B  21      -2.950   1.029  15.764  1.00 35.58      B    O
HETATM  417  N1  5fC B  21      -1.046   1.765  16.879  1.00 35.23      B    N
HETATM  418  C6  5fC B  21      -1.196   3.086  16.567  1.00 33.99      B    C
HETATM  419  C5A 5fC B  21      -0.352   5.299  16.088  1.00 40.73      B    C
HETATM  420  O5A 5fC B  21       0.557   6.134  15.970  1.00 46.68      B    O
HETATM  421  C5  5fC B  21      -0.130   3.874  16.374  1.00 35.86      B    C
HETATM  422  C4  5fC B  21       1.157   3.267  16.467  1.00 28.43      B    C
HETATM  423  N4  5fC B  21       2.253   3.995  16.299  1.00 32.91      B    N
HETATM  424  N3  5fC B  21       1.310   1.984  16.786  1.00 30.81      B    N
HETATM  425  C2  5fC B  21       0.227   1.200  16.963  1.00 28.51      B    C
HETATM  426  O2  5fC B  21       0.329   0.005  17.255  1.00 29.85      B    O
HETATM  427  OP2 5fC B  21      -6.418   4.564  14.764  1.00 42.85      B    O
ATOM    428  P    DG B  22      -5.590  -0.642  19.507  1.00 48.14      B    P
ATOM    429  OP1  DG B  22      -6.566  -1.738  19.540  1.00 48.52      B    O
ATOM    430  O5'  DG B  22      -4.361  -0.922  20.495  1.00 45.62      B    O
ATOM    431  C5'  DG B  22      -3.551  -2.115  20.397  1.00 43.78      B    C
ATOM    432  C4'  DG B  22      -2.314  -1.925  21.246  1.00 43.62      B    C
ATOM    433  C3'  DG B  22      -2.620  -1.421  22.657  1.00 46.92      B    C
ATOM    434  O3'  DG B  22      -1.778  -2.098  23.563  1.00 52.80      B    O
ATOM    435  C2'  DG B  22      -2.269   0.058  22.612  1.00 43.47      B    C
ATOM    436  C1'  DG B  22      -1.080  -0.020  21.682  1.00 37.63      B    C
ATOM    437  O4'  DG B  22      -1.450  -0.915  20.666  1.00 36.05      B    O
ATOM    438  N9   DG B  22      -0.584   1.200  21.066  1.00 33.86      B    N
ATOM    439  C4   DG B  22       0.746   1.411  20.810  1.00 31.62      B    C
ATOM    440  C5   DG B  22       0.823   2.674  20.281  1.00 35.96      B    C
ATOM    441  N7   DG B  22      -0.438   3.262  20.225  1.00 35.29      B    N
ATOM    442  C8   DG B  22      -1.239   2.351  20.707  1.00 32.23      B    C
ATOM    443  N3   DG B  22       1.726   0.501  20.961  1.00 30.03      B    N
ATOM    444  C2   DG B  22       2.910   0.991  20.638  1.00 27.87      B    C
ATOM    445  N2   DG B  22       3.988   0.234  20.784  1.00 27.26      B    N
ATOM    446  N1   DG B  22       3.099   2.219  20.071  1.00 29.46      B    N
ATOM    447  C6   DG B  22       2.099   3.162  19.863  1.00 28.03      B    C
ATOM    448  O6   DG B  22       2.381   4.253  19.367  1.00 31.42      B    O
ATOM    449  OP2  DG B  22      -5.925   0.752  19.835  1.00 43.73      B    O
ATOM    450  P  A DC B  23      -2.411  -2.835  24.828  0.70 63.05      B    P
```

```
ATOM    451  P  B DC B  23      -2.226  -2.215  25.065  0.30 57.21      B    P
ATOM    452  OP1A DC B  23      -2.905  -4.172  24.371  0.70 53.56      B    O
ATOM    453  OP1B DC B  23      -3.531  -2.928  25.082  0.30 56.27      B    O
ATOM    454  O5'  DC B  23      -1.129  -3.177  25.709  1.00 55.63      B    O
ATOM    455  C5'  DC B  23      -0.062  -3.886  25.031  1.00 53.67      B    C
ATOM    456  C4'  DC B  23       1.282  -3.291  25.374  1.00 51.17      B    C
ATOM    457  C3'  DC B  23       1.413  -2.797  26.816  1.00 47.47      B    C
ATOM    458  O3'  DC B  23       2.262  -3.653  27.576  1.00 45.15      B    O
ATOM    459  C2'  DC B  23       2.098  -1.453  26.691  1.00 46.78      B    C
ATOM    460  C1'  DC B  23       2.368  -1.279  25.215  1.00 44.28      B    C
ATOM    461  O4'  DC B  23       1.482  -2.132  24.549  1.00 46.14      B    O
ATOM    462  N1   DC B  23       2.081   0.078  24.767  1.00 37.76      B    N
ATOM    463  C6   DC B  23       0.809   0.571  24.805  1.00 37.20      B    C
ATOM    464  C5   DC B  23       0.546   1.824  24.436  1.00 32.17      B    C
ATOM    465  C4   DC B  23       1.629   2.605  23.957  1.00 31.87      B    C
ATOM    466  N4   DC B  23       1.421   3.851  23.563  1.00 33.91      B    N
ATOM    467  N3   DC B  23       2.880   2.136  23.927  1.00 32.45      B    N
ATOM    468  C2   DC B  23       3.135   0.879  24.351  1.00 32.11      B    C
ATOM    469  O2   DC B  23       4.286   0.408  24.374  1.00 32.87      B    O
ATOM    470  OP2A DC B  23      -3.323  -1.855  25.504  0.70 55.11      B    O
ATOM    471  OP2B DC B  23      -2.146  -0.861  25.675  0.30 53.52      B    O
ATOM    472  P    DG B  24       2.435  -3.376  29.155  1.00 50.50      B    P
ATOM    473  OP1  DG B  24       2.710  -4.660  29.813  1.00 52.60      B    O
ATOM    474  O5'  DG B  24       3.821  -2.588  29.291  1.00 34.59      B    O
ATOM    475  C5'  DG B  24       5.059  -3.083  28.827  1.00 32.11      B    C
ATOM    476  C4'  DG B  24       5.991  -1.913  28.629  1.00 29.05      B    C
ATOM    477  C3'  DG B  24       6.190  -1.013  29.843  1.00 31.74      B    C
ATOM    478  O3'  DG B  24       7.106  -1.670  30.743  1.00 33.44      B    O
ATOM    479  C2'  DG B  24       6.699   0.265  29.213  1.00 31.87      B    C
ATOM    480  C1'  DG B  24       6.008   0.285  27.852  1.00 32.26      B    C
ATOM    481  O4'  DG B  24       5.458  -1.032  27.634  1.00 32.57      B    O
ATOM    482  N9   DG B  24       4.911   1.240  27.746  1.00 30.42      B    N
ATOM    483  C4   DG B  24       4.965   2.488  27.160  1.00 29.08      B    C
ATOM    484  C5   DG B  24       3.684   2.989  27.250  1.00 28.01      B    C
ATOM    485  N7   DG B  24       2.841   2.078  27.883  1.00 28.82      B    N
ATOM    486  C8   DG B  24       3.610   1.061  28.161  1.00 29.69      B    C
ATOM    487  N3   DG B  24       6.070   3.084  26.664  1.00 28.49      B    N
ATOM    488  C2   DG B  24       5.799   4.269  26.145  1.00 28.65      B    C
ATOM    489  N2   DG B  24       6.778   4.986  25.584  1.00 30.60      B    N
ATOM    490  N1   DG B  24       4.554   4.836  26.145  1.00 30.24      B    N
ATOM    491  C6   DG B  24       3.415   4.272  26.707  1.00 30.60      B    C
ATOM    492  O6   DG B  24       2.337   4.883  26.657  1.00 33.90      B    O
ATOM    493  OP2  DG B  24       1.340  -2.476  29.617  1.00 46.76      B    O
TER     495       DG B  24
HETATM  494  O   HOH D   1      -1.235   3.732   4.967  1.00 33.32           O
HETATM  495  O   HOH D   2       7.433   1.866   6.580  1.00 35.46           O
HETATM  496  O   HOH D   3      -3.248   3.702   7.088  1.00 31.96           O
HETATM  497  O   HOH D   4      -0.686   5.471   3.037  1.00 33.80           O
HETATM  498  O   HOH D   5       1.299   3.980   1.638  1.00 29.48           O
HETATM  499  O   HOH D   6      -2.796   1.657   9.007  1.00 32.07           O
HETATM  500  O   HOH D   7      -2.389  -1.131  12.704  1.00 34.16           O
HETATM  501  O   HOH D   8      -4.168   0.271  10.872  1.00 32.19           O
HETATM  502  O   HOH D   9      -5.494   5.247   6.748  1.00 41.05           O
HETATM  503  O   HOH D  10      -7.753   6.312   3.107  0.50 34.38           O
HETATM  504  O   HOH D  11       2.364   5.060  -0.380  1.00 41.29           O
HETATM  505  O   HOH D  12       2.242   3.285  -2.738  1.00 28.85           O
```

```
HETATM  506  O   HOH D  13       4.898   4.355  -1.815  0.50 25.06           O
HETATM  507  O   HOH D  14       6.853  -0.930   7.245  1.00 38.82           O
HETATM  508  O   HOH D  15       6.843  -2.303   4.608  1.00 38.51           O
HETATM  509  O   HOH D  16       4.607  -4.679   7.406  0.50 28.91           O
HETATM  510  O   HOH D  17       0.869  -5.582   4.449  1.00 34.87           O
HETATM  511  O   HOH D  18       1.448  -5.739   1.661  1.00 43.62           O
HETATM  512  O   HOH D  19       1.415   7.262  21.771  0.50 30.85           O
HETATM  513  O   HOH D  21      16.991  -2.215  16.301  0.50 34.08           O
HETATM  514  O   HOH D  22      10.557  -0.035  13.154  0.50 34.41           O
HETATM  515  O   HOH D  23      11.050  -0.243   6.784  0.50 36.17           O
HETATM  516  O   HOH D  24       5.586  -4.164  10.221  0.50 27.55           O
HETATM  517  O   HOH D  25       0.292  -8.834   9.617  0.50 28.42           O
HETATM  518  O   HOH D  26      -6.129   0.688   0.238  0.50 26.28           O
HETATM  519  O   HOH D  27      -3.753   1.339  -6.468  1.00 40.42           O
HETATM  520  O   HOH D  28      -3.342  -3.599  -4.471  0.50 38.93           O
HETATM  521  O   HOH D  29      -3.410  11.131  11.508  0.50 36.07           O
HETATM  522  O   HOH D  30       6.549  -2.220   9.846  0.50 24.58           O
HETATM  523  O   HOH D  31      -5.264  -0.871  -5.608  0.50 37.78           O
HETATM  524  O   HOH D  32       1.654  -2.427  21.810  0.50 30.64           O
HETATM  525  O   HOH D  33      13.668  -4.092  15.692  0.50 37.66           O
HETATM  526  O   HOH D  34       4.002   8.540  19.726  0.50 29.90           O
HETATM  527  O   HOH D  35       7.435   1.468  12.226  0.50 32.35           O
HETATM  528  O   HOH D  36       4.958   5.388  12.916  0.50 24.22           O
HETATM  529  O   HOH D  37       7.379   4.989  12.247  0.50 31.89           O
HETATM  530  O   HOH D  38       2.911   8.402   9.179  0.50 32.48           O
HETATM  532  O   HOH D  40      -1.224   5.098  23.716  0.50 32.01           O
HETATM  533  O   HOH D  41       0.575   9.124  -2.874  0.50 37.11           O
HETATM  534  O   HOH D  42      -5.646  -3.658   4.484  0.50 33.30           O
HETATM  535  O   HOH D  43       9.996  -9.555  -0.480  0.50 42.10           O
HETATM  536  O   HOH D  45       6.760   6.132   6.242  0.50 24.09           O
HETATM  537  O   HOH D  46       0.474   9.434  11.141  0.50 27.15           O
HETATM  538  O   HOH D  47      -2.819  -5.561  -3.598  0.50 35.36           O
HETATM  539  O   HOH D  48      -9.076   5.897  -1.749  0.50 38.31           O
HETATM  540  O   HOH D  49      -2.498   4.972  -8.464  0.50 44.09           O
HETATM  541  O   HOH D  50      -3.105  11.098   6.472  0.50 47.78           O
HETATM  542  O   HOH D  51      -4.028   4.461  17.622  0.50 35.08           O
HETATM  543  O   HOH D  52     -10.123  -1.526  20.655  0.50 46.82           O
HETATM  544  O   HOH D  53       0.143   1.387  29.067  0.50 39.83           O
HETATM  545  O   HOH D  54      -1.704 -10.349  11.681  0.50 29.01           O
HETATM  546  O   HOH D  55      -6.503  -5.881  -8.818  0.50 41.85           O
HETATM  547  O   HOH D  56      -0.821  11.591  -7.683  1.00 55.79           O
END
```

File A-3: Dictionary file of 5-formyl-2’-deoxycytidine.

```
global_
_lib_name         ?
_lib_version      ?
_lib_update       ?
# ------------------------------------------------
#
# ---   LIST OF MONOMERS ---
#
data_comp_list
loop_
_chem_comp.id
_chem_comp.three_letter_code
_chem_comp.name
_chem_comp.group
_chem_comp.number_atoms_all
_chem_comp.number_atoms_nh
_chem_comp.desc_level
5fC      5fC  '2'-DEOXY-formylCYTIDINE-5'-MONOPHOSPHATE ' DNA          33  21 .
#
# --- DESCRIPTION OF MONOMERS ---
#
data_comp_5fC
#
loop_
_chem_comp_atom.comp_id
_chem_comp_atom.atom_id
_chem_comp_atom.type_symbol
_chem_comp_atom.type_energy
_chem_comp_atom.partial_charge
_chem_comp_atom.x
_chem_comp_atom.y
_chem_comp_atom.z
 5fC           P      P    P          0.000      0.001   -0.540    0.119
 5fC           OP3    O    OP        -0.660      0.227   -1.987    0.516
 5fC           OP1    O    OP        -0.660      0.381    0.441    1.194
 5fC           "O5'"  O    O2         0.000     -1.613   -0.460    0.002
 5fC           "C5'"  C    CH2        0.000     -2.412    0.695   -0.366
 5fC           "H5'"  H    H          0.000     -2.004    1.206   -1.241
 5fC           "H5''" H    H          0.000     -2.500    1.405    0.459
 5fC           "C4'"  C    CH1        0.000     -3.783    0.131   -0.701
 5fC           "H4'"  H    H          0.000     -3.606   -0.705   -1.391
 5fC           "C3'"  C    CH1        0.000     -4.862    0.991   -1.350
 5fC           "H3'"  H    H          0.000     -4.664    2.061   -1.199
 5fC           "O3'"  O    OH1        0.000     -4.928    0.664   -2.754
 5fC           "HO3'" H    H          0.000     -5.661    1.144   -3.165
 5fC           "C2'"  C    CH2        0.000     -6.185    0.558   -0.660
 5fC           "H2' " H    H          0.000     -6.865    0.104   -1.383
 5fC           "H2''" H    H          0.000     -6.673    1.415   -0.192
 5fC           "C1'"  C    CH1        0.000     -5.815   -0.478    0.421
 5fC           "H1'"  H    H          0.000     -6.089   -1.462    0.015
 5fC           "O4'"  O    O2         0.000     -4.377   -0.452    0.488
```

```
 5fC           N1     N    NR6       0.000     -6.434   -0.388    1.735
 5fC           C6     C    CR16      0.000     -7.038   -1.463    2.305
 5fC           H6     H    H         0.000     -7.005   -2.406    1.774
 5fC           H5C    H    H        -0.500      3.069    7.403    0.644
 5fC           C5A    C    C         0.000      1.797    7.837    0.471
 5fC           O5A    O    OC       -0.500      1.352    8.985    0.585
 5fC           C5     C    CR16      0.000     -7.709   -1.424    3.556
 5fC           H5     H    H         0.000     -8.160   -2.314    3.979
 5fC           C4     C    CR6       0.000     -7.767   -0.214    4.213
 5fC           N4     N    NH2       0.000     -8.312   -0.031    5.446
 5fC           H42    H    H         0.000     -8.712   -0.813    5.946
 5fC           H41    H    H         0.000     -8.318    0.889    5.866
 5fC           N3     N    NRD6      0.000     -6.825    0.681    3.769
 5fC           C2     C    CR6       0.000     -6.022    0.556    2.670
 5fC           O2     O    O         0.000     -4.961    1.193    2.528
 5fC           OP2    O    OP       -0.660      0.506   -0.216   -1.268
loop_
_chem_comp_tree.comp_id
_chem_comp_tree.atom_id
_chem_comp_tree.atom_back
_chem_comp_tree.atom_forward
_chem_comp_tree.connect_type
 5fC      "O5'"   n/a  "C5'"  START
 5fC      P      "O5'"     OP2 .
 5fC      OP3    P       .      .
 5fC      OP1    P       .      .
 5fC      "C5'"  "O5'"  "C4'"  .
 5fC      "H5'" "C5'"   .      .
 5fC      "H5''" "C5'"  .      .
 5fC      "C4'"  "C5'"  "C3'"  .
 5fC      "H4'"  "C4'"  .      .
 5fC      "C3'"  "C4'"  "C2'"  .
 5fC      "H3'"  "C3'"  .      .
 5fC      "O3'"  "C3'"  "HO3'" .
 5fC      "HO3'" "O3'"  .      .
 5fC      "C2'"  "C3'"  "C1'"  .
 5fC      "H2' " "C2'"  .      .
 5fC      "H2''" "C2'"  .      .
 5fC      "C1'"  "C2'"  N1     .
 5fC      "H1'"  "C1'"  .      .
 5fC      "O4'"  "C1'"  .      .
 5fC      N1     "C1'"  C6     .
 5fC      C5A    H5C    C5     .
 5fC      O5A    C5A    .      .
 5fC      C6     N1     C5     .
 5fC      H6     C6     .      .
 5fC      C5     C6     C4     .
 5fC      H5     C5     .      .
 5fC      C4     C5     N3     .
 5fC      N4     C4     H41    .
 5fC      H42    N4     .      .
 5fC      H41    N4     .      .
 5fC      N3     C4     C2     .
 5fC      C2     N3     O2     .
 5fC      O2     C2     .      END
 5fC      OP2    P      .     .
```

```
 5fC      "C4'"  "O4'"  .    ADD
 5fC      N1     C2     .    ADD
loop_
_chem_comp_bond.comp_id
_chem_comp_bond.atom_id_1
_chem_comp_bond.atom_id_2
_chem_comp_bond.type
_chem_comp_bond.value_dist
_chem_comp_bond.value_dist_esd
 5fC      OP3    P         deloc       1.485    0.017
 5fC      OP1    P         deloc       1.485    0.017
 5fC      OP2    P         deloc       1.485    0.017
 5fC      "O5'"  P         single      1.593    0.010
 5fC      "C5'"  "O5'"     single      1.440    0.016
 5fC      "C4'"  "C5'"     single      1.511    0.008
 5fC      "C4'"  "O4'"     single      1.446    0.011
 5fC      "C3'"  "C4'"     single      1.528    0.010
 5fC      "O4'"  "C1'"     single      1.420    0.013
 5fC      "O3'"  "C3'"     single      1.431    0.013
 5fC      "C2'"  "C3'"     single      1.518    0.010
 5fC      "C1'"  "C2'"     single      1.521    0.014
 5fC      N1     "C1'"     single      1.470    0.012
 5fC      N1     C2        single      1.397    0.010
 5fC      C6     N1        single      1.367    0.006
 5fC      O2     C2        double      1.240    0.009
 5fC      C2     N3        single      1.353    0.008
 5fC      N3     C4        single      1.335    0.007
 5fC      N4     C4        single      1.335    0.009
 5fC      C4     C5        single      1.425    0.008
 5fC      C5     C5A       single      1.452    0.020
 5fC      O5A    C5A       double      1.227    0.020
 5fC      C5     C6        single      1.339    0.008
 5fC      "H5'"  "C5'"     single      1.092    0.020
 5fC      "H5''" "C5'"     single      1.092    0.020
 5fC      "H4'"  "C4'"     single      1.099    0.020
 5fC      "H3'"  "C3'"     single      1.099    0.020
 5fC      "HO3'" "O3'"     single      0.967    0.020
 5fC      "H2' " "C2'"     single      1.092    0.020
 5fC      "H2''" "C2'"     single      1.092    0.020
 5fC      "H1'"  "C1'"     single      1.099    0.020
 5fC      H41    N4        single      1.010    0.020
 5fC      H42    N4        single      1.010    0.020
 5fC      H6     C6        single      1.083    0.020
loop_
_chem_comp_angle.comp_id
_chem_comp_angle.atom_id_1
_chem_comp_angle.atom_id_2
_chem_comp_angle.atom_id_3
_chem_comp_angle.value_angle
_chem_comp_angle.value_angle_esd
 5fC      OP3    P      OP1     107.400    3.200
 5fC      OP3    P      "O5'"   104.000    1.900
 5fC      OP1    P      "O5'"   108.100    2.900
 5fC      OP3    P      OP2     108.300    3.200
 5fC      OP1    P      OP2     119.600    1.500
 5fC      "O5'"  P      OP2     108.300    2.700
```

```
5fC      P      "O5'"  "C5'"   120.900    1.600
5fC      "O5'"  "C5'"  "H5'"  109.470    3.000
5fC      "O5'"  "C5'"  "H5''"  109.470    3.000
5fC      "O5'"  "C5'"  "C4'"   110.200    1.400
5fC      "H5'" "C5'"  "H5''"  107.900    3.000
5fC      "H5'" "C5'"  "C4'"   109.470    3.000
5fC      "H5''" "C5'"  "C4'"   109.470    3.000
5fC      "C5'"  "C4'"  "H4'"   108.340    3.000
5fC      "C5'"  "C4'"  "C3'"   114.700    1.500
5fC      "C5'"  "C4'"  "O4'"   109.400    1.600
5fC      "H4'"  "C4'"  "C3'"   108.340    3.000
5fC      "H4'"  "C4'"  "O4'"   109.470    3.000
5fC      "C3'"  "C4'"  "O4'"   105.600    1.000
5fC      "C4'"  "C3'"  "H3'"   108.340    3.000
5fC      "C4'"  "C3'"  "O3'"   110.300    2.200
5fC      "C4'"  "C3'"  "C2'"   103.200    1.000
5fC      "H3'"  "C3'"  "O3'"   109.470    3.000
5fC      "H3'"  "C3'"  "C2'"   108.340    3.000
5fC      "O3'"  "C3'"  "C2'"   110.600    2.700
5fC      "C3'"  "O3'"  "HO3'"  109.470    3.000
5fC      "C3'"  "C2'"  "H2' "  109.470    3.000
5fC      "C3'"  "C2'"  "H2''"  109.470    3.000
5fC      "C3'"  "C2'"  "C1'"   102.700    1.400
5fC      "H2' " "C2'"  "H2''"  107.900    3.000
5fC      "H2' " "C2'"  "C1'"   109.470    3.000
5fC      "H2''" "C2'"  "C1'"   109.470    3.000
5fC      "C2'"  "C1'"  "H1'"   108.340    3.000
5fC      "C2'"  "C1'"  "O4'"   106.100    1.100
5fC      "C2'"  "C1'"  N1      114.200    1.600
5fC      "H1'"  "C1'"  "O4'"   109.470    3.000
5fC      "H1'"  "C1'"  N1      109.470    3.000
5fC      "O4'"  "C1'"  N1      107.800    0.800
5fC      "C1'"  "O4'"  "C4'"   109.700    1.400
5fC      "C1'"  N1     C6      120.800    1.200
5fC      "C1'"  N1     C2      118.800    1.100
5fC      C6     N1     C2      120.300    0.400
5fC      N1     C6     H6      120.000    3.000
5fC      N1     C6     C5      121.000    0.500
5fC      H6     C6     C5      120.000    3.000
5fC      H5C    C5A    O5A     119.783    3.000
5fC      H5C    C5A    C5      114.626    3.000
5fC      O5A    C5A    C5      125.591    3.000
5fC      C5A    C5     C4      124.414    3.000
5fC      C5A    C5     C6      119.940    3.000
5fC      C6     C5     H5      120.000    3.000
5fC      C6     C5     C4      117.400    0.500
5fC      H5     C5     C4      120.000    3.000
5fC      C5     C4     N4      120.200    0.700
5fC      C5     C4     N3      121.900    0.400
5fC      N4     C4     N3      118.000    0.700
5fC      C4     N4     H42     120.000    3.000
5fC      C4     N4     H41     120.000    3.000
5fC      H42    N4     H41     120.000    3.000
5fC      C4     N3     C2      119.900    0.500
5fC      N3     C2     O2      121.900    0.700
5fC      N3     C2     N1      119.200    0.700
```

```
 5fC      O2     C2     N1      118.900    0.600
loop_
_chem_comp_tor.comp_id
_chem_comp_tor.id
_chem_comp_tor.atom_id_1
_chem_comp_tor.atom_id_2
_chem_comp_tor.atom_id_3
_chem_comp_tor.atom_id_4
_chem_comp_tor.value_angle
_chem_comp_tor.value_angle_esd
_chem_comp_tor.period
 5fC      var_1    OP2    P      "O5'"  "C5'"    -60.014   20.000   1
 5fC      var_2    P      "O5'"  "C5'"  "C4'"    179.972   20.000   1
 5fC      var_3    "O5'"  "C5'"  "C4'"  "C3'"   -179.996   20.000   3
 5fC      var_4    "C5'"  "C4'"  "O4'"  "C1'"    150.000   20.000   1
 5fC      var_5    "C5'"  "C4'"  "C3'"  "C2'"   -150.000   20.000   3
 5fC      var_6    "C4'"  "C3'"  "O3'"  "HO3'"   175.000   20.000   1
 5fC      var_7    "C4'"  "C3'"  "C2'"  "C1'"     30.000   20.000   3
 5fC      var_8    "C3'"  "C2'"  "C1'"  N1       120.000   20.000   3
 5fC      var_9    "C2'"  "C1'"  "O4'"  "C4'"    -30.000   20.000   1
 5fC      var_10   "C2'"  "C1'"  N1     C6       120.298   20.000   1
 5fC      CONST_1  "C1'"  N1     C2     N3       180.000    0.000   0
 5fC      CONST_2  "C1'"  N1     C6     C5       180.000    0.000   0
 5fC      CONST_3  N1     C6     C5     C4         0.000    0.000   0
 5fC      CONST_4  C6     C5     C4     N3         0.000    0.000   0
 5fC      CONST_5  C5     C4     N4     H41      180.000    0.000   0
 5fC      CONST_6  C5     C4     N3     C2         0.000    0.000   0
 5fC      CONST_7  C4     N3     C2     O2       180.000    0.000   0
loop_
_chem_comp_chir.comp_id
_chem_comp_chir.id
_chem_comp_chir.atom_id_centre
_chem_comp_chir.atom_id_1
_chem_comp_chir.atom_id_2
_chem_comp_chir.atom_id_3
_chem_comp_chir.volume_sign
 5fC      chir_01  "C4'"  "C5'"  "O4'"  "C3'"     negativ
 5fC      chir_02  "C3'"  "C4'"  "O3'"  "C2'"     negativ
 5fC      chir_03  "C1'"  "O4'"  "C2'"  N1        positiv
loop_
_chem_comp_plane_atom.comp_id
_chem_comp_plane_atom.plane_id
_chem_comp_plane_atom.atom_id
_chem_comp_plane_atom.dist_esd
 5fC      plan-1    N1        0.020
 5fC      plan-1    "C1'"     0.020
 5fC      plan-1    C2        0.020
 5fC      plan-1    C6        0.020
 5fC      plan-1    N3        0.020
 5fC      plan-1    C4        0.020
 5fC      plan-1    C5        0.020
 5fC      plan-1    O2        0.020
 5fC      plan-1    N4        0.020
 5fC      plan-1    H5        0.020
 5fC      plan-1    H6        0.020
 5fC      plan-1    H42       0.020
```

```
 5fC      plan-1    H41       0.020
 5fC      plan-1    C5A       0.020
 5fC      plan-2    N4        0.020
 5fC      plan-2    C4        0.020
 5fC      plan-2    H41       0.020
 5fC      plan-2    H42       0.020
# ------------------------------------------------------
# ------------------------------------------------------
# ------------------------------------------------------
```

File A-4: PDB coordinates for the crystal structure of 5-carboxyl-2'-deoxycytidine in B-type DNA, DDD[ca].

```
HEADER    ----
COMPND    ---
REMARK   3
REMARK   3 REFINEMENT.
REMARK   3   PROGRAM     : REFMAC 5.8.0049
REMARK   3   AUTHORS     : MURSHUDOV,SKUBAK,LEBEDEV,PANNU,
REMARK   3                 STEINER,NICHOLLS,WINN,LONG,VAGIN
REMARK   3
REMARK   3    REFINEMENT TARGET : MAXIMUM LIKELIHOOD
REMARK   3
REMARK   3  DATA USED IN REFINEMENT.
REMARK   3   RESOLUTION RANGE HIGH (ANGSTROMS) :   1.95
REMARK   3   RESOLUTION RANGE LOW  (ANGSTROMS) :  25.89
REMARK   3   DATA CUTOFF            (SIGMA(F)) : NONE
REMARK   3   COMPLETENESS FOR RANGE        (%) :  97.54
REMARK   3   NUMBER OF REFLECTIONS             :    4704
REMARK   3
REMARK   3  FIT TO DATA USED IN REFINEMENT.
REMARK   3   CROSS-VALIDATION METHOD          : THROUGHOUT
REMARK   3   FREE R VALUE TEST SET SELECTION  : RANDOM
REMARK   3   R VALUE     (WORKING + TEST SET) : 0.22492
REMARK   3   R VALUE           (WORKING SET) :  0.22129
REMARK   3   FREE R VALUE                     :  0.26721
REMARK   3   FREE R VALUE TEST SET SIZE   (%) :  8.0
REMARK   3   FREE R VALUE TEST SET COUNT      :   409
REMARK   3
REMARK   3  FIT IN THE HIGHEST RESOLUTION BIN.
REMARK   3   TOTAL NUMBER OF BINS USED           :      20
REMARK   3   BIN RESOLUTION RANGE HIGH           :    1.950
REMARK   3   BIN RESOLUTION RANGE LOW            :    2.000
REMARK   3   REFLECTION IN BIN     (WORKING SET) :      352
REMARK   3   BIN COMPLETENESS (WORKING+TEST) (%) :    99.48
REMARK   3   BIN R VALUE           (WORKING SET) :    0.331
REMARK   3   BIN FREE R VALUE SET COUNT          :       30
REMARK   3   BIN FREE R VALUE                    :    0.261
REMARK   3
REMARK   3  NUMBER OF NON-HYDROGEN ATOMS USED IN REFINEMENT.
REMARK   3   ALL ATOMS               :      508
REMARK   3
REMARK   3  B VALUES.
REMARK   3   FROM WILSON PLOT           (A**2) : NULL
REMARK   3   MEAN B VALUE      (OVERALL, A**2) :  45.709
REMARK   3   OVERALL ANISOTROPIC B VALUE.
REMARK   3    B11 (A**2) :     0.97
REMARK   3    B22 (A**2) :     1.63
REMARK   3    B33 (A**2) :    -2.60
REMARK   3    B12 (A**2) :     0.00
REMARK   3    B13 (A**2) :    -0.00
REMARK   3    B23 (A**2) :    -0.00
```

```
REMARK   3
REMARK   3  ESTIMATED OVERALL COORDINATE ERROR.
REMARK   3   ESU BASED ON R VALUE                            (A):   0.195
REMARK   3   ESU BASED ON FREE R VALUE                       (A):   0.178
REMARK   3   ESU BASED ON MAXIMUM LIKELIHOOD                 (A):   0.129
REMARK   3   ESU FOR B VALUES BASED ON MAXIMUM LIKELIHOOD (A**2):   4.712
REMARK   3
REMARK   3 CORRELATION COEFFICIENTS.
REMARK   3   CORRELATION COEFFICIENT FO-FC      :   0.967
REMARK   3   CORRELATION COEFFICIENT FO-FC FREE :   0.949
REMARK   3
REMARK   3  RMS DEVIATIONS FROM IDEAL VALUES        COUNT    RMS    WEIGHT
REMARK   3   BOND LENGTHS REFINED ATOMS        (A):   555 ; 0.009 ; 0.012
REMARK   3   BOND LENGTHS OTHERS               (A):   266 ; 0.002 ; 0.020
REMARK   3   BOND ANGLES REFINED ATOMS   (DEGREES):   855 ; 2.209 ; 1.307
REMARK   3   BOND ANGLES OTHERS          (DEGREES):   632 ; 3.717 ; 3.000
REMARK   3   CHIRAL-CENTER RESTRAINTS       (A**3):    72 ; 0.164 ; 0.200
REMARK   3   GENERAL PLANES REFINED ATOMS      (A):   290 ; 0.022 ; 0.020
REMARK   3   GENERAL PLANES OTHERS             (A):   112 ; 0.003 ; 0.020
REMARK   3
REMARK   3  ISOTROPIC THERMAL FACTOR RESTRAINTS.     COUNT   RMS    WEIGHT
REMARK   3   SIDE-CHAIN BOND REFINED ATOMS  (A**2):   555 ; 5.348 ; 4.691
REMARK   3   SIDE-CHAIN BOND OTHER ATOMS  (A**2)  :   554 ; 5.346 ; 4.692
REMARK   3   SIDE-CHAIN ANGLE OTHER ATOMS (A**2)  :   856 ; 7.305 ; 7.043
REMARK   3   LONG RANGE B REFINED ATOMS (A**2)    :   777 ; 8.229 ;42.681
REMARK   3   LONG RANGE B OTHER ATOMS (A**2)      :   777 ; 8.218 ;42.680
REMARK   3
REMARK   3  NCS RESTRAINTS STATISTICS
REMARK   3   NUMBER OF NCS GROUPS : NULL
REMARK   3
REMARK   3  TWIN DETAILS
REMARK   3   NUMBER OF TWIN DOMAINS  : NULL
REMARK   3
REMARK   3
REMARK   3  TLS DETAILS
REMARK   3   NUMBER OF TLS GROUPS  : NULL
REMARK   3
REMARK   3
REMARK   3  BULK SOLVENT MODELLING.
REMARK   3   METHOD USED :  MASK
REMARK   3   PARAMETERS FOR MASK CALCULATION
REMARK   3   VDW PROBE RADIUS   :   1.20
REMARK   3   ION PROBE RADIUS   :   0.80
REMARK   3   SHRINKAGE RADIUS   :   0.80
REMARK   3
REMARK   3  OTHER REFINEMENT REMARKS:
REMARK   3  HYDROGENS HAVE BEEN ADDED IN THE RIDING POSITIONS
REMARK   3  U VALUES      : REFINED INDIVIDUALLY
REMARK   3
LINK         O3'  DT A   8                 P   5CC A   9                 p
LINK         O3' 5CC A   9                 P    DG A  10                 p
LINK         O3'  DT B  20                 P   5CC B  21                 p
LINK         O3' 5CC B  21                 P    DG B  22                 p
CRYST1   24.250   41.340   66.410  90.00  90.00  90.00 P 21 21 21
SCALE1      0.041237  0.000000  0.000000        0.00000
SCALE2     -0.000000  0.024190  0.000000        0.00000
```

```
SCALE3      0.000000 -0.000000  0.015058        0.00000
ATOM      1  O5' DC  A   1      -7.052 -14.063  24.484  1.00 73.36      A    O
ATOM      2  C5' DC  A   1      -5.905 -13.257  24.182  1.00 62.09      A    C
ATOM      3  C4' DC  A   1      -6.351 -11.836  23.928  1.00 63.76      A    C
ATOM      4  C3' DC  A   1      -6.443 -11.252  22.509  1.00 61.30      A    C
ATOM      5  O3' DC  A   1      -7.673 -10.550  22.278  1.00 65.33      A    O
ATOM      6  C2' DC  A   1      -5.444 -10.113  22.580  1.00 61.59      A    C
ATOM      7  C1' DC  A   1      -5.574  -9.641  24.048  1.00 53.10      A    C
ATOM      8  O4' DC  A   1      -5.848 -10.802  24.824  1.00 53.30      A    O
ATOM      9  N1  DC  A   1      -4.355  -9.007  24.567  1.00 48.71      A    N
ATOM     10  C6  DC  A   1      -3.134  -9.581  24.349  1.00 58.64      A    C
ATOM     11  C5  DC  A   1      -2.007  -8.978  24.749  1.00 60.56      A    C
ATOM     12  C4  DC  A   1      -2.131  -7.709  25.388  1.00 51.04      A    C
ATOM     13  N4  DC  A   1      -1.042  -7.051  25.785  1.00 54.76      A    N
ATOM     14  N3  DC  A   1      -3.321  -7.145  25.619  1.00 45.32      A    N
ATOM     15  C2  DC  A   1      -4.447  -7.752  25.191  1.00 47.51      A    C
ATOM     16  O2  DC  A   1      -5.566  -7.259  25.389  1.00 45.69      A    O
ATOM     17  P  ADG  A   2      -8.351 -10.327  20.812  0.60 49.70      A    P
ATOM     18  P  BDG  A   2      -9.153 -11.222  22.408  0.40 59.25      A    P
ATOM     19  OP1ADG  A   2      -9.000 -11.582  20.404  0.60 50.59      A    O
ATOM     20  OP1BDG  A   2      -9.335 -11.666  23.809  0.40 58.44      A    O
ATOM     21  O5'ADG  A   2      -9.518  -9.334  21.278  0.60 54.51      A    O
ATOM     22  O5'BDG  A   2     -10.039  -9.931  22.140  0.40 58.67      A    O
ATOM     23  C5' DG  A   2      -9.415  -8.734  22.621  1.00 62.56      A    C
ATOM     24  C4' DG  A   2     -10.317  -7.546  22.797  1.00 55.32      A    C
ATOM     25  C3' DG  A   2     -10.782  -6.859  21.508  1.00 51.37      A    C
ATOM     26  O3' DG  A   2     -11.857  -5.953  21.672  1.00 52.05      A    O
ATOM     27  C2' DG  A   2      -9.616  -5.975  21.144  1.00 42.77      A    C
ATOM     28  C1' DG  A   2      -8.949  -5.668  22.484  1.00 47.50      A    C
ATOM     29  O4' DG  A   2      -9.434  -6.605  23.442  1.00 54.26      A    O
ATOM     30  N9  DG  A   2      -7.481  -5.744  22.467  1.00 36.58      A    N
ATOM     31  C4  DG  A   2      -6.637  -4.739  22.850  1.00 32.32      A    C
ATOM     32  C5  DG  A   2      -5.363  -5.227  22.631  1.00 40.20      A    C
ATOM     33  N7  DG  A   2      -5.407  -6.489  22.051  1.00 36.88      A    N
ATOM     34  C8  DG  A   2      -6.684  -6.752  21.966  1.00 43.75      A    C
ATOM     35  N3  DG  A   2      -7.009  -3.574  23.423  1.00 32.55      A    N
ATOM     36  C2  DG  A   2      -5.970  -2.825  23.730  1.00 34.57      A    C
ATOM     37  N2  DG  A   2      -6.163  -1.606  24.252  1.00 35.05      A    N
ATOM     38  N1  DG  A   2      -4.665  -3.214  23.532  1.00 32.46      A    N
ATOM     39  C6  DG  A   2      -4.266  -4.432  22.979  1.00 31.98      A    C
ATOM     40  O6  DG  A   2      -3.066  -4.703  22.873  1.00 35.68      A    O
ATOM     41  OP2ADG  A   2      -7.363  -9.714  19.886  0.60 48.06      A    O
ATOM     42  OP2BDG  A   2      -9.354 -12.172  21.289  0.40 61.54      A    O
ATOM     43  P   DC  A   3     -12.543  -5.279  20.386  1.00 61.14      A    P
ATOM     44  OP1 DC  A   3     -13.982  -5.060  20.730  1.00 60.08      A    O
ATOM     45  O5' DC  A   3     -11.833  -3.871  20.239  1.00 47.64      A    O
ATOM     46  C5' DC  A   3     -12.087  -2.899  21.230  1.00 42.68      A    C
ATOM     47  C4' DC  A   3     -11.068  -1.816  21.039  1.00 40.14      A    C
ATOM     48  C3' DC  A   3     -11.129  -1.226  19.639  1.00 38.58      A    C
ATOM     49  O3' DC  A   3     -11.931  -0.043  19.680  1.00 52.48      A    O
ATOM     50  C2' DC  A   3      -9.673  -0.960  19.280  1.00 43.61      A    C
ATOM     51  C1' DC  A   3      -8.866  -1.593  20.411  1.00 33.96      A    C
ATOM     52  O4' DC  A   3      -9.770  -2.425  21.129  1.00 34.90      A    O
ATOM     53  N1  DC  A   3      -7.760  -2.446  19.979  1.00 36.21      A    N
ATOM     54  C6  DC  A   3      -8.011  -3.649  19.379  1.00 39.52      A    C
ATOM     55  C5  DC  A   3      -7.016  -4.475  19.051  1.00 42.17      A    C
```

```
ATOM     56  C4  DC  A   3      -5.690  -4.053  19.354  1.00 37.92      A    C
ATOM     57  N4  DC  A   3      -4.652  -4.820  19.026  1.00 39.08      A    N
ATOM     58  N3  DC  A   3      -5.439  -2.886  19.939  1.00 31.81      A    N
ATOM     59  C2  DC  A   3      -6.454  -2.070  20.283  1.00 36.23      A    C
ATOM     60  O2  DC  A   3      -6.254  -0.969  20.820  1.00 33.44      A    O
ATOM     61  OP2 DC  A   3     -12.066  -5.955  19.157  1.00 53.75      A    O
ATOM     62  P   DG  A   4     -12.355   0.658  18.288  1.00 57.75      A    P
ATOM     63  OP1 DG  A   4     -13.630   1.365  18.529  1.00 61.40      A    O
ATOM     64  O5' DG  A   4     -11.231   1.770  18.124  1.00 53.26      A    O
ATOM     65  C5' DG  A   4     -11.068   2.713  19.204  1.00 52.38      A    C
ATOM     66  C4' DG  A   4      -9.860   3.589  18.975  1.00 56.18      A    C
ATOM     67  C3' DG  A   4      -9.904   4.367  17.663  1.00 51.80      A    C
ATOM     68  O3' DG  A   4      -9.301   5.645  17.866  1.00 62.96      A    O
ATOM     69  C2' DG  A   4      -9.116   3.493  16.709  1.00 52.63      A    C
ATOM     70  C1' DG  A   4      -8.069   2.874  17.630  1.00 46.75      A    C
ATOM     71  O4' DG  A   4      -8.652   2.787  18.919  1.00 46.49      A    O
ATOM     72  N9  DG  A   4      -7.628   1.542  17.268  1.00 44.50      A    N
ATOM     73  C4  DG  A   4      -6.351   1.057  17.419  1.00 43.27      A    C
ATOM     74  C5  DG  A   4      -6.405  -0.254  16.994  1.00 35.39      A    C
ATOM     75  N7  DG  A   4      -7.676  -0.572  16.539  1.00 39.85      A    N
ATOM     76  C8  DG  A   4      -8.370   0.520  16.726  1.00 47.00      A    C
ATOM     77  N3  DG  A   4      -5.294   1.745  17.897  1.00 40.90      A    N
ATOM     78  C2  DG  A   4      -4.188   1.017  17.875  1.00 36.18      A    C
ATOM     79  N2  DG  A   4      -3.039   1.533  18.309  1.00 35.32      A    N
ATOM     80  N1  DG  A   4      -4.137  -0.285  17.458  1.00 33.84      A    N
ATOM     81  C6  DG  A   4      -5.216  -1.017  16.987  1.00 37.56      A    C
ATOM     82  O6  DG  A   4      -5.059  -2.196  16.656  1.00 37.54      A    O
ATOM     83  OP2 DG  A   4     -12.244  -0.357  17.217  1.00 52.40      A    O
ATOM     84  P   DA  A   5      -9.212   6.689  16.649  1.00 73.45      A    P
ATOM     85  OP1 DA  A   5      -9.300   8.063  17.231  1.00 78.48      A    O
ATOM     86  OP2 DA  A   5     -10.136   6.219  15.562  1.00 64.08      A    O
ATOM     87  O5' DA  A   5      -7.707   6.556  16.153  1.00 57.34      A    O
ATOM     88  C5' DA  A   5      -6.685   6.812  17.124  1.00 60.56      A    C
ATOM     89  C4' DA  A   5      -5.339   6.490  16.528  1.00 61.87      A    C
ATOM     90  C3' DA  A   5      -5.085   7.143  15.171  1.00 60.08      A    C
ATOM     91  O3' DA  A   5      -3.762   7.635  15.348  1.00 62.97      A    O
ATOM     92  C2' DA  A   5      -5.181   5.986  14.195  1.00 55.49      A    C
ATOM     93  C1' DA  A   5      -4.641   4.842  15.041  1.00 50.61      A    C
ATOM     94  O4' DA  A   5      -5.183   5.068  16.324  1.00 53.78      A    O
ATOM     95  N9  DA  A   5      -5.028   3.496  14.610  1.00 39.67      A    N
ATOM     96  C4  DA  A   5      -4.212   2.386  14.555  1.00 35.52      A    C
ATOM     97  C5  DA  A   5      -5.016   1.348  14.116  1.00 33.44      A    C
ATOM     98  N7  DA  A   5      -6.311   1.794  13.877  1.00 36.91      A    N
ATOM     99  C8  DA  A   5      -6.272   3.067  14.207  1.00 43.13      A    C
ATOM    100  N3  DA  A   5      -2.907   2.313  14.874  1.00 34.17      A    N
ATOM    101  C2  DA  A   5      -2.439   1.076  14.666  1.00 35.80      A    C
ATOM    102  N1  DA  A   5      -3.087  -0.007  14.210  1.00 32.37      A    N
ATOM    103  C6  DA  A   5      -4.407   0.093  13.949  1.00 33.95      A    C
ATOM    104  N6  DA  A   5      -5.059  -0.986  13.516  1.00 31.51      A    N
ATOM    105  P   DA  A   6      -3.126   8.618  14.308  1.00 65.36      A    P
ATOM    106  OP1 DA  A   6      -2.767   9.852  15.053  1.00 65.12      A    O
ATOM    107  OP2 DA  A   6      -3.990   8.639  13.091  1.00 55.72      A    O
ATOM    108  O5' DA  A   6      -1.745   7.905  13.937  1.00 54.98      A    O
ATOM    109  C5' DA  A   6      -0.952   7.280  14.947  1.00 50.01      A    C
ATOM    110  C4' DA  A   6       0.056   6.374  14.281  1.00 49.55      A    C
ATOM    111  C3' DA  A   6       0.647   6.970  13.003  1.00 46.82      A    C
```

```
ATOM    112  O3' DA  A   6       2.049   6.773  13.203  1.00 53.59      A    O
ATOM    113  C2' DA  A   6      -0.045   6.196  11.885  1.00 44.53      A    C
ATOM    114  C1' DA  A   6      -0.330   4.833  12.529  1.00 41.49      A    C
ATOM    115  O4' DA  A   6      -0.562   5.119  13.897  1.00 46.07      A    O
ATOM    116  N9  DA  A   6      -1.506   4.113  12.019  1.00 35.31      A    N
ATOM    117  C4  DA  A   6      -1.593   2.766  11.740  1.00 34.49      A    C
ATOM    118  C5  DA  A   6      -2.905   2.543  11.355  1.00 29.46      A    C
ATOM    119  N7  DA  A   6      -3.642   3.720  11.419  1.00 35.23      A    N
ATOM    120  C8  DA  A   6      -2.775   4.614  11.838  1.00 36.22      A    C
ATOM    121  N3  DA  A   6      -0.609   1.854  11.801  1.00 36.17      A    N
ATOM    122  C2  DA  A   6      -1.047   0.661  11.408  1.00 36.35      A    C
ATOM    123  N1  DA  A   6      -2.267   0.303  10.995  1.00 31.68      A    N
ATOM    124  C6  DA  A   6      -3.238   1.238  10.961  1.00 30.29      A    C
ATOM    125  N6  DA  A   6      -4.456   0.867  10.580  1.00 31.37      A    N
ATOM    126  P   DT  A   7       3.154   7.216  12.109  1.00 58.14      A    P
ATOM    127  OP1 DT  A   7       4.426   7.459  12.845  1.00 53.73      A    O
ATOM    128  O5' DT  A   7       3.392   5.884  11.240  1.00 50.01      A    O
ATOM    129  C5' DT  A   7       3.829   4.630  11.835  1.00 48.02      A    C
ATOM    130  C4' DT  A   7       3.712   3.520  10.814  1.00 43.46      A    C
ATOM    131  C3' DT  A   7       4.361   3.864   9.472  1.00 48.85      A    C
ATOM    132  O3' DT  A   7       5.419   2.927   9.344  1.00 63.44      A    O
ATOM    133  C2' DT  A   7       3.259   3.716   8.426  1.00 42.75      A    C
ATOM    134  C1' DT  A   7       2.252   2.831   9.141  1.00 39.37      A    C
ATOM    135  O4' DT  A   7       2.329   3.226  10.512  1.00 43.92      A    O
ATOM    136  N1  DT  A   7       0.846   2.931   8.730  1.00 36.71      A    N
ATOM    137  C6  DT  A   7       0.187   4.141   8.759  1.00 40.24      A    C
ATOM    138  C5  DT  A   7      -1.103   4.287   8.446  1.00 38.65      A    C
ATOM    139  C5M DT  A   7      -1.798   5.613   8.470  1.00 40.91      A    C
ATOM    140  C4  DT  A   7      -1.852   3.100   8.086  1.00 36.66      A    C
ATOM    141  N3  DT  A   7      -1.133   1.925   8.108  1.00 31.96      A    N
ATOM    142  C2  DT  A   7       0.202   1.775   8.365  1.00 34.29      A    C
ATOM    143  O2  DT  A   7       0.749   0.684   8.354  1.00 34.03      A    O
ATOM    144  O4  DT  A   7      -3.052   3.067   7.849  1.00 36.53      A    O
ATOM    145  OP2 DT  A   7       2.552   8.248  11.185  1.00 43.35      A    O
ATOM    146  P   DT  A   8       6.412   3.030   8.138  1.00 67.95      A    P
ATOM    147  OP1 DT  A   8       7.785   2.884   8.702  1.00 74.13      A    O
ATOM    148  O5' DT  A   8       5.998   1.769   7.268  1.00 62.03      A    O
ATOM    149  C5' DT  A   8       5.835   0.484   7.892  1.00 56.42      A    C
ATOM    150  C4' DT  A   8       5.075  -0.405   6.938  1.00 57.99      A    C
ATOM    151  C3' DT  A   8       5.658  -0.469   5.526  1.00 52.21      A    C
ATOM    152  O3' DT  A   8       5.937  -1.856   5.336  1.00 60.60      A    O
ATOM    153  C2' DT  A   8       4.553   0.071   4.621  1.00 46.93      A    C
ATOM    154  C1' DT  A   8       3.302  -0.182   5.427  1.00 45.70      A    C
ATOM    155  O4' DT  A   8       3.716   0.052   6.783  1.00 46.28      A    O
ATOM    156  N1  DT  A   8       2.137   0.708   5.164  1.00 41.19      A    N
ATOM    157  C6  DT  A   8       2.322   2.064   5.322  1.00 37.35      A    C
ATOM    158  C5  DT  A   8       1.344   2.968   5.199  1.00 37.20      A    C
ATOM    159  C7  DT  A   8       1.573   4.438   5.358  1.00 41.50      A    C
ATOM    160  C4  DT  A   8       0.021   2.496   4.886  1.00 38.03      A    C
ATOM    161  N3  DT  A   8      -0.102   1.123   4.737  1.00 34.39      A    N
ATOM    162  C2  DT  A   8       0.893   0.184   4.871  1.00 35.72      A    C
ATOM    163  O2  DT  A   8       0.689  -1.007   4.681  1.00 34.97      A    O
ATOM    164  O4  DT  A   8      -0.962   3.211   4.760  1.00 35.24      A    O
ATOM    165  OP2 DT  A   8       6.040   4.213   7.308  1.00 73.03      A    O
HETATM  166  O5C 5CC A   9       4.207   1.575   1.530  1.00 54.21      A    O
HETATM  167  C5A 5CC A   9       3.049   1.984   1.695  1.00 52.02      A    C
```

```
HETATM  168  O5A 5CC A   9       2.734   3.175   1.909  1.00 46.46      A    O
HETATM  169  C5  5CC A   9       2.042   0.921   1.541  1.00 43.46      A    C
HETATM  170  C4  5CC A   9       0.638   1.279   1.382  1.00 37.98      A    C
HETATM  171  N4  5CC A   9       0.172   2.555   1.437  1.00 38.12      A    N
HETATM  172  N3  5CC A   9      -0.222   0.275   1.164  1.00 37.14      A    N
HETATM  173  C2  5CC A   9       0.213  -1.003   1.155  1.00 38.84      A    C
HETATM  174  O2  5CC A   9      -0.622  -1.888   0.954  1.00 40.39      A    O
HETATM  175  C6  5CC A   9       2.476  -0.414   1.487  1.00 40.05      A    C
HETATM  176  N1  5CC A   9       1.571  -1.383   1.250  1.00 45.45      A    N
HETATM  177  C1' 5CC A   9       1.891  -2.820   1.316  1.00 41.48      A    C
HETATM  178  O4' 5CC A   9       2.191  -2.944   2.699  1.00 45.28      A    O
HETATM  179  C2' 5CC A   9       3.122  -3.361   0.619  1.00 48.49      A    C
HETATM  180  C3' 5CC A   9       3.628  -4.428   1.591  1.00 53.39      A    C
HETATM  181  O3' 5CC A   9       3.155  -5.730   1.276  1.00 62.46      A    O
HETATM  182  C4' 5CC A   9       3.071  -4.057   2.947  1.00 57.68      A    C
HETATM  183  C5' 5CC A   9       4.233  -3.674   3.865  1.00 57.58      A    C
HETATM  184  O5' 5CC A   9       5.211  -2.881   3.178  1.00 57.19      A    O
HETATM  185  P   5CC A   9       6.539  -2.386   3.956  1.00 62.68      A    P
HETATM  186  OP2 5CC A   9       7.308  -3.617   4.286  1.00 62.34      A    O
HETATM  187  OP1 5CC A   9       7.099  -1.200   3.191  1.00 60.71      A    O
ATOM    188  P   DG  A  10       3.669  -6.457  -0.061  1.00 69.97      A    P
ATOM    189  OP1 DG  A  10       3.588  -7.931   0.147  1.00 75.89      A    O
ATOM    190  O5' DG  A  10       2.586  -6.046  -1.160  1.00 63.96      A    O
ATOM    191  C5' DG  A  10       1.183  -6.326  -1.005  1.00 65.19      A    C
ATOM    192  C4' DG  A  10       0.434  -5.640  -2.122  1.00 59.41      A    C
ATOM    193  C3' DG  A  10       1.061  -5.778  -3.515  1.00 55.44      A    C
ATOM    194  O3' DG  A  10      -0.019  -5.802  -4.431  1.00 64.52      A    O
ATOM    195  C2' DG  A  10       1.776  -4.452  -3.725  1.00 53.34      A    C
ATOM    196  C1' DG  A  10       0.769  -3.549  -3.045  1.00 46.38      A    C
ATOM    197  O4' DG  A  10       0.427  -4.221  -1.855  1.00 50.21      A    O
ATOM    198  N9  DG  A  10       1.160  -2.189  -2.694  1.00 43.07      A    N
ATOM    199  C4  DG  A  10       0.280  -1.150  -2.522  1.00 35.88      A    C
ATOM    200  C5  DG  A  10       1.056  -0.065  -2.195  1.00 39.24      A    C
ATOM    201  N7  DG  A  10       2.393  -0.428  -2.091  1.00 41.95      A    N
ATOM    202  C8  DG  A  10       2.407  -1.696  -2.408  1.00 42.06      A    C
ATOM    203  N3  DG  A  10      -1.053  -1.207  -2.711  1.00 34.13      A    N
ATOM    204  C2  DG  A  10      -1.631  -0.024  -2.551  1.00 34.68      A    C
ATOM    205  N2  DG  A  10      -2.954   0.091  -2.730  1.00 29.21      A    N
ATOM    206  N1  DG  A  10      -0.951   1.118  -2.193  1.00 33.88      A    N
ATOM    207  C6  DG  A  10       0.427   1.183  -1.969  1.00 40.73      A    C
ATOM    208  O6  DG  A  10       0.950   2.248  -1.637  1.00 38.33      A    O
ATOM    209  OP2 DG  A  10       4.918  -5.793  -0.545  1.00 61.30      A    O
ATOM    210  P   DC  A  11      -0.388  -7.136  -5.177  1.00 67.48      A    P
ATOM    211  OP1 DC  A  11      -0.244  -8.256  -4.200  1.00 77.79      A    O
ATOM    212  O5' DC  A  11      -1.956  -6.989  -5.434  1.00 66.75      A    O
ATOM    213  C5' DC  A  11      -2.882  -6.528  -4.416  1.00 60.75      A    C
ATOM    214  C4' DC  A  11      -3.916  -5.614  -5.041  1.00 60.57      A    C
ATOM    215  C3' DC  A  11      -4.246  -5.907  -6.508  1.00 52.98      A    C
ATOM    216  O3' DC  A  11      -5.669  -5.751  -6.541  1.00 51.48      A    O
ATOM    217  C2' DC  A  11      -3.410  -4.877  -7.270  1.00 47.97      A    C
ATOM    218  C1' DC  A  11      -3.413  -3.676  -6.324  1.00 44.23      A    C
ATOM    219  O4' DC  A  11      -3.469  -4.230  -5.011  1.00 46.22      A    O
ATOM    220  N1  DC  A  11      -2.262  -2.747  -6.338  1.00 40.13      A    N
ATOM    221  C6  DC  A  11      -0.989  -3.243  -6.375  1.00 43.73      A    C
ATOM    222  C5  DC  A  11       0.073  -2.434  -6.267  1.00 40.04      A    C
ATOM    223  C4  DC  A  11      -0.174  -1.054  -6.023  1.00 39.18      A    C
```

```
ATOM    224  N4  DC  A  11       0.849  -0.212  -5.888  1.00 36.48      A    N
ATOM    225  N3  DC  A  11      -1.418  -0.554  -5.962  1.00 33.83      A    N
ATOM    226  C2  DC  A  11      -2.484  -1.380  -6.082  1.00 40.39      A    C
ATOM    227  O2  DC  A  11      -3.661  -0.962  -5.964  1.00 39.89      A    O
ATOM    228  OP2 DC  A  11       0.389  -7.125  -6.458  1.00 69.13      A    O
ATOM    229  P   DG  A  12      -6.554  -6.181  -7.813  1.00 59.59      A    P
ATOM    230  OP1 DG  A  12      -7.981  -6.293  -7.340  1.00 55.36      A    O
ATOM    231  O5' DG  A  12      -6.457  -4.924  -8.788  1.00 62.07      A    O
ATOM    232  C5' DG  A  12      -7.284  -3.759  -8.598  1.00 54.25      A    C
ATOM    233  C4' DG  A  12      -6.816  -2.673  -9.534  1.00 52.34      A    C
ATOM    234  C3' DG  A  12      -6.676  -3.089 -10.995  1.00 47.72      A    C
ATOM    235  O3' DG  A  12      -7.864  -2.672 -11.658  1.00 49.53      A    O
ATOM    236  C2' DG  A  12      -5.445  -2.332 -11.481  1.00 49.88      A    C
ATOM    237  C1' DG  A  12      -4.947  -1.577 -10.250  1.00 46.38      A    C
ATOM    238  O4' DG  A  12      -5.496  -2.265  -9.142  1.00 46.34      A    O
ATOM    239  N9  DG  A  12      -3.503  -1.554 -10.067  1.00 43.47      A    N
ATOM    240  C4  DG  A  12      -2.731  -0.493  -9.658  1.00 37.26      A    C
ATOM    241  C5  DG  A  12      -1.439  -0.984  -9.604  1.00 40.22      A    C
ATOM    242  N7  DG  A  12      -1.410  -2.338  -9.905  1.00 38.98      A    N
ATOM    243  C8  DG  A  12      -2.652  -2.634 -10.181  1.00 45.61      A    C
ATOM    244  N3  DG  A  12      -3.169   0.761  -9.394  1.00 32.26      A    N
ATOM    245  C2  DG  A  12      -2.177   1.560  -9.036  1.00 36.77      A    C
ATOM    246  N2  DG  A  12      -2.413   2.849  -8.772  1.00 30.64      A    N
ATOM    247  N1  DG  A  12      -0.865   1.155  -8.923  1.00 35.53      A    N
ATOM    248  C6  DG  A  12      -0.395  -0.120  -9.211  1.00 40.62      A    C
ATOM    249  O6  DG  A  12       0.805  -0.374  -9.093  1.00 38.59      A    O
ATOM    250  OP2 DG  A  12      -5.843  -7.286  -8.556  1.00 53.88      A    O
TER
ATOM    251  O5' DC  B  13       4.030   9.050  -9.830  1.00 76.36      B    O
ATOM    252  C5' DC  B  13       3.184   9.719  -8.875  1.00 73.06      B    C
ATOM    253  C4' DC  B  13       2.101   8.783  -8.380  1.00 66.23      B    C
ATOM    254  C3' DC  B  13       1.734   8.711  -6.898  1.00 56.80      B    C
ATOM    255  O3' DC  B  13       0.341   8.972  -6.761  1.00 72.67      B    O
ATOM    256  C2' DC  B  13       1.929   7.244  -6.579  1.00 61.22      B    C
ATOM    257  C1' DC  B  13       1.448   6.620  -7.892  1.00 55.30      B    C
ATOM    258  O4' DC  B  13       2.004   7.436  -8.925  1.00 60.23      B    O
ATOM    259  N1  DC  B  13       1.870   5.223  -8.095  1.00 50.03      B    N
ATOM    260  C6  DC  B  13       3.185   4.867  -7.971  1.00 53.03      B    C
ATOM    261  C5  DC  B  13       3.564   3.587  -8.069  1.00 52.21      B    C
ATOM    262  C4  DC  B  13       2.543   2.611  -8.276  1.00 46.93      B    C
ATOM    263  N4  DC  B  13       2.859   1.328  -8.401  1.00 48.27      B    N
ATOM    264  N3  DC  B  13       1.255   2.944  -8.368  1.00 43.23      B    N
ATOM    265  C2  DC  B  13       0.882   4.236  -8.260  1.00 39.47      B    C
ATOM    266  O2  DC  B  13      -0.310   4.578  -8.318  1.00 42.49      B    O
ATOM    267  P   DG  B  14      -0.165  10.283  -5.990  1.00 74.71      B    P
ATOM    268  OP1 DG  B  14      -0.062  11.416  -6.949  1.00 65.40      B    O
ATOM    269  O5' DG  B  14      -1.714  10.008  -5.768  1.00 68.31      B    O
ATOM    270  C5' DG  B  14      -2.603   9.942  -6.887  1.00 61.31      B    C
ATOM    271  C4' DG  B  14      -3.718   9.007  -6.512  1.00 56.78      B    C
ATOM    272  C3' DG  B  14      -4.193   9.221  -5.078  1.00 55.19      B    C
ATOM    273  O3' DG  B  14      -5.608   9.320  -4.977  1.00 65.23      B    O
ATOM    274  C2' DG  B  14      -3.671   8.018  -4.322  1.00 59.93      B    C
ATOM    275  C1' DG  B  14      -3.502   6.963  -5.387  1.00 53.16      B    C
ATOM    276  O4' DG  B  14      -3.204   7.664  -6.579  1.00 51.71      B    O
ATOM    277  N9  DG  B  14      -2.396   6.047  -5.146  1.00 43.09      B    N
ATOM    278  C4  DG  B  14      -2.412   4.684  -5.314  1.00 35.28      B    C
```

```
ATOM    279  C5  DG  B  14      -1.135   4.262  -4.996  1.00 35.51      B    C
ATOM    280  N7  DG  B  14      -0.328   5.343  -4.668  1.00 38.78      B    N
ATOM    281  C8  DG  B  14      -1.110   6.380  -4.796  1.00 39.46      B    C
ATOM    282  N3  DG  B  14      -3.482   3.943  -5.661  1.00 30.87      B    N
ATOM    283  C2  DG  B  14      -3.190   2.650  -5.706  1.00 33.78      B    C
ATOM    284  N2  DG  B  14      -4.128   1.779  -6.037  1.00 30.07      B    N
ATOM    285  N1  DG  B  14      -1.946   2.128  -5.423  1.00 31.50      B    N
ATOM    286  C6  DG  B  14      -0.823   2.884  -5.085  1.00 39.15      B    C
ATOM    287  O6  DG  B  14       0.282   2.324  -4.900  1.00 33.23      B    O
ATOM    288  OP2 DG  B  14       0.441  10.297  -4.620  1.00 66.01      B    O
ATOM    289  P   DC  B  15      -6.271   9.590  -3.539  1.00 63.47      B    P
ATOM    290  OP1 DC  B  15      -7.522  10.358  -3.728  1.00 67.33      B    O
ATOM    291  O5' DC  B  15      -6.811   8.158  -3.100  1.00 54.40      B    O
ATOM    292  C5' DC  B  15      -7.562   7.326  -3.995  1.00 51.25      B    C
ATOM    293  C4' DC  B  15      -7.441   5.901  -3.509  1.00 46.61      B    C
ATOM    294  C3' DC  B  15      -7.792   5.746  -2.040  1.00 44.34      B    C
ATOM    295  O3' DC  B  15      -9.179   5.459  -2.064  1.00 50.96      B    O
ATOM    296  C2' DC  B  15      -6.931   4.590  -1.581  1.00 45.36      B    C
ATOM    297  C1' DC  B  15      -5.858   4.469  -2.634  1.00 39.26      B    C
ATOM    298  O4' DC  B  15      -6.069   5.472  -3.597  1.00 42.42      B    O
ATOM    299  N1  DC  B  15      -4.475   4.571  -2.179  1.00 40.04      B    N
ATOM    300  C6  DC  B  15      -3.914   5.750  -1.774  1.00 42.25      B    C
ATOM    301  C5  DC  B  15      -2.623   5.812  -1.414  1.00 44.06      B    C
ATOM    302  C4  DC  B  15      -1.867   4.599  -1.467  1.00 41.06      B    C
ATOM    303  N4  DC  B  15      -0.573   4.598  -1.137  1.00 41.98      B    N
ATOM    304  N3  DC  B  15      -2.401   3.454  -1.899  1.00 32.64      B    N
ATOM    305  C2  DC  B  15      -3.708   3.403  -2.226  1.00 38.81      B    C
ATOM    306  O2  DC  B  15      -4.246   2.358  -2.609  1.00 37.99      B    O
ATOM    307  OP2 DC  B  15      -5.184  10.006  -2.608  1.00 70.00      B    O
ATOM    308  P   DG  B  16      -9.973   5.257  -0.697  1.00 52.67      B    P
ATOM    309  OP1 DG  B  16     -11.411   5.558  -1.018  1.00 56.56      B    O
ATOM    310  O5' DG  B  16      -9.765   3.692  -0.440  1.00 47.18      B    O
ATOM    311  C5' DG  B  16     -10.466   2.729  -1.242  1.00 40.66      B    C
ATOM    312  C4' DG  B  16      -9.991   1.342  -0.889  1.00 40.99      B    C
ATOM    313  C3' DG  B  16     -10.429   0.853   0.490  1.00 50.57      B    C
ATOM    314  O3' DG  B  16     -10.613  -0.547   0.310  1.00 53.07      B    O
ATOM    315  C2' DG  B  16      -9.255   1.237   1.375  1.00 47.97      B    C
ATOM    316  C1' DG  B  16      -8.093   0.971   0.448  1.00 46.26      B    C
ATOM    317  O4' DG  B  16      -8.543   1.338  -0.859  1.00 44.71      B    O
ATOM    318  N9  DG  B  16      -6.887   1.731   0.755  1.00 43.12      B    N
ATOM    319  C4  DG  B  16      -5.599   1.271   0.673  1.00 37.18      B    C
ATOM    320  C5  DG  B  16      -4.808   2.327   1.069  1.00 40.97      B    C
ATOM    321  N7  DG  B  16      -5.580   3.449   1.334  1.00 42.08      B    N
ATOM    322  C8  DG  B  16      -6.808   3.036   1.181  1.00 44.90      B    C
ATOM    323  N3  DG  B  16      -5.229   0.015   0.344  1.00 43.74      B    N
ATOM    324  C2  DG  B  16      -3.921  -0.141   0.392  1.00 40.21      B    C
ATOM    325  N2  DG  B  16      -3.384  -1.332   0.079  1.00 35.64      B    N
ATOM    326  N1  DG  B  16      -3.045   0.851   0.766  1.00 34.17      B    N
ATOM    327  C6  DG  B  16      -3.411   2.149   1.123  1.00 39.04      B    C
ATOM    328  O6  DG  B  16      -2.543   2.974   1.440  1.00 33.50      B    O
ATOM    329  OP2 DG  B  16      -9.275   5.985   0.403  1.00 43.89      B    O
ATOM    330  P   DA  B  17     -10.873  -1.550   1.543  1.00 60.01      B    P
ATOM    331  OP1 DA  B  17     -12.021  -2.422   1.129  1.00 59.36      B    O
ATOM    332  OP2 DA  B  17     -10.942  -0.775   2.802  1.00 58.30      B    O
ATOM    333  O5' DA  B  17      -9.497  -2.354   1.620  1.00 58.53      B    O
ATOM    334  C5' DA  B  17      -8.789  -2.780   0.431  1.00 55.14      B    C
```

```
ATOM    335  C4' DA  B  17      -7.675  -3.722   0.823  1.00 56.49      B    C
ATOM    336  C3' DA  B  17      -8.026  -4.605   2.021  1.00 53.29      B    C
ATOM    337  O3' DA  B  17      -7.341  -5.822   1.756  1.00 59.36      B    O
ATOM    338  C2' DA  B  17      -7.490  -3.810   3.199  1.00 50.22      B    C
ATOM    339  C1' DA  B  17      -6.229  -3.160   2.612  1.00 46.71      B    C
ATOM    340  O4' DA  B  17      -6.455  -3.011   1.204  1.00 47.99      B    O
ATOM    341  N9  DA  B  17      -5.911  -1.842   3.151  1.00 37.32      B    N
ATOM    342  C4  DA  B  17      -4.658  -1.354   3.436  1.00 34.38      B    C
ATOM    343  C5  DA  B  17      -4.846  -0.048   3.848  1.00 35.45      B    C
ATOM    344  N7  DA  B  17      -6.197   0.263   3.882  1.00 39.75      B    N
ATOM    345  C8  DA  B  17      -6.783  -0.819   3.433  1.00 37.86      B    C
ATOM    346  N3  DA  B  17      -3.482  -1.977   3.263  1.00 29.34      B    N
ATOM    347  C2  DA  B  17      -2.472  -1.188   3.616  1.00 31.34      B    C
ATOM    348  N1  DA  B  17      -2.501   0.079   4.045  1.00 35.22      B    N
ATOM    349  C6  DA  B  17      -3.699   0.681   4.193  1.00 31.34      B    C
ATOM    350  N6  DA  B  17      -3.725   1.944   4.610  1.00 33.90      B    N
ATOM    351  P   DA  B  18      -7.303  -6.994   2.827  1.00 63.46      B    P
ATOM    352  OP1 DA  B  18      -7.241  -8.255   2.071  1.00 64.68      B    O
ATOM    353  OP2 DA  B  18      -8.370  -6.728   3.844  1.00 53.26      B    O
ATOM    354  O5' DA  B  18      -5.901  -6.785   3.565  1.00 54.25      B    O
ATOM    355  C5' DA  B  18      -4.671  -6.774   2.835  1.00 54.30      B    C
ATOM    356  C4' DA  B  18      -3.551  -6.636   3.834  1.00 60.05      B    C
ATOM    357  C3' DA  B  18      -3.744  -7.511   5.078  1.00 57.72      B    C
ATOM    358  O3' DA  B  18      -2.496  -8.194   5.192  1.00 67.11      B    O
ATOM    359  C2' DA  B  18      -4.102  -6.521   6.176  1.00 51.89      B    C
ATOM    360  C1' DA  B  18      -3.362  -5.260   5.720  1.00 48.44      B    C
ATOM    361  O4' DA  B  18      -3.453  -5.265   4.305  1.00 53.97      B    O
ATOM    362  N9  DA  B  18      -3.910  -3.989   6.189  1.00 35.82      B    N
ATOM    363  C4  DA  B  18      -3.190  -2.870   6.544  1.00 32.88      B    C
ATOM    364  C5  DA  B  18      -4.128  -1.902   6.843  1.00 28.49      B    C
ATOM    365  N7  DA  B  18      -5.412  -2.389   6.666  1.00 32.98      B    N
ATOM    366  C8  DA  B  18      -5.227  -3.623   6.257  1.00 31.86      B    C
ATOM    367  N3  DA  B  18      -1.852  -2.730   6.570  1.00 34.29      B    N
ATOM    368  C2  DA  B  18      -1.513  -1.502   6.968  1.00 37.34      B    C
ATOM    369  N1  DA  B  18      -2.302  -0.481   7.312  1.00 30.43      B    N
ATOM    370  C6  DA  B  18      -3.644  -0.644   7.228  1.00 32.56      B    C
ATOM    371  N6  DA  B  18      -4.435   0.371   7.544  1.00 32.31      B    N
ATOM    372  P   DT  B  19      -2.141  -9.149   6.471  1.00 74.44      B    P
ATOM    373  OP1 DT  B  19      -1.272 -10.207   5.964  1.00 64.28      B    O
ATOM    374  O5' DT  B  19      -1.178  -8.232   7.345  1.00 65.35      B    O
ATOM    375  C5' DT  B  19      -0.039  -7.620   6.735  1.00 55.18      B    C
ATOM    376  C4' DT  B  19       0.486  -6.590   7.703  1.00 55.11      B    C
ATOM    377  C3' DT  B  19       0.799  -7.186   9.080  1.00 51.36      B    C
ATOM    378  O3' DT  B  19       2.185  -6.891   9.230  1.00 59.86      B    O
ATOM    379  C2' DT  B  19      -0.143  -6.459  10.032  1.00 47.78      B    C
ATOM    380  C1' DT  B  19      -0.350  -5.145   9.304  1.00 42.93      B    C
ATOM    381  O4' DT  B  19      -0.470  -5.522   7.937  1.00 44.44      B    O
ATOM    382  N1  DT  B  19      -1.539  -4.383   9.648  1.00 36.97      B    N
ATOM    383  C6  DT  B  19      -2.776  -4.967   9.486  1.00 37.81      B    C
ATOM    384  C5  DT  B  19      -3.927  -4.333   9.728  1.00 32.22      B    C
ATOM    385  C5M DT  B  19      -5.263  -4.992   9.583  1.00 35.85      B    C
ATOM    386  C4  DT  B  19      -3.862  -2.926  10.092  1.00 33.44      B    C
ATOM    387  N3  DT  B  19      -2.589  -2.398  10.209  1.00 29.18      B    N
ATOM    388  C2  DT  B  19      -1.411  -3.045   9.974  1.00 33.38      B    C
ATOM    389  O2  DT  B  19      -0.341  -2.496  10.117  1.00 31.82      B    O
ATOM    390  O4  DT  B  19      -4.822  -2.221  10.313  1.00 32.29      B    O
```

```
ATOM    391  OP2 DT  B  19      -3.370  -9.421   7.271  1.00 58.04      B    O
ATOM    392  P   DT  B  20       2.987  -7.355  10.533  1.00 73.37      B    P
ATOM    393  OP1 DT  B  20       4.336  -7.803  10.068  1.00 74.65      B    O
ATOM    394  O5' DT  B  20       3.175  -5.996  11.345  1.00 61.52      B    O
ATOM    395  C5' DT  B  20       3.754  -4.841  10.707  1.00 55.26      B    C
ATOM    396  C4' DT  B  20       3.680  -3.682  11.670  1.00 53.84      B    C
ATOM    397  C3' DT  B  20       4.217  -3.996  13.068  1.00 50.90      B    C
ATOM    398  O3' DT  B  20       5.170  -2.965  13.327  1.00 62.87      B    O
ATOM    399  C2' DT  B  20       2.975  -4.020  13.957  1.00 49.74      B    C
ATOM    400  C1' DT  B  20       2.052  -3.048  13.241  1.00 44.10      B    C
ATOM    401  O4' DT  B  20       2.305  -3.283  11.860  1.00 43.92      B    O
ATOM    402  N1  DT  B  20       0.606  -3.214  13.424  1.00 40.32      B    N
ATOM    403  C6  DT  B  20       0.025  -4.437  13.167  1.00 40.76      B    C
ATOM    404  C5  DT  B  20      -1.297  -4.648  13.177  1.00 39.81      B    C
ATOM    405  C7  DT  B  20      -1.902  -6.001  12.979  1.00 41.83      B    C
ATOM    406  C4  DT  B  20      -2.166  -3.516  13.405  1.00 35.39      B    C
ATOM    407  N3  DT  B  20      -1.519  -2.316  13.646  1.00 35.41      B    N
ATOM    408  C2  DT  B  20      -0.160  -2.088  13.624  1.00 35.76      B    C
ATOM    409  O2  DT  B  20       0.319  -0.984  13.824  1.00 38.20      B    O
ATOM    410  O4  DT  B  20      -3.395  -3.562  13.433  1.00 34.95      B    O
ATOM    411  OP2 DT  B  20       2.102  -8.234  11.359  1.00 57.95      B    O
HETATM  412  O5C 5CC B  21       1.187  -5.336  16.493  1.00 46.32      B    O
HETATM  413  C5A 5CC B  21       0.033  -4.955  16.609  1.00 43.90      B    C
HETATM  414  O5A 5CC B  21      -0.946  -5.721  16.523  1.00 42.54      B    O
HETATM  415  C5  5CC B  21      -0.146  -3.499  16.873  1.00 36.64      B    C
HETATM  416  C4  5CC B  21      -1.454  -2.881  16.914  1.00 33.87      B    C
HETATM  417  N4  5CC B  21      -2.580  -3.525  16.750  1.00 38.01      B    N
HETATM  418  N3  5CC B  21      -1.587  -1.559  17.167  1.00 38.01      B    N
HETATM  419  C2  5CC B  21      -0.490  -0.827  17.355  1.00 39.91      B    C
HETATM  420  O2  5CC B  21      -0.602   0.420  17.593  1.00 41.01      B    O
HETATM  421  C6  5CC B  21       0.959  -2.689  16.990  1.00 34.81      B    C
HETATM  422  N1  5CC B  21       0.805  -1.414  17.288  1.00 40.67      B    N
HETATM  423  C1' 5CC B  21       1.965  -0.507  17.400  1.00 42.36      B    C
HETATM  424  O4' 5CC B  21       2.527  -0.579  16.053  1.00 46.37      B    O
HETATM  425  C2' 5CC B  21       3.135  -0.872  18.301  1.00 48.67      B    C
HETATM  426  C3' 5CC B  21       4.289  -0.241  17.532  1.00 50.17      B    C
HETATM  427  O3' 5CC B  21       4.546   1.048  18.023  1.00 54.36      B    O
HETATM  428  C4' 5CC B  21       3.871  -0.002  16.097  1.00 52.40      B    C
HETATM  429  C5' 5CC B  21       4.900  -0.599  15.145  1.00 51.53      B    C
HETATM  430  O5' 5CC B  21       5.208  -1.941  15.597  1.00 50.23      B    O
HETATM  431  P   5CC B  21       6.014  -2.994  14.670  1.00 59.94      B    P
HETATM  432  OP2 5CC B  21       7.400  -2.453  14.349  1.00 63.64      B    O
HETATM  433  OP1 5CC B  21       5.844  -4.384  15.242  1.00 60.90      B    O
ATOM    434  P   DG  B  22       5.523   1.234  19.278  1.00 63.24      B    P
ATOM    435  OP1 DG  B  22       6.228   2.511  19.089  1.00 53.84      B    O
ATOM    436  O5' DG  B  22       4.524   1.407  20.504  1.00 58.34      B    O
ATOM    437  C5' DG  B  22       3.641   2.538  20.528  1.00 57.93      B    C
ATOM    438  C4' DG  B  22       2.565   2.263  21.544  1.00 55.34      B    C
ATOM    439  C3' DG  B  22       3.099   1.703  22.866  1.00 48.75      B    C
ATOM    440  O3' DG  B  22       2.257   2.239  23.883  1.00 58.60      B    O
ATOM    441  C2' DG  B  22       2.807   0.220  22.754  1.00 50.28      B    C
ATOM    442  C1' DG  B  22       1.461   0.311  22.046  1.00 45.51      B    C
ATOM    443  O4' DG  B  22       1.660   1.253  21.020  1.00 49.05      B    O
ATOM    444  N9  DG  B  22       0.921  -0.908  21.460  1.00 42.92      B    N
ATOM    445  C4  DG  B  22      -0.410  -1.118  21.205  1.00 35.01      B    C
ATOM    446  C5  DG  B  22      -0.493  -2.405  20.735  1.00 34.09      B    C
```

```
ATOM    447  N7  DG  B  22       0.764  -2.998  20.690  1.00 37.78      B    N
ATOM    448  C8  DG  B  22       1.568  -2.077  21.143  1.00 37.91      B    C
ATOM    449  N3  DG  B  22      -1.402  -0.225  21.409  1.00 32.49      B    N
ATOM    450  C2  DG  B  22      -2.583  -0.727  21.096  1.00 32.72      B    C
ATOM    451  N2  DG  B  22      -3.672   0.011  21.259  1.00 29.60      B    N
ATOM    452  N1  DG  B  22      -2.769  -1.994  20.611  1.00 32.87      B    N
ATOM    453  C6  DG  B  22      -1.759  -2.926  20.396  1.00 33.11      B    C
ATOM    454  O6  DG  B  22      -2.032  -4.038  19.923  1.00 33.81      B    O
ATOM    455  OP2 DG  B  22       6.238  -0.054  19.526  1.00 50.97      B    O
ATOM    456  P   DC  B  23       2.833   2.874  25.222  1.00 69.16      B    P
ATOM    457  OP1 DC  B  23       3.533   4.135  24.860  1.00 67.95      B    O
ATOM    458  O5' DC  B  23       1.515   3.305  26.020  1.00 60.67      B    O
ATOM    459  C5' DC  B  23       0.540   4.225  25.456  1.00 55.51      B    C
ATOM    460  C4' DC  B  23      -0.848   3.744  25.799  1.00 49.85      B    C
ATOM    461  C3' DC  B  23      -0.963   3.281  27.254  1.00 46.91      B    C
ATOM    462  O3' DC  B  23      -1.810   4.080  28.095  1.00 62.85      B    O
ATOM    463  C2' DC  B  23      -1.531   1.881  27.169  1.00 47.23      B    C
ATOM    464  C1' DC  B  23      -1.892   1.663  25.711  1.00 44.13      B    C
ATOM    465  O4' DC  B  23      -1.124   2.582  24.965  1.00 52.32      B    O
ATOM    466  N1  DC  B  23      -1.575   0.295  25.246  1.00 42.25      B    N
ATOM    467  C6  DC  B  23      -0.311  -0.216  25.354  1.00 40.65      B    C
ATOM    468  C5  DC  B  23      -0.039  -1.469  24.976  1.00 36.20      B    C
ATOM    469  C4  DC  B  23      -1.124  -2.255  24.495  1.00 37.38      B    C
ATOM    470  N4  DC  B  23      -0.924  -3.510  24.102  1.00 32.57      B    N
ATOM    471  N3  DC  B  23      -2.367  -1.775  24.423  1.00 36.06      B    N
ATOM    472  C2  DC  B  23      -2.624  -0.511  24.807  1.00 36.95      B    C
ATOM    473  O2  DC  B  23      -3.770  -0.035  24.765  1.00 34.27      B    O
ATOM    474  OP2 DC  B  23       3.475   1.785  26.011  1.00 68.51      B    O
ATOM    475  P   DG  B  24      -1.895   3.756  29.706  1.00 65.98      B    P
ATOM    476  OP1 DG  B  24      -2.219   5.008  30.422  1.00 73.34      B    O
ATOM    477  O5' DG  B  24      -3.196   2.853  29.878  1.00 50.48      B    O
ATOM    478  C5' DG  B  24      -4.458   3.294  29.395  1.00 43.21      B    C
ATOM    479  C4' DG  B  24      -5.321   2.078  29.209  1.00 45.92      B    C
ATOM    480  C3' DG  B  24      -5.315   1.187  30.445  1.00 43.48      B    C
ATOM    481  O3' DG  B  24      -6.193   1.662  31.464  1.00 45.23      B    O
ATOM    482  C2' DG  B  24      -5.735  -0.142  29.874  1.00 41.50      B    C
ATOM    483  C1' DG  B  24      -5.165  -0.110  28.478  1.00 43.73      B    C
ATOM    484  O4' DG  B  24      -4.737   1.236  28.188  1.00 42.19      B    O
ATOM    485  N9  DG  B  24      -4.037  -1.019  28.308  1.00 40.63      B    N
ATOM    486  C4  DG  B  24      -4.104  -2.263  27.734  1.00 37.01      B    C
ATOM    487  C5  DG  B  24      -2.807  -2.735  27.730  1.00 38.88      B    C
ATOM    488  N7  DG  B  24      -1.944  -1.810  28.299  1.00 37.54      B    N
ATOM    489  C8  DG  B  24      -2.718  -0.815  28.643  1.00 42.88      B    C
ATOM    490  N3  DG  B  24      -5.223  -2.863  27.267  1.00 32.29      B    N
ATOM    491  C2  DG  B  24      -4.959  -4.057  26.772  1.00 34.86      B    C
ATOM    492  N2  DG  B  24      -5.951  -4.819  26.308  1.00 31.79      B    N
ATOM    493  N1  DG  B  24      -3.700  -4.605  26.728  1.00 33.42      B    N
ATOM    494  C6  DG  B  24      -2.559  -4.030  27.258  1.00 37.91      B    C
ATOM    495  O6  DG  B  24      -1.479  -4.624  27.181  1.00 42.12      B    O
ATOM    496  OP2 DG  B  24      -0.776   2.839  30.100  1.00 58.28      B    O
TER
HETATM  497  O   HOH C   1       0.588  -3.711   5.454  1.00 47.03      C    O
HETATM  498  O   HOH C   2      -4.986   3.972  19.066  1.00 48.74      C    O
HETATM  499  O   HOH C   3      -0.804  -6.736  21.651  1.00 62.93           O
HETATM  500  O   HOH C   4       3.639  -1.444  -6.010  1.00 43.57           O
HETATM  501  O   HOH C   5      -1.918  -3.898   1.851  1.00 42.95           O
```

```
HETATM  502  O   HOH C   6       9.979  -6.519   0.022  1.00 55.15           O
HETATM  503  O   HOH C   7       1.705   1.415  13.189  0.50 28.51           O
HETATM  504  O   HOH C   8      -1.325   3.757  16.259  0.50 29.57           O
HETATM  505  O   HOH C   9       2.674  -3.617   7.388  0.50 33.17           O
HETATM  506  O   HOH C  10      -7.307  -6.067   6.373  0.50 37.42           O
HETATM  507  O   HOH C  11     -15.049  -3.165  17.560  0.50 46.43           O
HETATM  508  O   HOH C  12       1.173   8.787  10.026  0.50 40.69           O
END
```

File A-4: Dictionary file for 5-carboxyl-2'-deoxycytidine.

```
global_
_lib_name         ?
_lib_version      ?
_lib_update       ?
# ------------------------------------------------
#
# ---   LIST OF MONOMERS ---
#
data_comp_list
loop_
_chem_comp.id
_chem_comp.three_letter_code
_chem_comp.name
_chem_comp.group
_chem_comp.number_atoms_all
_chem_comp.number_atoms_nh
_chem_comp.desc_level
5CC     5CC 'carboxycytosine.                    ' DNA                33  22 .
#
# --- DESCRIPTION OF MONOMERS ---
#
data_comp_5CC
#
loop_
_chem_comp_atom.comp_id
_chem_comp_atom.atom_id
_chem_comp_atom.type_symbol
_chem_comp_atom.type_energy
_chem_comp_atom.partial_charge
_chem_comp_atom.x
_chem_comp_atom.y
_chem_comp_atom.z
 5CC           O5C    O    OC       -0.500     17.156    1.546  133.298
 5CC           C5A    C    C         0.000     15.884    1.980  133.125
 5CC           O5A    O    OC       -0.500     15.439    3.128  133.239
 5CC           C5     C    CR6       0.000     14.843    0.926  132.900
 5CC           C4     C    CR6       0.000     13.391    1.297  132.751
 5CC           N4     N    N         0.000     13.023    2.590  132.778
 5CC           HN4    H    H         0.000     12.105    2.823  132.683
 5CC           N3     N    NRD6      0.000     12.462    0.326  132.532
 5CC           C2     C    CR6       0.000     12.852   -0.992  132.517
 5CC           O2     O    O         0.000     11.979   -1.885  132.293
 5CC           C6     C    CR16      0.000     15.218   -0.432  132.852
 5CC           H6     H    H         0.000     16.250   -0.723  133.004
 5CC           N1     N    NR6       0.000     14.260   -1.376  132.611
 5CC           "C1'"  C    CH1       0.000     14.634   -2.814  132.670
 5CC           "H1'"  H    H         0.000     13.766   -3.462  132.484
 5CC           "O4'"  O    O2        0.000     15.066   -2.902  134.016
 5CC           "C2'"  C    CH2       0.000     15.902   -3.362  131.989
 5CC           "H2'1" H    H         0.000     16.653   -2.580  131.857
 5CC           "H2'2" H    H         0.000     15.672   -3.811  131.021
```

```
 5CC           "C3'"  C    CH1       0.000     16.441   -4.434  132.934
 5CC           "H3'"  H    H         0.000     17.539   -4.486  132.929
 5CC           "O3'"  O    OH1       0.000     15.835   -5.694  132.688
 5CC           "HO3'" H    H         0.000     16.197   -6.347  133.300
 5CC           "C4'"  C    CH1       0.000     15.894   -4.063  134.270
 5CC           "H4'"  H    H         0.000     15.268   -4.883  134.650
 5CC           "C5'"  C    CH2       0.000     17.055   -3.800  135.257
 5CC           "H5'1" H    H         0.000     17.692   -4.688  135.254
 5CC           "H5'2" H    H         0.000     16.617   -3.672  136.249
 5CC           "O5'"  O    O2        0.000     17.854   -2.625  134.921
 5CC           P      P    P         0.000     19.461   -2.542  135.320
 5CC           HP     H    H         0.000     19.651   -1.364  136.020
 5CC           OP2    O    OP        0.000     19.773   -3.675  136.258
 5CC           OP1    O    OP        0.000     20.189   -2.409  134.015
loop_
_chem_comp_tree.comp_id
_chem_comp_tree.atom_id
_chem_comp_tree.atom_back
_chem_comp_tree.atom_forward
_chem_comp_tree.connect_type
 5CC      O5C    n/a    C5A    START
 5CC      C5A    O5C    C5     .
 5CC      O5A    C5A    .      .
 5CC      C5     C5A    C6     .
 5CC      C4     C5     N3     .
 5CC      N4     C4     HN4    .
 5CC      HN4    N4     .      .
 5CC      N3     C4     C2     .
 5CC      C2     N3     O2     .
 5CC      O2     C2     .      .
 5CC      C6     C5     N1     .
 5CC      H6     C6     .      .
 5CC      N1     C6     "C1'"  .
 5CC      "C1'"  N1     "C2'"  .
 5CC      "H1'"  "C1'"  .      .
 5CC      "O4'"  "C1'"  .      .
 5CC      "C2'"  "C1'"  "C3'"  .
 5CC      "H2'1" "C2'"  .      .
 5CC      "H2'2" "C2'"  .      .
 5CC      "C3'"  "C2'"  "C4'"  .
 5CC      "H3'"  "C3'"  .      .
 5CC      "O3'"  "C3'"  "HO3'" .
 5CC      "HO3'" "O3'"  .      .
 5CC      "C4'"  "C3'"  "C5'"  .
 5CC      "H4'"  "C4'"  .      .
 5CC      "C5'"  "C4'"  "O5'"  .
 5CC      "H5'1" "C5'"  .      .
 5CC      "H5'2" "C5'"  .      .
 5CC      "O5'"  "C5'"  P      .
 5CC      P      "O5'"  OP1    .
 5CC      HP     P      .      .
 5CC      OP2    P      .      .
 5CC      OP1    P      .      END
 5CC      "C4'"  "O4'"  .    ADD
 5CC      N1     C2     .    ADD
loop_
```

```
_chem_comp_bond.comp_id
_chem_comp_bond.atom_id_1
_chem_comp_bond.atom_id_2
_chem_comp_bond.type
_chem_comp_bond.value_dist
_chem_comp_bond.value_dist_esd
 5CC      OP2    P         deloc       1.510    0.020
 5CC      OP1    P         deloc       1.510    0.020
 5CC      P      "O5'"     single      1.600    0.020
 5CC      "O5'"  "C5'"     single      1.426    0.020
 5CC      "C5'"  "C4'"     single      1.524    0.020
 5CC      "C4'"  "O4'"     single      1.426    0.020
 5CC      "C4'"  "C3'"     single      1.524    0.020
 5CC      "O4'"  "C1'"     single      1.426    0.020
 5CC      "C1'"  N1        single      1.465    0.020
 5CC      "C2'"  "C1'"     single      1.524    0.020
 5CC      N1     C2        single      1.410    0.020
 5CC      N1     C6        single      1.337    0.020
 5CC      O2     C2        aromatic    1.250    0.020
 5CC      C2     N3        aromatic    1.350    0.020
 5CC      N3     C4        aromatic    1.350    0.020
 5CC      N4     C4        aromatic    1.355    0.020
 5CC      C4     C5        single      1.487    0.020
 5CC      C6     C5        aromatic    1.390    0.020
 5CC      C5     C5A       single      1.500    0.020
 5CC      O5A    C5A       deloc       1.250    0.020
 5CC      C5A    O5C       deloc       1.250    0.020
 5CC      "C3'"  "C2'"     single      1.524    0.020
 5CC      "O3'"  "C3'"     single      1.432    0.020
 5CC      HP     P         single      1.383    0.020
 5CC      "H5'1" "C5'"     single      1.092    0.020
 5CC      "H5'2" "C5'"     single      1.092    0.020
 5CC      "H4'"  "C4'"     single      1.099    0.020
 5CC      "H1'"  "C1'"     single      1.099    0.020
 5CC      HN4    N4        single      0.954    0.020
 5CC      H6     C6        single      1.083    0.020
 5CC      "H3'"  "C3'"     single      1.099    0.020
 5CC      "H2'1" "C2'"     single      1.092    0.020
 5CC      "H2'2" "C2'"     single      1.092    0.020
 5CC      "HO3'" "O3'"     single      0.967    0.020
loop_
_chem_comp_angle.comp_id
_chem_comp_angle.atom_id_1
_chem_comp_angle.atom_id_2
_chem_comp_angle.atom_id_3
_chem_comp_angle.value_angle
_chem_comp_angle.value_angle_esd
 5CC      O5C    C5A    O5A     123.000    3.000
 5CC      O5C    C5A    C5      120.000    3.000
 5CC      O5A    C5A    C5      120.000    3.000
 5CC      C5A    C5     C4      120.000    3.000
 5CC      C5A    C5     C6      120.000    3.000
 5CC      C4     C5     C6      120.000    3.000
 5CC      C5     C4     N4      120.000    3.000
 5CC      C5     C4     N3      120.000    3.000
 5CC      N4     C4     N3      120.000    3.000
```

```
 5CC      C4     N4     HN4     120.000    3.000
 5CC      C4     N3     C2      120.000    3.000
 5CC      N3     C2     O2      120.000    3.000
 5CC      N3     C2     N1      120.000    3.000
 5CC      O2     C2     N1      120.000    3.000
 5CC      C5     C6     H6      120.000    3.000
 5CC      C5     C6     N1      120.000    3.000
 5CC      H6     C6     N1      120.000    3.000
 5CC      C6     N1     "C1'"   120.000    3.000
 5CC      C6     N1     C2      120.000    3.000
 5CC      "C1'"  N1     C2      120.000    3.000
 5CC      N1     "C1'"  "H1'"   109.470    3.000
 5CC      N1     "C1'"  "O4'"   109.470    3.000
 5CC      N1     "C1'"  "C2'"   109.470    3.000
 5CC      "H1'"  "C1'"  "O4'"   109.470    3.000
 5CC      "H1'"  "C1'"  "C2'"   108.340    3.000
 5CC      "O4'"  "C1'"  "C2'"   109.470    3.000
 5CC      "C1'"  "O4'"  "C4'"   111.800    3.000
 5CC      "C1'"  "C2'"  "H2'1"  109.470    3.000
 5CC      "C1'"  "C2'"  "H2'2"  109.470    3.000
 5CC      "C1'"  "C2'"  "C3'"   111.000    3.000
 5CC      "H2'1" "C2'"  "H2'2"  107.900    3.000
 5CC      "H2'1" "C2'"  "C3'"   109.470    3.000
 5CC      "H2'2" "C2'"  "C3'"   109.470    3.000
 5CC      "C2'"  "C3'"  "H3'"   108.340    3.000
 5CC      "C2'"  "C3'"  "O3'"   109.470    3.000
 5CC      "C2'"  "C3'"  "C4'"   111.000    3.000
 5CC      "H3'"  "C3'"  "O3'"   109.470    3.000
 5CC      "H3'"  "C3'"  "C4'"   108.340    3.000
 5CC      "O3'"  "C3'"  "C4'"   109.470    3.000
 5CC      "C3'"  "O3'"  "HO3'"  109.470    3.000
 5CC      "C3'"  "C4'"  "H4'"   108.340    3.000
 5CC      "C3'"  "C4'"  "C5'"   111.000    3.000
 5CC      "C3'"  "C4'"  "O4'"   109.470    3.000
 5CC      "H4'"  "C4'"  "C5'"   108.340    3.000
 5CC      "H4'"  "C4'"  "O4'"   109.470    3.000
 5CC      "C5'"  "C4'"  "O4'"   109.470    3.000
 5CC      "C4'"  "C5'"  "H5'1"  109.470    3.000
 5CC      "C4'"  "C5'"  "H5'2"  109.470    3.000
 5CC      "C4'"  "C5'"  "O5'"   109.470    3.000
 5CC      "H5'1" "C5'"  "H5'2"  107.900    3.000
 5CC      "H5'1" "C5'"  "O5'"   109.470    3.000
 5CC      "H5'2" "C5'"  "O5'"   109.470    3.000
 5CC      "C5'"  "O5'"  P       120.500    3.000
 5CC      "O5'"  P      HP      109.500    3.000
 5CC      "O5'"  P      OP2     108.200    3.000
 5CC      "O5'"  P      OP1     108.200    3.000
 5CC      HP     P      OP2     109.500    3.000
 5CC      HP     P      OP1     109.500    3.000
 5CC      OP2    P      OP1     119.900    3.000
loop_
_chem_comp_tor.comp_id
_chem_comp_tor.id
_chem_comp_tor.atom_id_1
_chem_comp_tor.atom_id_2
_chem_comp_tor.atom_id_3
```

```
_chem_comp_tor.atom_id_4
_chem_comp_tor.value_angle
_chem_comp_tor.value_angle_esd
_chem_comp_tor.period
 5CC      var_1      O5C    C5A    C5     C6        -1.252          20.000   3
 5CC      CONST_1    C5A    C5     C4     N3         0.000           0.000   0
 5CC      CONST_2    C5     C4     N4     HN4      180.000           0.000   0
 5CC      CONST_3    C5     C4     N3     C2         0.000           0.000   0
 5CC      CONST_4    C4     N3     C2     O2         0.000           0.000   0
 5CC      CONST_5    C5A    C5     C6     N1         0.000           0.000   0
 5CC      CONST_6    C5     C6     N1     "C1'"      0.000           0.000   0
 5CC      CONST_7    C6     N1     C2     N3         0.000           0.000   0
 5CC      var_2      C6     N1     "C1'"  "C2'"    -46.191          20.000   3
 5CC      var_3      N1     "C1'"  "O4'"  "C4'"   -161.146          20.000   3
 5CC      var_4      N1     "C1'"  "C2'"  "C3'"    180.000          20.000   3
 5CC      var_5      "C1'"  "C2'"  "C3'"  "C4'"    180.000          20.000   3
 5CC      var_6      "C2'"  "C3'"  "O3'"  "HO3'"   180.000          20.000   3
 5CC      var_7      "C2'"  "C3'"  "C4'"  "C5'"    180.000          20.000   3
 5CC      var_8      "C3'"  "C4'"  "O4'"  "C1'"     22.569          20.000   3
 5CC      var_9      "C3'"  "C4'"  "C5'"  "O5'"    180.000          20.000   3
 5CC      var_10     "C4'"  "C5'"  "O5'"  P       -150.542          20.000   3
loop_
_chem_comp_chir.comp_id
_chem_comp_chir.id
_chem_comp_chir.atom_id_centre
_chem_comp_chir.atom_id_1
_chem_comp_chir.atom_id_2
_chem_comp_chir.atom_id_3
_chem_comp_chir.volume_sign
 5CC      chir_01          "C4'"  "C5'"  "O4'"  "C3'"     negativ
 5CC      chir_02          "C1'"  "O4'"  N1     "C2'"     negativ
 5CC      chir_03          "C3'"  "C4'"  "C2'"  "O3'"     positiv
loop_
_chem_comp_plane_atom.comp_id
_chem_comp_plane_atom.plane_id
_chem_comp_plane_atom.atom_id
_chem_comp_plane_atom.dist_esd
 5CC      plan-1            N1        0.020
 5CC      plan-1            "C1'"     0.020
 5CC      plan-1            C2        0.020
 5CC      plan-1            C6        0.020
 5CC      plan-1            N3        0.020
 5CC      plan-1            C4        0.020
 5CC      plan-1            C5        0.020
 5CC      plan-1            O2        0.020
 5CC      plan-1            N4        0.020
 5CC      plan-1            HN4       0.020
 5CC      plan-1            C5A       0.020
 5CC      plan-1            H6        0.020
 5CC      plan-2            C5A       0.020
 5CC      plan-2            C5        0.020
 5CC      plan-2            O5A       0.020
 5CC      plan-2            O5C       0.020
#
# ---------------------------------------------------
# ---------------------------------------------------
```

REFERENCES

1. Dahm, R. (2005) Friedrich Miescher and the discovery of DNA, *Developmental Biology 278*, 274-288.
2. Avery, O. T., MacLeod, C. M., and McCarty, M. (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III, *The Journal of Experimental Medicine 79*, 137-158.
3. Chargaff, E., and Lipshitz, R. (1953) Composition of Mammalian Desoxyribonucleic Acids, *Journal of the American Chemical Society 75*, 3658-3661.
4. Chargaff, E. (1951) Structure and function of nucleic acids as cell constituents, *Federation Proceedings*, *10*, 654-659.
5. Wilkins, M. (2003) *The Third Man of the Double Helix: The Autobiography of Maurice Wilkins*, Oxford University Press Oxford, UK.
6. Watson, J. D., and Crick, F. H. (1953) Molecular structure of nucleic acids, *Nature 171*, 737-738.
7. Watson, J. D. (2011) *The Double Helix: A personal Account of the Discovery of the Structure of DNA*, Scribner, New York, NY.
8. Maddox, B. (2003) *Rosalind Franklin: The dark lady of DNA*, HarperCollins UK.
9. Crick, F. H. (1966) The genetic code, yesterday, today, and tomorrow, In *Cold Spring Harbor Symposia on Quantitative Biology*, 3-9, Cold Spring Harbor Laboratory Press.
10. Meselson, M., and Stahl, F. W. (1958) The replication of DNA in *Escherichia coli*, *Proceedings of the National Academy of Sciences U. S. A. 44*, 671-682.
11. Nirenberg, M. W. (1963) The genetic code: II. *Scientific American 208*, 80-94.
12. Nirenberg, M. (2004) Historical review: Deciphering the genetic code - a personal account, *Trends in Biochemical Sciences, 29*, 46-54.
13. Khorana, H. G. (1968) Nucleic acid synthesis in the study of the genetic code, *Nobel Lectures: Physiology or Medicine (1963-1970)*, 341-369.

14. Dickerson, R. E., Drew, H. R., Conner, B. N., Wing, R. M., Fratini, A. V., and Kopka, M. L. (1982) The anatomy of A-, B-, and Z-dna, *Science 216*, 475-485.
15. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer, New York.
16. Bonasio, R., Tu, S., and Reinberg, D. (2010) Molecular signals of epigenetic states, *Science 330*, 612-616.
17. Jones, P. A., and Takai, D. (2001) The role of DNA methylation in mammalian epigenetics, *Science 293*, 1068-1070.
18. Santi, D. V., Garrett, C. E., and Barr, P. J. (1983) On the mechanism of inhibition of DNA-cytosine methyltransferases by cytosine analogs, *Cell 33*, 9-10.
19. Goll, M. G., and Bestor, T. H. (2005) Eukaryotic cytosine methyltransferases, *Annual Review of Biochemistry 74*, 481-514.
20. Yebra, M. J., and Bhagwat, A. S. (1995) A cytosine methyltransferase converts 5-methylcytosine in DNA to thymine, *Biochemistry 34*, 14752-14757.
21. Shen, J.-C., Rideout III, W. M., and Jones, P. A. (1992) High frequency mutagenesis by a DNA methyltransferase, *Cell 71*, 1073-1080.
22. Liutkeviciute, Z., Liutkeviciute, G., Masevicius, V., Daujotyte, D., and Klimasauskas, S. (2009) Cytosine-5-methyltransferases add aldehydes to DNA, *Nature Chemical Biology 5*, 400-402.
23. Meissner, A. (2010) Epigenetic modifications in pluripotent and differentiated cells, *Nature Biotechnology 28*, 1079-1088.
24. Feng, S., Jacobsen, S. E., and Reik, W. (2010) Epigenetic reprogramming in plant and animal development, *Science 330*, 622-627.
25. Nabel, C. S., Manning, S. A., and Kohli, R. M. (2011) The Curious Chemical Biology of Cytosine: Deamination, Methylation,and Oxidation as Modulators of Genomic Potential, *ACS Chemical Biology 7*, 20-30.
26. Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haff, T. (2000) Embriogenesis: Demethylation of the zygotic paternal genome, *Nature 403*, 501-502.
27. Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., and Fundele, R., Dean, W., Reik, W., Walter, J. (2000) Active demethylation of the paternal genome in the mouse zygote, *Current Biology 10*, 475-478.

28. Paroush, Z., Keshet, I., and Yisraeli, J. (1990) Dynamics of demethylation and activation of the alpha-actin gene in myoblasts, *Cell 63*, 1229-1237.
29. Zhang, F., Pomerantz, J. H., Sen, G., Palermo, A. T., and Blau, H. M. (2007) Active tissue-specific DNA demethylation conferred by somatic cell nuclei in stable heterokaryons, *Proceedings of the National Academy of Sciences U. S. A. 104*, 4395-4400.
30. Rai, K., Huggins, I. J., James, S. R., Karpf, A. R., Jones, D. A., and Cairns, B. R. (2008) DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and Gadd45, *Cell 135*, 1201-1212.
31. Bhutani, N., Brady, J. J., Damian, M., Sacco, A., Corbel, S. Y., and Blau, H. M. (2010) Reprogramming towards pluripotency requires AID-dependent DNA demethylation, *Nature 463*, 1042-1047.
32. Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., and Honjo, T., (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme, *Cell 102*, 553-563.
33. Kriaucionis, S., and Heintz, N. (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain, *Science 324*, 929-930.
34. Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., and Aravind, L. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1, *Science 324*, 930-935.
35. Iyer, L. M., Tahiliani, M., Rao, A., and Aravind, L. (2009) Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids, *Cell Cycle 8*, 1698-1710.
36. Loenarz, C., and Schofield, C. J. (2009) Oxygenase catalyzed 5-methylcytosine hydroxylation, *Chemistry & biology 16*, 580-583.
37. Globisch, D., Muenzel, M., Muller, M., Michalakis, S., Wagner, M., Koch, S., Brueckl, T., Biel, M., and Carell, T. (2010) Tissue distribution of 5-

hydroxymethylcytosine and search for active demethylation intermediates, *PLoS One 5*, e15367.

38. Guo, J., Su, Y., Zhong, C., Ming, G. L., and Song, H. (2011) Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain, *Cell 145*, 423-434.
39. Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., He, C., and Zhang, Y. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine, *Science 333*, 1300-1303.
40. He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., and Li, L. (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA, *Science 333*, 1303-1307.
41. Valinluck, V., Tsai, H. H., Rogstad, D. K., Burdzy, A., Bird, A., and Sowers, L. C. (2005) Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2), *Nucleic Acid Research 33,* 3057-3064.
42. Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., and Cigliano, A., Le Coz, M., Devarajan, K., Wessels, K., Soprano, D., Abramowitz, L. K., Bartolomei, M. S., Rainbow, F., Bassi, M. R., Bruno, T., Fanciulli, M., Renner, C., Klein-Szanto A. J., Matsumoto Y., Kobi, D., Davidson I., Alberti, C., Laure, L., and Bellacosa, A. (2011) Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair., *Cell 146*, 67-79.
43. He, Y. F., Li, B. Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., Sun, Y., Li, X., Dai, Q., Song, C., Zhang, K., He, C., and Xu, G. (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA., *Science 333*, 1303-1307.
44. Fortini, B., Pascucci, B., Parlanti, E., D'Errico, M., Simonelli, V., and Dogliotto, E. (2003) The base excision repair: mechanisms and its relevance for cancer susceptibility, *Biochimie 85*, 1053-1071.
45. Wilson, D. M., and Bohr, V. A. (2007) THe miechanisc of base excision repair, and its relationship to aging and disease, *DNA repair 6*, 544-559.

46. Mishina, Y., Duguit, E., and He, C. (2006) Direct reversal of DNA alkylation damage, *Chemical Reviews 106*, 215-232.
47. Stivers, J. T. (2008) Extrahelical damage base recognition by DNA glycosylase enzymes., *Chemistry A European Journal 14*, 786-793.
48. Gueron, M., and Leroy, J.-L. (1995) Studies of base pair kinetics by NMR measurement of proton exchange, *Methods in Enzymology 16*, 383-413.
49. Snoussi, K., and Leroy, J. L. (2001) Imino Proton Exchange and Base-Pair Kinetics in RNA Duplexes, *Biochemistry 40*, 8898-8904.
50. Chen, C., and Russu, I. M. (2004) Sequence-Dependence of the Energetics of Opening of AT Basepairs in DNA, *Biophysical Journal 87*, 2545-2551.
51. Moe, J. G., and Russu, I. M. (1990) Proton exchange and base-pair opening kinetics in 5'-(CGCGAATTCGCG)-3' and related dodecamers, *Nucleic acids research 18*, 821-827.
52. Moe, J. G., and Russu, I. M. (1992) Kinetics and energetics of base-pair opening in 5'-d(CGCGAATTCGCG)-3' and a substituted dodecamer containing G/T mismatches, *Biochemistry 31*, 8421-8428.
53. Priyakumar, U. D., and MacKerell, A. D. (2005) NMR Imino Proton Exchange Experiments on Duplex DNA Primarily Monitor the Opening of Purine Bases, *Journal of the American Chemical Society 128*, 678-679.
54. Cao, C., Jiang, Y. L., Stivers, J. T., and Song, F. (2004) Dynamic opening of DNA during the enzymatic search for a damaged base, *Nature Structural & Molecular Biology 11*, 1230-1236.
55. Krosky, D. J., Song, F., and Stivers, J. T. (2005) The Origins of High-Affinity Enzyme Binding to an Extrahelical DNA Base., *Biochemistry 44*, 5949-5959.
56. Krosky, D. J., Schwarz, F. P., and Stivers, J. T. (2004) Linear Free Energy Correlations for Enzymatic Base Flipping: How Do Damaged Base Pairs Facilitate Specific Recognition?, *Biochemistry 43*, 4188-4195.
57. Hardeland, U., Bentele, M., Lettieri, T., Steinacher, R., Jiricny, J., and Schar, P. (2001) Thymine DNA glycosylase, *Progress in Nucleic Acid Research and Molecular Biology 68*, 235-253.

58. Rabi, I. I., Zacharias, J. R., Millman, S., and Krusch, P. (1938) A New Method of Measuring Nuclear Magnetic Moment, *Physical Review 53*, 318.
59. Bloch, F. (1946) Nuclear induction, *Physical review 70*, 460-474.
60. Purcell, E. M., Torrey, H., and Pound, R. V. (1946) Resonance absorption by nuclear magnetic moments in a solid, *Physical review 69*, 37.
61. Ernst, R. R., and Anderson, W. (1965) Sensitivity Enhancement in Magnetic Resonance. II. Investigation of Intermediate Passage Conditions., *Review of Scientific Instruments 36,* 1696-1706.
62. Jeener, J. (1971) Lecture presented at Ampere International Summer School II, *Basko Polje, Yugoslavia*.
63. Aue, W., Bartholdi, E., and Ernst, R. R. (1976) 2Dimensional spectroscopy. Application to nuclear magnetic resonance, *The Journal of Chemical Physics 64*, 2229-2246.
64. Wuthrich, K., Wider, G., Wagner, G., and Braun, W. (1982) Sequential resonance assignments as a basis for determination of spatial protein structures by high resolution proton nuclear magnetic resonance, *Journal of Molecular Biology 155*, 311-319.
65. Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons Inc. USA.
66. Rance, M., Sorensen, O., Bodenhausen, G., Wagner, G., Ernst, R., and Wuthrich, K. (1983) Improved spectral resolution in COSY ^{1}H NMR spectra of proteins via double quantum filtering, *Biochemical and Biophysical Research Communications 117*, 479-485.
67. Zhou, N., Manogaran, S., Zon, G., and James, T. L. (1988) Deoxyribose ring conformation of [d-(GGTATACC)]$_2$: an analysis of vicinal proton-proton coupling constants from two-dimensional proton nuclear magnetic resonance, *Biochemistry 27*, 6013-6020.
68. Broido, M. S., Zon, G., and James, T. L. (1984) Complete assignment of the non-exchangeable proton NMR resonances of [d-(GGAATTCC)]$_2$ using two-dimensional nuclear overhauser effect spectra, *Biochemical and biophysical research communications 119*, 663-670.

69. Boelens, R., Scheek, R. M., Dijkstra, K., and Kaptein, R. (1985) Sequential assignment of imino- and amino-proton resonances in ^{1}H NMR spectra of oligonucleotides by two-dimensional NMR spectroscopy. Application to a *lac* operator fragment, *Journal of the Magnetic Resonance 62*, 378-386.

70. Marion, D., and Wuthrich, K. (1983) Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^{1}H-^{1}H spin-spin coupling constants in proteins, *Biochemical and Biophysical Research communications 113*, 967-974.

71. Piantini, U., Sorensen, O., and Ernst, R. R. (1982) Multiple quantum filters for elucidating NMR coupling networks, *Journal of the American Chemical Society 104*, 6800-6801.

72. Saenger, W. (1984) *Principles of nucleic acid structure*, Vol. 7, Springer-Verlag New York.

73. Griesinger, C., Otting, G., Wuthrich, K., and Ernst, R. R. (1988) Clean TOCSY for proton spin system identification in macromolecules, *Journal of the American Chemical Society 110*, 7870-7872.

74. Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R., Wyckoff, H., and Phillips, D. (1958) A three-dimensional model of the myoglobin molecule obtained by X-ray analysis, *Nature 181*, 662-666.

75. Kendrew, J., Dickerson, R., Strandberg, B., and Hart, R. (1997) Structure of myoglobin, *Science Is Not a Quiet Life: Unravelling the Atomic Mechanism of Haemoglobin 4*, 169-175.

76. Dickerson, R. E. (1980) Crystal structure analysis of a complete turn of B-DNA, *Nature 287*, 755-758.

77. Tereshko, V., Minasov, G., and Egli, M. (1999) The Dickerson-Drew B-DNA dodecamer revisited at atomic resolution, *Journal of the American Chemical Society 121*, 470-471.

78. Maiti, A., and Drohat, A. C. (2011) Thymine DNA Glycosylase Can Rapidly Excise 5-Formylcytosine and 5-Carboxylcytosine, *Journal of Biological Chemistry 286*, 35334-35338.

79. Tainer, J. A. (2001) Structural implications of BER enzymes: Dragons dancing--the structural biology of DNA base excision repair, *Progress in Nucleic Acid Research and Molecular Biology 68*, 299-304.

80. Stivers, J. T. (2004) Site-specific DNA damage recognition by enzyme-induced base flipping, *Progress in Nucleic Acid Research and Molecular Biology 77*, 37-65.

81. Fromme, J. C., and Verdine, G. L. (2004) Base excision repair, *Advances in Protein Chemistry 69*, 1-41.

82. Stivers, J. T. (2004) Site-specific DNA damage recognition by enzyme-induced base flipping, *Progress in Nucleic Acid Research and Molecular Biology 77*, 37-65.

83. Marky, L. A., and Breslauer, K. J. (1987) Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves, *Biopolymers 26*, 1601-1620.

84. Kaushik, M., Suehl, N., and Marky, L. A. (2007) Calorimetric unfolding of the bimolecular and i-motif complexes of the human telomere complementary strand, $d(C3TA2)_4$, *Biophysical Chemistry 126*, 154-164.

85. Gueron, M., and Leroy, J. L. (1995) Studies of base pair kinetics by NMR measurement of proton exchange, *Methods in Enzymology 261*, 383-413.

86. Parker, J. B., and Stivers, J. T. (2011) Dynamics of uracil and 5-fluorouracil in DNA, *Biochemistry 50*, 612-617.

87. Crenshaw, C. M., Wade, J. E., Arthanari, H., Frueh, D., Lane, B. F., and Nunez, M. E. (2011) Hidden in plain sight: Subtle effects of the 8-oxoguanine lesion on the structure, dynamics, and thermodynamics of a 15-base pair oligodeoxynucleotide duplex, *Biochemistry 50*, 8463-8477.

88. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J. A., Jr.; , Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J.,

Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, N. J., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Ö., Foresman, J. B., Ortiz, J. V., Cioslowski, J., and Fox, D. J. (2009) Gaussian 09, Revision D.01.

89. Rosenbaum, G., Alkire, R. W., Evans, G., Rotella, F. J., Lazarski, K., Zhang, R. G., Ginell, S. L., Duke, N., Naday, I., and Lazarz, J. (2005) The Structural Biology Center 19ID undulator beamline: facility specifications and protein crystallographic results, *Journal of Synchrotron Radiation 13*, 30-45.
90. Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution-from diffraction images to an initial model in minutes, *Acta Crystallographica Section D: Biological Crystallography 62*, 859-866.
91. Kabsch, W. (2010) XDS, *Acta Crystallographica Section D 66*, 125-132.
92. Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement, *Acta Crystallographica Section D 66*, 133-144.
93. Evans, P. (2005) Scaling and assessment of data quality, *Acta Crystallographica Section D: Biological Crystallography 62*, 72-82.
94. Collaborative Computational Project, N. (1994) The CCP4 suite: programs for protein crystallography, *Acta Crystallographica Section D 50*, 760-763.
95. Vagin, A. (1989) New translation and packing functions, *Newsletter on protein crystallography., Daresbury Laboratory 24*.
96. Vagin, A., and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement, *Journal of Applied Crystallography 30*, 1022-1025.
97. Emsley, P., Lohkamp, B., Scott, W., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallographica Section D: Biological Crystallography 66*, 486-501.

98. Vagin, A. A., Steiner, R. A., Lebedev, A. A., Potterton, L., McNicholas, S., Long, F., and Murshudov, G. N. (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use, *Acta Crystallographica Section D: Biological Crystallography 60*, 2184-2195.

99. Winn, M. D., Murshudov, G. N., and Papiz, M. Z. (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions, *Methods in Enzymology 374*, 300-321.

100. Lavery, R., Moakher, M., Maddocks, J., Petkeviciute, D., and Zakrzewska, K. (2009) Conformational analysis of nucleic acids revisited: Curves+, *Nucleic acids research 37*, 5917-5929.

101. Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. (1988) Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells: The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases, *Journal of Molecular Biology 203*, 971-983.

102. Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999) DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development, *Cell 99*, 247-257.

103. Meissner, A. (2010) Epigenetic modifications in pluripotent and differentiated cells, *Nature Biotechnology 28*, 1079-1088.

104. Wu, S. C., and Zhang, Y. (2010) Active DNA demethylation: Many roads lead to Rome, *Nature Reviews Molecular Cell Biology 11*, 607-620.

105. Zhang, P., Su, L., Wang, Z., Zhang, S., Guan, J., Chen, Y., Yin, Y., Gao, F., Tang, B., and Li, Z. (2012) The involvement of 5-hydroxymethylcytosine in active DNA demethylation in mice, *Biology of Reproduction 104*, 1-9.

106. Wossidlo, M., Arand, J., Sebastiano, V., Lepikhov, K., Boiani, M., Reinhardt, R., Scholer, H., and Walter, J. (2010) Dynamic link of DNA demethylation, DNA strand breaks and repair in mouse zygotes, The *EMBO Journal 29*, 1877-1888.

107. Kriaucionis, S., and Heintz, N. (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain, *Science 324*, 929-930.

108. Munzel, M., Globisch, D., and Carell, T. (2011) 5-Hydroxymethylcytosine, the sixth base of the genome, *Angewandte Chemie International Edition 50*, 6460-6468.

109. Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L., and Rao, A. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1, *Science 324*, 930-935.

110. Ito, S., D'Alessio, A. C., Taranova, O. V., Hong, K., Sowers, L. C., and Zhang, Y. (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification, *Nature 466*, 1129-1133.

111. Cadet, J., and Wagner, J. R. (2013) TET enzymatic oxidation of 5-methylcytosine, 5-hydroxymethylcytosine and 5-formylcytosine, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis.*

112. Bienvenu, C., Wagner, J. R., and Cadet, J. (1996) Photosensitized oxidation of 5-methyl-2'-deoxycytidine by 2-methyl-1,4-naphthoquinone: Characterization of 5-(hydroperoxymethyl)-2'-deoxycytidine and stable methyl group oxidation products, *Journal of the American Chemical Society 118*, 11406-11411.

113. Yang, H., Liu, Y., Bai, F., Zhang, J. Y., Ma, S. H., Liu, J., Xu, Z. D., Zhu, H. G., Ling, Z. Q., Ye, D., Guan, K. L., and Xiong, Y. (2013) Tumor development is associated with decrease of TET gene expression and 5-methylcytosine hydroxylation, *Oncogene 32*, 663-669.

114. Globisch, D., Munzel, M., Muller, M., Michalakis, S., Wagner, M., Koch, S., Bruckl, T., Biel, M., and Carell, T. (2010) Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates, *PLoS ONE 5*, e15367.

115. Maiti, A., and Drohat, A. C. (2011) Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: Potential implications for active demethylation of CpG sites, *Journal of Biological Chemistry 286*, 35334-35338.

116. Renciuk, D., Blacque, O., Vorlickova, M., and Spingler, B. (2013) Crystal structures of B-DNA dodecamer containing the epigenetic modifications 5-

hydroxymethylcytosine or 5-methylcytosine, *Nucleic Acids Research 41,* 9891-9900.

117. Maiti, A., Morgan, M. T., Pozharski, E., and Drohat, A. C. (2008) Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition, *Proceedings of the National Academy of Sciences U.S.A. 105*, 8890-8895.

118. Maiti, A., Morgan, M. T., and Drohat, A. C. (2009) Role of two strictly conserved residues in nucleotide flipping and N-glycosylic bond cleavage by human thymine DNA glycosylase, *Journal of Biological Chemistry 284*, 36680-36688.

119. Hashimoto, H., Hong, S., Bhagwat, A. S., Zhang, X., and Cheng, X. (2012) Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation, *Nucleic Acids Research 40*, 10203-10214.

120. Drew, H. R., Wing, R. M., Takano, T., Broka, C., Tanaka, S., Itakura, K., and Dickerson, R. E. (1981) Structure of a B-DNA dodecamer: Conformation and dynamics, *Proceedings of the National Academy of Sciences U.S.A. 78*, 2179-2183.

121. Drew, H. R., and Dickerson, R. E. (1981) Structure of a B-DNA dodecamer. III. Geometry of hydration, *Journal of Molecular Biology 151*, 535-556.

122. Patel, D. J., Shapiro, L., and Hare, D. (1987) DNA and RNA: NMR studies of conformations and dynamics in solution, *Quarterly Reviews of Biophysics 20*, 35-112.

123. Reid, B. R. (1987) Sequence-specific assignments and their use in NMR studies of DNA structure, *Quarterly Reviews of Biophysics 20*, 2-28.

124. Tereshko, V., Minasov, G., and Egli, M. (1999) The Dickerson-Drew B-DNA dodecamer revisited at atomic resolution, *Journal of the American Chemical Society 121*, 6970-6970.

125. Loenarz, C., and Schofield, C. J. (2009) Oxygenase catalyzed 5-methylcytosine hydroxylation, *Chemical Biology 16*, 580-583.

126. Wagner, J. R., and Cadet, J. (2010) Oxidation reactions of cytosine DNA components by hydroxyl radical and one-electron oxidants in aerated aqueous solutions, *Accounts of Chemical Research 43*, 564-571.

127. Bennett, M. T., Rodgers, M. T., Hebert, A. S., Ruslander, L. E., Eisele, L., and Drohat, A. C. (2006) Specificity of human thymine DNA glycosylase depends on N-glycosidic bond stability, *Journal of the American Chemical Society 128*, 12510-12519.

128. Suzuki, M. M., and Bird, A. (2008) DNA methylation landscapes: provocative insights from epigenomics, *Nature Reviews Genetics 9*, 465-476.

129. Zhang, H., Zhang, X., Clark, E., Mulcahey, M., Huang, S., and Shi, Y. G. (2010) TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine, *Cell Research 20*, 1390-1393.

130. Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A. J., Zheng, L., Zhang, H., and Huang, S. (2011) Genome-wide regulation of 5hmC, 5mC, and Cene expression by Tet1 hydroxylase in mouse embryonic stem cells, *Molecular Cell 42*, 451-464.

131. Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C. J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J. r. (2011) 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming, *Nature Communications 2*, 241.

132. Hajkova, P., Jeffries, S. J., Lee, C., Miller, N., Jackson, S. P., and Surani, M. A. (2010) Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway, *Science 329*, 78-82.

133. Pfaffeneder, T., Hackner, M., Truss, M., Muenzel, M., Mueller, M., Deiml, C. A., Hagemeier, C., and Carell, T. (2011) The discovery of 5-formylcytosine in embryonic stem cell DNA, *Angewandte Chemie International Edition 50*, 7008-7012.

134. Raiber, E.-A., Beraldi, D., Ficz, G., Burgess, H. E., Branco, M. R., Murat, P., Oxley, D., Booth, M. J., Reik, W., and Balasubramanian, S. (2012) Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with

transcription and depends on thymine DNA glycosylase, *Genome Biology 13*, R69.

135. Kellinger, M. W., Song, C.-X., Chong, J., Lu, X.-Y., He, C., and Wang, D. (2012) 5-formylcytosine and 5-carboxylcytosine reduce the rate and substrate specificity of RNA polymerase II transcription, *Nature Structural & Molecular Biology 19*, 831-833.

136. Raiber, E., Beraldi, D., Ficz, F., Burgess, H., Branco, M., Murat, P., Oxley, D., Booth, M., Reik, W., and Balasubramanian, S. (2012) Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase., *Genome Biology 13*, R69.

137. Nabel, C., Jia, H., Ye, Y., Shen, L., Goldschmidt, H., Stivers, J., Zhang, Y., and Kohli, R. (2012) AID/APOBEC deaminases disfavor modified cytosines implicates in DNA demethylation, *Nature Chemical Biology 8*, 751-758.

138. Song, C., Szulwach, K., Dai, Q., Fu, Y., Mao, S., Lin, L., Street, C., Li, Y., Poidevin, M., Wu, H., Gao, J., Liu, P., Li, L., Xu, G., Jin, P., and He, C. (2013) Genome-wide profiling of 5-forylcytosine reveals its roles in epigenetic priming, *Cell 153,* 678-691.

139. Shen, L., Wu, H., Diep, D., Yamaguchi, S., D'Alessio, A., Fung, H., Zhang, K., and Zhang, Y. (2013) Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics, *Cell 153*, 692-706.

140. Schiesser, S., Pfaffeneder, T., Sadeghian, K., Hackner, B., Steigenberger, B., Schroeder, A. S., Steinbacher, J., Kashiwazaki, G., Hoefner, G., Wanner, K. T., Ochsenfeld, C., and Carell, T. (2013) Deamination, Oxidation, and C-C Bond Cleavage Reactivity of 5-Hydroxymethylcytosine, 5-Formylcytosine, and 5-Carboxycytosine, *Journal of the American Chemical Society 135*, 14593-14599.

141. Maiti, A., Michelson, A. Z., Armwood, C. J., Lee, J. K., and Drohat, A. C. (2013) Divergent Mechanisms for Enzymatic Excision of 5-Formylcytosine and 5-Carboxylcytosine from DNA, *Journal of the American Chemical Society 135*, 15813-15822.

142. Hashimoto, H., Liu, Y., Upadhyay, A. K., Chang, Y., Howerton, S. B., Vertino, P. M., Zhang, X., and Cheng, X. (2012) Recognition and potential mechanisms for

replication and erasure of cytosine hydroxymethylation, *Nucleic Acids Research 40*, 4841-4849.

143. Hashimoto, H., Zhang, X., and Cheng, X. (2013) Selective Excision of 5-Carboxylcytosine by a Thymine DNA Glycosylase Mutant, *Journal of Molecular Biology 425*, 971-976.

144. LaFrancois, C. J., Fujimoto, J., and Sowers, L. C. (1998) Synthesis and Characterization of Isotopically Enriched Pyrimidine Deoxynucleoside Oxidation Damage Products, *Chemical Research in Toxicology 11*, 75-83.

145. Cortazar, D., Kunz, C., Selfridge, J., Lettieri, T., Saito, Y., MacDougall, E., Wirz, A., Schuermann, D., Jacobs, A. L., and Siegrist, F. (2011) Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability, *Nature 470*, 419-423.

146. Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., and Soprano, D. (2011) Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair, *Cell 146*, 67-79.

147. Friedman, J. I., and Stivers, J. T. (2010) Detection of damaged DNA bases by DNA glycosylase enzymes, *Biochemistry 49*, 4957-4967.

148. Parker, J. B., Bianchet, M. A., Krosky, D. J., Friedman, J. I., Amzel, L. M., and Stivers, J. T. (2007) Enzymatic capture of an extrahelical thymine in the search for uracil in DNA, *Nature 449*, 433-437.

149. Neidigh, J. W., Darwanto, A., Williams, A. A., Wall, N. R., and Sowers, L. C. (2009) Cloning and characterization of Rhodotorula glutinis thymine hydroxylase, *Chemical Research in Toxicology 22*, 885-893.

150. Smiley, J., Kundracik, M., Landfried, D., and Barnes Sr, V. (2005) Genes of the thymidine salvage pathway: Thymine-7-hydroxylase from a Rhodotorula glutinis cDNA library and iso-orotate decarboxylase from *Neurospora crassa*, *Biochimica et Biophysica Acta 1723*, 256-264.

151. Tardy-Planechaud, S. v., Fujimoto, J., Lin, S. S., and Sowers, L. C. (1997) Solid phase synthesis and restriction endonuclease cleavage of oligodeoxynucleotides containing 5-(hydroxymethyl)-cytosine, *Nucleic Acids Research 25*, 553-558.

152. Schiesser, S., Hackner, B., Pfaffeneder, T., Müller, M., Hagemeier, C., Truss, M., and Carell, T. (2012) Mechanism and Stem-Cell Activity of 5-Carboxycytosine Decarboxylation Determined by Isotope Tracing, *Angewandte Chemie International Edition 51*, 6516-6520.

153. Szulik, M. W., Voehler, M. W., Ganguly, M., Gold, B., and Stone, M. P. (2013) Site-Specific Stabilization of DNA by a Tethered Major Groove Amine, 7-Aminomethyl-7-deaza-2'-deoxyguanosine, *Biochemistry 52*, 7659-7668.

154. Lau, A. Y., Scharer, O. D., Samson, L., Verdine, G. L., and Ellenberger, T. (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: mechanisms for nucleotide flipping and base excision, *Cell 95*, 249-258.

155. Hollis, T., Ichikawa, Y., and Ellenberger, T. (2000) DNA bending and a flip-out mechanism for base excision by the helix-hairpin-helix DNA glycosylase, Escherichia coli AlkA, *The EMBO Journal 19*, 758-766.

156. Bruner, S. D., Norman, D. P., and Verdine, G. L. (2000) Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA, *Nature 403*, 859-866.

157. Fromme, J. C., and Verdine, G. L. (2002) Structural insights into lesion recognition and repair by the bacterial 8-oxoguanine DNA glycosylase MutM, *Nature Structural & Molecular Biology 9*, 544-552.

158. Gilboa, R., Zharkov, D. O., Golan, G., Fernandes, A. S., Gerchman, S. E., Matz, E., Kycia, J. H., Grollman, A. P., and Shoham, G. (2002) Structure of formamidopyrimidine-DNA glycosylase covalently complexed to DNA, *Journal of Biological Chemistry 277*, 19811-19816.

159. Krosky, D. J., Song, F., and Stivers, J. T. (2005) The origins of high-affinity enzyme binding to an extrahelical DNA base, *Biochemistry 44*, 5949-5959.

160. Xu, S., Li, W., Zhu, J., Wang, R., Li, Z., Xu, G.-L., and Ding, J. (2013) Crystal structures of isoorotate decarboxylases reveal a novel catalytic mechanism of 5-carboxyl-uracil decarboxylation and shed light on the search for DNA decarboxylase, *Cell Research 23*, 1296-1309.

161. Hashimoto, H., Hong, S., Bhagwat, A. S., Zhang, X., and Cheng, X. (2012) Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA

glycosylase domain: its structural basis and implications for active DNA demethylation, *Nucleic Acids Research 40*, 10203-10214.

162. Zhang, L., Lu, X., Lu, J., Liang, H., Dai, Q., Xu, G.-L., Luo, C., Jiang, H., and He, C. (2012) Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA, *Nature Chemical Biology 8*, 328-330.
163. Muenzel, M., Globisch, D., and Carell, T. (2011) 5-Hydroxymethylcytosine, the Sixth Base of the Genome, *Angewandte Chemie International Edition 50*, 6460-6468.
164. Eltejaye, A. (2009) Thymine DNA Glycosylase (TDG), *Stanislaus Journal of Biochemical Reviews*.
165. Jin, S.-G., Jiang, Y., Qiu, R., Rauch, T. A., Wang, Y., Schackert, G., Krex, D., Lu, Q., and Pfeifer, G. P. (2011) 5-Hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations, *Cancer Research 71*, 7360-7365.
166. Fromme, J. C., and Verdine, G. L. (2003) Base excision repair, *Advances in Protein Chemistry 69*, 1-41.
167. Fu, D., Calvo, J. A., and Samson, L. D. (2012) Balancing repair and tolerance of DNA damage caused by alkylating agents, *Nature Reviews Cancer 12*, 104-120.
168. Hosfield, D. J., Daniels, D. S., Mol, C. D., Putnam, C. D., Parikh, S. S., and Tainer, J. A. (2001) DNA damage recognition and repair pathway coordination revealed by the structural biochemistry of DNA repair enzymes, *Progress in Nucleic Acid Research and Molecular Biology 68*, 315-347.
169. Yang, W. (2006) Poor base stacking at DNA lesions may initiate recognition by many repair proteins, *DNA Repair 5*, 654-666.
170. Plum, G., and Breslauer, K. J. (1994) DNA lesions: a thermodynamic perspective, *Annals of the New York Academy of Sciences 726*, 45-56.
171. Ganguly, M., Szulik, M. W., Donahue, P. S., Clancy, K., Stone, M. P., and Gold, B. (2012) Thermodynamic Signature of DNA Damage: Characterization of DNA with a 5-Hydroxy-2'-deoxycytidine : 2'-deoxyguanosine Base Pair, *Biochemistry 51*, 2018-2027.

172. Szulik, M. W., Nocek, B., Joachimiak, A., and Stone, M. P. (2013) Crystal structure of 5-hydroxy-2'-deoxycytidine base paired with 2'-deoxyguanosine in Dickerson Drew Dodecamer, RCSB Protein Data Bank.