INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

DNA Structure and Environment : Crystallographic Studies of Monovalent Cation - Z-DNA Interactions and of an AHMA - DNA Complex

by

XIANGCHAO DONG

A dissertation submitted to the Graduate Faculty in Chemistry in partial fulfillment of the requirement for the degree of Doctor of Philosophy, The City University of New York

1998

This manuscript has been read and accepted for the Graduate Faculty in Chemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

July 22 1998
Date

Chair of Examining Committee

Sept. 15, 1998
Date

Executive Officer

Supervisory Committee

The City University of New York

Abstract

DNA Structure and Environment : Crystallographic Studies of Monovalent Cation - Z-DNA Interactions and of an AHMA - DNA Complex

by

XIANGCHAO DONG

Adviser: Professor Gary Joseph Quigley

DNA, the molecule of heredity, is the basic element in the replication process. Environment (for instance: counterions, water, drug molecules) can affect the structure and the function of DNA. The relations between the structural changes of the DNA molecules under different conditions and the function of the DNA are important research subjects in life science.

Interactions between rubidium, a monovalent cation, and Z-DNA have been studied using X-ray crystallography. Crystals of an oligonucleotide with the sequence of CGCGCG were crystallized with Rb^+ as counterion. Data to 1.76 Å resolution has been collected under cryogenic condition, to reduce the disorder of the ions in the structure. A total of sixteen rubidium ions have been located with thirteen of them having partial occupancy. The final R value is 17.5%. The result shows that in the Z-DNA, the structures of most monovalent counterions are dynamic — the atoms having fast exchange with water molecules in the crystal.

K-Z-DNA crystals were also obtained from crystallization of an oligonucleotide with the sequence of CGCGCG and with K^+ as counterion. Data was collected using an X-ray diffractometer. Molecular dynamic optimization for the disordered K^+ ions in the K-Z-DNA crystal structure was attempted. The K^+ position was not reproducible in simulation runs with different conditions. This result indicates that either the method we used is not suitable for solvent disorder calculation or that the ions and solvent are highly disordered that the simulation must involve so many superpositions that it is impossible to distinguish true simulation from noise fitting.

Crystals of an AHMA (3-(9-acridinylamino)-5-(hydroxymethyl) aniline) – d(CGTACG) complex have been obtained after an intensive crystallization condition search. Data with resolution of 2.9 Å has been collected using an area detector. The unit cell was found to be a=b=57.52 Å and c=122.17 Å with space group class p321. Attempts have been made to make a heavy metal derivative of the AHMA-DNA complex. Crystals obtained by soaking the AHMA-DNA complex crystal with Sm^{3+} showed a unit cell of a=b=28.54 Å and c=118.41 Å with the same space group as the native crystal. Patterson maps for both crystals suggested that the possible base stacking is in the xy plane. Further work has to be done to solve the structure.

To my family and friends

Acknowledgments

I would like to express my sincere thanks to many people behind my success in obtaining Ph.D. I specially thank the following individuals:

Professor Gary Quigley, my supervisor and mentor, for providing me the opportunity to learn x-ray crystallography, supervising my Ph.D. research project.

Professors Thomas Strekas, William Grossman, and Dixie Goss for their support as my committee member. I am especially grateful for Professor Grossman's effort of reading and correcting my thesis.

Special thanks to Professor Louis Massa and Professor Michael Drain. Their help and encouragement are greatly appreciated.

I thank my friends Lulu Huang, Alex Shafir, Mayumi Noto and other graduate student colleagues. Their friendship gave me tremendous support during the course of this work.

I thank Dr. Shun-le Chen and Mr. Waldemar Cieniewicz for their help in x-ray program and computer facility.

Finally, I would like to express my deep gratitude to my husband, my daughter, my mother and brother, their love and support are the bases for my accomplishment of this work.

Table of Contents

List of Tables

List of Figures

Abbreviations

a,b,c	unit cell lengths of the real lattice
α, β, γ	unit cell angels of the real lattice
AHMA	3-(9-acridinylamino)-5-(hydroxymethyl) aniline
m-AMSA	4'-(9-acridinylamino)methane-sulfon-m-anisidide
ATP	adenosine 5'-triphosphate
C	cytidine
DMTr	dimethoxytrityl
DNA	deoxyribonucleic acid
Fo	observed structure factors
Fc	calculated structure factors
G	guanosine
HPLC	high performance liquid chromatography
mmol	millimole
mM	millimolar
MPD	2-methyl-2,4-pentanediol
mRNA	messenger RNA
PEG	polyethylene glycol
RNA	ribonucleic acid
TEAA	triethylamine acetate
TFA	trifluoroacetic acid
tRNA	transfer RNA
VP-16	4'-demethylepipodophyllotoxin-9-(4,6-O-2-ethylidene-b-D-glucopyranoside

Chapter 1

Introduction

Among all the myriad properties of living organisms, one is absolutely essential for the continuance of life : a living system must be able to replicate itself. DNA - the molecule which carries the genetic code, or the description of heredity, is the basic element for the replication process.

Due to the central role of DNA to life, the structure - function relationship of DNA molecules has long been a major subject in the biological research. Since the environment (water, ions and drug molecules) around DNA plays a significant role in determining the conformation and functions of DNA. interaction between them at the molecular level is an important part of this subject.

In this thesis, several aspects of the interactions of DNA and its environment are examined to gain further insight into this subject. The specific areas of inquiry include examining the interactions of monovalent cations with DNA and that of a specific DNA binding drug with DNA.

1. 1 Structure and Function of Deoxyribonucleic Acid

1.1.1 History of the Discovery of DNA

The nucleic acids were first recognized in 1869 by Friedrich Miescher, a Swiss physician when he studied white blood cell in Tubingen, Germany (1). In 1935 a Russian born chemist, P. A. T. Levene, showed that sugars were connected to each other by a phosphodiester bond in both DNA and RNA (2). In the late 1940's Erwin Chargaff studied the base composition of nucleic acid with paper chromatography and found that the ratios of adenine to thymine and of guanine to cytosine were nearly one (3). Since the development of x-ray crystallography in 1912 it has been possible to study the detailed three-dimensional structure of molecules. A number of researchers applied this technique to the study of DNA structure in the late 1940's and 1950's including James Watson, Francis Crick, Rosalind Franklin, Maurice Wilkins and L. Pauling. In 1953 James Watson and Francis Crick proposed Watson-Crick base pairing (adenine pairs with thymine and guanine pairs with cytosine) and deduced the three-dimensional structure of DNA (4). This discovery established a foundation for the study of the mechanism of the biological function of DNA.

1.1.2 DNA is the Molecule of Heredity, and Genetic Information is Encoded in the Sequence of the DNA Structure

The DNA structure and its functions are explicitly reviewed in detail by Rich (5), Saenger (6), and Stryer (7) among others.

DNA is a double stranded polymer with two antiparallel chains running in opposite directions. The complementary property of the two strands plays a key role in the mechanism of DNA replication and in the expression of genetic information. In DNA replication, the double strands unwind and separate. With each single strand serving as a template, two new DNA double strands are synthesized which are identical to the parent DNA molecule. The replication fidelity is very high with an error frequency less than one in 100 million nucleotide. This is facilitated by a cellular repair system which has the function of fixing the resulting mismatched sequence.

On the other hand, genetic information is transferred through two steps: transcription and translation. In transcription, RNA polymerase binds to DNA and synthesizes a mRNA which is complementary to the DNA template (except T is replaced by U). In the translation process, through the machinery of the ribosome, transfer RNA (tRNA) reads each base triplet (codons) on the mRNA and translates it into a specific amino acid or as a termination. Thus, biological information is generally stored in DNA and flows from DNA to RNA to protein.

1.1.3 General Structure Profile of DNA

DNA is a polymer of deoxyribonucleotide units (Fig. 1-1). A nucleotide consists of a nitrogenous base, a sugar, and a phosphate group. The backbone of DNA is formed by 3' hydroxyl of sugar moiety of one deoxyribonucleotide and the 5'-hydroxyl of the subsequent sugar being joined by a phosphodiester bridge. The four bases are adenine, thymine, guanine and cytosine (Fig. 1-2). Bases are bonded to the C1' of deoxyribose through a β-N-glycosidic linkage.

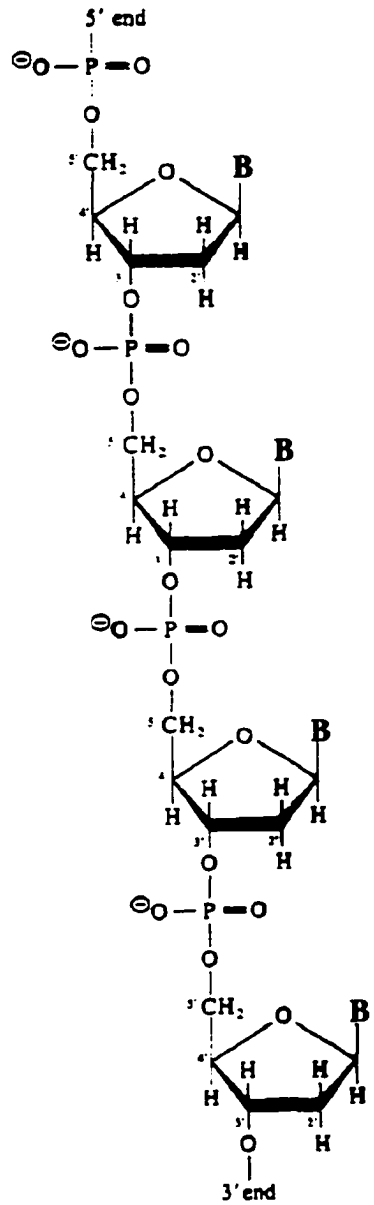


Fig. 1-1 Primary Structure of DNA
(B represents the bases) (after David Rawn (8))

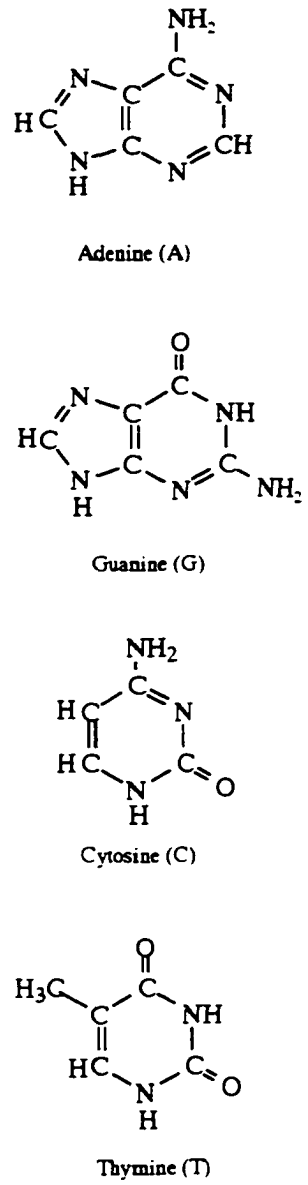


Fig. 1-2 Four Bases in DNA Structures

By convention a sequence of DNA is written and numbered in the direction from 5' to 3'. Sugar atoms are numbered with prime to distinguish them from the bases (Fig. 1-1). The naming of the torsion angles is illustrated in Fig. 1-3.

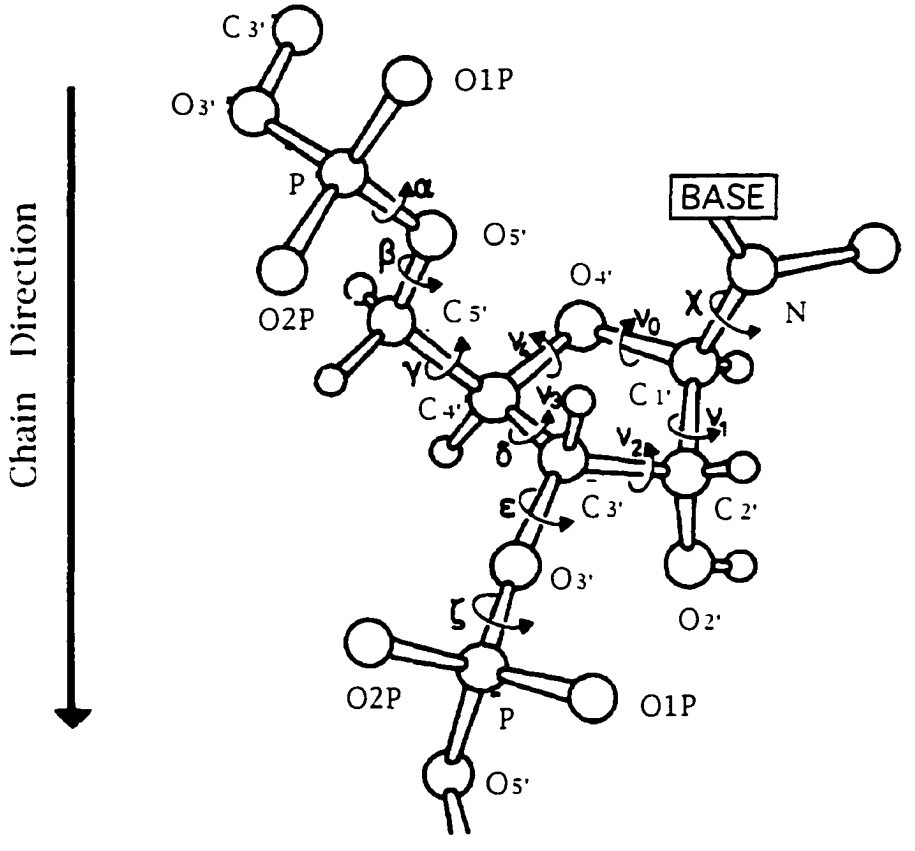


Fig. 1-3 Torsion Angles in DNA Structure (after Saenger (6))

1.1.4 Three Dimensional Structures of DNA

DNA is a double strand helix with two anti-parallel single strands interwound around each other. The hydrogen bonding between bases (G pairs with C and A pairs with T) holds the two chains together. The GC base pair is shown in Fig. 1-4.

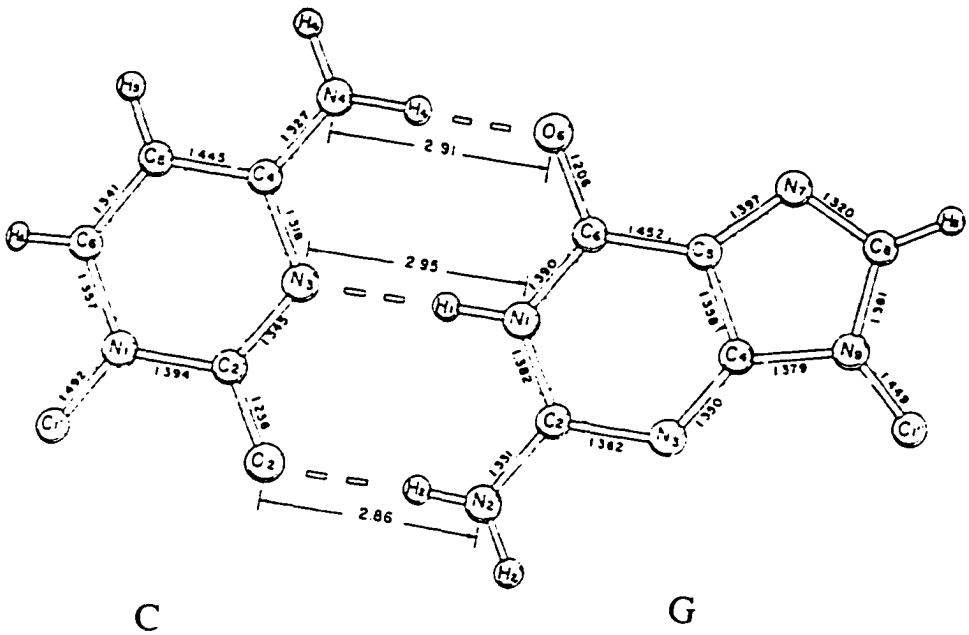


Fig. 1-4 The Watson-Crick CG base pair in the DNA Structure (after Saenger (6))

There are two grooves in the DNA structure. One is the major groove and the other is the minor groove (Fig. 1-5). The major groove is slightly deeper than the minor groove.

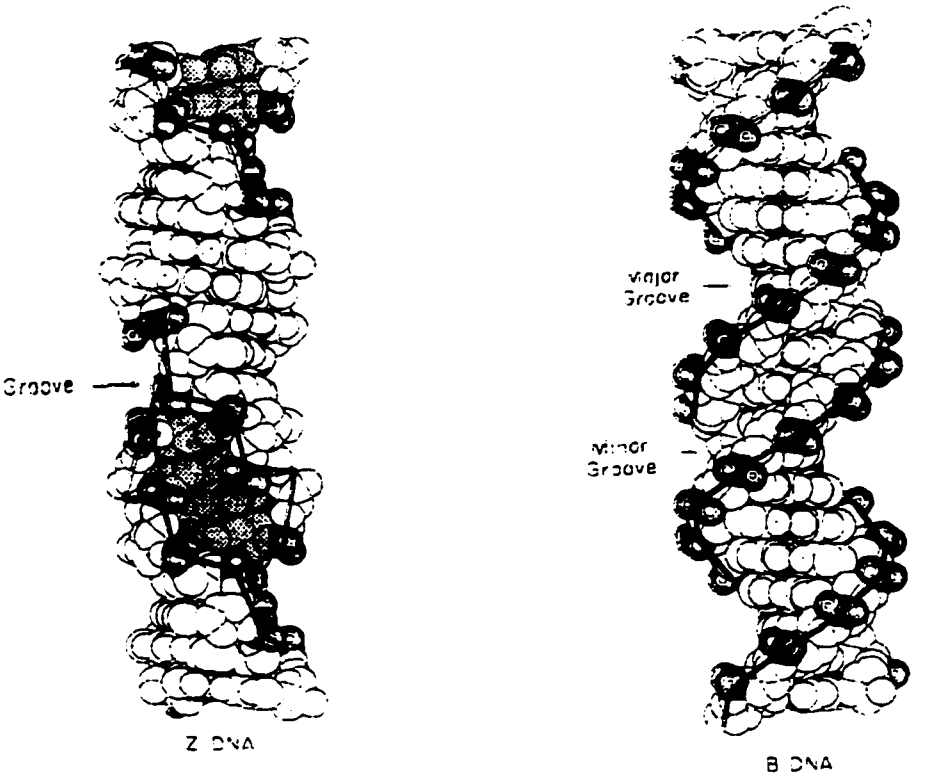
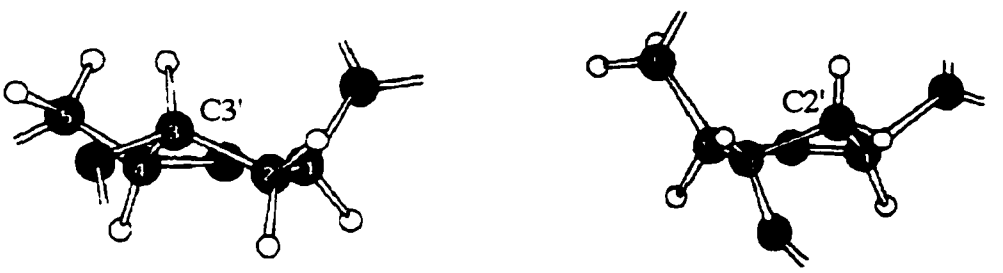


Fig. 1-5 Grooves in B-DNA and Z-DNA (From Gary Quigley)

Two major sugar puckering classes have been found in DNA: C3' endo and C2' endo (Fig. 1-6). In this nomenclature, the endo means the atom is out of the plane of ribose ring and on the C5' side.



C3'- endo C2'- endo

Fig. 1-6 Sugar Puckering in DNA Structure (after Stryer (7))

Since DNA is a negatively charged polymer, counterions are necessary factors for stabilizing the structure and determing its conformation. Water molecules normally surround and form hydrogen bonds with DNA and are thus an integral part of its structure.

1.1.5 Different Conformations in the DNA Family

DNA structure is dynamic and can have several conformations depending on the sequence and environmental conditions such as counterions and humidity. There are three major double helical forms of DNA : A-DNA, B-DNA and Z-DNA. The structural features of these three forms are listed on Table 1-1.

Table 1-1 Structural Comparison of A-, B- and Z-DNA

	A-DNA	B-DNA	Z-DNA
Helix handedness	right	right	left
Helix diameter	25.5 Å	23.7 Å	18.4 Å
Rise per base pair	2.3 Å	3.4 Å	3.8 Å
Pitch	25.3 Å	35.4 Å	45.6 Å
Base pairs per turn	11	10.4	12
Tilt of base pairs from normal to helix axis	19°	1°	9°
Conformation of glycosidic bond	anti	anti	anti at C,T syn at G
Major groove	narrow and very deep	wide and quite deep	flat
Minor groove	very broad and shallow	narrow and quite deep	very narrow and deep

B-DNA and **A-DNA**

The first published three-dimensional picture of DNA was the right handed, double helical B-DNA structure proposed by Watson and Crick in 1953 (4). In the B-DNA all furanose rings have characteristic C2'-endo and all nucleotides have an anti conformation. The planes of the bases are nearly perpendicular to the helix axis. The major groove is wider than the minor groove while the two grooves have a similar depth.

There are two fold axes perpendicular to the helical axis. One is in the plane of the base pairs and the other is half way between consecutive base pairs. It should be noted that these symmetries apply to the deoxyribose phosphate backbone and ignore differences in bases.

A-DNA was obtained by dehydration of the B-DNA. Like B-DNA, A-DNA is a right handed double helix. With 11 base pairs per turn, the diameter of A-DNA is slightly larger than B-DNA. The sugar pucker in A-DNA is C3'-endo and the tilt of the base pair from the normal to the helix in A-DNA is the largest in the DNA family (19°). All the crystal structures which have been determined as A form have a GC in their sequence. A-DNA also has anti conformation for nucleotide in its structure.

Z-DNA

The structure of Z-DNA was first reported by A. Wang *et al* in 1979 (9). The surprising result was that unlike A-DNA and B-DNA, Z-DNA is a left handed duplex.

The characteristic feature of Z-DNA is the zigzag backbone caused by the alternating nature of the deoxyribose phosphate chain in which there are two distinct residues in the repeating motif with very distinct conformations (10, 11).

The first Z-DNA crystals had the sequence of CGCGCG. The CpG sequence has basically one conformation in all Z-DNA structures examined. The GpC sequence can exist in two closely related conformations with the primary difference being a change in the position of the phosphorus and two non-ribose oxygen atoms. The dominant form is referred to as Z_I which is the gauche(-)-trans for the C5'-O5'-P-O3'-C3', phosphodiester linkage, while the less common Z_{II} is gauche(+)-trans (Fig. 1-7).

The syn-conformation of the deoxyguanosines and anti-conformation of the deoxycytidines are shown in Fig. 1-8.

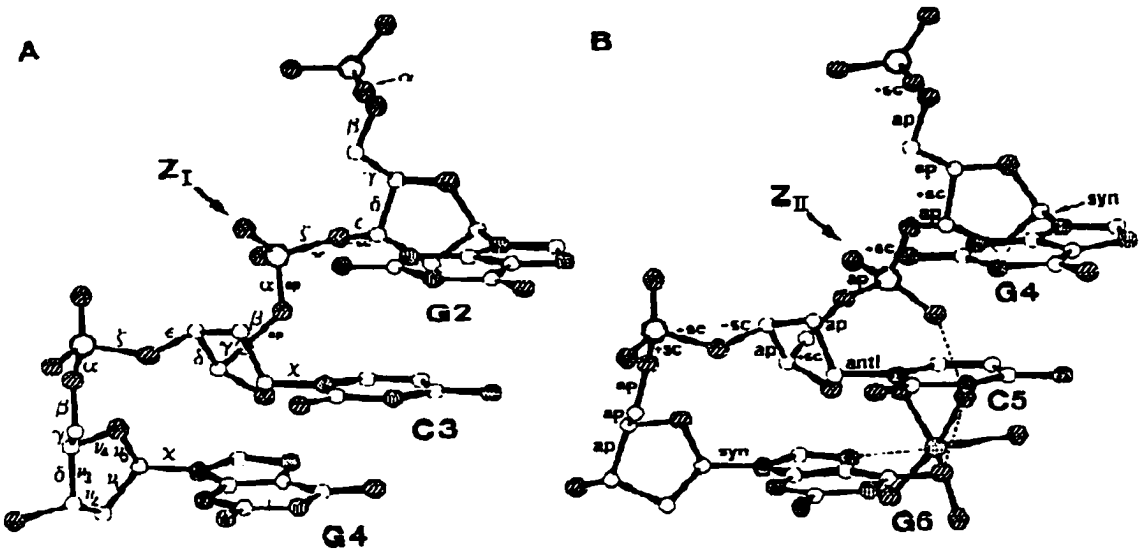
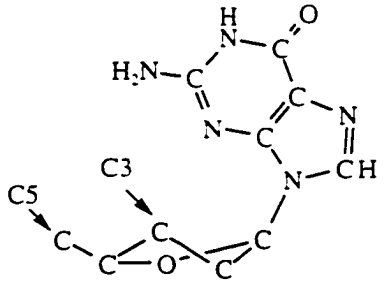
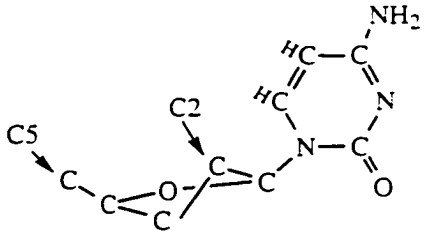


Fig. 1-7 Z_I and Z_{II} Conformations in Z-DNA Structure (From Gessner (11))



Syn-conformation of guanine
with C3 endo of sugar pucker



anti-conformation of Cytosine
with C2 endo of sugar pucker

Fig. 1-8 Base Conformations in Z-DNA

Compared with the A and B DNA, Z-DNA is slimmer and more compact. The tilt of base pairs from normal of the helix axis is larger than B-DNA but smaller than A-DNA (Table 1-1). The repeating unit in Z-DNA is not a mononucleotide as in B-DNA but a dinucleotide. The minor groove of Z-DNA is deeper than that of the B form, while the major groove has been transformed into a convex surface (Fig. 1-5).

Z-DNA has a 6-fold symmetry axis (parallel to the helix axis) (Fig. 1-9) and one dyad axis perpendicular to it and between the base pairs.

Unlike the B-DNA the base pairs of the Z-DNA are close to the outer surface of the duplex. The helix axis is beyond the inner hydrogen bonds of GC base pairs (Fig. 1-9).

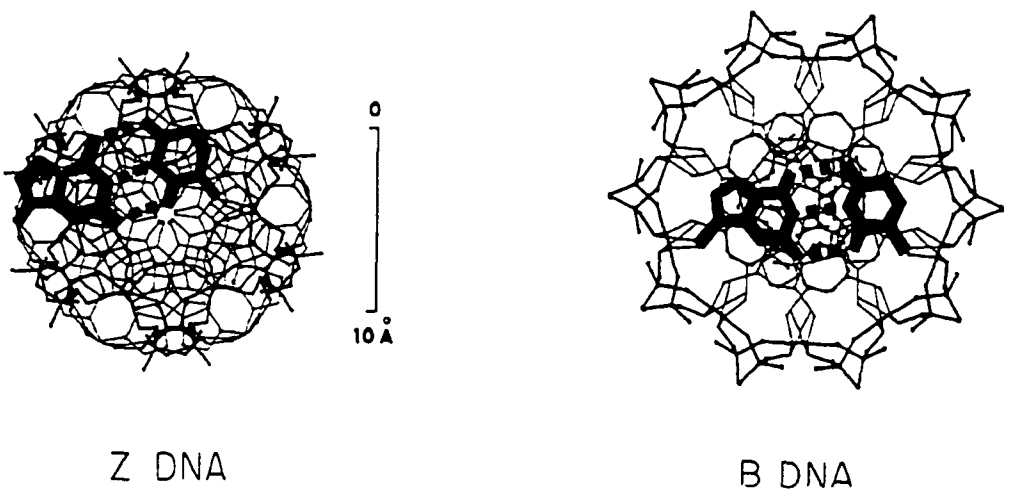


Fig. 1-9 Base pair positions in the Z-DNA and B-DNA (viewed down the helix axis)
(From Gary Quigley)
(Bold line shows the base pairs)

The first Z-DNA crystal was obtained with a sequence of the heteropolymer poly(dG-dC). Later, oligos with sequences: CGCGCGCG and CGCATGCG were crystallized in the Z-form (12). Since the distance of the closest approach of the phosphate group in the Z-DNA is 7.7 Å compared to 11.7 Å in B-DNA, a high cation concentration is needed to neutralize charge-charge repulsion. The Z conformation can be stabilized by supercoiling (13), some chemical modifications (for instance, methylation (14, 15) or bromination at C5 of cytosine (16, 17) and certain couterions such as $Co(NH_3)_6^{+3}$ (18).

1.1.6 The Significance of the Z-DNA Structure in the Biological Studies

The existence of the Z-DNA under the physiological conditions has been discussed (6), (19). The study show that in the eukaryotic DNA, the sequence (dm^5C-dG) occurs frequently (20) and the poly (dG-dm^5C)·poly(dG-dm^5C) can go into the Z-form under physiological conditions. If a poly (dG-dC) sequence is inserted into plasmid DNA, the B $\rightarrow$ Z conversion can be induced in the $(dG\text{-}dC)_n$ blocks under physiological salt condition (6). The biological function of Z-DNA is still under study.

Since Z-DNA crystals can diffract x-rays to atomic or near atomic resolution, it has been used to study the interactions of water and counterion with DNA. Z-DNA was used in monovalent-DNA interaction studies in this thesis project.

1. 2 Nucleic Acid — Metal Ion Interactions

The significance of metal ion on the function of nucleic acids has long been recognized. Metal ions are not only involved in the processes related to DNA replication, transcription, messenger RNA translation but also play important role in determining the stability of nucleic acid tertiary structures (21 - 23).

1.2.1 Function of Metal Ions in Biological Processes

Divalent metal ions are believed to be necessary activators for DNA and RNA polymerases which are enzymes in DNA replication, transcription and translation. Replication of the polynucleotide catalyzed by DNA polymerases requires a template-primer complex, four deoxynucleotide substrates and a divalent cation activator such as Mg^{2+}, Mn^{2+}, Ni^{2+} or Co^{2+} (21), (24, 25). All RNA polymerases require a divalent cation for activity (23). The role of the divalent cation on the function of DNA polymerase is to form a metal bridge between the enzyme and the terminal phosphoryl group of the substrate, facilitating the departure of the pyrophosphate leaving group (26).

In transcription, metal ions are needed for polymerase function since the reaction also has phosphate bond cleavage and formation. In translation, Mg^{2+} is normally required for *in vitro* protein synthesis. Ca^{2+}, Mn^{2+} can be used as substitutes (27).

Zn^{2+} ions are also involved in the biological processes. DNA polymerase contains two tightly bound atoms of zinc(II). Many RNA polymerases contain multiple Zn^{2+} binding positions which suggests multiple functions including purely structural one (23).

Monovalent cations are also necessary for human life. K^+ and Na^+ exist in human body in millimolar concentrations (see chapter 3). Many DNA polymerase are stimulated as much as 3- to 5- fold by monovalent cations, particularly K^+ and NH_4^+ (28). At higher concentration of monovalent cations, most DNA polymerase are inhibited (25). This may be due to shifts in conformational stability and complex stability as the monovalent cations displace divalent cations.

1.2.2 Function of Metal Ions on the Stabilizing the Structures of Nucleic Acids

Cations are required for the stabilization of nucleic acid structure since the repulsive negative charges of the phosphate groups of the helix has to be neutralized. Concentration of cations can change the conformation of the DNA, for instance, according to circular dichroism and Raman spectroscopy, higher concentration of sodium chloride and magnesium chloride can convert poly(dG-dC) from B-DNA to Z-DNA (29, 30).

For the cations which bind to the oxygen of the phosphate of the DNA, an increase of the ion concentration shifts the helix-coil transition of DNAs to higher temperature (T_m) which means the helix is stabilized. In binding to the bases, metal ions often hinder the intra-base pair hydrogen interactions, destabilizing the structure and lowering the T_m (6). For instance, increase of Mg^{2+} concentration can raise the T_m while addition of Cu^{2+} will decrease the T_m of DNA.

Many x-ray crystallographic studies have been carried out on the interactions between cations and DNA. The major concern of the studies is how these metal ions stabilize (or destabilize) the DNA structure. The questions to be answered are: what are the binding positions or binding patterns of the cations, and the effect of these interactions on the DNA structure.

The published studies involved divalent cations including Mg^{2+} (11, 18), Co^{2+}, Cu^{2+}, Ba^{2+} (31) and metal-NH_3 complex such as $Co(NH_3)_6^{3+}$ (18), $Ru(NH_3)_6^{3+}$ (32). Mg^{2+} ions are found to bind to the N or O of the bases directly or through the water bridge in the DNA (18) and to cross-link the phosphate group in tRNA (33, 34). Co^{2+} and Cu^{2+} bind exclusively to the N7 of guanines with direct coordinations. The coordination geometry is octahedral for the Co^{2+}, and trigonal bipyramid for Cu^{2+}. Both $Co(NH_3)_6^{3+}$ and $Ru(NH_3)_6^{3+}$ are found to be effective on stabilizing the Z-DNA structures due to their higher charge and specific fitting of more hydrogen bonds. $Co(NH_3)_6^{3+}$ bind to the phosphate oxygen of Cytidine 9, N7 (base nitrogen) of Guanosine 10 (G10) and O6 (carbonyl oxygen) of G10 with NH_3 as H-bonding donor. On the study for the $Ru(NH_3)_6^{3+}$ and Z-DNA structure, one $Ru(NH_3)_6^{3+}$ was found to have the same position as $Co(NH_3)_6^{3+}$. This $Ru(NH_3)_6^{3+}$ bridges the adjacent DNA strands through hydrogen binding with N7 and O6 of the guanine bases. $Ru(NH_3)_6^{3+}$ also bridges two duplex through the H-bonding between ammonia ligand and phosphate oxygen.

1.2.3 Metal Binding Sites in the Nucleic Acids

Potential metal binding sites on the nucleic acids include phosphate oxygen atoms, nitrogen on the base rings and exocyclic base keto groups. Both oxygen and nitrogen atoms are hard "bases" due to their low polarizability. Since hard bases tend to bind with hard acids, the preferential binding sites of cations depends on their polarizability or "handness".

Teeter and Quigley *et al* classified DNA binding metal cations into three groups (35). The first group consists of hard metals including alkali, alkaline earth and rare earth metals. These metals have high electron density and bind more strongly with the phosphates than with the bases. The second group consists of soft metals such as Pt(II), Os(VI), Os(VIII), and Hg(II). These metals have polarizable electron shell and often bind to the bases of the nucleic acid. The third group includes the first row transition metals. These metals are effective in neutralizing charge at low concentrations and can bind both phosphate and base. At high concentration they destabilize the helix conformation probably due to the direct binding to the bases whose electronegative sites are buried in the duplex.

Polyamines such as spermine, spermidine and putrescine can be classified into another nucleic acid binding group. The polyamines occur naturally in the cell of many organisms and have been implicated in protein synthesis. Spermine binds primarily to the phosphate groups, but may also bind to the bases.

1. 3 Interactions Between DNA and Intercalating "Drugs"

The intercalating anti-cancer drugs belong to one of the drug groups which has been most intensively studied and visualized at the molecular level. With the polycyclic chromophore in their structure, drugs in this group are known or postulated to intercalate between the consecutive base pairs of DNA duplex with their planar polycyclic aromatic ring (36).

The earliest crystal structure solved for an intercalating drug and DNA complex is the structure of the 9-Aminoacridine complexing with the dinucleotide: deoxyguanylyl-deoxyiodocytidine (37). Later, other structures including daunomycin (38) (Fig. 1-10), (39 - 41), triostain A (42, 43), adriamycin (44), nogolamacin (45, 46), ditercalinium (47) complexed with DNA have also been reported from crystallographic studies.

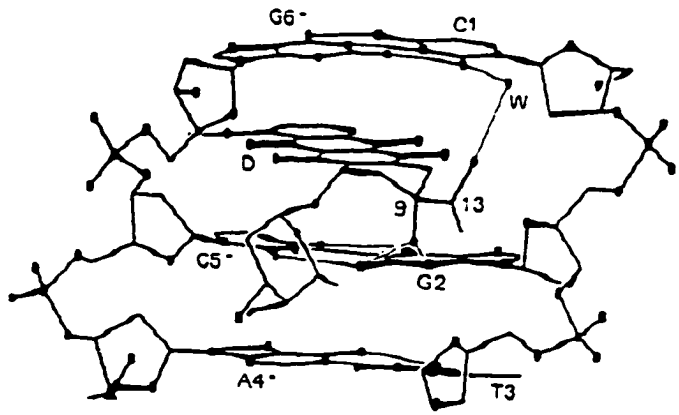


Fig. 1-10 Diagram of Daunomycin (D) Intercalated into d(CGTACG)
The thin lines show the hydrogen bonding (From Quigley (38))

The investigations of the mechanism of the anti-cancer activity suggested that the binding of intercalators with DNA induces unwinding, lengthening and stiffening of the DNA helices, thus introduces changes which affect both the condensation of chromatin (48, 49) and its interaction with DNA associated enzymes (36). The second postulated mechanism is that since DNA topoisomerase II is a prime target for some intercalating drugs, formation of ternary complex including DNA, enzyme and drug impairs the function of topoisomerase II and so to the division process (36). A crystal structure of complex of DNA and AHMA (3-(9-acridinylamino)-5-(hydroxymethyl) aniline, a topoisomerase II inhibitor) has been studied in this work.

References

1. F. Miescher, *Med. Chem. Unt.* **4**, 441-460 (1871).
2. I. Rosenfield, E. Ziff, *DNA for Beginners* (Writers and Readers Publishing, Inc, 1983).
3. E. Chargaff, *Experientia* **6**, 201-209 (1950).
4. J. D. Watson, F. H. Crick, *Nature* **171**, 737-738 (1953).
5. A. Rich, G. J. Quigley, A. H. J. Wang, *Biomolecular Stereodynamics* (Adenine Press, Guilderland, N.Y., USA, 1981), vol. I and II, pp. 35-52.
6. W. Saenger, *Principles of Nucleic Acid Structure* (Springer-Verlag New York Inc., Berlin, 1984).
7. L. Stryer, *Biochemistry* (W.H. Freeman and Company, New York. Fourth ed., 1995).
8. D. Rawn, *Biochemistry* (Neil Patterson Publishers, 1989).
9. A. H. J. Wang, G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel, A. Rich, *Nature* **282**, 680-686 (1979).
10. A. H. J. Wang, G. J. Quigley, F. J. Kolpak, G. van der Marel, A. Rich, J. H. van Boom, *Science* **211**, 171-176 (1981).
11. R. V. Gessner, C. A. Frederick, G. J. Quigley, A. Rich, A. H. Wang, *J Biol Chem* **264**, 7921-35 (1989).
12. S. Fujii, A. H. J. Wang, G. J. Quigley, H. Westerink, G. van der Marel, J. H. van Boom, A. Rich, *Biopolymers* **24**, 243-250 (1985).
13. A. Rich, A. Wang, A. Nordhiem, *Nucleic Acid Research* , 11-35 (1983).

14. S. Fujii, A. H. J. Wang, G. van der Marel, J. H. van Boom, A. Rich, *Nuc. Acids Res.* **10**, 7879-7892 (1982).

15. A. H. J. Wang, R. V. Gessner, G. A. van der Marel, J. H. van Boom, A. Rich, *P.N.A.S.* **82**, 3611-3615 (1985).

16. E. M. Lafer, A. Moller, A. Nordheim, B. D. Stollar, A. Rich, *Proc. Natl. Acad. Sci. USA* **78**, 3546-3550 (1981).

17. A. Moller, A. Nordheim, S. A. Kozlowski, D. Patel, A. Rich, *Biochemistry* **23**, 54-62 (1984).

18. R. V. Gessner, G. J. Quigley, A. H. J. Wang, G. A. van der Marel, J. H. van Boom, A. Rich, *Biochemistry* **24**, 237-240 (1985).

19. A. Rich, A. Nordheim, A. Wang, *Annual Review of Biochemistry* **53**, 791-846 (1984).

20. A. Razin, A. D. Riggs, *Scinece* **210**, 604-610 (1980).

21. M. A. Sirover, L. A. Loeb, *Biochemical and Biophysical Research Communications* **70**, 812 (1976).

22. T. Spiro, *Nucleic Acid-Metal Ion Interaction* (John Wiley & Sons, Inc, 1980).

23. A. S. Mildvan, L. A. Loeb, *CRC Critical Review in Biochemistry* **6**, 219-244 (1979).

24. T. Kornberg, A. Kornberg, *The Enzymes* **10**, 173 (1974).

25. L. A. Loeb, *The Enzymes* **10**, 173 (1974).

26. D. L. Sloan, L. A. Loeb, A. S. Mildvan, R. J. Feldmann, *J. Biol. Chem.* **250**, 8913 (1975).

27. J. P. Slater, I. Tamir, L. A. Loeb, A. S. Mildvan, *J. Biol. Chem.* **247**, 2784 (1972).

28. L. A. Leob, *Pharm, Ther. A* **2**, 117 (1977).

29. F. M. Pohl, T. M. Jovin, *J. Mol. Biol.* **67**, 375-396 (1972).

30. T. J. Thamann, R. C. Lord, A. Wang, A. Rich, *Nucleic Acids Res.* **9**, 5443-5457 (1981).

31. Y.-G. Gao, M. Sriram, A. Wang, *Nucleic Acids Res* **21**, 4093-4101 (1993).

32. P. S. Ho, C. A. Frederick, D. Saal, A. H. J. Wang, A. Rich, *J. Biomol. Struct. Dyn.* **4**, 521-534 (1987).

33. S. R. Holbrook, J. L. Sussman, R. W. Warrant, G. M. Church, S. H. Kim, *Nucl. Acids Res.* **4**, 2811-2820 (1977).

34. G. J. Quigley, M. M. Teeter, A. Rich, *Proc. Nat. Acad. Sci. USA* **75**, 64-68 (1978).

35. M. M. Teeter, G. J. Quigley, A. Rich, *Nucleic Acid-Metal Ion Interactions* (John Wiley and Sons, New York, 1980) pp. 145-177.

36. B. C. Baguley, *Anticancer Drug Des* **6**, 1-35 (1991).

37. T. D. Sakore, B. S. Reddy, H. M. Sobell, *J. Mol. Biol.* **135**, 763-786 (1979).

38. G. J. Quigley, A. H. J. Wang, G. Ughetto, G. van der Marel, J. H. van Boom, A. Rich, *P.N.A.S.* **77**, 7204-7208 (1980).

39. A. H. Wang, G. Ughetto, G. J. Quigley, A. Rich, *Biochemistry* **26**, 1152-63 (1987).

40. M. H. Moore, W. N. Hunter, B. L. d'Estaintot, O. Kennard, *J Mol Biol* **206**, 693-705 (1989).

41. C. M. Nunn, L. Van Meervelt, S. D. Zhang, M. H. Moore, O. Kennard, *J Mol Biol* **222**, 167-77 (1991).

42. G. J. Quigley, G. Ughetto, G. A. van der Marel, J. H. van Boom, A. H. J. Wang, A. Rich, *Science* **232**, 1255-1258 (1986).

43. A. H. J. Wang, G. Ughetto, G. J. Quigley, A. Rich, *J. Biomol. Struct. Dyn.* **4**, 319-342 (1986).

44. M. Cirilli, F. Bachechi, G. Ughetto, F. P. Colonna, M. L. Capobianco, *J Mol Biol* **230**, 878-89 (1993).

45. L. D. Williams, M. Egli, G. Qi, P. Bash, G. A. van der Marel, J. H. van Boom, A. Rich, C. A. Frederick, *Proc Natl Acad Sci U S A* **87**, 2225-9 (1990).

46. M. Egli, L. D. Williams, C. A. Frederick, A. Rich, *Biochemistry* **30**, 1364-72 (1991).

47. Q. Gao, L. D. Williams, M. Egli, D. Rabinovich, S. L. Chen, G. J. Quigley, A. Rich, *Proc Natl Acad Sci U S A* **88**, 2422-2426 (1991).

48. J. Markovits, Y. Pommier, M. R. Mattern, C. Esnault, B. P. Roques, J. B. Le Pecq, K. W. Kohn, *Cancer Res* **46**, 5821-6 (1986).

49. L. Larue, M. Quesne, J. Paoletti, *Biochem Pharmacol* **36**, 3563-9 (1987).

Chapter 2

X-ray Crystallographic Method and Experiment

2.1 Study of Macromolecular Structure by X-ray Crystallography

X-ray crystallographic methods were first used on molecular structure studies by W. L. Bragg in 1913 when he showed that in sodium chloride crystals, each sodium is surrounded by six chlorine atoms and each chlorine by six sodium. Later with the same method Kathleen Lonsdale found that the benzene ring is a regular hexagon instead of having alternating single and double bond in its structure. Due to its advantages in the determination of molecular structures, X-ray diffraction has become a popular technology in molecular structure analysis and many advanced techniques continue to be developed (1, 2). The principles of x-ray crystallography and its application to molecular structure determination has been elucidated by Glusker (3), Ladd (4), and Blundell (5) among others.

2.1.1 Basic Principles of X-ray Crystallography

X-ray Diffraction from Crystals

Crystals are made of a three dimensional array of regularly arranged atoms. In order for interference of the diffracted beams to produce large changes in the amount of scattering in different directions, the wavelength of the radiation should be close to the distance between atoms. X-rays produced by bombarding elements such as Cu and Mo fall into this range. In the x-ray crystallographic studies, the radiation most often used are Cu Kα (λ = 1.54 Å) and Mo Kα (λ = 0.71 Å).

In 1913, W.L. Bragg showed his famous equation :

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

where, λ is the wavelength of the radiation, d is the spacing between the lattice plane in the crystal, θ is the complement of the angle of the incidence X-ray beam, and n is a whole number.

This equation relates the reflection beam angle with lattice plane distance and x-ray wavelength. It tells us that the diffracted beams combine only if the path difference between the scattering beams (d $\sin\theta$) is the whole number multiple of the wavelength.

When an x-ray beam is reflected from a crystal, the direction of the diffracted beam depends on the repeating distance of the pattern of the atomic arrangement and can be used to measure the size and the shape of the unit cell. The intensities of the reflected beams are proportional to the arrangement of atoms in the unit cell - the smallest translationally repeating motif in the crystal.

X-ray Sources

In the x-ray source, electrons produced at a cathode are accelerated in a potential field and bombard the metal target (Cu, for instance). Some electrons in the orbital near the metal nucleus are removed by the bombardment. Then the electrons in the outer shell transit into the vacated orbital and emit a characteristic x-radiation. K series radiation is formed by electron transition from outer shell to K shell. Cu Kα radiation with wavelength of 1.54 Å was used for my thesis work.

2.1.2 Phase Problem

To be able to solve the three dimensional structure of a molecule the position of each atom in the unit cell has to be determined. The image of the scattering matter in the point x, y, z can be represented by a fourier summation :

$$\rho(x,y,z) = \frac{1}{V_c} \sum_{\text{all hkl}} \sum \sum F(hkl) \exp [-2\pi i(hx+ky+lz)]$$

where, V_c - volume of the unit cell

F(hkl) - structure factor for the particular set of h,k,l. (h,k,l are the Miller indices with which the crystal plane or face make intercepts a/h, b/k, c/l with the edge of the unit cell of lengths a, b, and c), x,y,z are the fractional coordinates which are the atomic coordinates expressed as the fractions of the unit cell length.

and

$$F(hkl) = |F(hkl)| e^{i \alpha (hkl)}$$

Where, $|F(hkl)|$ is the amplitude of the structure factor, $|F(hkl)|^2$ is proportional to the intensity of the reflections. α is the phase of the scattered beam.

We can measure the amplitude |F| which is proportional to the square root of the intensity of the reflections, but the direct measurement of the

relative phase is not possible. So the trick of crystallography lies in the finding of the phase.

There are four methods that have been used to solve the phase problem.

1. Direct Methods

When the number of structure amplitudes measured is more then ten times the number of the parameters to be determined, direct methods, which utilize the sign and phase relations, can be used to solve the structure. One such relation applicable to centrosymmetric system, where phase can only be 0° or 180°, is that the sign of F(hkl) is probably equal to the product of sign of F(h'k'l') and F(h-h',k-k',l-l'). Possible phases can be derived from systematically searching for sets of reflections (with high intensities) whose indices are related in this way. An electron density map can then be calculated. Peaks are assigned for the atoms in the structure to get the trial structure. From the trial structure the new phases are calculated. The structure can be solved by repeating this cycle. This method is mainly used for small molecules with up to roughly 50 non-hydrogen atoms. In my thesis work, direct methods was used for solving the drug molecule (AHMA) structure and searching for the heavy metal positions in the AHMA-DNA heavy metal derivative crystals.

2. The Patterson Method

The Patterson method consists of evaluating Patterson function which can be represented with the following equation :

$$P(u,v,w) = \frac{1}{V_c} \sum_{\text{all h.k.l}} \sum \sum |F|^2 \cos 2\pi (hu+kv+lw)$$

In the equation, V_c is the unit cell volume, |F| is the amplitude of the structure factor, h, k, l are the Miller indices and u, v, w are the fractional coordinates of the unit cell in three dimensions.

It can be shown that the Patterson is also given by :

$$P(u,v,w) = \int \int \int \rho(x+u,y+v,z+w)\ \rho(x,y,z)\ dx\ dy\ dz$$

x, y, z and u,v,w are two independent sets of fractional coordinates within the unit cell.

That is at each point u, v, w, the value for the Patterson function can be obtained by multiplying the electron density at point x, y, z with that at x+u, y+v, z+w and adding the resulting products for all value of x,y,z. If two atoms are separated by a vector (u,v,w), then there will be a peak in the Patterson map at (u,v,w).

The difficulty with the Patterson method is that a unit cell with N independent atoms will have N^2 inter-atomic vectors (peaks in the Patterson map). When N is large as for a macromolecule, the peak overlap make it difficult to solve the structure directly using this method. The molecular

replacement below is closely related. Patterson map was calculated to get information of the AHMA-DNA structure in my work.

3. Molecular Replacement Method

The molecular replacement and the heavy atom methods are two frequently used methods for macromolecular structure determination.

Molecular replacement is the first choice if one can find a molecule suspected of having the same or similar structure as the unknown structure. With the molecular replacement method the phase from the known structure and reflection intensity from unknown are used for calculating the electron density map, difference electron density map and 2Fo-Fc maps.

The equations for these three maps are as following :

Electron density map :

$$\rho(x,y,z) = \frac{1}{V} \sum_{\text{all } h,\ k,\ l} \left(\sum \left(\sum |F_o(h,k,l)| e^{i \alpha_c} e^{-2\pi i (hx+ky+lz)} \right) \right)$$

Difference density map :

$$\Delta\rho(x,y,z) = \frac{1}{V} \sum_{\text{all } h,\ k,\ l} \left(\sum \left(\sum \left(|F_o(h,k,l)| - |F_c(h,k,l)| \right) \left(e^{i \alpha_c} e^{-2\pi i(hx+ky+lz)} \right) \right) \right)$$

2Fo-Fc map :

$$\rho(x,y,z) = \frac{1}{V}\sum_{\text{all } h,k,l}\left(\sum \left(\sum \left(2\,|F_o(h,k,l)| - |F_c(h,k,l)| \right) \left(e^{i\,\alpha_c}\, e^{-2\pi i(hx+ky+lz)} \right) \right) \right)$$

In the above equations, F_o is the observed structure factor and F_c is the calculated structure factor.

The difference map and 2Fo-Fc map can be used to correct the atomic position. With the new trial structure the phases can be recalculated. Map and refinement procedures can be repeated until the best fit between the trial and real structure is obtained.

In order to use the known model, its position in the unit cell of the unknown structure must be established. In cases where this is not trivial or obvious, this is done by calculating the Patterson maps of 1): the isolated known structure, and 2): the unknown from its intensity ($|F|^2$) data. In AHMA-DNA complex structure project, computerized procedures known as rotation search and translation search were carried out. Rotation search is bases on the fact that the Patterson of the unknown, near the origin, is primarily composed of the sum of Pattersons of the individual molecules appropriately rotated. From this, the appropriate rotation is selected. Translation search is based on the fact that the Patterson of more remote portions and closer portions which are not accounted for by the vectors within the molecules are due to the Patterson vectors between molecules.

4. Heavy Atom Replacement Method

Heavy atom method with isomorphous replacement is another important method for solving macromolecular structures. The heavy atom method is based on the phenomena that if one atom with several times as many electrons as other atoms is present in the molecule, the magnitudes and phases of many of the reflections of the crystal are dominated by it. The isomorphous replacement method consists of replacing an atom in the structure with an other one with significantly greater scattering power. Assuming the only change is due to the replacement, it is possible to use the change in diffracted intensity to first identify the location of the heavy atom and subsequently to gain information on the phase of the remaining structure.

A heavy atom can be soaked or crystallized into the crystal lattice. The "derivative crystal" must be isomorphous with the native one, i.e. all other atoms should remain in essentially the same position as the original crystal. The position of the heavy atom can be found with the use of a difference Patterson map or by applying direct methods to the difference data ($|F_{Derivative}| - |F_{Native}|$). As illustrated by Fig. 2-1, two possible phase angles can be found from the two intersection points made by two circles : one is a circle with radius $|F_N|$ (the amplitude of the native data), the other is drawn at a center derived from the position of heavy atom and with a radius of $|F_D|$ (the amplitude of the derivative data). If two heavy atom derivatives are used, an approximation of the real phase can be derived by the common intersections of two derivatives with the native structures.

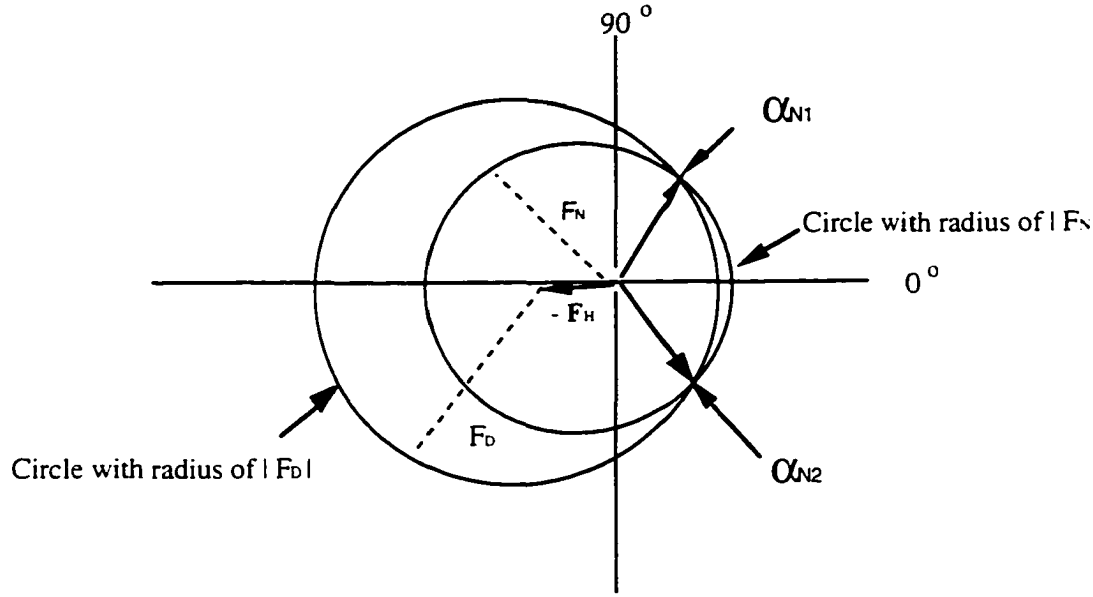


Fig. 2-1 Heavy Atom Derivative Method for Searching Phase

| F_D | - the amplitude of the structure factor of derivative
| F_N | - the amplitude of the structure factor of native
F_H - structure factor of heavy atom

2.1.3 Structure Refinement

Refinement with Electron Density Map and Potential Energy of the Structure

After the trial structure is obtained, refinement is normally conducted to get a better structure. There are two methods of structure refinement. The first one, which has been mentioned above, is to repetitively calculate the difference map and make correction based on the peaks in the map. The second one is the least square refinement which minimize the difference between observed structure amplitude and calculated one. The minimized value can be represented as :

$$Q = \Sigma W (hkl) (|F_o(hkl)| - |F_c(hkl)|)^2$$

Where w, the weight, is commonly the reciprocal of the square of the standard deviation of the experimental value of $|F_o|$. $|F_o(hkl)|$ is the observed structure factor amplitude and $|F_c(hkl)|$ is the calculated structure factor amplitude.

In some crystallographic programs, another term is involved in the structure refinement. This term minimizes the potential energy of the structure. Three methods were used in this work : positional, individual B and simulated annealing. A refinement and computer modeling program : X-plor (6) was used for all the macromolecular refinement. The target energy for minimization is :

$$E = E_{structure} + E_{xray}$$

where, $E_{structure}$ include bond length, bond angle, dihedral angle, nonbonding (including both electrostatic and van der Waals) energy.

and $$E_{xray} = W_x \Sigma (w|F_o| - |F_c|)^2$$

In the equation, W_X is the overall weight assigned for x-ray data and w is the weight for individual refinement. F_o, F_c are observed and calculated structure factors respectively.

For the small molecule (AHMA), least square method with Molen program (7) was employed for the refinement.

Conventional Residual R

A general method for evaluating the correctness of the trial structure is to calculate the discrepancy index or conventional residual R, defined as :

$$R = \frac{\sum |(|F_o| - |F_c|)|}{\sum |F_o|}$$

The R value is a measure of the precision of the trial model to the experimental data but not the accuracy of the structure.

For macromolecules, the acceptable R value is about 20%. Since small molecule can diffract x-ray to higher resolution and there is less disorder in their structure the general acceptable R for small molecule is below 10%.

2.2 Crystallization of DNA Oligonucleotide and DNA – Drug Complex

2.2.1 General Principles of Crystallization

Crystallization is a process in which molecules are going from a liquid solution into a regular three dimensional array solid.

The driving force for the crystallization is the difference between the chemical potential of a molecule in the supersaturated solution and that in saturated solution. It can be expressed in an equation as (8):

$$\Delta \mu = k_B T \ln (C/Cs)$$

Where, k_B - Boltzmann constant
T - absolute temperature
C - concentration before any crystallization
Cs - the solubility of the molecule

The phase diagram for crystallization is shown in Fig. 2-2 (9).

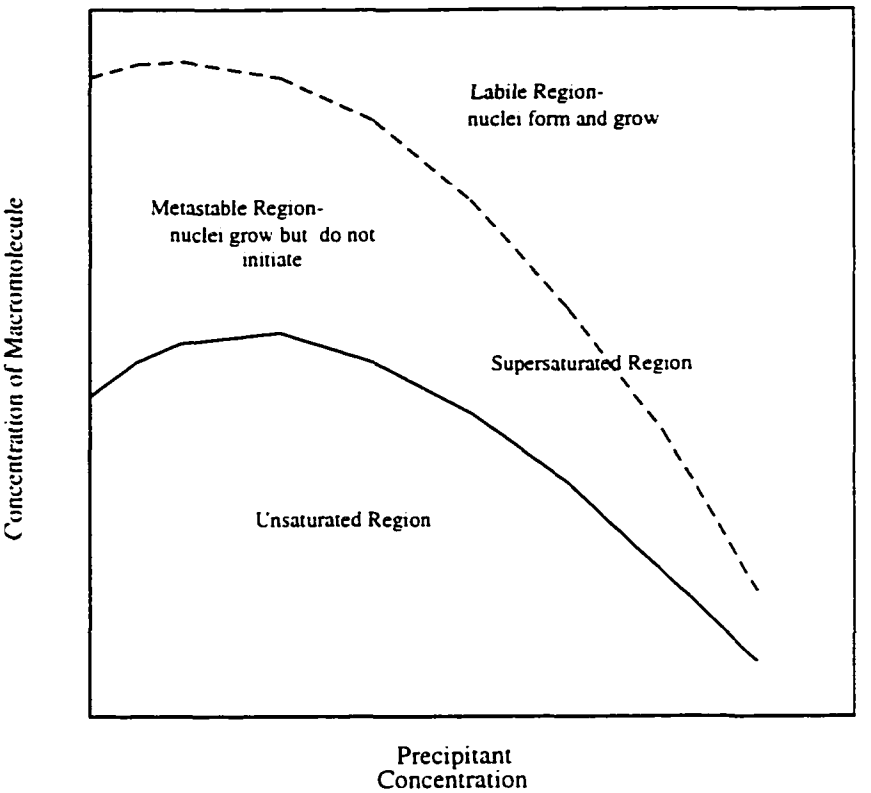


Fig. 2-2 A Phase Diagram for Crystallization (after Mcpherson (9))

Crystallization process can be divided into three steps : nucleation, growth and cessation of the crystallization.

In nucleation, a new phase originates from a supersaturated solution when the free energy of the system is higher than activation free energy of germination ΔG_g (10).

ΔG_g is composed by two terms as following :

$$\Delta G_g = [\ -kT(4\pi r^3)V\ln\beta\ ] + 4\pi\gamma r^2$$

In the equation, k - Boltmann constant, T - absolute temperature,
β - supersaturation, γ - interfacial free energy
r - radius of nucleus, V - volume of a molecule inside the crystal.

The generation of nuclei will decrease the level of supersaturation. In the metastable region of supersaturation, molecules continue to incorporate on the surface of nuclei. At certain stage crystals stop growing. The cessation of crystal growth can be due to the depletion of the solute, poisoning of the crystal surface or growth defects.

The ideal situation is to slowly bring the solution to a supersaturation that just reaches the labile region. Then the generation of several nuclei will lower the saturation to the metastable stage. The initial crystal grows until cessation.

Macromolecule crystallization is a multiparameter process. The main factors involved in the crystal growth include : purity of the macromolecule, conformational heterogeneities, supersaturation, temperature, pH, ionic strength and viscosity of the solution (11-13).

For crystallization of DNA, the most commonly used precipitating agents are MPD and PEG. Since nucleic acids are polyelectrolytes, counter-ions such as polyamines and divalent cations are important additives of the crystallization medium.

2.2.2 Screening of the Crystallization Conditions

To be able to choose an optimized condition for growing good single crystal, effect of the variation of each parameters on the growth of the crystal needs to be investigated.

The parameters which affect the crystallization in this work include :

Concentration of oligonucleotide

Molar ratio of drug to DNA (for drug - DNA complex)

Salt concentration ($MgCl_2$ for drug-DNA complex, RbCl, KCl for monovalent cations and Z-DNA)

Concentration of precipitant in the mother liquor droplet and in equilibration reservoir

pH and Temperature

The Grid Search method (14, 15) was employed for screening the condition in this work. The idea of grid search method is schematically shown in Fig. 2-3.

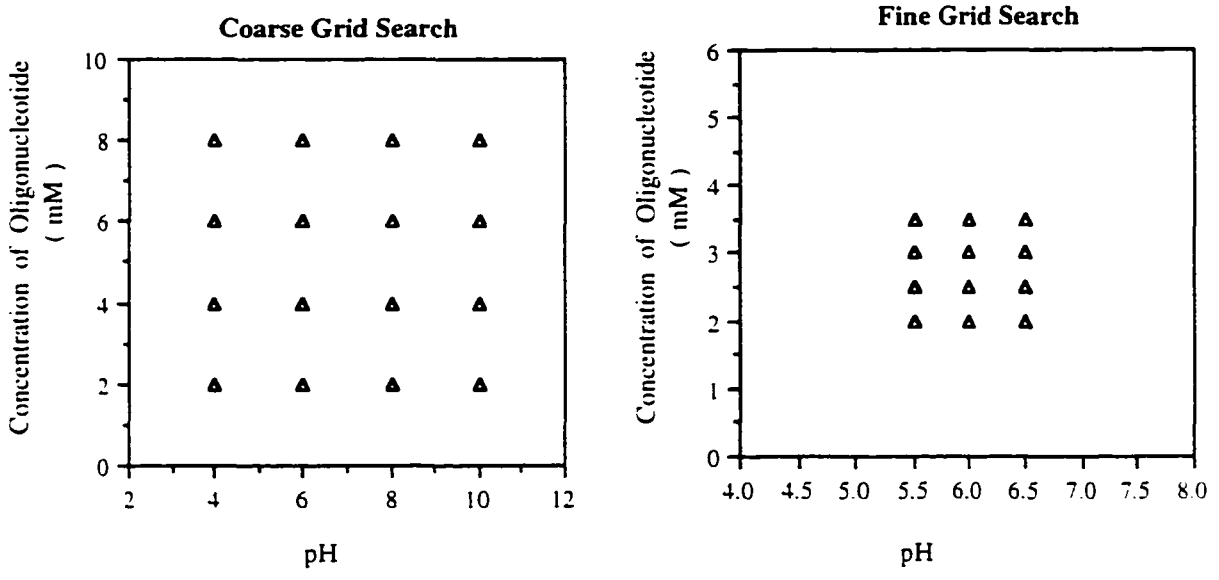


Fig. 2-3 Grid Search Method for the Crystallization Condition Screen
(Each triangle represents a trial condition)

The initial screen has larger grid interval. After the first crystal appear, finer grid screen was conducted to optimize the condition to obtain suitable size of crystals.

2. 3 Cryocrystallography

It is usually assumed that x-ray radiation produces radicals in the liquid regions of the crystal which causes specific chemical changes on the surface of molecules and results in disorder of the crystal lattice (16).

crystals of biological macromolecule suffer from disorder of some fragment, being sensitive to x-ray irradiation and rapid decrease in diffracted intensity with increasing $\sin\theta/\lambda$ due to the high concentration of solvent in the crystals and weak interactions between molecules (17, 18).

Disorder results from fast movement of a fragment between more than one positions. An approximate equation for transition time (τ_R) of a atom or a fragment between two positions which are separated by a molar potential energy barrier H_0 is given as (19):

$$\tau_R = \tau_0 \exp (H_0 / RT)$$

Where,
- R - molar gas constant (= 8.314 J/K·mol)
- T - absolute temperature
- H_0 - potential energy barrier between two transition positions

Lower temperature will increase the transition time τ_R. If the τ_R is larger than the X-ray data collection time, the two positions can be distinguished.

The primary benefit of cryocrystallography is to reduce the rate of radiation damage which often substantially extends crystal lifetime. Since

the data is collected under cryogenic temperature, thermal disorder is also reduced. In many cases, higher resolution data can be recorded. Cryocrystallographic investigation of macromolecular structure shows that at low temperature, the mean-square displacement of atom positions has a large decrease from that at room temperature, and more water molecule and interstitial ion can be found in the structure. Because of these advantages cryogenic data collection is getting more and more popular for the macromolecular crystallography (20-22).

Crystal mounting

The recent commonly used tool for the crystal mounting in cryocrystallography is a small fiber loop made by rayon or glass. The crystal can be picked up from crystallization solution with the loop due to surface tension and flash cooled to a desired temperature. Fig. 2-4 shows the fiber loop mounted on goniometer head with magnetic base which was used for Rb-Z-DNA low temperature data collection. Compared with capillary tube mounting, fiber loop mounting can facilitate faster and isotropic crystal cooling.

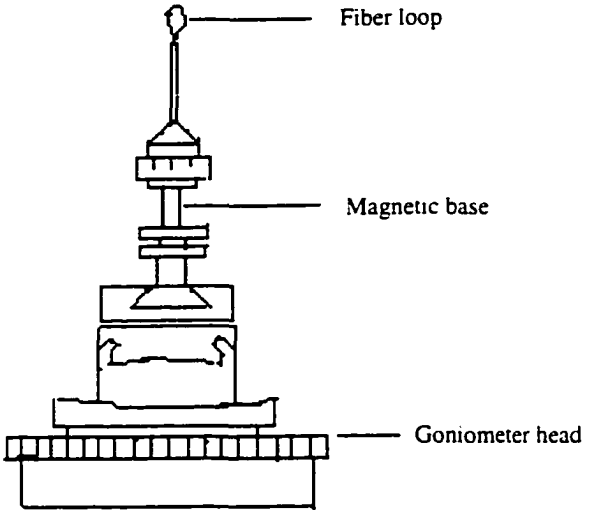


Fig. 2-4 Fiber Loop and Magnetic Base for Crystal Mounting

Two techniques have been used to prevent ice formation in crystal lattice in cryocrystallogaphy.

1. Flash - cooling

Flash - cooling (23, 24) is a technique which rapidly lowers the temperature so that ice nucleation dose not occur before the viscosity of the solution is high enough to make the water molecules immobile. Vitreous water is formed in flash cooling.

2. Cryoprotectant

Using cryoprotectant to substitute for the mother liquor is another method to prevent ice formation. Cryoprotectant can raise the viscosity of the solution, so less time is available for ice formation. Cryoprotectants form a rigid glass encasing the crystal in flash cooling. Suitable cryoprotectant should have properties of glass-formation and not dissolving the crystal or denaturing the molecule. Five commonly used cryoprotectants are glycerol, ethylene glycol, sucrose, 2-methyl-2,4-pentanediol (2-MPD) and polyethylene glycol 400.

2.4 Materials and Methods

DNA Synthesis

The oligonucleotide (with sequence CGCGCG or CGTACG) were synthesized by the RCMI SSS (Research Center in Minority Institutions Program, Synthesis, Sequencing and Separation Laboratory) facility at Hunter College using the phosphotriester method with a DNA synthesizer (Applied Biosystem). With the solid phase synthesis, DNA is built in the 3' to 5' direction. The schematic drawing of the general structure of oligomer after the completion of the coupling is shown in Fig. 2-5.

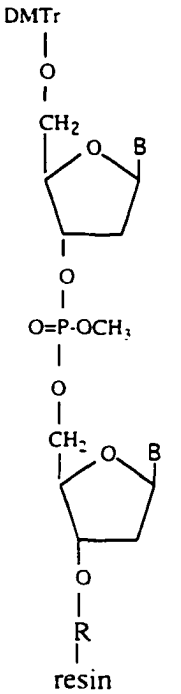


Fig. 2-5 Schematic Drawing of Structure of Oligomer after Synthesis

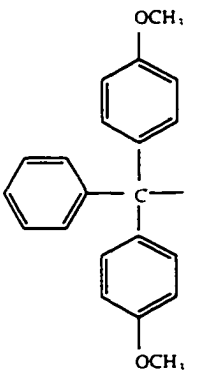


Fig. 2-6 DMTr Group

At this stage, the 5' -OH group is protected by a dimethoxytrityl (DMTr or trityl) group (Fig. 2-6). 3' phosphorous is protected by a methyl group and bases are protected by benzoyl group on the exocyclic amines of Adenine and Cytosine, an isobutyryl protecting group on the exocyclic amine of Guanine. Before the oligo is cleaved from the support, the methyl group on the phosphate is removed by the thiophenol, then the cleavage is done with ammonium hydroxide.

After the synthesis, the oligo samples were incubated with concentrated ammonia in a 55°C water bath for about 18 hours to remove the base protecting groups. The detritylation was done with 0.1% trifluoroacetic acid (TFA) in the process of the DNA purification by HPLC.

Oligonucleotide Purification and Desalting

The trityl-on oligonucleotide was purified with HPLC (Beckman) on a PureDNA™ column (Rainin, 21.4mm ID × 50mm L).

HPLC profile for oligo CGTACG semipreparative purification is shown in Fig. 2-7.

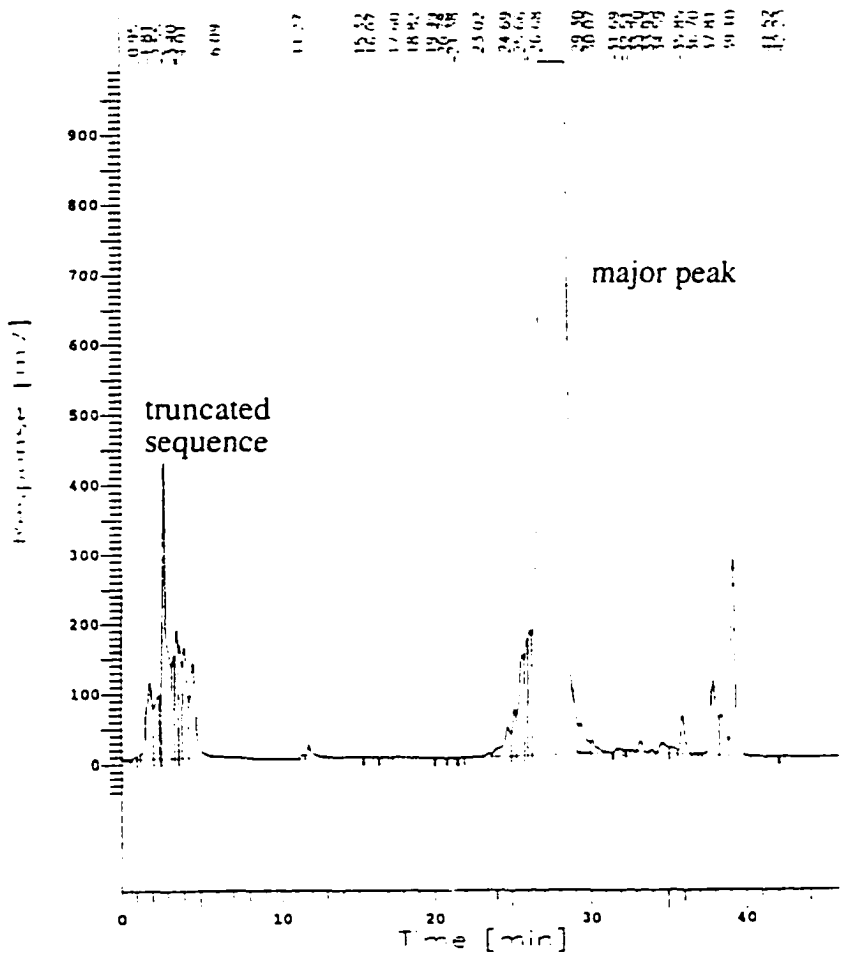


Fig. 2-7 HPLC Profile of Oligomer (CGTACG) Semipreparative Purification

The major peak from the purification was collected. TEAA salt from the HPLC mobile phase was removed by three repeated lyophilization. After purification, the purity of the oligomer was checked by an analytical column. The purity for all the samples is higher than 98%.

Crystallization

The vapor diffusion method was used for the Rb-Z-DNA and DNA-AHMA complex crystallization. In the experiments a droplet containing DNA oligomer, buffer, additive and precipitating agent is equilibrated against a reservoir containing a solution with concentration of precipitating agent higher than the droplet. Linbro boxes with sitting drops (volume from 4 ul to 20 ul) were used for crystallization.

The following x-ray crystallographic programs have been used in the work:

1. CAD4 - program for CDA4 diffractometer operation and the data collection (25).
2. Xips - program for data collection with MarResearch image plate (26).
3. X-plor - program for structure refinement and electron density map calculation (6).
4. CCP4 - package with various utility for crystallography, mainly used for seeking the heavy metal position in AHMA-DAN complex structure (27).

5. Molen - program for solving small molecule structure (7).
6. Shelxs - program for structure determination and refinement (28).
7. Denzo (29) and scalepack (30) - program for data process and reduction. Mainly Used for data collected with area detector.
8. Chain - program used for structure and electron density map display and structure manipulation (31).

References

1. L. Sir Bragg, *Sci. Amer.* **219**, 58-70 (1968).
2. M. M. Harding, *Chem. Brit.* **4**, 548-553 (1968).
3. J. P. Glusker, K. N. Trueblood, *Crystal Structure Analysis: a primer, (2nd edn.* 1985).
4. M. F. C. Ladd, R. A. Palmer, *Structure Determination by X-ray Crystallography* (1985).
5. T. L. Blundell, L. N. Johnson, *Protein Crystallography* (Academic press, 1976).
6. A. T. Brunger, *X-PLOR - A System for Crystallography and NMR* (The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, 1992).
7. Enraf-Nonius, *Molen-Structure Determination System* (1990).
8. R. Boistelle, J. P. Astier, *J. of Crystal Growth* **90**, 14-30 (1988).
9. A. McPHERSON, *Eur. J. Biochem.* **189**, 1-23 (1990).
10. A. Ducruix, R. Giege, *Crystallization of Nucleic Acids and Proteins* (Oxford University Press, New York, 1992).
11. A. McPHERSON, *Methods in Enzymology* (1985), vol. 114.
12. A. H. J. Wang, M. K. Teng, *J. Cryst. Growth* **90**, 295-310 (1988).
13. A. K. Aggarwal, *Methods:A companion to methods in enzymology* (1990), vol. 1.
14. M. J. CoX, P. C. Weber, *J. of Crystal Growth* **90**, 318-324 (1988).
15. P. C. Weber, *Methods: A Companion to Methods in Enzymology* **1**, 31-37 (1990).
16. D. J. Haas, M. Rossmann, *Acta Cryst.* **B26**, 998-1040 (1970).

17. G. Petsko, *J. Mol. Biol.* **96**, 381-392 (1975).

18. H. Hope, *Acta Cryst.* **B44**, 22-26 (1988).

19. H. Hartmann, F. Parak, W. Steigemann, G. A. Petsko, D. Ringe Ponzi, H. Frauenfelder, *Proc. Natl. Acad. Sci. USA* **79**, 4967-4971 (1982).

20. H. Hope, F. Frolow, K. von Bohlen, I. Makowski, C. Kratky, Y. Halfon, H. Danz, P. Webster, K. S. Bartels, H. G. Wittmann, A. Yonath, *Acta Cryst.* **B45**, 190-199 (1989).

21. W. Watt, A. Tulinsky, R. P. Swenson, K. D. Watenpaugh, *J. Mol. Biol.* **218**, 195-208 (1991).

22. T. Earnest, *Proteins: Structure, Function and Genetic* **10**, 171-187 (1991).

23. R. Sutton, *J. Chem. Soc. Faraday Trans* **87**, 101-105 (1991).

24. D. W. Rodgers, *Structure* **2**, 1135-1140 (1994).

25. Enraf-Nonius, *The CAD4 Manaul* (1988).

26. C. Klein, *The Xpis Manaul* (X-ray Research G.m.b.H., Germany, 1993).

27. N. Collaborative Computational Project, *The CCP4 Suit : Programs for Protein Crystallography* (Daresbury Laboratary, UK, 1996).

28. G. Sheldrick, *SHELXL-93* (Institut fuer Anorg. Chemie, Goettingen, Germany, 1993).

29. D. Gewirth, Z. Otwinowski, W. Minor, *The HKL Manaul* (Department of Molecular Biophysics and Biochemistry, Yale University, Edition 4, 1995).

30. D. Gewirth, Z. Otwinowski, *The Scalepack Manaul* (Department of Molecular Biophysics and Biochemistry, Yale University, 1995).

31. Baylor College of Medicine, *Chain Manaul* (Version 5.0, 1992).

Chapter 3

Low Temperature Crystal Structure of Rb-Z-DNA - Possible Binding Sites of Monovalent Cation in the Z-DNA Structure

3. 1 Significance of the Problem

As shown in Table 3-1, several monovalent metal ions are necessary cations in the human body.

Table 3-1 Concentration of Metal Ions in the Human Body, Blood Plasma, and in Intracellular Fluid

Cation	Total concentration in body (g/70 kg)	Blood plasma (mmole/liter)	Intracellular (mmole/liter)
Na	100	142	10
K	140	4	160
Ca	1100	3	1
Mg	35	1	13
Fe	4	0.018	
Cu	0.15	0.016	
Zn	3	0.018	
Mn	0.02		
Co	0.001		
Mo	< 0.001		

From Saenger : "Principles of Nucleic Acid Structure" (1)

Due to the significance of metal ions on the biological activity and in stabilizing the structure of the polynucleotides, many crystallographic studies of the interaction between metal ions and DNA (or RNA) have been conducted. Most published work involved alkaline earth and transition metals including Mg^{2+} (2-4), Ba^{2+} (5), Co^{2+} (6), Cu^{2+} (7), $Co(NH_3)_6^{3+}$ (3, 8), and $Ru(NH_3)_6^{3+}$ (9) (see chapter 1). Young *et al.* (10) used molecular dynamic simulations with the new AMBER 4.1 force field to investigate the

'spine of hydration' of B-DNA. Their result revealed that in over half of the trajectory, Na^+ ions are found in the ApT step of the minor groove of a CGCGAATTCGCG sequence — the position noted to be an electronegative region termed the 'ApT pocket'. Unfortunately, little crystallographic work has been reported for the interaction of monovalent cations with DNA.

As part of my research projects, interaction between monovalent cations and DNA by x-ray crystallography has been studied. Since : a) Z-DNA diffracts to a consistently higher resolution than the A and B forms; b) we have extensive data on other cations interacting with Z-DNA; Z-DNA was selected as a structure of choice. Due to its propensity to assume a Z-DNA conformation, the sequence CGCGCG was selected for the studies.

Considering various monovalent cations for x-ray examination, Lithium — Li(I) interacting with DNA is interesting but due to its low atomic number (and accordingly low number of scattering electrons) it is difficult to visualize reliably by x-ray diffraction. Sodium — Na(I) is certainly biologically relevant and interesting, unfortunately, its atomic number and size make it difficult to distinguish reliably from the ubiquitous water. Rb(I) and K(I) have larger atomic numbers than oxygen which makes it easier to distinguish them from the oxygen in the aqueous solvent environment; Rb^+ and K^+ were accordingly chosen for the cations in this study.

Crystals of Rb-Z-DNA and K-Z-DNA with suitable size were obtained from intensive screening of the crystallization conditions. Initial data,

collected with the diffractometer at room temperature, failed to reveal ordered Rb^+ and K^+ positions; accordingly cryocrystallography, using the area detector for data collection, was employed in order to reduce the disorder of the structure. A total of sixteen possible Rb^+ binding sites were revealed in low temperature Rb-Z-DNA structure.

3. 2 Experimental

DNA Synthesis and Purification

The DNA synthesis and purification is described in Chapter 2.

Crystallization

The vapor diffusion method with sitting drops was used for the crystallization. The crystals of Rb-Z-DNA and K-Z-DNA were obtained after extensive crystallization condition searches including variations of pH, temperature (room temperature and 4 °C) and concentration of DNA, salt, buffer and MPD (2-methyl-2,4-pentanediol). Under optimum conditions, crystals with the shape of pseudo-hexagonal blocks (Fig. 3-1) were grown in 3 to 6 days at room temperature. For the Rb-Z-DNA crystals, the crystallizing drop was initially a mixture of 4.2 mM oligonucleotide, 40 mM rubidium cacodylate (pH 7.0), 200 mM rubidium chloride and 5% (by volume) (2-MPD). The equilibrating solution contained 300 mM rubidium chloride and 60% MPD. The dimensions of good crystals were approximately $0.5 \times 0.4 \times 0.2$ mm^3.

Fig. 3-1 Photograph of Rb-Z-DNA Crystal
The dimensions of the bigger crystal are about 0.5×0.4×0.2 mm^3

The same crystallization conditions were used for the K-Z-DNA crystallization except rubidium chloride was replaced by potassium chloride and rubidium cacodylate was replaced by potassium cacodylate. K-Z-DNA crystals were obtained with similar shape and size to the Rb-Z-DNA crystals.

The initial data collections for both K-DNA and Rb-DNA crystals were carried out with diffractometer using Cu Kα radiation at room temperature. Crystals with dimensions of about 0.5×0.3×0.3 mm^3 were mounted in a capillary tube with a droplet of mother liquor. An Enraf-Nonius diffractometer equipped with rotating anode generator was used for the data collection. The best data set is the K-Z-DNA data which has a total of 5391 reflections with resolution of 1.09 Å. The Molen program (see chapter 2) was used for the intensity decay and Lorentz-Polarization factor

correction. The crystal unit cell was found to have space group $P2_12_12_1$ and similar dimensions to the Mg-Z-DNA structure (see chapter 5). Molecular replacement was used for the structure determination with Mg-Z-DNA structure as the trial model. The X-plor program (see chapter 2) was utilized for the refinement and the Chain program (see chapter 2) was used for electron density map display and structure manipulation and analysis. Unfortunately, no Rb^+ or K^+ ions can be clearly located in the structure.

Cryogenic Data Collection

In a typical cryogenic data collection, a Rb-Z-DNA crystal with dimensions of $0.3\times0.2\times0.2$ mm^3 was picked up from the crystallization dip with a rayon loop (mounted on a magnetic base) and plunged in and out of a 60% MPD (MPD/H_2O, v/v) solution quickly to rinse off the salt solution from the crystal surface, then immediately frozen with liquid propane (inside a Hampton Research Crystalcap). The magnetic base with the frozen crystal was mounted on the goniometer head. The crystal was centered under a -190 °C stream of nitrogen.

A cryostat (Enraf Nonius FR558S) was used for the cold nitrogen stream supply and temperature regulation. Data collection was done on the 18mm MarResearch image plate area detector using monochromatized Cu Kα radiation generated from an Enraf-Nonius rotating anode generator. The x-ray power employed was 50 mA at 45 kV voltage. The distance between the crystal and image plate was 70 mm. The exposure time was 120 seconds per frame.

A data set with a total of 90° of phi at increments of 1.5° in each scan was collected. The data set has total of 2556 unique reflections with 2474 reflections having intensity larger than 2 sigma.

Data Processing and Unit Cell Determination

Image files were processed with the Denzo program (11) and reduced with Scalepack (12). The crystal was found to have unit cell dimensions of a=17.92 Å, b=31.12 Å and c= 43.96 Å with space group $P2_12_12_1$. The resolution of the data set is 1.76 Å with an R_{merge} of 0.061.

$$R_{merge} = \Sigma (| I - < I > |) / \Sigma I$$

In the equation, the sums are over all reflections with multiple equivalences. I is the intensity of the reflections. $< I >$ is the average of the equivalent reflections for the particular reflection being considered.

The statistics after the data processing are listed in Table 3-2.

Table 3-2 Rb-Z-DNA Structure Data Statistics

A. Summary of I/Sigma in Resolution Shells

Shell Lower limit	Upper limit	% of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
30.00	3.66	1.2	1.8	3.2	3.8	4.1	7.3	33.6	60.2	93.9
3.66	2.91	0.6	0.6	0.6	1.9	2.8	5.4	24.3	66.9	91.2
2.91	2.54	0.0	0.0	1.0	1.3	2.7	6.7	24.4	69.9	94.3
2.54	2.31	0.3	1.0	1.3	2.0	4.6	9.6	28.5	64.9	93.4
2.31	2.14	0.0	0.3	0.7	1.7	4.8	11.2	32.7	62.9	95.6
2.14	2.02	0.4	1.8	2.5	3.5	6.0	13.0	34.4	60.4	94.7
2.02	1.91	0.7	1.4	3.1	5.8	10.3	19.6	44.3	50.2	94.5
1.91	1.83	0.0	1.8	3.2	4.7	8.3	18.0	48.6	46.4	95.0
1.83	1.76	1.7	5.8	11.3	14.7	21.5	34.1	65.9	27.3	93.2
1.76	1.70	0.7	0.7	1.0	1.4	2.4	3.1	5.5	0.7	6.1
All hkl		0.6	1.5	2.8	4.0	6.6	12.6	34.0	51.3	85.3

B. Summary of Reflection Intensities and R-factors by Shells

R linear = SUM (ABS(I - <I>)) / SUM (I)
R square = SUM ((I - <I>) ** 2) / SUM (I ** 2)
Chi**2 = SUM ((I - <I>) ** 2) / (Error ** 2 * N / (N-1)))
In all sums single measurements are excluded

Shell limit	Lower Angstrom	Upper	Average I	Average error	stat.	Norm Chi**2	Linear R-fac	Square R-fac
30.00	3.66		55950.4	2477.0	1838.3	2.091	0.073	0.102
3.66	2.91		53584.3	2021.7	1284.8	1.938	0.050	0.062
2.91	2.54		32795.2	1255.5	843.2	1.984	0.050	0.059
2.54	2.31		21361.9	798.8	537.3	2.106	0.056	0.072
2.31	2.14		18484.3	745.3	517.2	2.327	0.055	0.063
2.14	2.02		15259.9	594.7	418.3	2.518	0.060	0.068
2.02	1.91		11824.7	490.5	350.9	2.578	0.066	0.068
1.91	1.83		10236.0	462.7	344.3	2.784	0.081	0.093
1.83	1.76		5754.6	334.0	276.5	2.502	0.087	0.074
1.76	1.70		4361.7	351.2	304.8	2.124	0.093	0.096
All reflections			25689.3	1050.2	734.8	2.300	0.061	0.082

Both R-linear and R-square are less than 0.1 for the overall data set which means the agreement between symmetry related reflections is good. Up to 1.76 Å, over 90% of the reflections were obtained and the data were reliably measured with two third of the reflections in the high resolution shell (1.83-1.76Å) having $I/\sigma(I)$ of 10 or greater. χ^2 (the goodness-of-fit) is about 2 which also showed that the quality of the data is very acceptable.

Structure Refinement

Since the unit cell of Rb-Z-DNA is similar to that of the Mg-Z-DNA structure, the Mg-Z-DNA (13) structure was used for the initial trial model and the molecular replacement method was employed for the structure determination. The starting structure (including solvent H_2O) was refined with the X-plor program against Rb-Z-DNA data. After several cycles of the positional and constrained isotropic B-factor refinement and one of simulated annealing the R factor dropped to 23.0%. H_2Os with very low B values, some as low as 2.0 $Å^2$ (the minimum the program allows), were changed to Rb^+ ions. Electron density maps (2Fo-Fc and difference Fo-Fc) were used to locate new solvent molecules and to adjust the positions of the atoms. Solvent molecules were deleted if they had very high B values and there was negligible 2Fo-Fc density at their position. With the Chain program the molecular structure, electron density map and coordination of Rb^+ ions were examined and new water molecules were added. Multiple cycles of positional refinement, individual B refinement, simulated annealing and water and ion reassessment were carried out on the structure until convergence. For disordered structures, two partially occupied sites

for each atom were assigned to alternate structures so that they did not interact with each other in the refinement.

For comparison, a data set of a Rb-Z-DNA crystal was also collected at room temperature with the area detector. The structure from this data set was refined with the same method as above.

3. 3 Results and Discussion

As a result of the structure refinement sixteen likely Rb^+ positions have been located. An electron density peak which we believe to belong to a chloride ion has also been identified. The final residual factor (R factor) for the structure is 17.5%. In addition, use of the low temperature Rb^+ structure allowed a reinterpretation of the room temperature data and credible assignment of several monovalent cations in the structure. The room temperature structure has final R factor of 18.4% in which four Rb^+ ions were located but with higher B values than in the low temperature structure.

Portions of the electron density map of the low temperature structure are shown in Fig. 3-2 and 3-3.

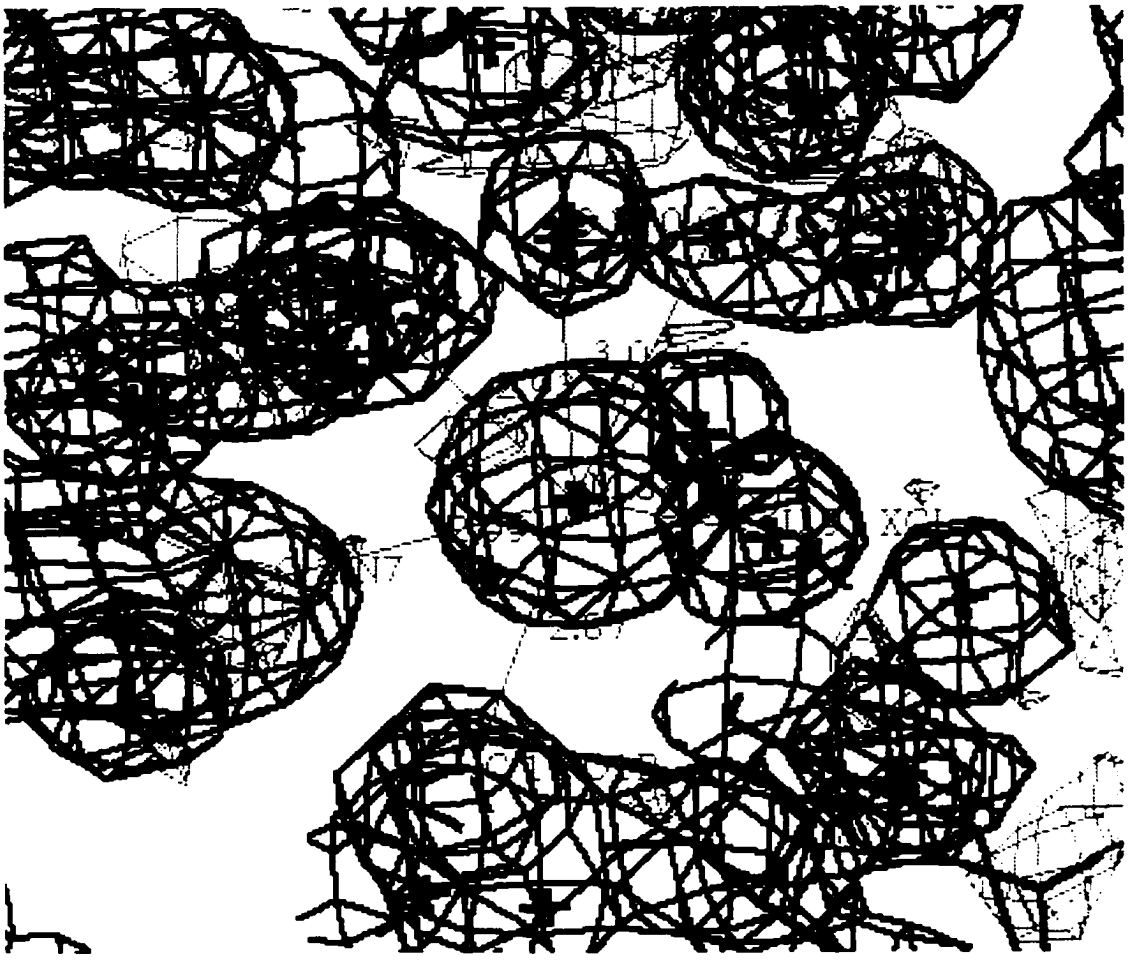


Fig. 3-2 Electron Density Map of Rb-Z-DNA Structure Centered on Rb no. 88

2Fo-Fc map is shown as blue solid lines and Fo-Fc map in Cyan.
Rb^+ ions are drawn as green stars. Water molecules are shown as red stars and Chloride ion as a black star.

(In microfilm, 2Fo-Fc map is shown in dark gray solid lines and Fo-Fc map in light gray solid lines)

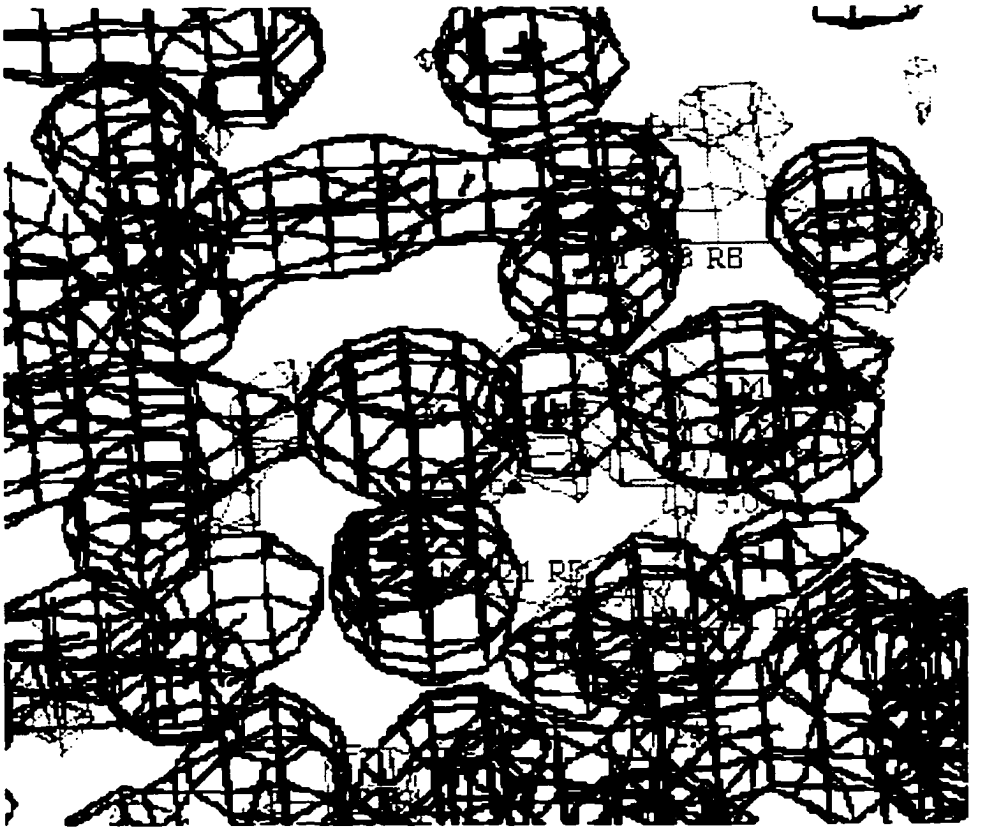


Fig. 3-3 Electron Density Map of Rb^+ Ions Cluster in Rb-Z-DNA Structure including Rb's no. 210, 222, 212, 321, 322, 323

2Fo-Fc map is shown as blue solid lines and Fo-Fc map in Cyan.
Rb^+ ions are drawn as green stars.

3.3.1 Low Temperature DNA structure in the Rb-Z-DNA Crystal

1). Changes in the Unit Cell as a Function of Temperature

Compared with the Rb-Z-DNA room temperature structure, the unit cell of the Rb-Z-DNA low temperature structure is 1.4% smaller. The Mg^{2+} salt structure shrinks by a similar amount. It is interesting to note that in both cases the major change is in the c-axis, the major stacking or helical direction.

Table 3-3 lists several unit cells of the Z form of $d(CGCGCG)_2$ at different temperature.

Table 3-3 Unit Cell Dimensions of Selected Z-DNA Structures

crystallization salt		spermine		magnesium		rubidium	
Temperature of data collection		room T[14] ~ 298K	low T[15] 163K	room T[13] ~ 298K	low T[16] 150K	room T ~ 298K	low T 83K
unit cell dimension (Å)[4]	a	18.41	18.27	18.01	17.85	17.90	17.92
	b	30.77	30.69	31.03	31.10	31.10	31.12
	c	43.15	42.46	44.80	44.18	44.69	43.96
Cell Volume (Å^3)		24,443	23,808	25,036	24,526	24,878	24,515

2). The Structure of $d(CGCGCG)_2$

The DNA part of the Rb-Z-DNA structure is very similar to the magnesium form except that the phosphate group of Cytidine 5 is found to be disordered with a partial Z_I/Z_{II} conformation (Fig. 3-4). The disordered structure was revealed by two very close electron density peaks (with distance of 1.36 Å) near the position of the phosphorus site. These two peaks were assigned as phosphorus atoms and refined in alternate conformations with half occupancy. After refinement, electron density peaks of the oxygen atoms which are connected to both phosphorus atoms also appeared. The whole phosphate group (P, O1P, O2P, O5', C5' H5', H5")* was then refined with half occupancy in each structure. The final B factors for the atoms in both groups were about 6.6 $Å^2$ indicating that the assignment of half occupancy for each was appropriate.

* Note : For the naming of the atoms refer to chapter 1, Fig. 1-3 and Fig. 1-4.

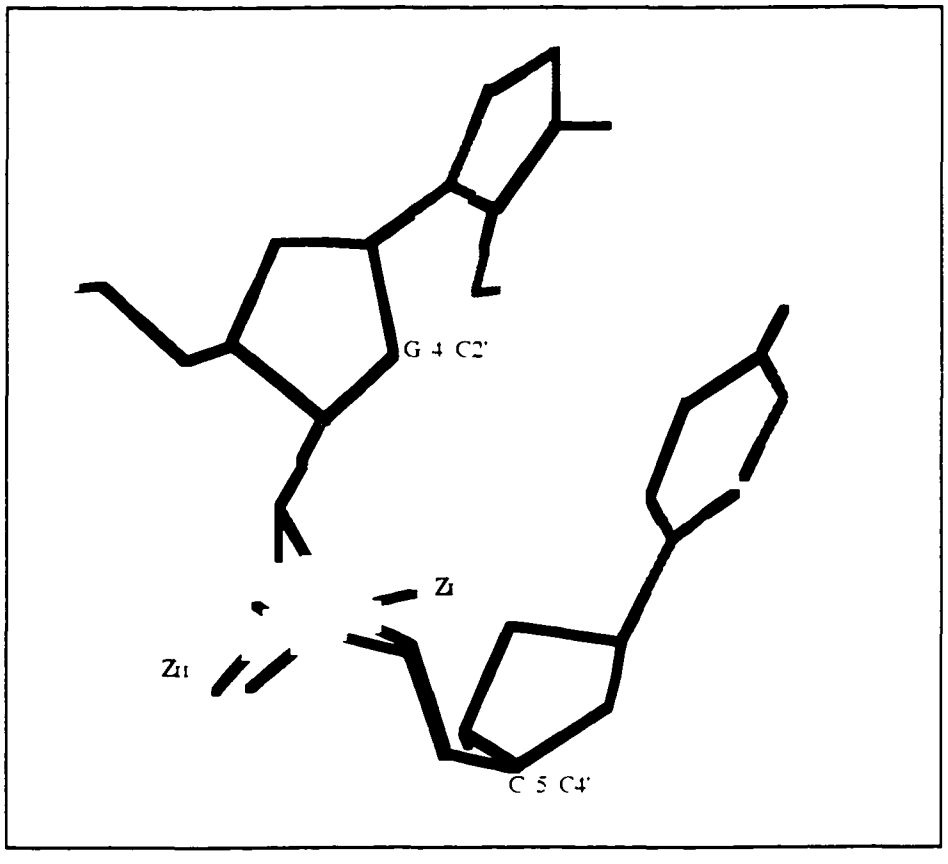


Fig. 3-4 Disorder of Phosphate Group at Cytidine 5

It is possible that the low temperature "freezes" the two conformations which have similar potential energy and have fast exchange at room temperature.

In the structure all the purine rings are in the syn-conformation which is consistent with the characteristics of Z-DNA. The sugar ring in the cytidine residues are C2' endo and the guanosines are C3' endo except for the two termini and G6 residues. Table 3-4 lists the torsion angles of the sugar phosphate backbone.

Table 3-4a Torsion Angles of the Ribose Phosphate Backbone for the Cytidine Residues

residue	Angle							conformation	sugar pucker
	α	β	γ	δ	ε	ζ	χ		
C1	-	-	53	147	-96	77	-150		C2' endo
C3	-158	-127	66	138	-99	75	-155		C2' endo
C5	161	177	54	145	-92	69	-149	Z_{II}	C2' endo
C5	-161	-116	57	142				Z_I	
C7	-	-	61	142	-85	77	-154		C2' endo
C9	-153	-125	61	148	-91	71	-149		C2' endo
C11	-157	-124	58	137	-93	65	-154		C2' endo

Table 3-4b Torsion Angles of the Ribose Phosphate Backbone for the Guanosine Residues

residue	Angle							conformation	sugar pucker
	α	β	γ	δ	ε	ζ	χ		
G2	60	-167	178	92	-123	-68	58		C3' endo
G4	68	-176	178	92	-175	43	63	Z_{II}	C3' endo
				90	-115	-73		Z_I	
G6	81	177	-178	149	-	-	76		C2' endo, C3' exo
G8	64	-170	174	89	-123	-70	64		C3' endo
G10	67	-178	-180	96	-121	-66	60		C3' endo
G12	86	179	-178	158	-	-	75		C2' endo

3.3.2 Identification of the Rb^+ Ions in the Rb-Z-DNA Structure

As with most Alkali atoms, rubidium is soft and doesn't have a regular hydration coordination geometry. The Rb-O distance (2.93 - 3.02 Å) is close to the H_2O hydrogen bonding distance (2.8 Å) (17), so it is difficult to make an unambiguous position assignment. An interesting consequence of this may be that Rb^+ fits into the H_2O "lattice" rather isomorphically minimizing the disruption to the remaining water structure.

In the Rb-Z-DNA low temperature structure, all the electron density peaks in the solvent area were initially assigned as H_2O molecules. At a certain refinement stage some H_2O molecules had the temperature factor as low as 2.0 $Å^2$. This B value* is set as the lowest value which the X-plor program will allow and is less than the average B of the DNA which is about 4.5 $Å^2$. These low B values indicate that there is a high electron density in the position suggesting the atom should have more electrons than oxygen. After checking the position and coordinations, the electron density peaks of these atoms were assigned as Rb^+. After more cycles of positional, B refinement and simulated annealing, some Rb^+'s had B values of more than 40 when they were assigned full occupancy but very low B when assigned as water. These were, therefore, considered as disordered Rb^+ ions. Some of these ions were also close to other Rb^+ ions. Partial occupancies were

* $B=8\pi^2\mu^2$ where μ is the RMS displacement of the atom from its center location. B factor is the measurement of the vibration of the atom. The lower the B is, the less vibration the atom has.

assigned to these Rb^+ ions. After further refinement, almost all Rb^+ ions have B value below 30.0 $Å^2$ (except Rb^+ no.31 which has a B of 31.76 $Å^2$). The final structure has three fully occupied Rb^+ ions and thirteen Rb^+ with partial occupancy (Table 3-5). One chloride ion was also found in the structure (see below).

The total positive charge is 9.75 per duplex for a net charge of -1.25 (one $d(CGCGCG)_2$ duplex plus the chloride ion has total charge of -11 (five phosphate groups per DNA strand and one chloride ion at pH 7).

Most Rb^+ ions assigned in the structure have five or more coordination ligands and no regular coordination geometry. The coordination distance of Rb^+ ions in the structure are listed in Table 3-7 *.

In the room temperature structure of Rb-Z-DNA, Rb no. 31, 32, 88, 284 were also identifiable (Table 3-6).

*Note: The rubidium numbering is the same as that used in the PDB file. The rubidium ions with full occupancy and those partially occupied which do not have obvious alternative site have numbers below 200. The rubidium atoms with numbers in the 200's and 300's represent alternative positional arrangements.

Table 3-5 Occupancy and Temperature Factor of Rb^+ Ions In the Low Temperature Rb-Z-DNA Structure

Atom	Residue Number*	Coords x	y	z	Occupancy	Temperature Factor
Rb^+	32	19.637	23.294	10.441	1.00	13.61
Rb^+	88	5.917	10.408	2.193	1.00	17.87
Rb^+	111	8.208	24.192	1.145	1.00	21.82
Rb^+	34	4.699	21.697	1.186	.75	26.06
Rb^+	31	9.701	22.527	16.228	.50	31.76
Rb^+	40	18.731	19.194	5.085	.50	23.75
Rb^+	41	8.400	9.965	-1.245	.50	23.77
Rb^+	210	8.062	18.816	-7.928	.50	26.61
Rb^+	211	14.563	11.374	14.272	.50	21.92
Rb^+	212	12.877	17.304	-4.035	.50	20.44
Rb^+	320	13.969	11.715	11.475	.50	24.84
Rb^+	321	14.304	13.767	15.472	.50	23.80
Rb^+	322	9.677	20.821	-6.691	.50	16.57
Rb^+	323	11.629	19.400	-3.118	.50	27.72
Rb^+	379	5.527	16.484	9.004	.50	26.39
Rb^+	284	17.853	15.213	6.508	.50	19.66

Table 3-6 Occupancy and Temperature Factor of Rb^+ Ions in the Room Temperature Rb-Z-DNA Structure

Atom	Residue Number*	Coords. x	y	z	Occupancy	Temperature Factor
Rb^+	31	9.75	22.51	17.26	1.00	32.95
Rb^+	32	19.97	22.47	10.84	1.00	41.43
Rb^+	284	17.25	15.29	7.06	.50	27.75
Rb^+	88	5.79	10.51	2.04	.50	30.12

*Note: The residue number is the number in the PDB file in Table 3-5, 3-6.

Table 3-7 Coordination Distance of Rb^+ in the Low Temperature Rb-Z-DNA Structure

atom	residue number	coordinated atoms	distance (Å)
Rb	32	G(6) O2P	3.19
		O5'	3.33
		C(9)# O1P	3.17
		O2P	2.82
		W35	2.74
		W76	2.79
		W162	3.50
		W237	2.93
		W338	2.73
Rb	88	G(2) N7	2.99
		C(11)# O1P	2.87
		CL93	3.09
		W24	3.30
		W270	3.03
		W51	2.93
		W263	3.27
		W363	2.66
Rb	111	C(3) O1P	2.79
		W46	2.90
		W52	3.06
		W96	3.19
		W110	2.71
Rb	41	G(2) O6	2.71
		W49	2.90
		W98	2.89
		W370	3.33
		CL93	2.90
Rb	34	G(2) O3'	3.10
		C(3) O1P	2.83
		W44	3.15
		W48	3.03
		W58	3.44
		W118	2.76
		W330	2.98
		W230	3.50
Rb	31	G(6) O6	2.72
		G(12)# O6	2.91
		W25	3.28
		W27	2.85
		W55	2.96
		W67	2.74
		W263	2.71
		W363	3.00
Rb	40	W44	3.42
		W81	2.89
		W87	2.88
		W168	3.05
		W366	3.33
		W266	2.82
		W356	2.98
		W368	3.22

* The residues with # signs are residues in the symmetry related molecules.

Table 3-7 *continue*

atom	residue number	coordinated atoms	distance (Å)
Rb	210	G(2) O2P	2.77
		Rb322	2.86
		W33	3.02
		W71	2.70
		W78	3.09
		W105	3.03
		W160	3.34
		W237	2.91
		W337	3.05
Rb	211	C(7) O3'	3.22
		G(8) O2P	2.74
		Rb 320	2.88
		Rb 321	2.69
		Rb 322	3.03
		W71	3.40
		W237	3.48
		W337	2.47
Rb	212	C(1) O2	2.60
		Rb321	2.49
		Rb323	2.61
		W47	2.94
Rb	320	G(8) N2	2.94
		C(9) O2P	2.62
		Rb211	2.88
		W106	2.80
		W337	2.79
Rb	321	C(7) O2	2.80
		Rb211	2.69
		Rb212	2.49
		W71	2.80
Rb	322	G(2) O1P	2.94
		Rb210	2.86
		Rb211	3.03
		W115	3.36
		W337	2.85
		W338	3.04
		W331	3.21
Rb	323	C(1) O3'	3.47
		G(2) O1P	2.93
		Rb212	2.61
		W47	3.39
		W113	2.81
		W330	3.33
		W230	2.83
		W331	3.01
Rb	284	C(9) O3'	3.32
		G(10) O2P	2.84
		W21	3.32
		W116	2.82
		W158	3.11
		W366	3.12
Rb	379	G(4) O6	2.90
		G(8) O6	2.97
		W66	3.09
		W277	3.33

3.3.3 Assignment of the Disordered Rb^+ Ions

In the Rb-Z-DNA structure several Rb^+ ions and water molecules around them are found to be disordered. In some of these positions certain strong electron density peaks appear too close for multiple simultaneous Rb^+ sites. At the same time they appeared too strong to be considered as H_2O but less dense than the more isolated Rb^+ sites. Accordingly, these positions were assigned as alternate partially occupied Rb^+ sites. In the structure file the sites in one pattern were labelled with the number 200's and those in the alternate positions labelled with 300's. Rb^+ in one group (with number 200's) do not interact with Rb^+ in another group (with number 300's) in the refinement.

The partial occupancy of these rubidium ions shows that in these locations the rubidium may only be present for a certain fraction of the time.

3.3.4 Distribution and Binding Position of the Rb^+ Ions in the Rb-Z-DNA Structure

Ten of the rubidium ions may be considered to be coordinated to at least one phosphate oxygen while only one is coordinated to an N7* of G. Five rubidium ions bind to the carbonyl oxygen of the bases (three Rb^+ ions bind to O6 and two bind to O2). Four Rb^+ ions are coordinated to the O3' (oxygen of the phosphodiester bond) of the nucleotide.

*Note : For the naming of the atoms, refer to chapter 1, Fig. 1-3 and Fig. 1-4.

1). Rubidium Ions on the Surface of the DNA

Two of the rubidium ions were found to bind both the primary DNA and a symmetry related molecule stabilizing the lattice (primary DNA is simply the specific duplex DNA which has been chosen as the reference duplex to examine and describe the interaction). Rb no.32 is bound to a phosphate oxygen (O5' and O2P) of the Guanosine 6 (G6) and phosphate oxygen atoms of the Cytidine 9 (C9) on the symmetry related DNA (Fig. 3-5)*.

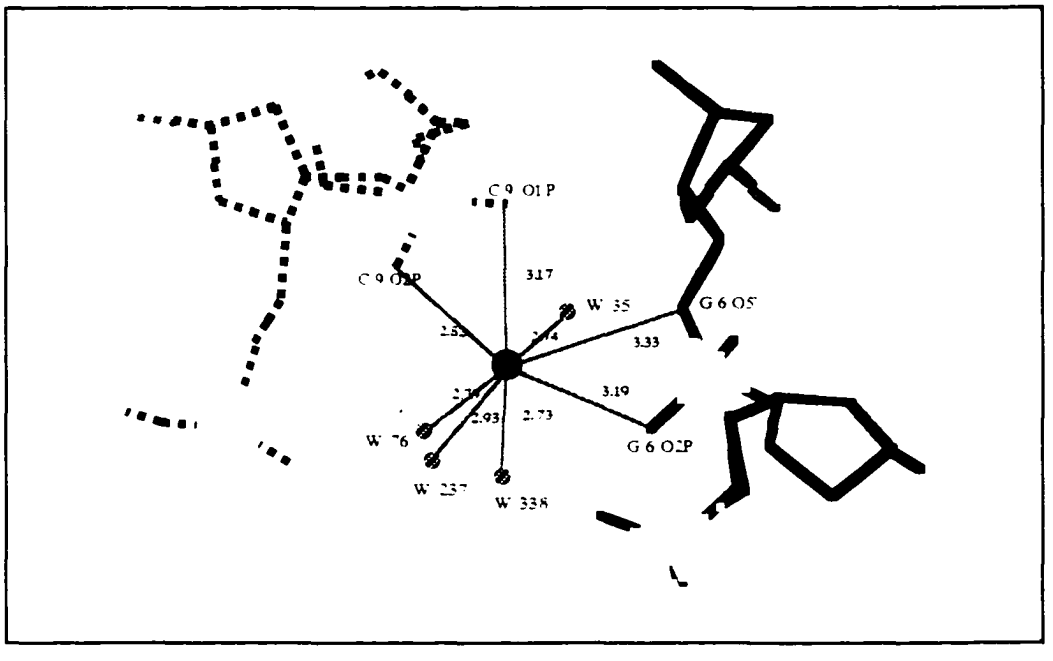


Fig. 3-5 Coordination of Rb no.32 in the Z-DNA Structure

*In the Fig. 3-4 to 3-19, molecules are colored by atom type (Carbon-black, Nitrogen-blue, Oxygen-red, Phosphorus-yellow). Central DNA molecule is drawn in solid stick and symmetry related molecule in dashed line. Rb^+ ion is shown in solid circle and oxygen atoms of water in diagonally striped. Chloride ion is drawn in small solid circle. Interaction distance is shown in angstrom.

Rb no.88 coordinates the N7 (base nitrogen) of the Guanosine 2 (G2) and O1P of the Cytidine 11 (C11) on the symmetry related structure. It is also connected to O6 (carbonyl oxygen of the base) of Guanosine 10 (G10) (interstrand interaction) through a water bridge (Fig. 3-6).

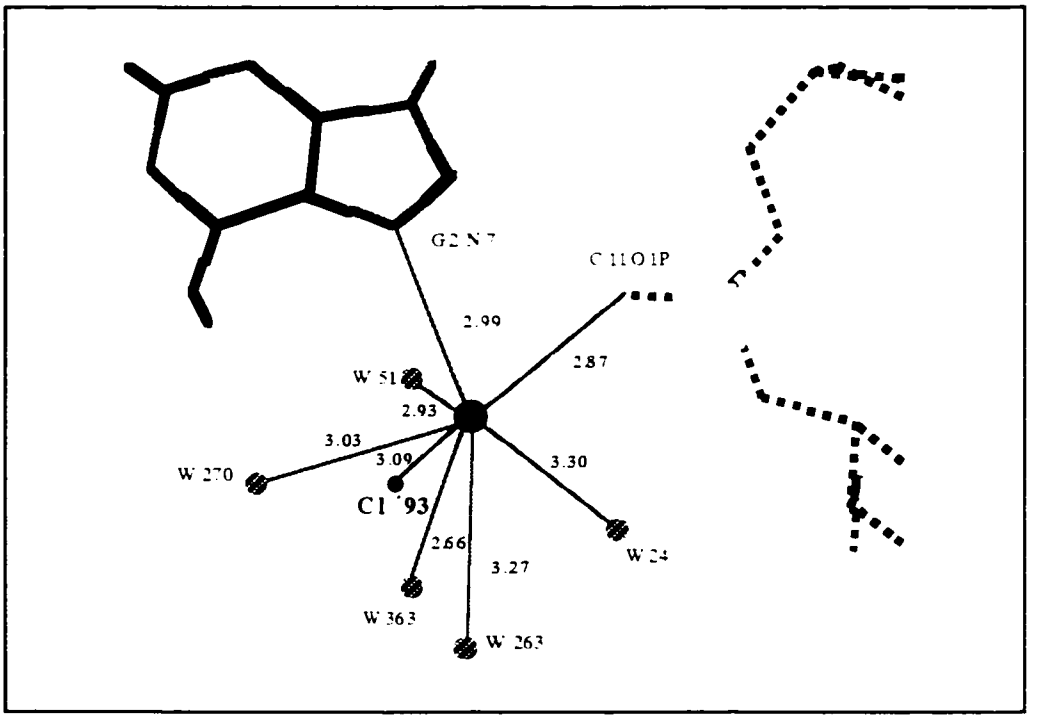


Fig. 3-6 Coordination of Rb no.88 in the Z-DNA Structure

Rb no.111 (Fig. 3-7) and no.34 (Fig. 3-8) are bound to oxygen of phosphates in the central DNA and to a symmetry-related molecule through H_2O.

Rb no.31 coordinates O6 (base oxygen) of G6 and O6 of a symmetry-related Guanosine 12 (G12) so it bridges the duplex and another duplex stacking on it (Fig. 3-9). These two residues are consecutive base pairs of the continuous Z-helix running through the crystal with every 6^{th} phosphate missing.

Rb no.379 connects O6 of Guanosine 4 (G4) and O6 of Guanosine 8 (G8) which connects two strands in the duplex (Fig. 3-10) and is similar to the interactions of Rb no.31 listed above except the connection is within one molecule.

Rb no.41 coordinates O6 of G2 and connects other strand and symmetry related molecule through an H_2O bridge (Fig. 3-11). Rb no.40 doesn't directly connect DNA but connects two strands through a water bridge (Fig. 3-12).

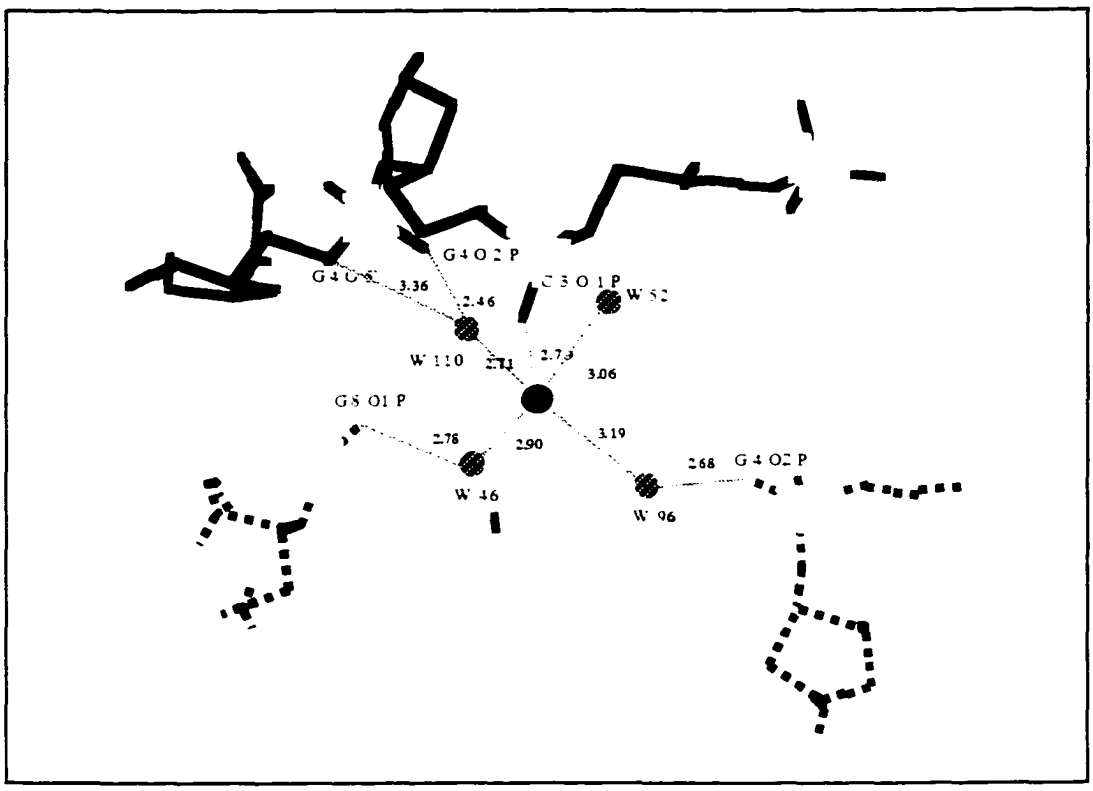


Fig. 3-7 Coordination of Rb no.111 in the Z-DNA Structure

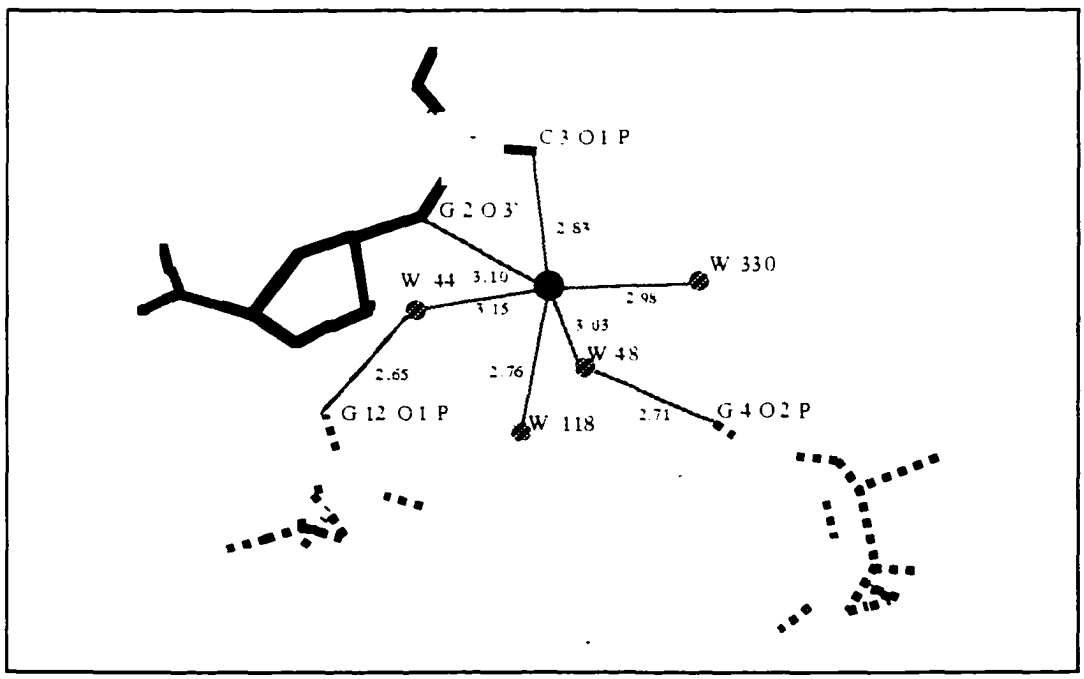


Fig. 3-8 Coordination of Rb no.34 in the Z-DNA Structure

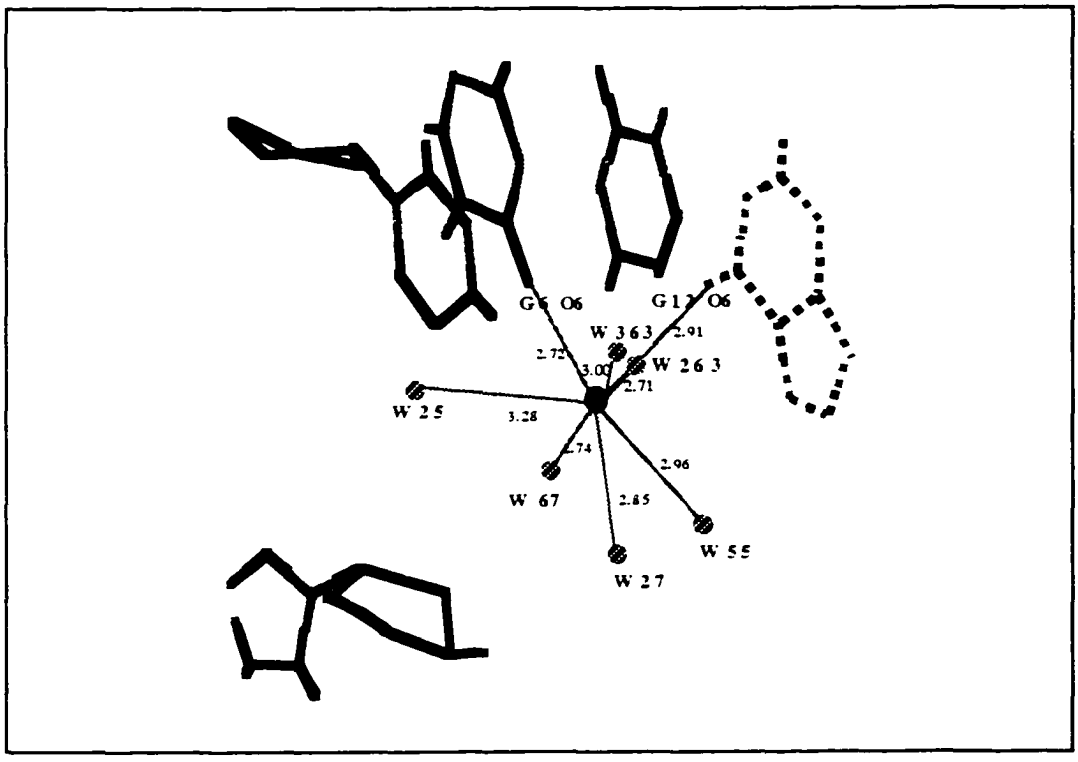


Fig. 3-9 Coordination of Rb no.31 in the Z-DNA Structure

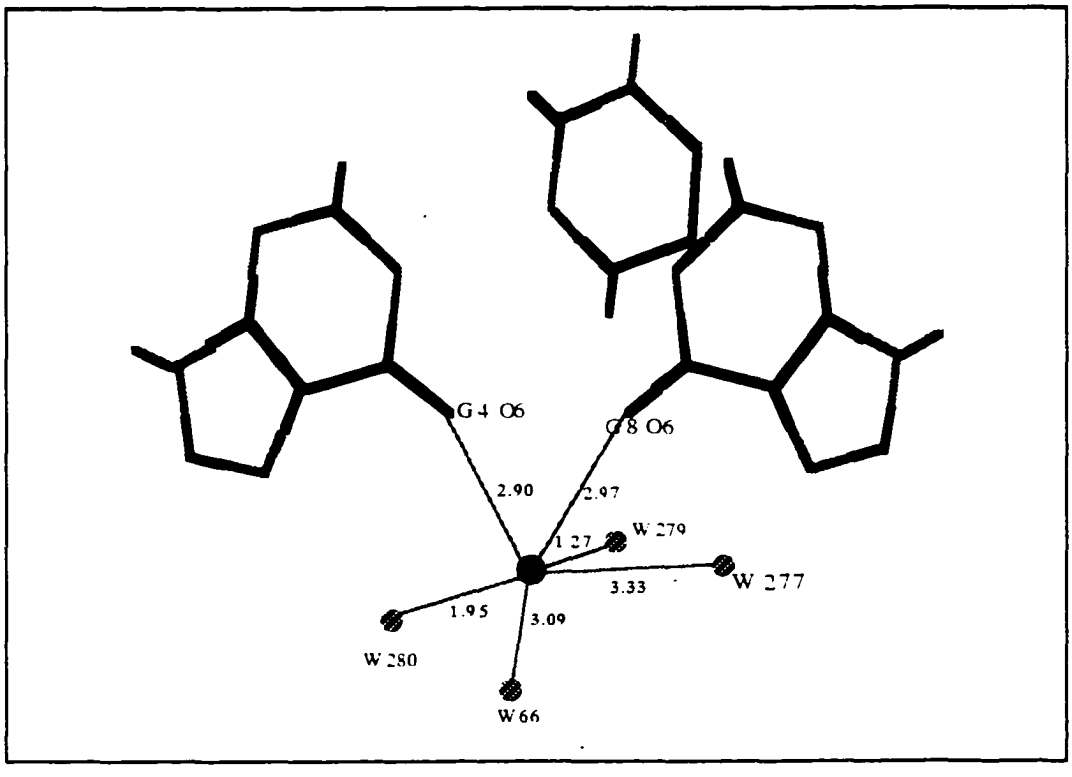


Fig. 3-10 Coordination of Rb no.379 in the Z-DNA Structure

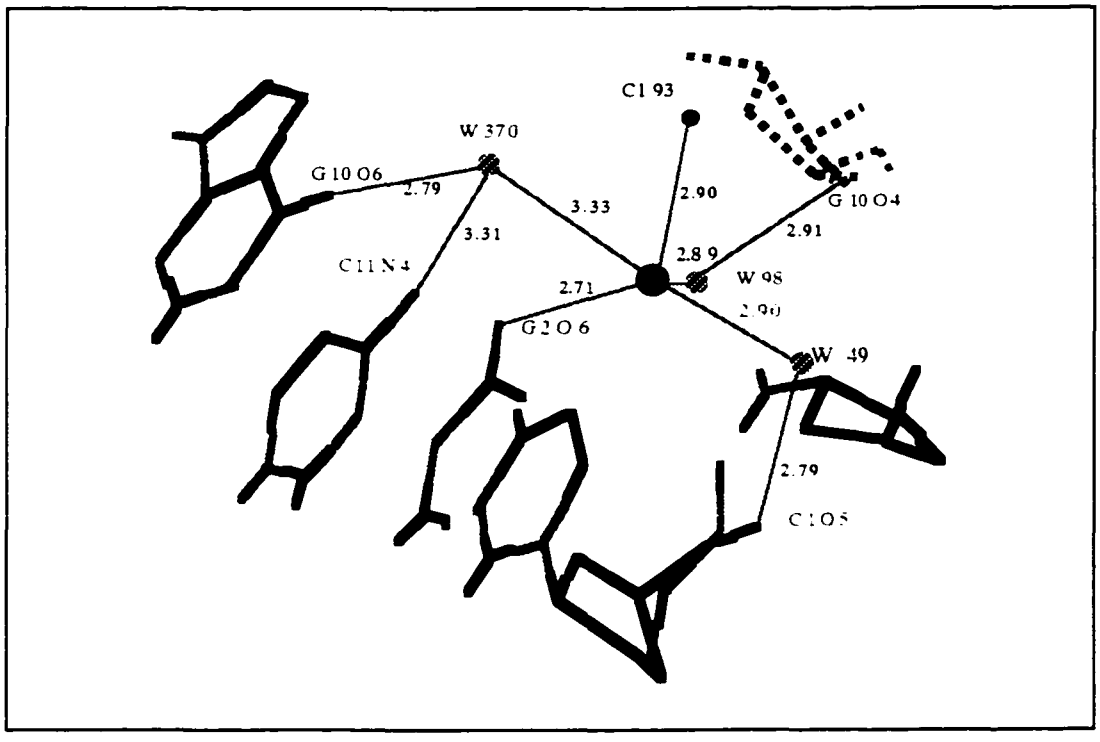


Fig. 3-11 Coordination of Rb no.41 in the Z-DNA Structure

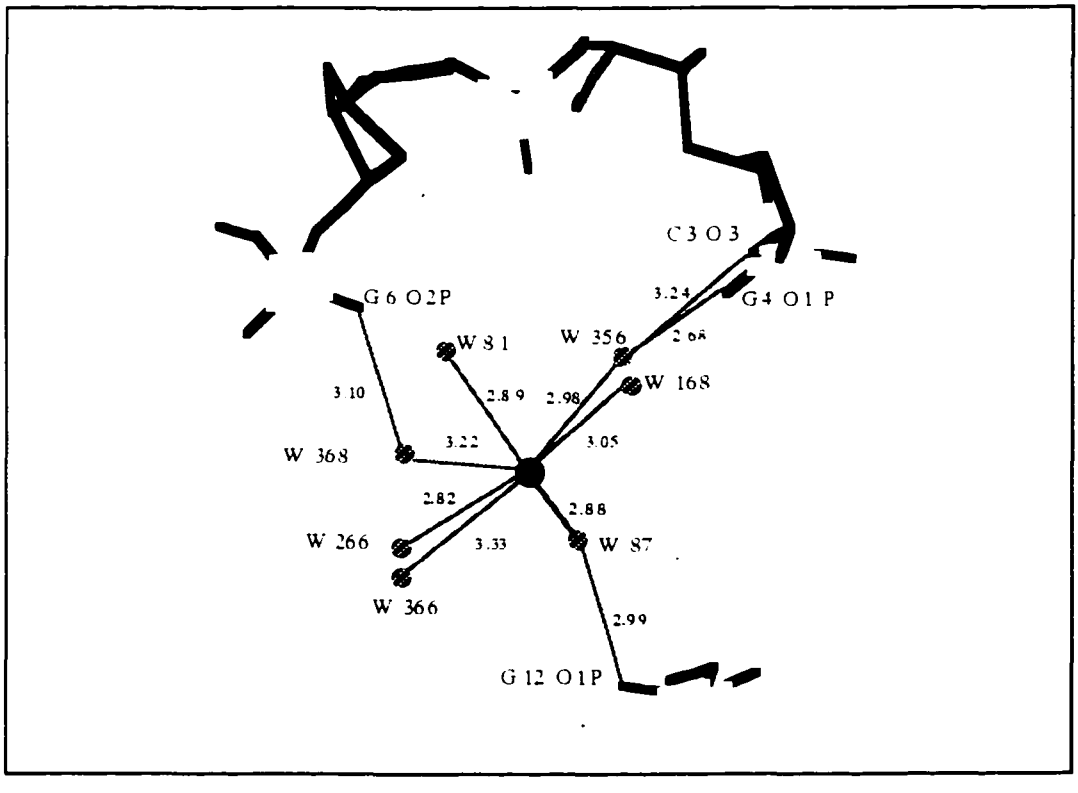


Fig. 3-12 Coordination of Rb no.40 in the Z-DNA Structure

2). Partially Occupied Rubidium Ions Near the Groove of the Duplex

In addition to the convex surface of DNA, a Rb^+ cluster and Rb no.284 are found within or near the groove of the Z-DNA structure. They are all partially occupied. The cluster involves Rb no.210, 211, 212, 320, 321, 322, and 323. Figure 3-13 shows the geometry and distance of the Rb^+ ions in the cluster.

The rubidium cluster is in the electronegative pocket of the DNA. This pocket is formed by phosphate oxygen and nitrogen of C7, G8, C9 and phosphate groups of symmetry related molecules.

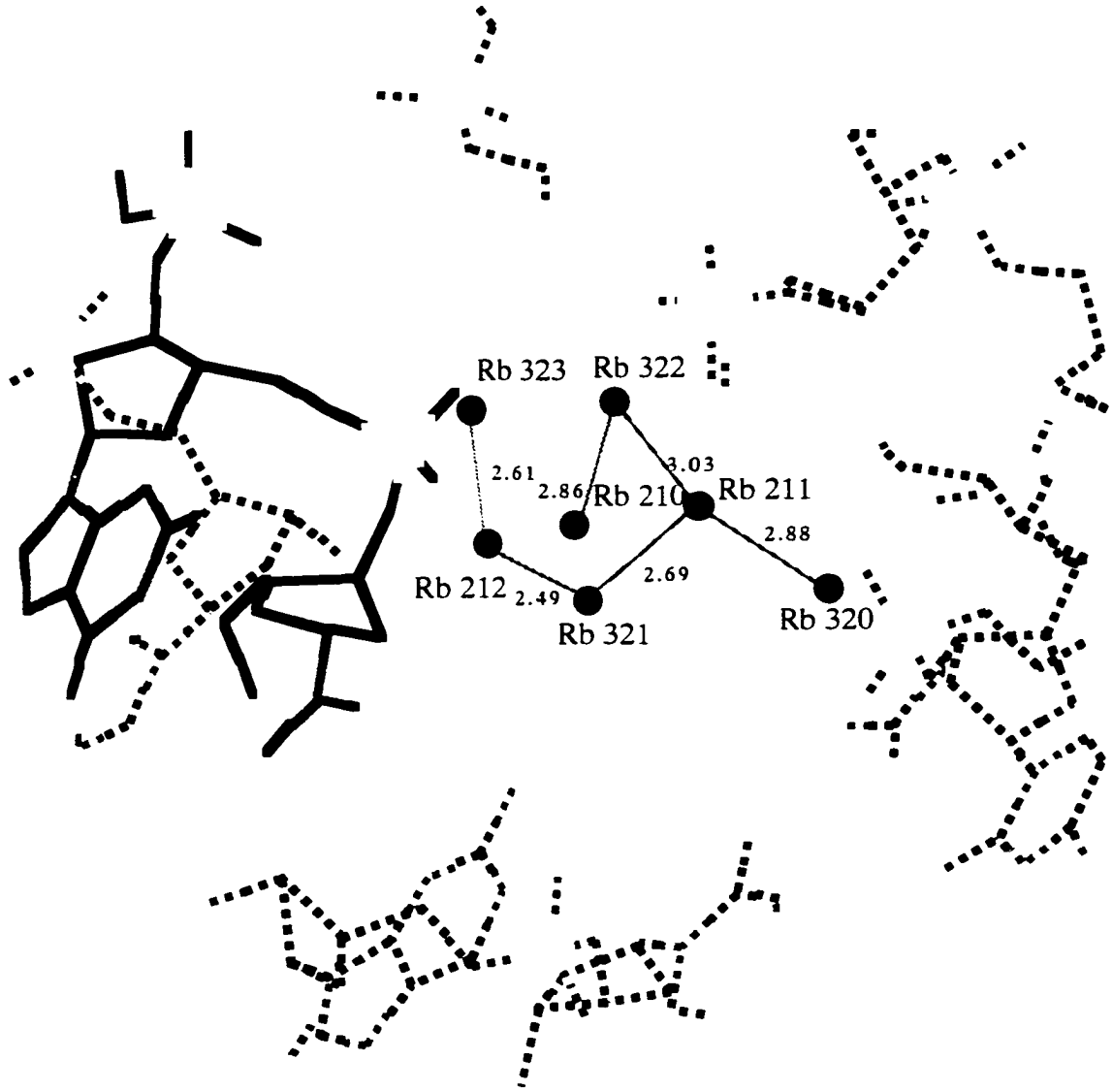


Fig. 3-13 Rb^+ Ion Cluster in the Electronegtive Pocket of the Z-DNA Structure

The Rb no.321 and 211 bridge O2 (carbonyl oxygen) of C7 and O2P of G8 (Fig. 3-14). Rb no.323 and 212 bridge O2 of C1 and O1P of G2 (Fig. 3-15). The distances between the two Rb^+ ions in this two pair are only about 2.6 Å.

Rb no.320 connects N2 (base nitrogen) of G8 and O2P of C9. The distance between Rb no.320 and O2P on the C9 is only 2.62 Å (Fig. 3-16).

Rubidium no.210 and 322 coordinate O2P of G2 and O1P of G2 (Fig. 3-17). Rubidium no.210 also connects O1P of G10 , O3' of G12 and C5 phosphate oxygen (both Z_I and Z_{II}) through water molecules.

Rb no.284 is inside the groove of the Z-DNA. It coordinates the G10 O2P and C9 O3' in the DNA molecule (Fig. 3-18).

In summary, rubidium ions in the low temperature structure were found to be able to directly coordinate to the oxygen of the phosphate, the oxygen and nitrogen of the base groups or to connect these atoms through water bridges. Rb^+ connects the two DNA double strands (Rb no.31) and stabilized the duplex though inter-strand coordination (Rb no.88 and no.379). The partially occupied Rb^+ sites could alternately be occupied by Rb^+ ions and H_2Os. The relatively large number of partially occupied sites and the inability to distinguish more well defined sites in the room temperature structure suggested that there are numerous relatively low energy sites for monovalent cation binding to Z-DNA and that many of these may also be mutually exclusive. That is when an ion is in one position, other nearby are no longer favorable.

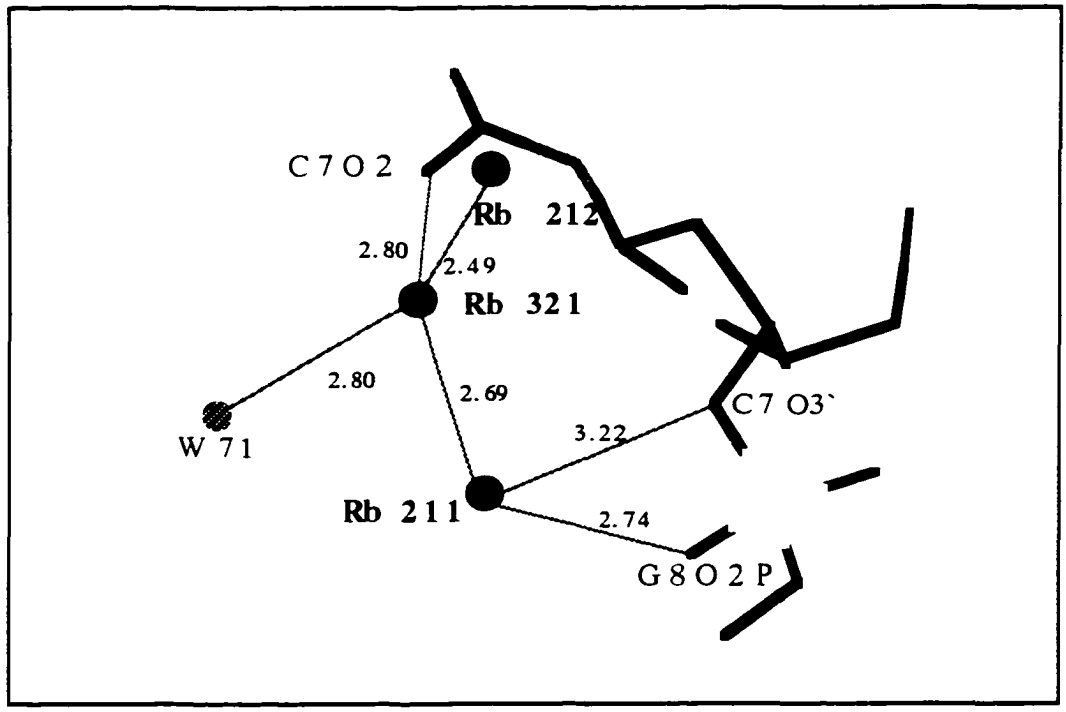


Fig. 3-14 Coordination of Rb no.321 and 211 in the Z-DNA Structure

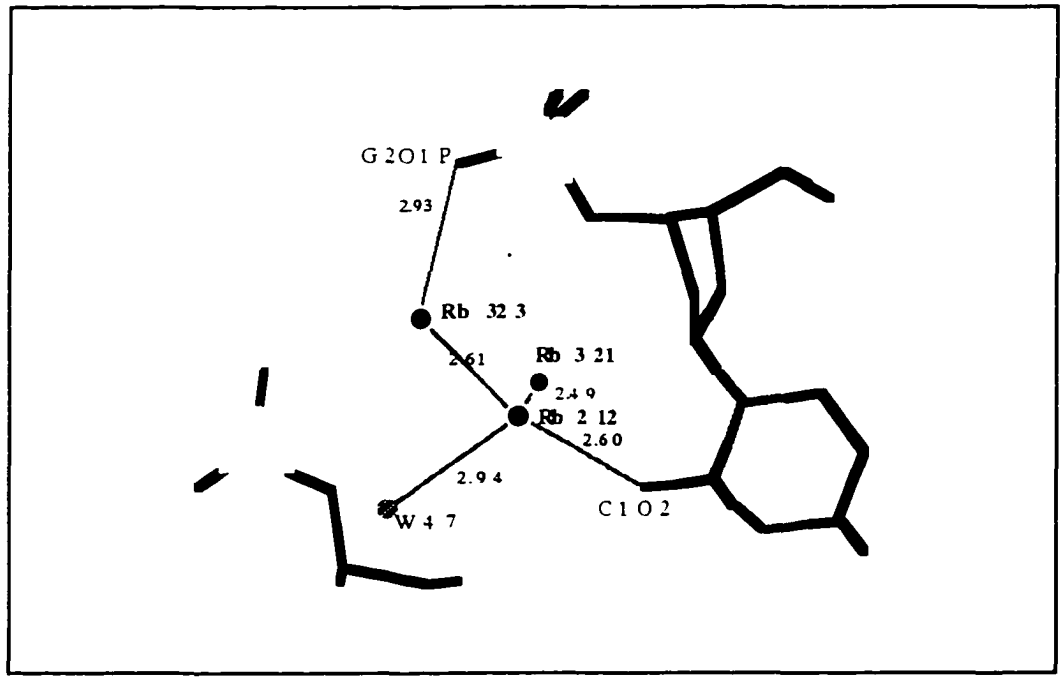


Fig. 3-15 Coordination of Rb no.323 and 212 in the Z-DNA Structure

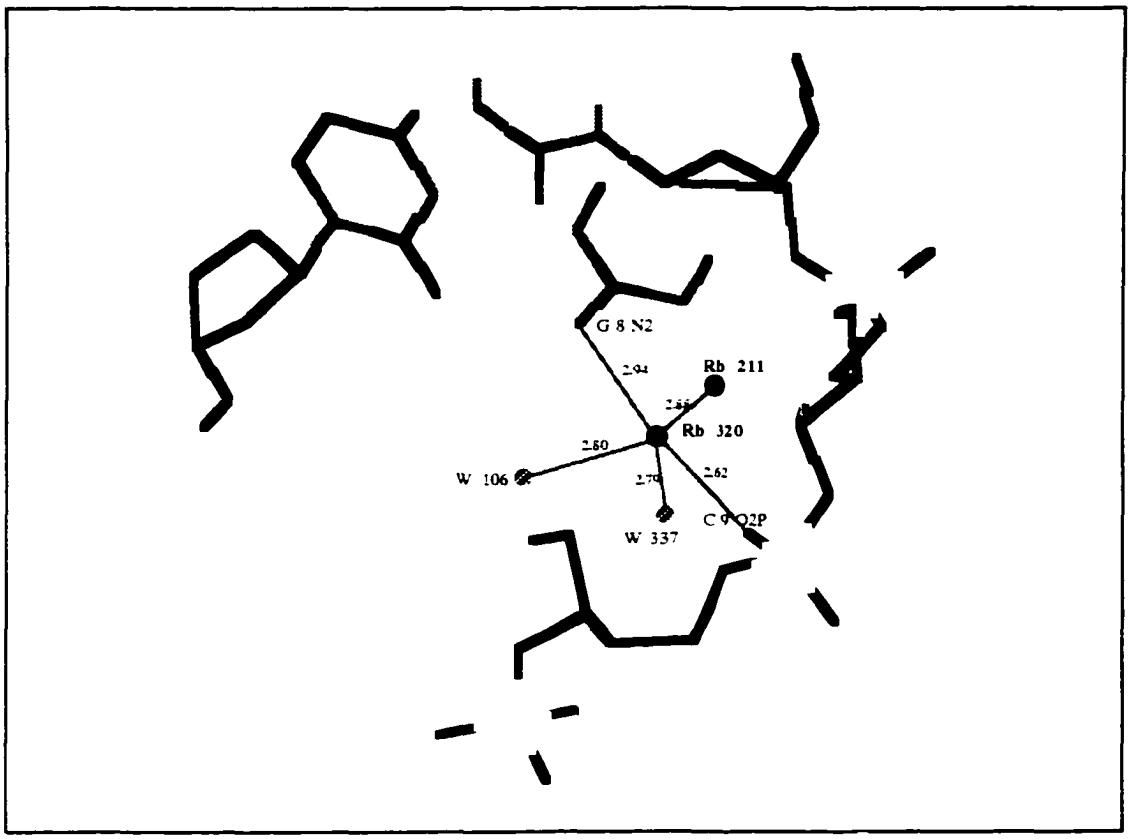


Fig. 3-16 Coordination of Rb no.320 in the Z-DNA Structure

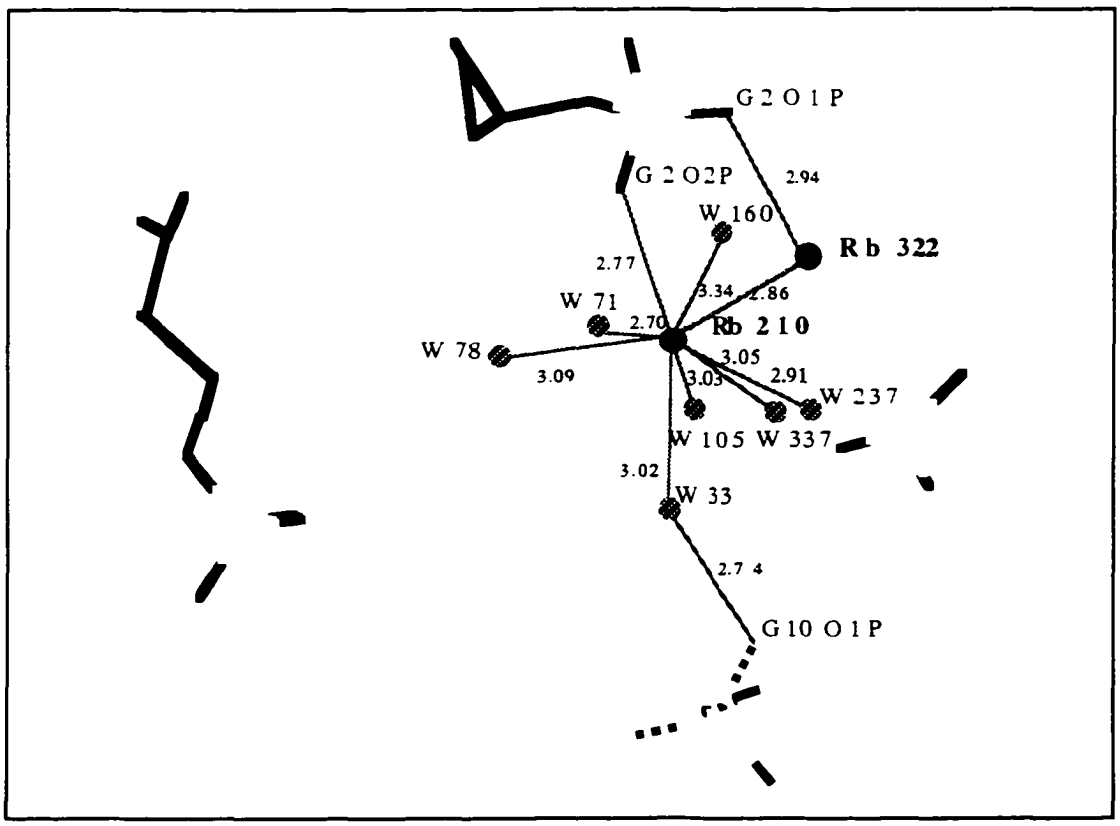


Fig. 3-17 Coordination of Rb no.210 in the Z-DNA Structure

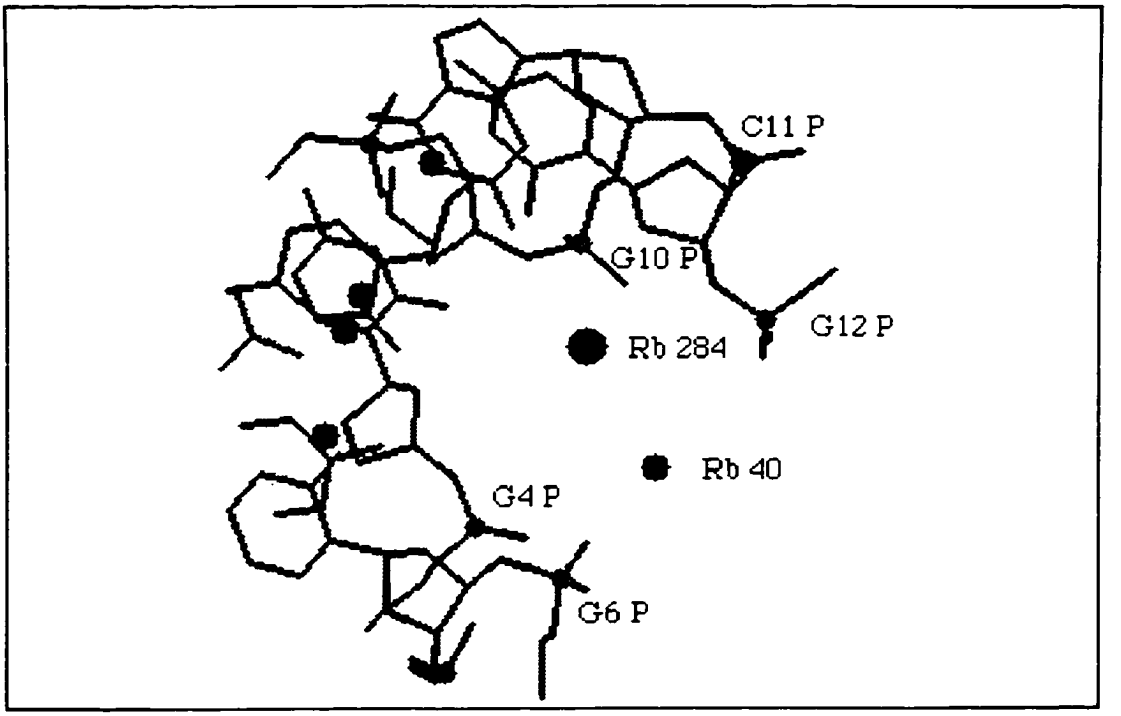


Fig. 3-18A Rb no. 284 inside the Groove of DNA Duplex

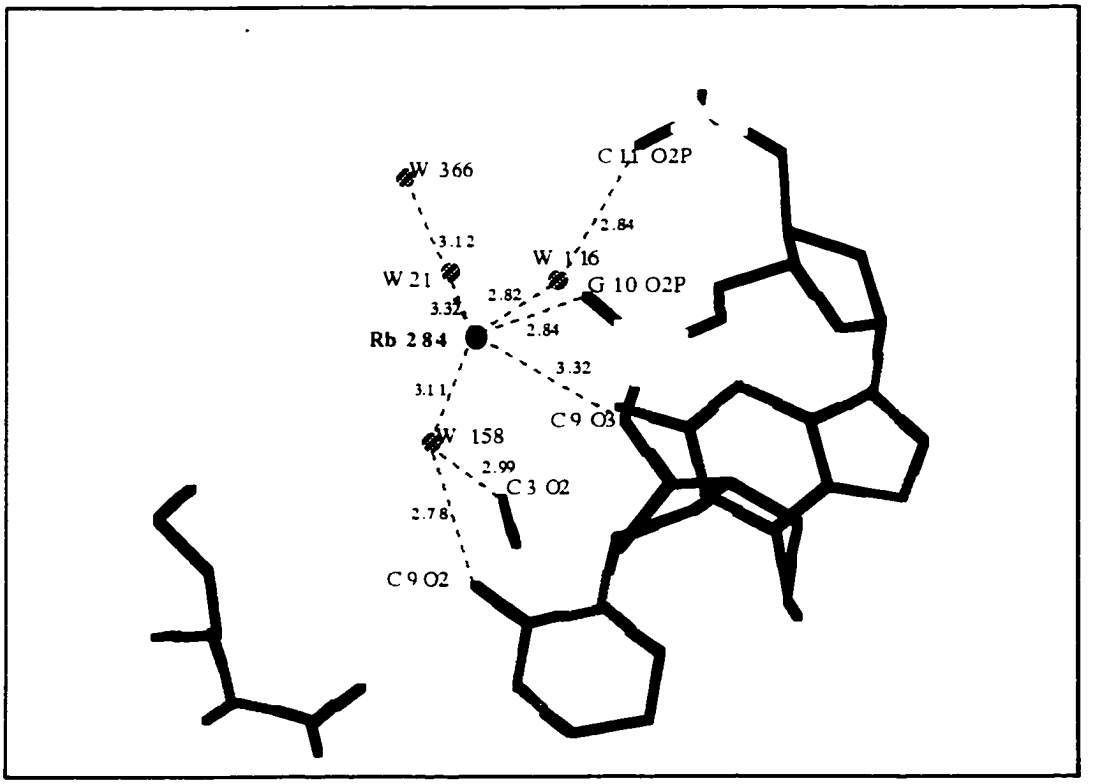


Fig. 3-18B Coordination of Rb no. 284 in Z-DNA structure

3). The Chloride Ion Assignment

One strong electron density peak was found to be within a distance of about 3.0 Å from both Rb no.88 and Rb no.41 (Fig. 3-6 and Fig. 3-19). Since Rb no.88 is fully occupied this peak can not be assigned as another partial occupied Rb^+. Considering that RbCl was used in the crystallization and the Cl^- can be involved the crystal, chloride ion was assigned to this peak. After refinement the B factor for Cl^- is 26.49 $Å^2$ with full occupancy.

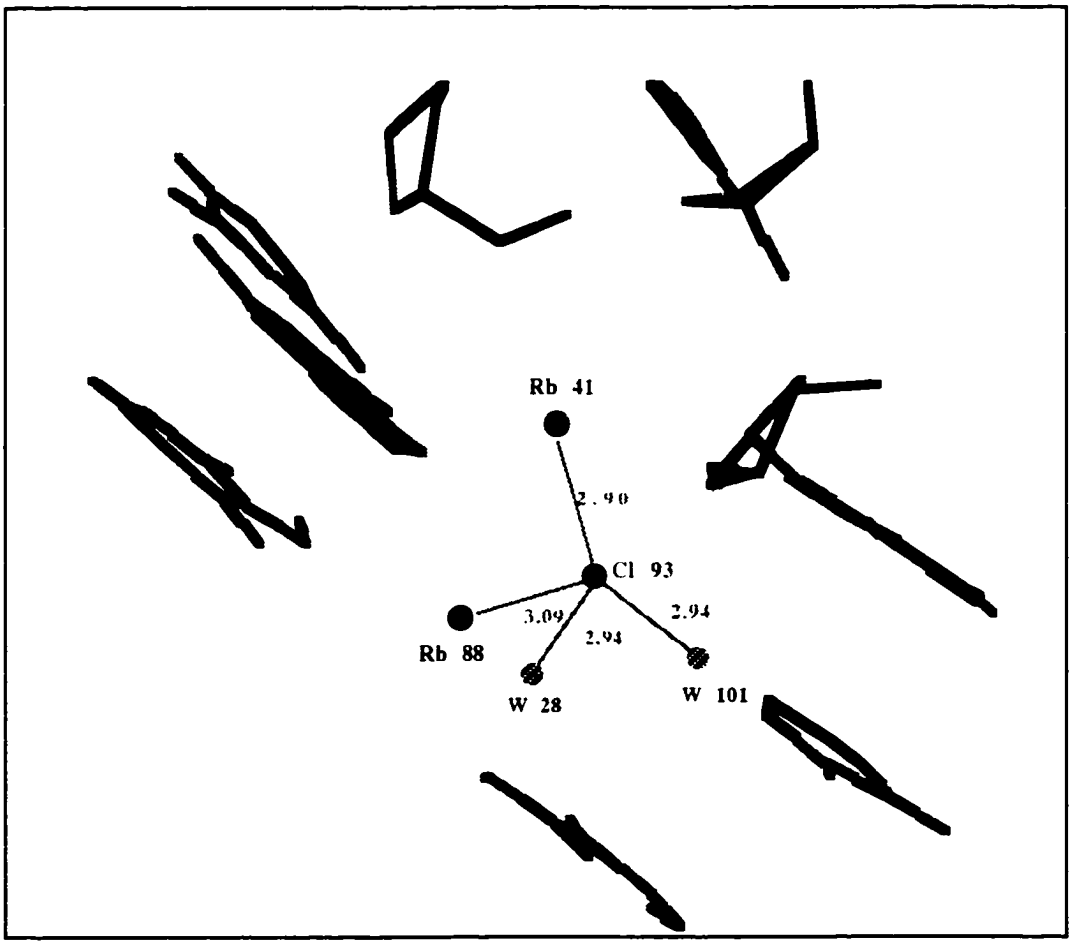


Fig. 3-19 Chloride Ion in the Rb-Z-DNA Structure

The result of this work indicates that in the Z-DNA structure most mono-valent cations are mobile. The mobility of the ions depends on the DNA structure, the position of the ions and temperature. This result agrees with the conclusion of the Matthew Young' molecular dynamic simulations which is the structures of counterion in DNA are not static but dynamic. The cations exchange with water molecules, so most of them can only have fractional occupancy.

References

1. W. Saenger, *Principles of Nucleic Acid Structure*, (Springer-Verlag New York Inc., 1984).
2. G. J. Quigley, M. M. Teeter, A. Rich, *Proc. Nat. Acad. Sci. USA* **75**, 64-68 (1978).
3. R. V. Gessner, G. J. Quigley, A. H. J. Wang, G. A. van der Marel, J. H. van Boom, A. Rich, *Biochemistry* **24**, 237-240 (1985).
4. C. A. Frederick, M. Coll, D. M. G. A. Van, B. J. H. Van, A. H. J. Wang, *Biochemistry* **27**, 8350-8361 (1988).
5. Y.-C. Jean, Y.-G. Gao, A. Wang, *Biochemistry* **32**, 381-388 (1993).
6. Y.-G. Gao, M. Sriram, A. Wang, *Nucleic Acids Res* **21**, 4093-4101 (1993).
7. T. F. Kagawa, B. H. Geierstanger, A. H. Wang, P. S. Ho, *J. Biol. Chem.* **266**, 20175-20184 (1991).

8. R. G. Brennan, E. Westhof, M. Sundaralingam, *J. Biomol. Struct. Dyn.* **3**, 649-666 (1986).

9. P. S. Ho, C. A. Frederick, D. Saal, A. H. J. Wang, A. Rich, *J. Biomol. Struct. Dyn.* **4**, 521-534 (1987).

10. M. A. Young, B. Jayaram, D. L. Beveridge, *J. Am. Chem. Soc.* **119**, 56-69 (1997).

11. D. Gewirth, Z. Otwinowski, W. Minor, *The HKL Manaul* (Department of Molecular Biophysics and Biochemistry, Yale University, Edition 4, 1995).

12. D. Gewirth, Z. Otwinowski, *The Scalepack Manaul* (Department of Molecular Biophysics and Biochemistry, Yale University, 1995).

13. R. V. Gessner, C. A. Frederick, G. J. Quigley, A. Rich, A. H. Wang, *J Biol Chem* **264**, 7921-35 (1989).

14. M. Egli, L. D. Williams, Qi Gao, A. Rich, *Biochemistry* **30**, 11388-11402 (1991).

15. D. Bancroft, L. D. Williams, A. Rich, M. Egli, *Biochemistry* **33**, 1073-1086 (1994).

16. Shun-le Chen, A Dissertation for the Degree of Doctor of Philosophy (CUNY, 1993)

17. J. L. Fulton, D. M. Pfund, S. L. Wallen, M. Newville, E. A. Stern, Yanjun Ma, *J. Chem. Phys.* **105**, 2161-2166 (1996).

Chapter 4

I. Crystallization and Preliminary X-ray Analysis of AHMA-d(CGTACG) Complex

II. Structure Analysis of AHMA

I. Crystallization and Preliminary X-ray Analysis of AHMA-d(CGTACG)Complex

4. 1 Background

3-(9-acridinylamino)-5-(hydroxymethyl) aniline (AHMA) is an antitumor drug designed and synthesized by Dr. Tsann-Long Su as part of the research group of Dr. K. Watanabe at Sloan-Kettering Institute for Cancer Research (1).

As a derivative of the clinically used antileukemia drug m-AMSA (4'-(9-acridinylamino)methane-sulfon-m-anisidide or amsacrine), AHMA (Fig. 4-1) eliminates the metabolic degradation pathway of AMSA. This is a consequence of its meta substitution on the anilino ring which prevents it from forming an iminoquinone intermediate upon oxidation. Biological study has shown that AHMA not only has greater efficiency against murine leukemia and solid tumors than m-AMSA or the related drug VP-16 (4'-demethylepipodophyllotoxin-9-(4,6-O-2-ethylidene-β-D-glucopyranoside), but also has longer half-life in human plasma (1, 2). AHMA was also found to inhibit Topo II (topoisomerase II) - mediated DNA decatenation and relaxation (3).

AHMA m-AMSA

Fig. 4-1 Structure of AHMA and m-AMSA

Amsacrine analogs are among the most conspicuous clinical drug in the series of antitumor acridines. Since m-AMSA was reported in 1974 with high activity against L1210 leukemia (4, 5), structure-function relationship of it and its analog have been extensively studied (6-12). In the 1980's Nelson and Pommier proposed that the antitumor activity of m-AMSA is related to the topoisomerase II mediated cleavage of DNA (13, 14). Studies by Bernd Granzen *et al.* indicated that the effect of AMSA on the formation of the DNA-topoisomerase II complex is biphasic — first increasing then decreasing the formation of the complex with increasing drug concentration (11). The hypotheses for the inhibitory action of AMSA for the enzyme include: 1. The inhibitor(AMSA) induces and stabilizes the cleavable complex formed by DNA and the Topo II enzyme and interrupts the DNA strand scission, passing, and resealing (15-18). 2. The intercalation of the drug in the DNA unwinds the DNA strands, making the recognition between DNA and enzyme more difficult (11).

Numerous studies have shown that 9-anilinoacridine derivatives bind to DNA in vitro (7, 19). Some studies indicate that the binding of m-AMSA has a preference for alternating purine-pyrimidine polymers (20) with no obvious difference between poly(dA-dT).poly(dA-dT) and poly(dG-dC).poly(dG-dC) (9, 20, 21). Surface-enhanced Raman spectroscopy studies indicate two kinds of interaction of AMSA with the DNA-enzyme complex: non-specific (via the acridine moiety) and specific (via the side chain of the drug) (22). Since the size of the substituents on the NH_2 of the anilino ring has little effect on the cytotoxicity and since the mutational alteration of type II topoisomerases can cause drug resistance, the AMSA analog is postulated to intercalate in the DNA base stack with its acridine moiety and interact with the enzyme with its anilino ring (1, 9, 23). Computer modeling has been used to search for the possible binding mode of m-AMSA (9, 24). But no crystal or NMR structure of a DNA-drug complex has been reported to date.

X-ray crystallographic analysis of the AHMA-DNA complex will contribute to the structure-function relationship study of acridine analogs and to the development of new antineoplastic reagents.

4. 2 Experimental

DNA Synthesis and Purification

AHMA•HCl salt was provide by Dr. Tsann-long Su and showed a single spot on TLC analysis. A DNA oligomer with sequence of CGTACG was synthesized by the Synthesis and Sequencing Laboratory at Hunter College and purified and detritylated with a PureDNA column (Rainin Instrument).

Crystallization

The vapor diffusion with sitting drop method was employed for the crystallization. Using the Grid Search method, extensive screening for the crystallization condition was conducted. Table 4-1 lists the conditions which have been studied.

Table 4 -1 The Parameters Screened for AHMA-DNA complex Crystallization

parameter	variation
Ratio of DNA & AHMA	1: 1, 2 : 1, 1 : 2
Conc. of DNA & AHMA (mM)	1.5 2.0 2.5 3.0 3.5 4.0 6.0
Conc. of $MgCl_2$	0 10 15 20 25 30 35 40 45 50 55 60
Conc. of Spermine	0 5 10
Buffer pH	4.5 5.0 5.5 6.0 6.5 7.0 7.5 8.0
Conc. of MPD (v/v) in equilibration solution	20% to 55% with interval of 5% or finer
Temperature	room temperature and 4 °C

PEG-400 and isopropanol have also been tried as precipitants. PEG-400 resulted in a precipitation of the solute in the crystallization drops. Isopropanol had similar or worse results (depends on other conditions) on the DNA-drug complex crystallization compared with MPD.

More than 300 conditions involving combinations of the different parameters have been tested. Without spermine the crystals grew most extensively in one or two dimensions forming long thin needles or thin plates. The best crystals that were obtained from solution without spermine were thin plates with dimensions about 0.4 x 0.3 mm^2 in area from a drop with 2 mM DNA and AHMA, 20 mM Na-cacodylate (pH 5.0 and pH 5.5), 20 or 40 mM $MgCl_2$ and equilibrated with 36% 2-methyl 2-4-pentanediol (MPD). Spermine was later used to improve the crystallization. Combinations of different concentration of spermine, salt ($MgCl_2$), DNA and AHMA and different pH's were again screened.

Under optimized conditions yellow pseudo-hexagonal crystals with size of $0.4 \times 0.4 \times 0.1$ mm^3 were grown in two to three weeks at room temperature (Fig. 4-2). The starting mother liquor contained 2.0 mM of oligonucleotide, 2.0 mM of AHMA, 20 mM of sodium cacodylate buffered to pH 6.0, 5 mM of spermine, 10 mM of magnesium chloride and 5% (by volume) MPD. The drops were equilibrated against 35% MPD.

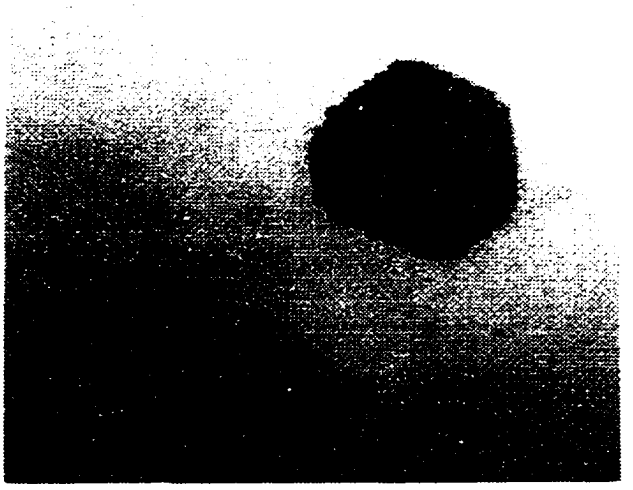


Fig. 4-2 Photograph of Crystals of AHMA-DNA Complex

The bigger crystal dimensions are about 0.35 mm × 0.35 mm × 0.07 mm

Data Collection and Unit Cell Determination

Crystals were mounted and sealed in a glass capillary tube with a droplet of mother liquor. X-ray diffraction data sets were collected with monochromatic Cu Kα radiation on an Enraf-Nonius rotating anode x-ray generator equipped with a MarResearch image plate area detector. Phi-scans of 180° in increment of 0.5° were used for data collection. The distance between the crystal and the image plate was 120 mm. Fig. 4-3 shows one image "film" of the data set.

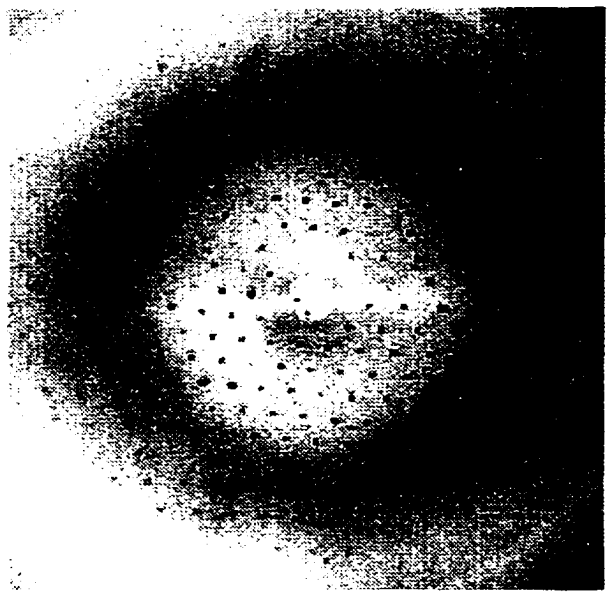


Fig. 4-3 Photograph of One Image Film in the AHMA-DNA Data Set

A total of 5074 unique reflections were collected with 4314 reflections having intensity larger than 2 sigma. The image files were processed with the Denzo program (25). The scalepack program (26) was utilized for data reduction.

The unit cell dimension of the DNA-AHMA crystal was found to be a=b=57.52 Å, c=122.17 Å with space group of the p321 class (P321, $P3_121$, $P3_221$).

Table 4-2 shows the statistics of the data set.

Table 4-2 AHMA-DNA Complex Crystal Data Statistics

A. Summary of I/Sigma in resolution shells

Shell Lower limit	Shell Upper limit	I/Sigma in resolution shells: % of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
30.00	6.02	1.9	3.9	5.6	6.3	7.6	10.7	28.1	63.9	92.1
6.02	4.78	0.2	1.3	1.8	2.1	3.0	6.5	28.3	67.9	96.2
4.78	4.18	0.0	0.6	1.3	2.5	3.3	8.7	39.0	56.2	95.2
4.18	3.80	0.8	1.8	2.6	4.0	5.5	16.1	58.0	37.5	95.5
3.80	3.53	0.7	2.3	4.8	7.4	14.0	34.0	76.9	14.5	91.4
3.53	3.32	1.6	4.0	10.0	14.2	23.8	46.8	78.0	6.5	84.5
3.32	3.15	3.1	9.7	16.9	25.8	41.2	67.3	85.8	1.9	87.6
3.15	3.02	4.5	16.6	31.2	43.8	65.3	82.9	87.8	0.0	87.8
3.02	2.90	12.5	29.7	7.5	57.3	73.2	81.4	82.6	0.0	82.6
2.90	2.80	0.7	2.3	3.3	3.8	3.8	3.8	3.8	0.0	3.8
All hkl		2.6	7.2	12.3	16.4	23.6	35.1	56.2	25.7	81.9

B. Summary of reflection intensities and R-factors by shells

R linear = SUM (ABS(I - <I>)) / SUM (I)
R square = SUM ((I - <I>) ** 2) / SUM (I ** 2)
Chi**2 = SUM ((I - <I>) ** 2) / (Error ** 2 * N / (N-1)))
In all sums single measurements are excluded

Shell Lower limit	Shell Upper Angstrom	Average I	Average error	Norm stat.	Chi**2	Linear R-fac	Square R-fac
30.00	6.02	21213.0	839.2	616.7	3.887	0.088	0.102
6.02	4.78	16850.3	664.3	498.9	3.563	0.092	0.111
4.78	4.18	14627.8	650.4	516.8	3.525	0.101	0.121
4.18	3.80	12637.9	647.3	533.6	3.293	0.109	0.115
3.80	3.53	6991.3	505.4	450.5	2.803	0.141	0.141
3.53	3.32	6259.5	538.1	485.0	2.504	0.174	0.163
3.32	3.15	3495.4	420.9	395.4	2.154	0.232	0.147
3.15	3.02	1350.8	357.9	353.6	1.738	0.459	0.550
3.02	2.90	779.4	352.5	350.5	1.449	0.708	0.578
2.90	2.80	409.2	476.0	475.5	0.893	0.000	0.981
All reflections		9853.7	564.2	472.8	2.919	0.112	0.112

The crystals diffracted to a resolution of 2.9 Å. However, beyond 3.1 Å, 40% of the reflections have intensity less than 2 sigma. The data was 88% complete at 3.15 Å resolution with an R_{merge} of 0.108. In the resolution shell of 3.32 to 3.15 Å, 83% of the reflections have intensities greater than 2 σ.

Attempts to process the data based on the assumptions of other related symmetries such as a P6 or P312 clearly excluded these possibilities by their much poorer and clearly unacceptable statistics. The lower symmetry P3 did not show significantly better statistics than the P321. Consideration of crystal packing and the apparent presence of only 0,0,ℓ reflections with ℓ = 3n (integer multiples of 3) suggest that the space group is either $P3_121$ or $P3_221$ which are indistinguishable from the data alone. These space groups have 6 asymmetric units per cell. Considerations of cell volume. likely density and volume of the similar drug-DNA complexes lead to the conclusions that each asymmetric unit contains 3 to 5 drug-DNA duplex unit (see section 4.4 for more detail).

4. 3 Heavy-atom Method Trials

Due to the multimolecular nature of the asymmetric unit and the complexity of initial molecular replacement results, it was decided to attempt the heavy-atom method for solving the structure. In this method a heavy atom has to be soaked into the crystal or bound to the crystal structure during the crystallization. The first heavy atom that was tried was

samarium(III). The first attempt was to involve Sm^{+3} in the initial crystallization solution. Different Sm^{+3} concentrations (from 5 to 80 mM) in the initial crystallization drops were tested. Crystals were obtained with similar shape and size as the crystal without Sm^{+3}. Data was collected with the same condition mentioned above. After data collection and processing, isomorphous replacement and anomalous scattering methods from the CCP4 program (see chapter 2) were used to try to locate the heavy atoms. Unfortunately, no Sm^{+3} were visualized. Soaking the native crystals with Sm^{+3}, Cu^{+2}, Au^{+2} (concentration from about 1.0 to 4.0 mM) was then attempted. After soaking with Sm^{+3}, the crystal appeared somewhat striated. A crystal was selected and mounted for the data collection. The data collection and processing methods were the same as for the native crystals. After data processing, a smaller unit cell (a=b=28.54 Å, c=118.41 Å), but apparently the same symmetry as the native one was obtained. Due to the high mosaicity of the data set, probably resulting from the transformation of the unit cell, and resultant spot overlap, it was impossible to extract high quality merged data.

Soaking with Cu^{+2} once produced the small unit cell but the second time the crystals did not diffract. The Au^{+2} soaking destroyed the crystal diffraction.

4. 4 Preliminary X-ray Analysis

Patterson maps were calculated with the data from the native DNA-AHMA complex and from the Sm^{3+} soaked crystals (small unit cell) using the CCP4 program. The Chain program was used to examine the Patterson map. A dummy molecule with distance of 14.0 Å between tow atoms (four times the base stacking space) was used to check base stacking in the structure. A class of strong reflections around 3.4 Å resolution and with small *h* and *k* values lead to a pattern of strong peaks in the Patterson map suggested that the drug and base stacking directions are nearly perpendicular to the c-axis of the unit cell with the DNA axis lying more or less parallel to the xy plane (Fig. 4-4).

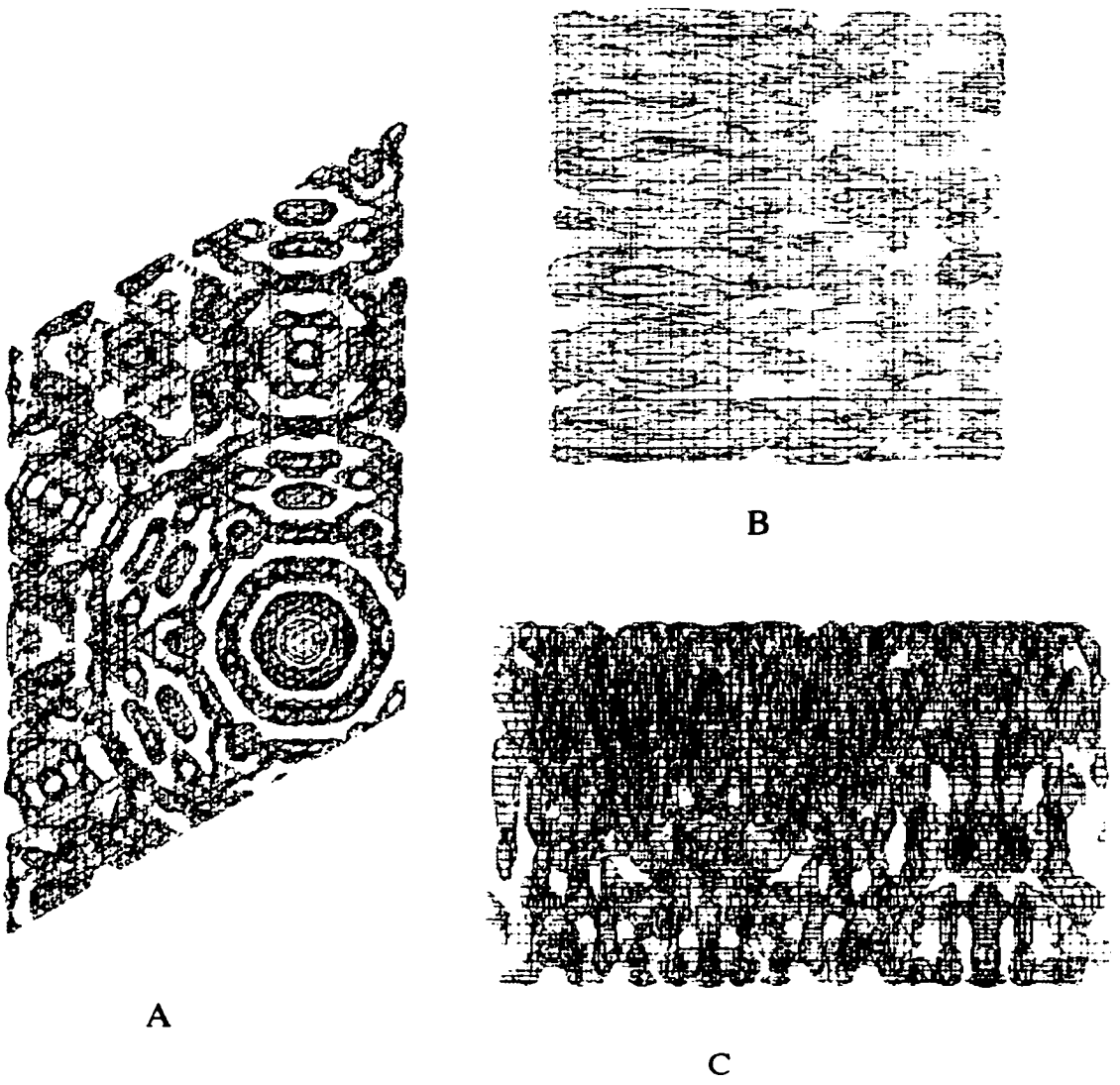


Fig. 4-4 Patterson map of AHMA-DNA complex structure (native crystal)
A. Map viewed down z-axis B. Map viewed down x-axis
C. Map viewed down y-axis

The red dots in the A represent the atoms which have distance of 14.00 Å from each other in a dummy molecule

As mentioned previously the P321 class of space groups has six asymmetric units per unit cell. If we assume a crystal density in the range of 1.1 to 1.4 g/ml and a solvent content of 40 - 70% by volume, or if we use the volume of other similar drug-DNA complexes (such as daunomycin-DNA (27) which is 10230 $Å^3$), we are led to the conclusion that there should be between three and five drug-DNA complex per asymmetric unit in the large unit cell but only one in the small cell. The fact that the transformations from the large to smaller unit cell takes place without gross disruption of the crystals or a change in the space group suggests that the large cell has four molecules per asymmetric unit and that in the large unit cell they are packed in such a manner that with only slight rearrangement, each of two pairs aligns to form perfect repeats leading to the observed smaller unit cell with a and b dimensions half of those of the large unit cell. The schematic drawing of the dimensions of one possible AHMA-DNA complex is shown in Fig. 4-5.

A crystallographic search procedures (using the X-plor and CCP4 programs) were used to attempt to find the suitable starting structure. The search model was constructed by replacing the daunomycin of the daunomycin-DNA structure with the AHMA molecule and carrying out local energy minimization to optimize the drug-DNA model. No satisfying result has been obtained to date.

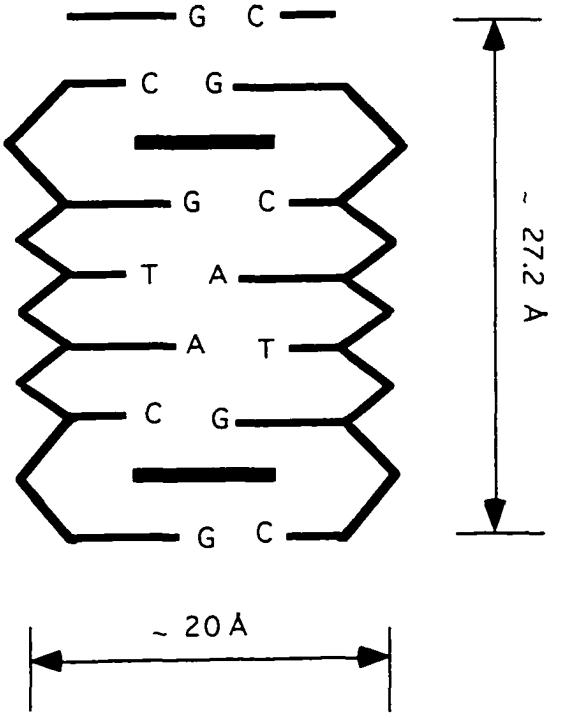


Fig. 4-5 Schematic Diagram of Estimated Dimension of One Duplex of AHMA-d(CGTACG)

The intercalating drugs are shown in black bars. DNA is drawn in thinner lines. The diameter of one duplex is about 20 Å. The space between two base pair stacking is about 3.4 Å. Eight spacing (including intercalating drug) is about 27.2 Å.

II. Crystallization and Structure Determination of AHMA by X-ray Crystallography

To assist in understanding the conformation of AHMA binding to DNA, AHMA was crystallized and its structure was analyzed by x-ray crystallography.

4. 5 Experimental

Crystallization of AHMA

AHMA·HCl was provided by Dr. Tsann-Long Su and showed a single spot on a TLC plate.

Conditions for crystallization were screened using both slow evaporation of solvent and the vapor diffusion methods. Variation of buffer pH, concentration of salt and concentration of AHMA and precipitant were examined.

Single crystals with marginal size for data collection were grown with the vapor diffusion method from a solution of 3.0 mM AHMA·HCl, 20 mM Na-Cacodylate (pH 8.0), 5% isopropanol (by volume) and equilibrated against 8% isopropanol.

Data Collection and Unit Cell Determination

An amber colored diamond shaped crystal with dimensions of about 0.30×0.30×0.10 mm^3 was mounted on a glass fiber. An Enraf-Nonius diffractometer and rotating anode x-ray generator with Cu Kα radiation was used for data collection. The data was collected at room temperature.

A total of 3490 reflections were collected, of which 2648 were unique. The unit cell was found to be triclinic with dimensions of a = 8.04 Å, b = 9.95 Å, c = 11.63 Å and α = 69.78°, β = 84.02°, γ = 65.70°. The space group is $P\bar{1}$.

Data Reduction

The Molen package (see chapter 2) was used for data processing including data correction for decay and absorption effects. The slope of the least-squares line through a plot of standard reflection intensities versus time indicated a total loss in intensity of 0.4%. Linear decay, Lorentz and polarization corrections were applied to the data set. An empirical absorption correction based on a series of psi-scans was also applied to the data. Intensities of equivalent reflections were averaged.

Structure Solution and Refinement

The Multan program (within the Molen package), a direct methods procedure, was employed for structure determination. After the first Multan run, an E-map with a "chicken wire" pattern appeared. A total of 25 electron density peaks were located. Appropriate peaks were selected as

certain atoms. The structure was refined with full-matrix least-square techniques. The minimized function is :

$$\sum \text{weight}_i * (| F_{o,i} | - | F_{c,i} |)^2$$

In the equation, $F_{o,i}$ is the ith observed structure factor, $F_{c,i}$ is the ith calculated structure factor.

New E-maps and difference Fourier maps were calculated revealing additional atoms till the structure was complete except for the hydrogen atoms. After all the non-hydrogen atoms in the structure were located, one significant peak remained which was assigned as an oxygen of a water molecule.

Hydrogen atoms were included in the later refinement and restrained to ride on the atoms to which they are bonded.

Atom names used in the structure file are shown in Fig. 4-6.

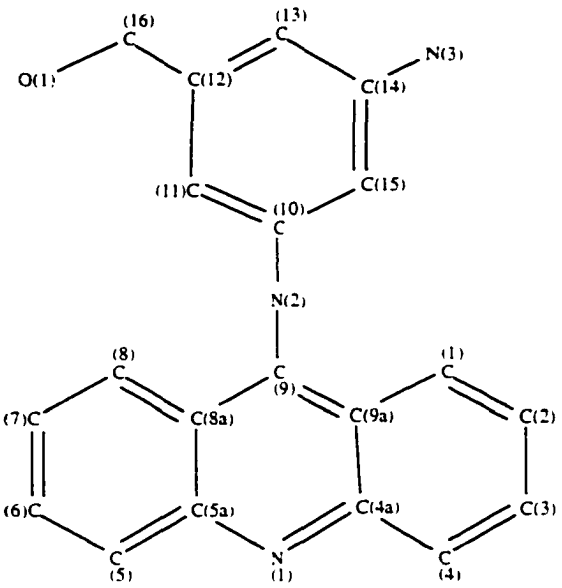
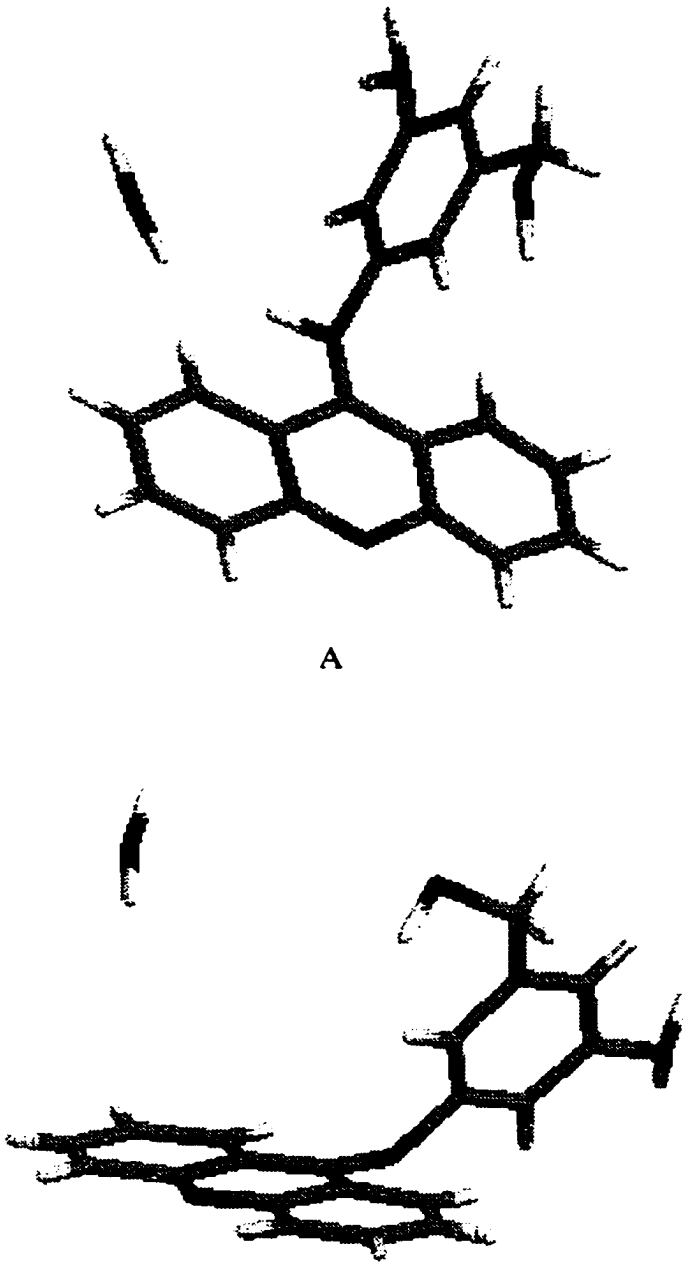


Fig. 4-6 Atom Names in the AHMA Structure File

4. 6 Results and Discussion

The final solution is a structure with an unweighted R value of 10.6%. Although it is a somewhat high R value there is no significant difference electron density in the electron density map. The higher R value may result from the disordered water molecule in the structure as well as the poor crystal quality. The marginal crystal quality itself may reflect a nonuniformity of hydration within the crystal lattice.

The 3-Dimensional structure and packing picture of AHMA are shown in Fig. 4-7 and Fig. 4-8.



A

B

Fig. 4-7 Structure of AHMA

A. View 1 B. View 2
In the figure, carbon atoms are drawn in dark gray, nitrogen in dark blue, oxygen in red and hydrogen in light gray.

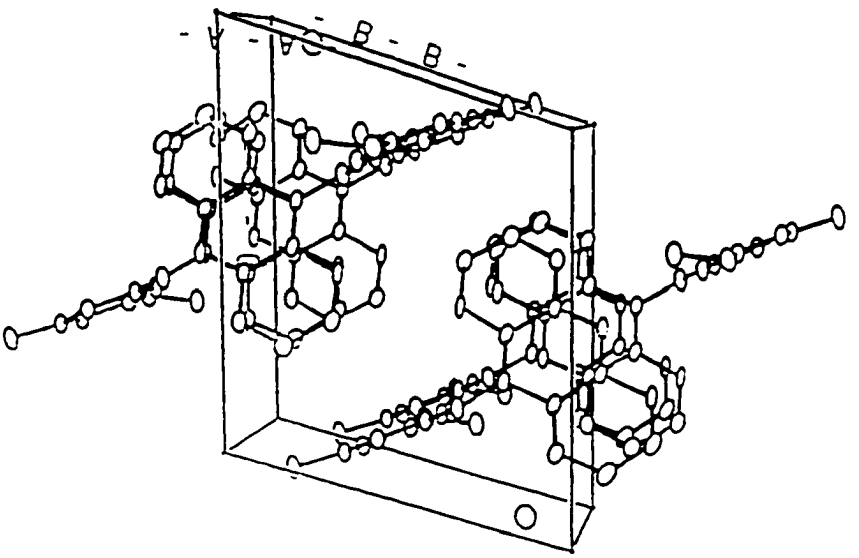


Fig. 4.8 Packing of AHMA in one Unit Cell

The bond distance, bond angles and dihedral angle of two planes are listed in Table 4-3, Table 4-4 and Table 4-5 respectively.

Table 4-3 Bond Distance in Angstroms in the AHMA Structure

Atom1	Atom2	Distance	Atom1	Atom2	Distance
O(1)	C(16)	1.462(8)	C(5)	C(6)	1.316(9)
N(1)	C(4a)	1.390(6)	C(6)	C(7)	1.388(8)
N(1)	C(5a)	1.326(7)	C(7)	C(8)	1.408(8)
N(2)	C(9)	1.387(6)	C(8a)	C(8)	1.400(8)
N(2)	C(10)	1.421(6)	C(8a)	C(9)	1.461(7)
N(3)	C(14)	1.395(8)	C(9a)	C(9)	1.378(8)
C(1)	C(2)	1.322(9)	C(10)	C(11)	1.40 (1)
C(1)	C(9a)	1.464(8)	C(10)	C(15)	1.371(8)
C(2)	C(3)	1.386(8)	C(11)	C(12)	1.401(7)
C(3)	C(4)	1.398(9)	C(12)	C(13)	1.392(8)
C(4)	C(4a)	1.387(9)	C(12)	C(16)	1.50 (1)
C(4a)	C(9a)	1.420(7)	C(13)	C(14)	1.40 (1)
C(5a)	C(5)	1.474(7)	C(14)	C(15)	1.393(7)
C(5a)	C(8a)	1.407(7)			
O(1)	H(01)	1.170(4)	C(5)	H(5)	0.95 (4)
O(2)	H(21)	1.15 (7)	C(6)	H(6)	0.949(6)
O(2)	H(22)	0.989(6)	C(7)	H(7)	0.96 (6)
N(2)	H(05)	1.015(5)	C(8)	H(8)	0.95 (4)
N(3)	H(31)	0.875(6)	C(11)	H(11)	0.95 (4)
N(3)	H(32)	1.056(6)	C(13)	H(13)	0.96 (7)
C(1)	H(1)	0.95 (5)	C(15)	H(15)	0.95 (8)
C(2)	H(2)	0.94 (7)	C(16)	H(161)	0.96 (6)
C(3)	H(3)	0.95 (6)	C(16)	H(162)	0.96 (5)
C(4)	H(4)	0.96 (4)			

* numbers in parentheses adjacent to the distances are estimated standard deviations in the least significant digits.

Table 4-4 Bond Angles in Degrees in the AHMA Structure

Atom 1	Atom2	Atom3	Angle	Atom1	Atom2	Atom3	Angle
C(4a)	N(1)	C(5a)	117.2(4)	C(1)	C(9a)	C(4a)	115.7(5)
C(9)	N(2)	C(10)	125.0(5)	C(1)	C(9a)	C(9)	123.3(5)
C(2)	C(1)	C(9a)	120.7(5)	C(4a)	C(9a)	C(9)	121.0(5)
C(1)	C(2)	C(3)	123.5(6)	N(2)	C(9)	C(8a)	119.3(5)
C(2)	C(3)	C(4)	118.2(6)	N(2)	C(9)	C(9a)	122.7(5)
C(3)	C(4)	C(4a)	120.6(5)	C(8a)	C(9)	C(9a)	117.9(4)
N(1)	C(4a)	C(4)	117.3(4)	N(2)	C(10)	C(11)	120.1(5)
N(1)	C(4a)	C(9a)	121.3(5)	N(2)	C(10)	C(15)	118.6(6)
C(4)	C(4a)	C(9a)	121.3(5)	C(11)	C(10)	C(150	121.2(5)
N(1)	C(5a)	C(5)	117.9(4)	C(10)	C(11)	C(12)	118.7(5)
N(1)	C(5a)	C(8a)	126.2(5)	C(11)	C(12)	C(13)	119.7(7)
C(5)	C(5a)	C(8a)	115.9(5)	C(11)	C(12)	C(16)	120.7(5)
C(5a)	C(5)	C(6)	120.9(5)	C(13)	C(12)	C(16)	119.4(5)
C(5)	C(6)	C(7)	123.1(6)	C(12)	C(13)	C(14)	120.9(5)
C(6)	C(7)	C(8)	119.0(6)	N(3)	C(14)	C(13)	119.5(5)
C(5a)	C(8a)	C(8)	121.7(5)	N(3)	C(14)	C(15)	121.5(7)
C(5a)	C(8a)	C(9)	116.5(5)	C(13)	C(14)	C(15)	118.8(5)
C(8)	C(8a)	C(9)	121.8(4)	C(10)	C(15)	C(14)	120.6(6)
C(7)	C(8)	C(8a)	119.4(5)	O(1)	C(16)	C(12)	115.1(4)

* numbers in parentheses adjacent to the angles are estimated standard deviations in the least significant digits.

Table 4-5 Dihedral Angle Between Planes

Plane No.	Plane No.	Dihedral angle
1	2	66.69 +- 0.16

* Plane No. 1 is formed by acridine ring and plane No. 2 is formed by anilino ring.

In the structure the bond distance between N(2) and C(9) is 1.387 Å and between N(2) and C(10) is 1.421 Å (Fig. 4-6 and Table 4-2) suggesting that these bonds both have double bond characters.

The hydroxyl group of the AHMA is hydrogen bonding to atoms N(1) and N(2) in a symmetry related molecule and with the H_2O molecules. N(3) is hydrogen bonded with a symmetry related H_2O.

The acridine and aniline plane are almost perpendicular to each other and the dihedral angle between the two planes is 66.5° which is similar to the other two published amsacrine analog structures (28, 29). This structure was used as the replacement for the daunomycin in constructing the search model (see section 4.4). It was shown from the energy minimization that the observed structure could easily intercalate into a B-DNA duplex with only minor rotation of the side chain to optimize packing into the DNA groove.

References

1. T. L. Su, T. C. Chou, J. Y. Kim, J. T. Huang, G. Ciszewska, W. Y. Ren, G. M. Otter, F. M. Sirotnak, K. A. Watanabe, *J Med Chem* **38**, 3226-35 (1995).
2. T.-L. Su, J.-T. Huang, K. A. Watanabe, G. M. Otter, F. M. Sirotnak, X. -B. Kong, T.-C. Chou, *Proc. Am. Assoc. Cancer Pes.*, 2241 (1993).

3. T. L. Chou, F. Leteurte, T.-L. Su, K. A. Watanabe, X. B. Kong, W. T. Beck, Y. Prommier, *Proc. Am. Assoc. Cancer Res.* , 368 (1994).
4. B. F. Cain, G. J. Atwell, *Eur J Cancer* **10**, 539-49 (1974).
5. B. F. Cain, G. J. Atwell, W. A. Denny, *J. Med. Chem.* **18**, 1110-1117 (1975).
6. B. C. Baguley, W. A. Denny, G. J. Atwell, B. F. Cain, *J Med Chem* **24**, 170-7 (1981).
7. F. Hudecz, J. Kajtar, M. Szekerke, *Nucleic Acids Res* **9**, 6959-73 (1981).
8. B. Marshall, R. K. Ralph, *Advances in Cancer Research* (Academic Presss, Inc, 1985), vol. 44.
9. K. X. Chen, N. Gresh, B. Pullman, *Nucleic Acids Res* **16**, 3061-73 (1988).
10. R. M. Wadkins, D. E. Graves, *Biochemistry* **30**, 4277-83 (1991).
11. B. Granzen, D. E. Graves, B. C. Baguley, M. K. Danks, W. T. Beck, *Oncol Res* **4**, 489-96 (1992).
12. D. P. Figgitt, W. A. Denny, S. A. Gamage, R. K. Ralph, *Anticancer Drug Des* **9**, 199-208 (1994).
13. E. M. Nelson, K. M. Tewey, L. F. Liu, *Proc Natl Acad Sci U S A* **81**, 1361-5 (1984).
14. Y. Pommier, R. E. Schwartz, L. A. Zwelling, K. W. Kohn, *Biochemistry* **24**, 6406-10 (1985).
15. E. Schneider, Y.-H. Hsiang, L. F. Liu, *Adv. Pharmacol.* **21**, 149-181 (1990).
16. M. J. Robinson, N. Osheroff, *Biochemistry* **29**, 2511-5 (1990).
17. M. J. Robinson, N. Osheroff, *Biochemistry* **30**, 1807-13 (1991).
18. J. Cummings, J. F. Smyth, *Ann Oncol* **4**, 533-43 (1993).

19. M. J. Waring, *Eur J Cancer* **12**, 995-1001 (1976).

20. W. R. Wilson, B. C. Baguley, L. P. Wakelin, M. J. Waring, *Mol Pharmacol* **20**, 404-14 (1981).

21. C. Bailly, W. Denny, L. Mellor, L. Wakelin, M. Waring, *Biochemistry* **31**, 3514-3524 (1992).

22. I. Chourpa, H. Morjani, J. F. Riou, M. Manfait, *FEBS Lett* **397**, 61-4 (1996).

23. A. Maxwell, *J Antimicrob Chemother* **30**, 409-14 (1992).

24. B. S. Glisson, A. M. Killary, P. Merta, W. E. Ross, J. Siciliano, M. J. Siciliano, *Cancer Chemother Pharmacol* **31**, 131-8 (1992).

25. D. Gewirth, Z. Otwinowski, W. Minor, *The HKL Manaul* (Department of Molecular Biophysics and Biochemistry, Yale University, Edition 4, 1995).

26 D. Gewirth, Z. Otwinowski, *The Scalepack Manaul* (Department of Molecular Biophysics and Biochemistry, Yale University, 1995).

27. G. J. Quigley, A. H. Wang, G. Ughetto, G. van der Marel, J. H. van Boom, A. Rich, *Proc Natl Acad Sci U S A* **77**, 7204-8 (1980).

28. Z. H. L. Abraham, Shirley D. Cutbush, Reiko Kuroda, Stephen Neidle, R. Morrin Acheson, Grahame N. Taylor, *Chem. Soc. Perkin Trans II* , 461-465 (1985).

29. J. M. Karle, R. L. Cysyk, I. L. Karle, *Acta Cryst.* **B36**, 3012-3016 (1980).

Chapter 5

X-Ray Restrained Molecular Dynamic Calculation for Simulation of Disordered K^+ Ions in the K-Z-DNA Structure

5.1 Method of Disorder Calculation by X-ray Molecular Dynamics

Atoms in macromolecules are in constant dynamic motion even in the crystal lattice. Disorder or multiple conformations of the structure represent the multiple potential energy wells. When there is no free energy barrier for the motion the atom will have an anisotropic distribution. When there is a well defined energy barrier the atom will have discrete displacement. In the general x-ray calculation the model we use is a static molecule which represents the average position of the atoms. Disorder weakens the diffracted x-ray intensities. The static model will not give a good fit to the electron density map and the residual value (R value) becomes higher when there are many multiple conformation sites in the structure. Disorder or dynamic motion of macromolecules has been studied by scientists including Frauenfelder (1), Northrup (2), Mao (3), Karplus (4), Smith (5), Ichiye (6), Kuriyan (7, 8) and Burling (9). Dr. John Kuriyan and his co-author (10) proposed a method for the exploration of the disorder in the protein crystal structure. They created a duplicated model in which the second molecule is related to the first one by translating the molecule by 3 or more unit cell lengths along one unit cell axis. The x-ray restrained dynamic optimization (simulated annealing refinement) was then carried out on the structure. Alternative positions or multiple conformations can be extracted from reproducible displacement of the atoms from different dynamic runs with different conditions.

Data from the K-Z-DNA and Rb-Z-DNA crystals collected at room temperature with diffractometer did not reveal K^+ or Rb^+ ions whereas there are many disordered solvent areas in the structure. Accordingly Dr. Kuriyan's method was used on the K-Z-DNA data in an attempt to get information for the distribution of the monovalent cations in Z-DNA structure.

5. 2 Crystallization of K-Z-DNA Crystal and Data Collection

The vapor diffusion method with sitting drops was used for the crystallization. The starting mother liquor contained 6.0 mM oligonucleotide (with sequence of CGCGCG), 40 mM potassium cacodylate (pH 7.0), 300 mM potassium chloride, 5%(v/v) MPD, equilibrated with a equilibration solution containing 400 mM potassium chloride and 60% MPD. Single crystals with a suitable size for data collection were grown in three to six days at room temperature.

A crystal with dimensions of 0.5×0.3×0.3 mm^3 was mounted in a capillary tube with a droplet of mother liquor. An Enraf-Nonius CAD4 (computerized Automatic Diffractometer with four independent angular motions) diffractometer with Cu Kα radiation was used for the data collection. A total of 5391 reflections with resolution of 1.09 Å were collected. The Molen program (see chapter 2) was used for the intensity decay and Lorentz-Polarization factor correction. The crystal unit cell was

found to have dimension of a = 17.98 Å, b = 31.09 Å, c = 44.91 Å with space group $P2_12_12_1$.

5.3 Structure Refinement

The Mg-Z-DNA structure was used for the initial trial model. After the positional and isotropic B factor refinement the R value was 25.2%. Then the overall anisotropic B refinement was conducted and the result was use for scaling the intensity data. With the scaled data the R value dropped to 20.3%. The scaled intensity was used for the rest of the refinement.

The average B value for the DNA in the structure is about 7.5 $Å^2$. There are two H_2Os having B values of about 10. These two positions correlate to positions of Rb no.88 and Rb no.32 in the Rb-Z-DNA structure (Table 5-1).

Table 5-1 Comparison of the Positions of Two H_2Os with Low B factors in the K-Z-DNA Crystal with Rb^+ Positions in the Rb-Z-DNA Structure

structure	atom & residue no.		coordination			B factor	coordinated atoms
K-Z-DNA	H_2O	35	18.44	23.68	12.69	10.23	G6O4' & G10O1P#
(room T)	H_2O	88	5.73	10.70	2.35	9.94	G2N7 & C11O1P#
Rb-Z-DNA	Rb	32	19.64	23.29	10.45	13.55	G6O5' & C9O1(2)P#
(low T)	Rb	88	5.92	10.39	2.21	17.80	G2N7 & C11O1P#

* The atoms with # sign are those in symmetry related molecules.

There are some disordered solvent areas in the K-Z-DNA structure. In these areas two or three H_2Os have distance less than 2.0 Å. Many of these locations are close to the positions of disordered rubidium ions in the low temperature Rb-Z-DNA structure.

5. 4 X-ray Restrained Dynamic Optimization

Determination of the Initial K^+ Ion Trial Positions in the Structure

The initial coordination of ten K^+ ions were calculated so that they are within the hydrogen bonding distance of the ten negatively charged oxygen of the phosphate groups in one duplex. The comparison of the position of the trial K^+ and Rb^+ in the low temperature structure are listed in the Table 5-2.

Table 5-2 Initial Positions of the K^+ ions and Correlated Rb^+ in the Rb-Z-DNA Structure

atom	coordination			coordinated atoms	position related Rb^+	distance between two position(Å)
K13	7.98	20.60	-6.17	G2OP		
K14	8.28	23.12	0.95	C3OP	Rb34	3.83
K15	14.56	22.50	2.01	G4OP		
K16	13.39	25.81	12.78	C5OP		
K17	20.09	18.53	9.37	G6OP	Rb32	4.91
K18	13.56	8.06	16.82	G8OP		
K19	14.75	5.26	9.03	C9OP	Rb32	3.7
K20	19.99	12.24	9.51	G10OP	Rb284	4.73
K21	22.40	13.39	2.56	C11OP	Rb88	3.29
K22	19.97	19.44	1.94	G12OP		

One thing to be mentioned is the real K^+ cation may not be in the same position of the Rb^+ since :

1. The unit cell are little different between two structures.
2. The cation radii are somewhat different (1.33 Å for the K^+ and 1.48 Å for the Rb^+).

Creation of the Multiple Structure Models for Dynamic Optimization

The K-Z-DNA structure from the refinement above with the trial K^+ positions was used for the initial structure. Two models were created : twin and quartet. Table 5-3 lists the method for creating the models.

Table 5-3 Creation of the Multiple Structure Models

model	number of molecules in one model	method for creating multiple structure models
twin	2	translating the first molecule 3 times the c dimension length along c axis.
quartet	4	molecule 2: translating the first molecule 3 times the c dimension length along c axis. molecule 3: translating the first molecule 3 times the b dimension length along b axis. molecule 4: translating the first molecule 3 times the c dimension length along c axis and 3 times the b dimension length along b axis.

Dynamic Optimization

Dynamic simulation was carried out on the single, twin and quartet structures with the X-plor program. Coordination of DNA (except the phosphate groups) was fixed with constraints during the simulation. The velocity rescaling option was employed for temperature control.

Initial velocities were randomly chosen compatible with the desired temperature. Different duration times (1ps, 2ps and 3ps) were used in the heating stage in different dynamic run with 0.001 ps in each step. Different heating temperature (2000, 6000 K) were tested. Each heating stage was followed by fast cooling, then positional and B-factor refinement.

In the optimizations the molecules in one structure didn't interact with the others.

The optimization target function is :

$$E = E_{structure} + E_{x\text{-}ray}$$

and $$E_{x\text{-}ray} = W_A \Sigma (|F_o| - |F_c|)^2$$ (see chapter 2)

5. 5 Results and Discussion

Unfortunately, different dynamic runs did not yield reproducible K^+ locations. Within one model, each K^+ also has different positions in positionally translated molecules.

Disordered Phosphate Groups

Two disordered phosphate groups (C3 and C5) were found in the calculation. In the quartet structure, three molecules are in one conformation and one molecule in the other conformation for both C3 and C5 residues (Fig. 5-1).

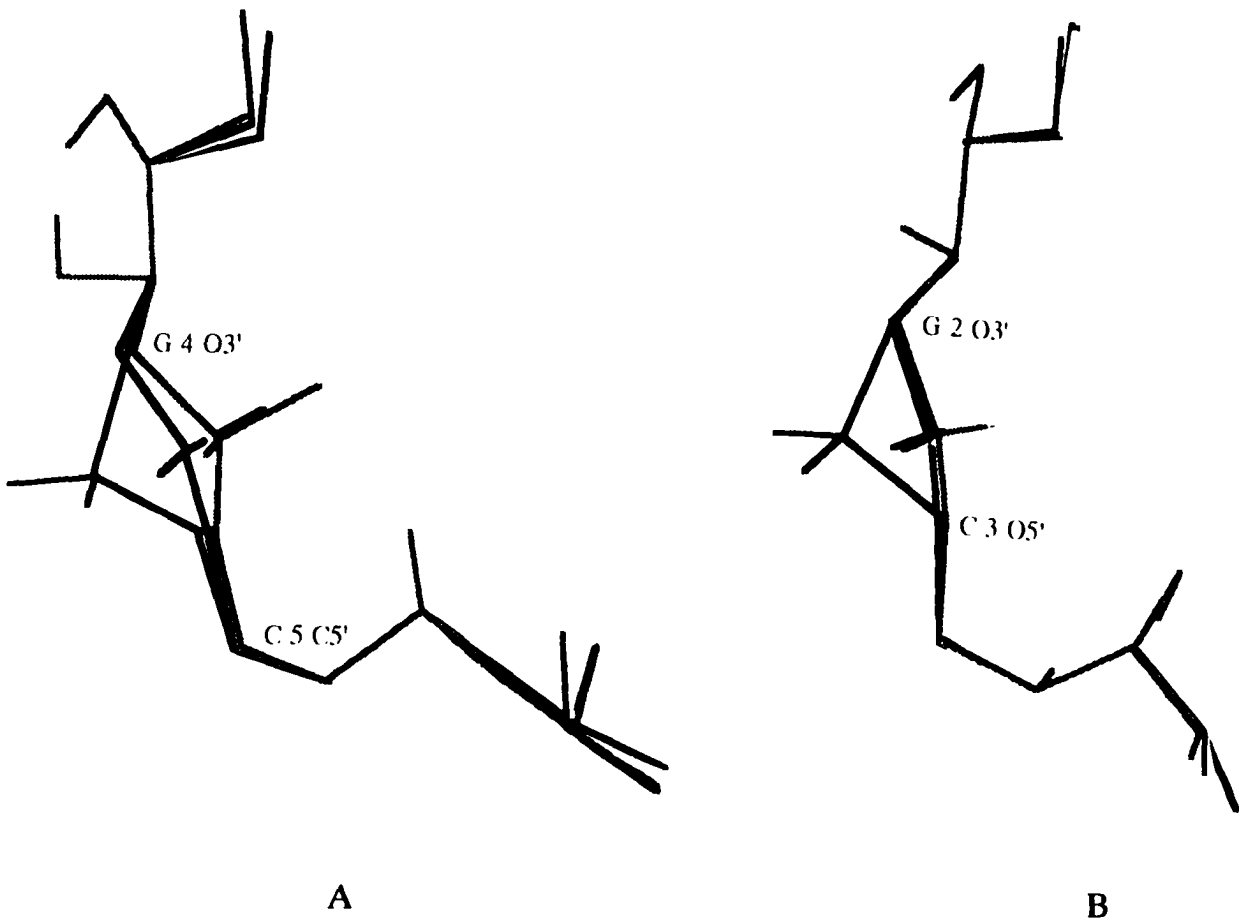


Fig. 5-1 Disorder of the Phosphate Groups in the K-Z-DNA Structure
A. Phosphate group on Cytidine 5 B. Phosphate group on Cytidine 3

Comparison of K^+ positions from Simulation in the K-Z-DNA Structure with the Rb^+ Positions in the low Temperature Rb-Z-DNA Crystal

In the single structure model only one K^+ (K14) has a low B value in all dynamic runs. The position of this K^+ is close to Rb no.111 in the Rb-Z-DNA structure. For the twin structure model, position of K17 is close to the position of Rb no.284 in the Rb-Z-DNA structure. In the quartet structure two K^+ positions with lower B factor were found which correlate with positions of Rb no.32 and Rb no.88 in the Rb-Z-DNA structure (Table 5-4).

Table 5-4 Coordinations of K17, K21 and related Rb^+ Positions in the Rb-Z-DNA Structure

atom	coordination			occupancy	B factor	correlated position in Rb structure
K17	18.347	22.061	9.336	1.00	72.95	Rb 32
K21	23.780	10.636	2.068	1.00	31.25	Rb 88*

* K 21 is related to Rb 88 by translating one unit cell length along the a axis.

K 21 has a low B factor and occupies the previous H_2O no.88 position which had B value of 9.0 Å. Most probably this is the location of one ordered potassium ion.

Solvent Structure of the K-Z-DNA

There are several disordered H_2O areas in the K-Z-DNA structure. These disordered H_2O areas could result from the disturbing of the disordered K^+ ions. Table 5-5 lists these areas and their positionally correlated Rb^+ in the low temperature Rb-Z-DNA structure.

Table 5-5 Correlation of Disordered Solvent Area with the Rb^+ Positions in Rb-Z-DNA Structure

H_2O residue number and distance (Å)					
A	A-B	B	A-C	C	related position in Rb-Z-DNA structure
201	1.32	122			Rb 32
30	1.12	205			Rb 88
208	1.49	161			Rb 111. Rb 34
210	0.92	78			Rb 379
155	1.65	202			Rb 379
160	1.56	214	1.36	215	Rb 41
203	1.13	84	1.58	86	Rb 284
40	1.42	171	1.15	223	Rb 40
129	1.02	206			Rb 321

The result indicates that either the method we used is not suitable for solvent disorder calculation or that the ions and solvent are highly disordered. In this case the simulation must involve so many superpositions that it may be impossible to distinguish true simulation from noise fitting.

References

1. H. Frauenfelder, G. A. Petsko, D. Tsernoglou, *Nature* **280**, 558-563 (1979).
2. S. H. Northrup, M. R. Pear, J. A. McCammon, M. Karplus, T. Takano, *Nature* **287**, 659-660 (1980).
3. B. Mao, M. R. Pear, J. A. McCammon, *Biopolymers* **21**, 1979-1989 (1982).
4. M. Karplus, J. A. McCammon, *Ann. Re. Biochem.* **53**, 263-300 (1983).
5. J. L. Smith, W. A. Hendrickson, R. B. Honzatko, S. Sheriff, *Biochemistry* **25**, 5018-5027 (1986).
6. T. Ichiye, M. Karplus, *Proteins* **2**, 236-259 (1987).
7. J. Kuriyan, M. Karplus, G. A. Petsko, *Proteins* **2**, 1-12 (1987).
8. J. Kuriyan, W. I. Weis, *Proc. Natl. Acad. Sci. USA* **88** (1991).
9. F. T. Burling, W. I. Weis, K. M. Flaherty, A. T. Brunger, *Science* **271**, 72-77 (1996).
10. John Kuriyan, Klara Osapay, Stephen K Burley, Axel T. Brunger, Wayne A. Hendrickson, Martin Kplus, *Proteins* **10**, 340-358 (1991).

Appendix

A. Rb-Z-DNA Condensed Protein Data Bank (pdb) Listing*

Atom No.	Atom Name	Residue Name	Residue No.	x	y	z	Q	B	Seg. ID
1	H5T	CYT	1	8.252	12.766	-3.818	1.00	.00	ALT0
2	O5'	CYT	1	7.451	13.218	-4.092	1.00	4.60	ALT0
3	C5'	CYT	1	7.119	14.151	-3.078	1.00	4.63	ALT0
4	H5'	CYT	1	6.936	13.636	-2.136	1.00	.00	ALT0
5	H5"	CYT	1	6.210	14.665	-3.389	1.00	.00	ALT0
6	C4'	CYT	1	8.231	15.175	-2.905	1.00	4.70	ALT0
7	H4'	CYT	1	7.894	15.927	-2.194	1.00	.00	ALT0
8	O4'	CYT	1	9.409	14.527	-2.370	1.00	4.67	ALT0
9	C1'	CYT	1	10.542	14.773	-3.212	1.00	4.71	ALT0
10	H1'	CYT	1	11.094	15.642	-2.854	1.00	.00	ALT0
11	N1	CYT	1	11.402	13.542	-3.257	1.00	4.73	ALT0
12	C6	CYT	1	10.829	12.296	-3.131	1.00	4.69	ALT0
13	H6	CYT	1	9.747	12.197	-3.060	1.00	.00	ALT0
14	C2	CYT	1	12.784	13.685	-3.440	1.00	4.71	ALT0
15	O2	CYT	1	13.309	14.794	-3.522	1.00	4.79	ALT0
16	N3	CYT	1	13.531	12.560	-3.560	1.00	4.75	ALT0
17	C4	CYT	1	12.989	11.338	-3.426	1.00	4.75	ALT0
18	N4	CYT	1	13.758	10.245	-3.566	1.00	4.74	ALT0
19	H41	CYT	1	13.365	9.328	-3.453	1.00	.00	ALT0
20	H42	CYT	1	14.739	10.347	-3.784	1.00	.00	ALT0
21	C5	CYT	1	11.598	11.185	-3.077	1.00	4.76	ALT0
22	H5	CYT	1	11.260	10.293	-2.565	1.00	.00	ALT0
23	C2'	CYT	1	9.891	15.082	-4.584	1.00	4.71	ALT0
24	H2'	CYT	1	9.545	14.154	-5.041	1.00	.00	ALT0
25	H2"	CYT	1	10.575	15.600	-5.257	1.00	.00	ALT0
26	C3'	CYT	1	8.666	15.877	-4.208	1.00	4.72	ALT0
27	H3'	CYT	1	7.882	15.855	-4.966	1.00	.00	ALT0
28	O3'	CYT	1	9.062	17.223	-3.952	1.00	4.85	ALT0
29	P	GUA	2	8.070	18.461	-4.144	1.00	4.92	ALT0
30	O1P	GUA	2	8.878	19.700	-4.094	1.00	4.85	ALT0
31	O2P	GUA	2	7.264	18.195	-5.347	1.00	4.92	ALT0
32	O5'	GUA	2	7.102	18.409	-2.857	1.00	4.92	ALT0
33	C5'	GUA	2	7.610	18.506	-1.534	1.00	4.90	ALT0
34	H5'	GUA	2	8.469	17.842	-1.427	1.00	.00	ALT0
35	H5"	GUA	2	7.931	19.532	-1.348	1.00	.00	ALT0
36	C4'	GUA	2	6.560	18.125	-.530	1.00	4.94	ALT0
37	H4'	GUA	2	5.698	18.790	-.599	1.00	.00	ALT0
38	O4'	GUA	2	6.176	16.782	-.766	1.00	4.82	ALT0
39	C1'	GUA	2	5.917	16.179	.502	1.00	4.74	ALT0
40	H1'	GUA	2	4.844	15.993	.561	1.00	.00	ALT0
41	N9	GUA	2	6.623	14.906	.636	1.00	4.56	ALT0
42	C4	GUA	2	7.982	14.688	.547	1.00	4.45	ALT0
43	N3	GUA	2	8.965	15.631	.499	1.00	4.37	ALT0
44	C2	GUA	2	10.200	15.118	.363	1.00	4.33	ALT0

45	N2	GUA	2	11.247	15.916	.173	1.00	4.29	ALT0
46	H21	GUA	2	12.170	15.522	.066	1.00	.00	ALT0
47	H22	GUA	2	11.118	16.916	.144	1.00	.00	ALT0
48	N1	GUA	2	10.455	13.759	.383	1.00	4.31	ALT0
49	H1	GUA	2	11.410	13.434	.348	1.00	.00	ALT0
50	C6	GUA	2	9.480	12.766	.463	1.00	4.42	ALT0
51	O6	GUA	2	9.811	11.578	.415	1.00	4.36	ALT0
52	C5	GUA	2	8.138	13.326	.506	1.00	4.44	ALT0
53	N7	GUA	2	6.891	12.683	.511	1.00	4.42	ALT0
54	C8	GUA	2	6.076	13.664	.687	1.00	4.47	ALT0
55	H8	GUA	2	5.016	13.505	.881	1.00	.00	ALT0
56	C2'	GUA	2	6.264	17.209	1.565	1.00	4.89	ALT0
57	H2'	GUA	2	6.687	16.764	2.466	1.00	.00	ALT0
58	H2"	GUA	2	5.362	17.766	1.824	1.00	.00	ALT0
59	C3'	GUA	2	7.154	18.192	.862	1.00	5.01	ALT0
60	H3'	GUA	2	8.192	17.860	.925	1.00	.00	ALT0
61	O3'	GUA	2	6.880	19.504	1.354	1.00	5.34	ALT0
62	P	CYT	3	8.009	20.506	1.853	1.00	5.54	ALT0
63	O1P	CYT	3	7.340	21.767	2.212	1.00	5.56	ALT0
64	O2P	CYT	3	9.101	20.546	.848	1.00	5.55	ALT0
65	O5'	CYT	3	8.564	19.827	3.192	1.00	5.65	ALT0
66	C5'	CYT	3	9.231	20.654	4.151	1.00	5.83	ALT0
67	H5'	CYT	3	9.340	21.680	3.798	1.00	.00	ALT0
68	H5"	CYT	3	8.620	20.661	5.052	1.00	.00	ALT0
69	C4'	CYT	3	10.593	20.113	4.507	1.00	6.01	ALT0
70	H4'	CYT	3	10.992	20.716	5.319	1.00	.00	ALT0
71	O4'	CYT	3	10.508	18.751	4.965	1.00	5.94	ALT0
72	C1'	CYT	3	11.464	17.940	4.295	1.00	5.92	ALT0
73	H1'	CYT	3	12.400	17.893	4.853	1.00	.00	ALT0
74	N1	CYT	3	10.883	16.612	4.046	1.00	5.85	ALT0
75	C6	CYT	3	9.544	16.440	3.947	1.00	5.79	ALT0
76	H6	CYT	3	8.878	17.303	3.928	1.00	.00	ALT0
77	C2	CYT	3	11.759	15.548	3.909	1.00	5.77	ALT0
78	O2	CYT	3	12.964	15.742	3.793	1.00	5.75	ALT0
79	N3	CYT	3	11.251	14.290	3.888	1.00	5.77	ALT0
80	C4	CYT	3	9.941	14.093	3.849	1.00	5.70	ALT0
81	N4	CYT	3	9.495	12.856	3.748	1.00	5.69	ALT0
82	H41	CYT	3	8.507	12.681	3.651	1.00	.00	ALT0
83	H42	CYT	3	10.145	12.083	3.771	1.00	.00	ALT0
84	C5	CYT	3	9.028	15.198	3.868	1.00	5.78	ALT0
85	H5	CYT	3	7.955	15.041	3.792	1.00	.00	ALT0
86	C2'	CYT	3	11.670	18.645	2.979	1.00	6.03	ALT0
87	H2'	CYT	3	10.828	18.441	2.316	1.00	.00	ALT0
88	H2"	CYT	3	12.605	18.358	2.499	1.00	.00	ALT0
89	C3'	CYT	3	11.583	20.128	3.353	1.00	6.16	ALT0
90	H3'	CYT	3	11.244	20.751	2.524	1.00	.00	ALT0
91	O3'	CYT	3	12.830	20.562	3.827	1.00	6.41	ALT0
92	P	GUA	4	13.334	22.039	3.661	1.00	6.72	ALT0
93	O1P	GUA	4	14.745	21.996	4.078	1.00	6.75	ALT0
94	O2P	GUA	4	12.931	22.603	2.362	1.00	6.68	ALT0
95	O5'	GUA	4	12.447	22.764	4.777	1.00	6.70	ALT0
96	C5'	GUA	4	12.757	22.431	6.131	1.00	6.72	ALT0
97	H5'	GUA	4	12.696	21.352	6.276	1.00	.00	ALT0
98	H5"	GUA	4	13.767	22.761	6.372	1.00	.00	ALT0
99	C4'	GUA	4	11.797	23.098	7.053	1.00	6.71	ALT0

100	H4'	GUA	4	11.848	24.182	6.941	1.00	.00	ALT0
101	O4'	GUA	4	10.502	22.635	6.726	1.00	6.71	ALT0
102	C1'	GUA	4	9.742	22.561	7.917	1.00	6.71	ALT0
103	H1'	GUA	4	8.995	23.356	7.892	1.00	.00	ALT0
104	N9	GUA	4	9.066	21.270	7.870	1.00	6.71	ALT0
105	C4	GUA	4	9.635	19.997	7.854	1.00	6.71	ALT0
106	N3	GUA	4	10.952	19.709	7.936	1.00	6.68	ALT0
107	C2	GUA	4	11.193	18.399	7.745	1.00	6.67	ALT0
108	N2	GUA	4	12.450	17.977	7.539	1.00	6.60	ALT0
109	H21	GUA	4	12.634	16.995	7.392	1.00	.00	ALT0
110	H22	GUA	4	13.211	18.640	7.534	1.00	.00	ALT0
111	N1	GUA	4	10.190	17.452	7.637	1.00	6.65	ALT0
112	H1	GUA	4	10.439	16.475	7.582	1.00	.00	ALT0
113	C6	GUA	4	8.848	17.727	7.605	1.00	6.70	ALT0
114	O6	GUA	4	8.033	16.814	7.581	1.00	6.70	ALT0
115	C5	GUA	4	8.601	19.126	7.629	1.00	6.71	ALT0
116	N7	GUA	4	7.379	19.817	7.572	1.00	6.76	ALT0
117	C8	GUA	4	7.733	21.055	7.669	1.00	6.73	ALT0
118	H8	GUA	4	7.013	21.871	7.599	1.00	.00	ALT0
119	C2'	GUA	4	10.695	22.854	9.088	1.00	6.69	ALT0
120	H2'	GUA	4	10.504	22.222	9.955	1.00	.00	ALT0
121	H2"	GUA	4	10.592	23.903	9.368	1.00	.00	ALT0
122	C3'	GUA	4	12.073	22.723	8.507	1.00	6.70	ALT0
123	H3'	GUA	4	12.432	21.700	8.635	1.00	.00	ALT0
124	O3'	GUA	4	12.944	23.724	8.997	.50	6.73	ALT1
125	P	CYT	5	13.635	23.723	10.444	.50	6.78	ALT1
126	O1P	CYT	5	14.843	24.577	10.261	.50	6.82	ALT1
127	O2P	CYT	5	12.632	24.015	11.477	.50	6.76	ALT1
128	O5'	CYT	5	14.166	22.276	10.771	.50	6.64	ALT1
129	C5'	CYT	5	15.146	22.196	11.804	.50	6.48	ALT1
130	H5'	CYT	5	16.045	22.723	11.484	.50	.00	ALT1
131	H5"	CYT	5	14.781	22.630	12.736	.50	.00	ALT1
132	C4'	CYT	5	15.476	20.764	12.023	1.00	6.40	ALT0
133	H4'	CYT	5	16.281	20.657	12.748	1.00	.00	ALT0
134	O4'	CYT	5	14.320	20.067	12.497	1.00	6.26	ALT0
135	C1'	CYT	5	14.032	18.961	11.608	1.00	6.22	ALT0
136	H1'	CYT	5	14.564	18.070	11.945	1.00	.00	ALT0
137	N1	CYT	5	12.572	18.715	11.513	1.00	6.13	ALT0
138	C6	CYT	5	11.703	19.789	11.575	1.00	6.05	ALT0
139	H6	CYT	5	12.084	20.792	11.770	1.00	.00	ALT0
140	C2	CYT	5	12.122	17.410	11.294	1.00	6.05	ALT0
141	O2	CYT	5	12.883	16.450	11.177	1.00	6.10	ALT0
142	N3	CYT	5	10.796	17.219	11.194	1.00	6.05	ALT0
143	C4	CYT	5	9.939	18.243	11.239	1.00	6.03	ALT0
144	N4	CYT	5	8.645	17.967	11.167	1.00	6.03	ALT0
145	H41	CYT	5	7.963	18.709	11.211	1.00	.00	ALT0
146	H42	CYT	5	8.341	17.009	11.067	1.00	.00	ALT0
147	C5	CYT	5	10.384	19.605	11.388	1.00	6.03	ALT0
148	H5	CYT	5	9.696	20.448	11.346	1.00	.00	ALT0
149	C2'	CYT	5	14.606	19.415	10.269	1.00	6.28	ALT0
150	H2'	CYT	5	13.959	20.171	9.823	1.00	.00	ALT0
151	H2"	CYT	5	14.752	18.586	9.577	1.00	.00	ALT0
152	C3'	CYT	5	15.895	20.113	10.701	1.00	6.36	ALT0
153	H3'	CYT	5	16.240	20.832	9.956	1.00	.00	ALT0
154	O3'	CYT	5	16.918	19.146	10.933	1.00	6.47	ALT0

155	P	GUA	6	18.441	19.596	10.825	1.00	6.43	ALT0
156	O1P	GUA	6	19.288	18.404	11.009	1.00	6.55	ALT0
157	O2P	GUA	6	18.553	20.411	9.607	1.00	6.56	ALT0
158	O5'	GUA	6	18.650	20.564	12.076	1.00	6.33	ALT0
159	C5'	GUA	6	18.901	20.016	13.391	1.00	6.13	ALT0
160	H5'	GUA	6	18.051	19.411	13.707	1.00	.00	ALT0
161	H5"	GUA	6	19.790	19.386	13.349	1.00	.00	ALT0
162	C4'	GUA	6	19.135	21.131	14.415	1.00	5.98	ALT0
163	H4'	GUA	6	19.956	21.754	14.060	1.00	.00	ALT0
164	O4'	GUA	6	17.952	21.964	14.520	1.00	5.92	ALT0
165	C1'	GUA	6	17.487	21.956	15.871	1.00	5.93	ALT0
166	H1'	GUA	6	17.868	22.838	16.387	1.00	.00	ALT0
167	N9	GUA	6	16.026	21.959	15.787	1.00	5.90	ALT0
168	C4	GUA	6	15.160	20.903	15.465	1.00	5.89	ALT0
169	N3	GUA	6	15.475	19.597	15.261	1.00	5.90	ALT0
170	C2	GUA	6	14.400	18.785	15.080	1.00	5.90	ALT0
171	N2	GUA	6	14.556	17.446	14.991	1.00	5.85	ALT0
172	H21	GUA	6	13.752	16.852	14.853	1.00	.00	ALT0
173	H22	GUA	6	15.477	17.040	15.071	1.00	.00	ALT0
174	N1	GUA	6	13.105	19.275	14.995	1.00	5.86	ALT0
175	H1	GUA	6	12.340	18.640	14.819	1.00	.00	ALT0
176	C6	GUA	6	12.771	20.619	15.142	1.00	5.90	ALT0
177	O6	GUA	6	11.600	20.959	15.082	1.00	5.88	ALT0
178	C5	GUA	6	13.912	21.461	15.342	1.00	5.90	ALT0
179	N7	GUA	6	13.979	22.833	15.527	1.00	5.90	ALT0
180	C8	GUA	6	15.222	23.040	15.827	1.00	5.89	ALT0
181	H8	GUA	6	15.603	24.038	16.037	1.00	.00	ALT0
182	C2'	GUA	6	18.099	20.680	16.473	1.00	5.93	ALT0
183	H2'	GUA	6	17.557	19.804	16.113	1.00	.00	ALT0
184	H2"	GUA	6	18.114	20.680	17.562	1.00	.00	ALT0
185	C3'	GUA	6	19.462	20.638	15.850	1.00	5.93	ALT0
186	H3'	GUA	6	19.950	19.663	15.861	1.00	.00	ALT0
187	O3'	GUA	6	20.321	21.610	16.440	1.00	5.93	ALT0
188	H3T	GUA	6	19.854	22.448	16.441	1.00	.00	ALT0
189	H5T	CYT	7	7.448	13.292	14.800	1.00	.00	ALT0
190	O5'	CYT	7	7.348	12.436	15.223	1.00	4.10	ALT0
191	C5'	CYT	7	7.933	11.458	14.348	1.00	4.16	ALT0
192	H5'	CYT	7	7.803	10.472	14.794	1.00	.00	ALT0
193	H5"	CYT	7	7.395	11.483	13.401	1.00	.00	ALT0
194	C4'	CYT	7	9.419	11.685	14.062	1.00	4.12	ALT0
195	H4'	CYT	7	9.768	10.914	13.378	1.00	.00	ALT0
196	O4'	CYT	7	9.692	12.982	13.485	1.00	4.11	ALT0
197	C1'	CYT	7	10.629	13.716	14.296	1.00	4.08	ALT0
198	H1'	CYT	7	11.646	13.540	13.941	1.00	.00	ALT0
199	N1	CYT	7	10.294	15.163	14.284	1.00	4.07	ALT0
200	C6	CYT	7	8.996	15.585	14.109	1.00	4.04	ALT0
201	H6	CYT	7	8.208	14.845	13.967	1.00	.00	ALT0
202	C2	CYT	7	11.308	16.078	14.502	1.00	3.99	ALT0
203	O2	CYT	7	12.455	15.696	14.648	1.00	3.99	ALT0
204	N3	CYT	7	10.999	17.393	14.542	1.00	3.95	ALT0
205	C4	CYT	7	9.733	17.820	14.339	1.00	3.97	ALT0
206	N4	CYT	7	9.441	19.111	14.455	1.00	3.88	ALT0
207	H41	CYT	7	8.488	19.426	14.374	1.00	.00	ALT0
208	H42	CYT	7	10.178	19.779	14.631	1.00	.00	ALT0
209	C5	CYT	7	8.675	16.897	14.113	1.00	3.96	ALT0

210	H5	CYT	7	7.649	17.227	13.956	1.00	.00	ALT0
211	C2'	CYT	7	10.460	13.091	15.684	1.00	4.11	ALT0
212	H2'	CYT	7	9.540	13.456	16.143	1.00	.00	ALT0
213	H2"	CYT	7	11.311	13.289	16.336	1.00	.00	ALT0
214	C3'	CYT	7	10.242	11.634	15.358	1.00	4.18	ALT0
215	H3'	CYT	7	9.718	11.099	16.152	1.00	.00	ALT0
216	O3'	CYT	7	11.487	10.992	15.127	1.00	4.25	ALT0
217	P	GUA	8	11.694	9.400	15.177	1.00	4.44	ALT0
218	O1P	GUA	8	10.888	8.813	16.262	1.00	4.48	ALT0
219	O2P	GUA	8	13.148	9.194	15.153	1.00	4.45	ALT0
220	O5'	GUA	8	11.093	8.862	13.802	1.00	4.54	ALT0
221	C5'	GUA	8	11.652	9.213	12.535	1.00	4.68	ALT0
222	H5'	GUA	8	11.749	10.299	12.491	1.00	.00	ALT0
223	H5"	GUA	8	12.635	8.759	12.411	1.00	.00	ALT0
224	C4'	GUA	8	10.723	8.754	11.438	1.00	4.76	ALT0
225	H4'	GUA	8	10.668	7.665	11.419	1.00	.00	ALT0
226	O4'	GUA	8	9.423	9.307	11.626	1.00	4.82	ALT0
227	C1'	GUA	8	8.840	9.569	10.348	1.00	4.80	ALT0
228	H1'	GUA	8	8.033	8.854	10.195	1.00	.00	ALT0
229	N9	GUA	8	8.277	10.923	10.373	1.00	4.84	ALT0
230	C4	GUA	8	8.969	12.098	10.534	1.00	4.81	ALT0
231	N3	GUA	8	10.314	12.210	10.696	1.00	4.84	ALT0
232	C2	GUA	8	10.707	13.495	10.795	1.00	4.83	ALT0
233	N2	GUA	8	11.986	13.832	10.982	1.00	4.84	ALT0
234	H21	GUA	8	12.244	14.803	11.086	1.00	.00	ALT0
235	H22	GUA	8	12.698	13.121	10.986	1.00	.00	ALT0
236	N1	GUA	8	9.837	14.551	10.733	1.00	4.86	ALT0
237	H1	GUA	8	10.203	15.492	10.727	1.00	.00	ALT0
238	C6	GUA	8	8.456	14.448	10.662	1.00	4.83	ALT0
239	O6	GUA	8	7.804	15.488	10.627	1.00	4.94	ALT0
240	C5	GUA	8	8.026	13.101	10.595	1.00	4.83	ALT0
241	N7	GUA	8	6.737	12.577	10.535	1.00	4.80	ALT0
242	C8	GUA	8	6.967	11.311	10.383	1.00	4.81	ALT0
243	H8	GUA	8	6.154	10.593	10.276	1.00	.00	ALT0
244	C2'	GUA	8	9.956	9.272	9.327	1.00	4.80	ALT0
245	H2'	GUA	8	9.938	9.941	8.468	1.00	.00	ALT0
246	H2"	GUA	8	9.823	8.254	8.993	1.00	.00	ALT0
247	C3'	GUA	8	11.246	9.257	10.105	1.00	4.79	ALT0
248	H3'	GUA	8	11.666	10.264	10.144	1.00	.00	ALT0
249	O3'	GUA	8	12.116	8.221	9.680	1.00	4.86	ALT0
250	P	CYT	9	13.648	8.476	9.345	1.00	4.95	ALT0
251	O1P	CYT	9	14.271	7.163	9.090	1.00	4.92	ALT0
252	O2P	CYT	9	14.222	9.362	10.361	1.00	4.87	ALT0
253	O5'	CYT	9	13.551	9.315	8.002	1.00	4.88	ALT0
254	C5'	CYT	9	14.650	9.268	7.062	1.00	4.92	ALT0
255	H5'	CYT	9	15.470	8.660	7.450	1.00	.00	ALT0
256	H5"	CYT	9	14.284	8.799	6.148	1.00	.00	ALT0
257	C4'	CYT	9	15.207	10.651	6.696	1.00	4.88	ALT0
258	H4'	CYT	9	16.019	10.491	5.990	1.00	.00	ALT0
259	O4'	CYT	9	14.245	11.557	6.080	1.00	4.95	ALT0
260	C1'	CYT	9	14.038	12.701	6.956	1.00	4.91	ALT0
261	H1'	CYT	9	14.648	13.540	6.617	1.00	.00	ALT0
262	N1	CYT	9	12.602	13.059	6.988	1.00	4.90	ALT0
263	C6	CYT	9	11.643	12.064	6.875	1.00	4.90	ALT0
264	H6	CYT	9	11.946	11.036	6.676	1.00	.00	ALT0

265	C2	CYT	9	12.240	14.401	7.183	1.00	4.88	ALT0
266	O2	CYT	9	13.062	15.290	7.399	1.00	4.79	ALT0
267	N3	CYT	9	10.914	14.695	7.199	1.00	4.85	ALT0
268	C4	CYT	9	9.977	13.738	7.151	1.00	4.90	ALT0
269	N4	CYT	9	8.692	14.086	7.264	1.00	4.90	ALT0
270	H41	CYT	9	7.974	13.377	7.254	1.00	.00	ALT0
271	H42	CYT	9	8.440	15.059	7.358	1.00	.00	ALT0
272	C5	CYT	9	10.329	12.350	7.009	1.00	4.94	ALT0
273	H5	CYT	9	9.570	11.568	6.993	1.00	.00	ALT0
274	C2'	CYT	9	14.538	12.229	8.316	1.00	4.90	ALT0
275	H2'	CYT	9	13.802	11.565	8.770	1.00	.00	ALT0
276	H2"	CYT	9	14.768	13.058	8.986	1.00	.00	ALT0
277	C3'	CYT	9	15.742	11.371	7.913	1.00	4.87	ALT0
278	H3'	CYT	9	16.061	10.681	8.695	1.00	.00	ALT0
279	O3'	CYT	9	16.844	12.219	7.534	1.00	4.88	ALT0
280	P	GUA	10	18.354	11.690	7.551	1.00	4.91	ALT0
281	O1P	GUA	10	18.559	10.907	8.784	1.00	4.85	ALT0
282	O2P	GUA	10	19.232	12.847	7.268	1.00	4.88	ALT0
283	O5'	GUA	10	18.458	10.663	6.335	1.00	4.89	ALT0
284	C5'	GUA	10	18.348	11.126	4.998	1.00	4.84	ALT0
285	H5'	GUA	10	17.401	11.649	4.859	1.00	.00	ALT0
286	H5"	GUA	10	19.170	11.810	4.786	1.00	.00	ALT0
287	C4'	GUA	10	18.426	9.950	4.065	1.00	4.79	ALT0
288	H4'	GUA	10	19.373	9.425	4.183	1.00	.00	ALT0
289	O4'	GUA	10	17.327	9.076	4.324	1.00	4.75	ALT0
290	C1'	GUA	10	16.790	8.634	3.080	1.00	4.73	ALT0
291	H1'	GUA	10	16.873	7.554	3.001	1.00	.00	ALT0
292	N9	GUA	10	15.357	8.888	3.045	1.00	4.76	ALT0
293	C4	GUA	10	14.729	10.115	3.163	1.00	4.72	ALT0
294	N3	GUA	10	15.316	11.333	3.164	1.00	4.67	ALT0
295	C2	GUA	10	14.422	12.341	3.267	1.00	4.71	ALT0
296	N2	GUA	10	14.831	13.603	3.333	1.00	4.60	ALT0
297	H21	GUA	10	14.156	14.348	3.427	1.00	.00	ALT0
298	H22	GUA	10	15.814	13.818	3.272	1.00	.00	ALT0
299	N1	GUA	10	13.057	12.125	3.368	1.00	4.72	ALT0
300	H1	GUA	10	12.428	12.913	3.409	1.00	.00	ALT0
301	C6	GUA	10	12.464	10.859	3.415	1.00	4.77	ALT0
302	O6	GUA	10	11.261	10.700	3.572	1.00	4.92	ALT0
303	C5	GUA	10	13.395	9.819	3.291	1.00	4.75	ALT0
304	N7	GUA	10	13.156	8.450	3.148	1.00	4.77	ALT0
305	C8	GUA	10	14.346	7.964	3.010	1.00	4.70	ALT0
306	H8	GUA	10	14.519	6.912	2.798	1.00	.00	ALT0
307	C2'	GUA	10	17.599	9.289	1.979	1.00	4.77	ALT0
308	H2'	GUA	10	16.991	9.564	1.118	1.00	.00	ALT0
309	H2"	GUA	10	18.392	8.613	1.661	1.00	.00	ALT0
310	C3'	GUA	10	18.310	10.462	2.643	1.00	4.79	ALT0
311	H3'	GUA	10	17.671	11.343	2.550	1.00	.00	ALT0
312	O3'	GUA	10	19.651	10.618	2.134	1.00	4.81	ALT0
313	P	CYT	11	20.210	12.039	1.622	1.00	4.76	ALT0
314	O1P	CYT	11	21.605	11.861	1.112	1.00	4.78	ALT0
315	O2P	CYT	11	19.926	13.065	2.639	1.00	4.82	ALT0
316	O5'	CYT	11	19.297	12.296	.340	1.00	4.68	ALT0
317	C5'	CYT	11	19.735	13.167	-.720	1.00	4.48	ALT0
318	H5'	CYT	11	20.697	13.617	-.470	1.00	.00	ALT0
319	H5"	CYT	11	19.854	12.579	-1.629	1.00	.00	ALT0

320	C4'	CYT	11	18.743	14.271	-.983	1.00	4.33	ALT0
321	H4'	CYT	11	19.145	14.918	-1.761	1.00	.00	ALT0
322	O4'	CYT	11	17.488	13.734	-1.454	1.00	4.24	ALT0
323	C1'	CYT	11	16.424	14.265	-.635	1.00	4.12	ALT0
324	H1'	CYT	11	16.030	15.176	-1.091	1.00	.00	ALT0
325	N1	CYT	11	15.364	13.249	-.453	1.00	4.02	ALT0
326	C6	CYT	11	15.652	11.894	-.421	1.00	3.94	ALT0
327	H6	CYT	11	16.671	11.556	-.611	1.00	.00	ALT0
328	C2	CYT	11	14.094	13.707	-.209	1.00	3.95	ALT0
329	O2	CYT	11	13.865	14.914	-.077	1.00	3.94	ALT0
330	N3	CYT	11	13.120	12.763	-.080	1.00	3.94	ALT0
331	C4	CYT	11	13.368	11.463	-.006	1.00	3.85	ALT0
332	N4	CYT	11	12.359	10.620	.219	1.00	3.88	ALT0
333	H41	CYT	11	12.535	9.632	.304	1.00	.00	ALT0
334	H42	CYT	11	11.416	10.973	.299	1.00	.00	ALT0
335	C5	CYT	11	14.696	10.980	-.164	1.00	3.92	ALT0
336	H5	CYT	11	14.929	9.921	-.076	1.00	.00	ALT0
337	C2'	CYT	11	17.089	14.615	.698	1.00	4.20	ALT0
338	H2'	CYT	11	17.264	13.702	1.268	1.00	.00	ALT0
339	H2"	CYT	11	16.504	15.321	1.287	1.00	.00	ALT0
340	C3'	CYT	11	18.460	15.135	.264	1.00	4.27	ALT0
341	H3'	CYT	11	19.219	14.998	1.036	1.00	.00	ALT0
342	O3'	CYT	11	18.445	16.503	-.162	1.00	4.30	ALT0
343	P	GUA	12	19.756	17.412	-.112	1.00	4.28	ALT0
344	O1P	GUA	12	20.359	17.297	1.235	1.00	4.27	ALT0
345	O2P	GUA	12	19.365	18.730	-.626	1.00	4.29	ALT0
346	O5'	GUA	12	20.750	16.725	-1.149	1.00	4.30	ALT0
347	C5'	GUA	12	20.714	17.040	-2.524	1.00	4.37	ALT0
348	H5'	GUA	12	19.750	16.760	-2.942	1.00	.00	ALT0
349	H5"	GUA	12	20.873	18.110	-2.649	1.00	.00	ALT0
350	C4'	GUA	12	21.791	16.266	-3.261	1.00	4.38	ALT0
351	H4'	GUA	12	22.758	16.448	-2.792	1.00	.00	ALT0
352	O4'	GUA	12	21.463	14.863	-3.174	1.00	4.44	ALT0
353	C1'	GUA	12	21.237	14.387	-4.520	1.00	4.44	ALT0
354	H1'	GUA	12	22.180	14.011	-4.920	1.00	.00	ALT0
355	N9	GUA	12	20.258	13.282	-4.492	1.00	4.53	ALT0
356	C4	GUA	12	18.881	13.361	-4.300	1.00	4.56	ALT0
357	N3	GUA	12	18.123	14.488	-4.315	1.00	4.59	ALT0
358	C2	GUA	12	16.815	14.258	-4.150	1.00	4.63	ALT0
359	N2	GUA	12	15.938	15.288	-4.029	1.00	4.58	ALT0
360	H21	GUA	12	14.954	15.102	-3.902	1.00	.00	ALT0
361	H22	GUA	12	16.272	16.241	-4.062	1.00	.00	ALT0
362	N1	GUA	12	16.311	12.982	-4.028	1.00	4.64	ALT0
363	H1	GUA	12	15.315	12.861	-3.917	1.00	.00	ALT0
364	C6	GUA	12	17.054	11.812	-4.047	1.00	4.63	ALT0
365	O6	GUA	12	16.489	10.730	-3.900	1.00	4.73	ALT0
366	C5	GUA	12	18.447	12.063	-4.188	1.00	4.59	ALT0
367	N7	GUA	12	19.519	11.159	-4.290	1.00	4.56	ALT0
368	C8	GUA	12	20.533	11.943	-4.449	1.00	4.50	ALT0
369	H8	GUA	12	21.549	11.555	-4.544	1.00	.00	ALT0
370	C2'	GUA	12	20.858	15.625	-5.315	1.00	4.41	ALT0
371	H2'	GUA	12	19.857	15.968	-5.049	1.00	.00	ALT0
372	H2"	GUA	12	20.934	15.459	-6.390	1.00	.00	ALT0
373	C3'	GUA	12	21.867	16.589	-4.773	1.00	4.42	ALT0
374	H3'	GUA	12	21.669	17.638	-4.992	1.00	.00	ALT0

375	O3'	GUA	12	23.182	16.240	-5.220	1.00	4.45	ALT0
376	H3T	GUA	12	23.806	16.775	-4.724	1.00	.00	ALT0
377	RB	RB	32	19.637	23.294	10.441	1.00	13.61	ALT0
378	RB	RB	88	5.917	10.408	2.193	1.00	17.87	ALT0
379	RB	RB	111	8.208	24.192	1.145	1.00	21.82	ALT0
380	RB	RB	34	4.699	21.697	1.186	.75	26.06	ALT0
381	RB	RB	31	9.701	22.527	16.228	.50	31.76	ALT0
382	RB	RB	40	18.731	19.194	5.085	.50	23.75	ALT0
383	RB	RB	41	8.400	9.965	-1.245	.50	23.77	ALT0
384	RB	RB	210	8.062	18.816	-7.928	.50	26.61	ALD1
385	RB	RB	211	14.563	11.374	14.272	.50	21.92	ALD1
386	RB	RB	212	12.877	17.304	-4.035	.50	20.44	ALD1
387	RB	RB	320	13.969	11.715	11.475	.50	24.84	ALD2
388	RB	RB	321	14.304	13.767	15.472	.50	23.80	ALD2
389	RB	RB	322	9.677	20.821	-6.691	.50	16.57	ALD2
390	RB	RB	323	11.629	19.400	-3.118	.50	27.72	ALD2
391	RB	RB	379	5.527	16.484	9.004	.50	26.39	ALD2
392	RB	RB	284	17.853	15.213	6.508	.50	19.66	ALD1
393	OH2	H2O	20	3.934	13.093	10.663	1.00	9.76	ALT0
394	H1	H2O	20	3.930	12.179	10.377	1.00	.00	ALT0
395	H2	H2O	20	4.608	13.129	11.342	1.00	.00	ALT0
396	OH2	H2O	21	2.352	15.002	8.765	1.00	11.89	ALT0
397	H1	H2O	21	2.634	14.472	9.510	1.00	.00	ALT0
398	H2	H2O	21	2.696	14.542	7.999	1.00	.00	ALT0
399	OH2	H2O	23	12.782	25.096	14.141	1.00	14.44	ALT0
400	H1	H2O	23	12.154	25.040	13.420	1.00	.00	ALT0
401	H2	H2O	23	12.534	24.382	14.729	1.00	.00	ALT0
402	OH2	H2O	24	13.912	25.834	17.103	1.00	18.19	ALT0
403	H1	H2O	24	13.046	25.442	17.220	1.00	.00	ALT0
404	H2	H2O	24	14.116	25.695	16.178	1.00	.00	ALT0
405	OH2	H2O	25	11.068	22.871	13.269	1.00	19.05	ALT0
406	H1	H2O	25	11.961	22.532	13.198	1.00	.00	ALT0
407	H2	H2O	25	10.649	22.611	12.448	1.00	.00	ALT0
408	OH2	H2O	27	8.134	24.851	16.766	1.00	15.62	ALT0
409	H1	H2O	27	8.392	25.010	17.674	1.00	.00	ALT0
410	H2	H2O	27	7.691	25.657	16.498	1.00	.00	ALT0
411	OH2	H2O	28	9.330	22.112	19.694	1.00	10.90	ALT0
412	H1	H2O	28	10.049	21.567	20.015	1.00	.00	ALT0
413	H2	H2O	28	9.125	22.693	20.427	1.00	.00	ALT0
414	OH2	H2O	29	11.064	27.223	14.934	1.00	11.56	ALT0
415	H1	H2O	29	11.744	26.566	14.787	1.00	.00	ALT0
416	H2	H2O	29	10.409	27.046	14.258	1.00	.00	ALT0
417	OH2	H2O	270	9.060	25.291	19.495	.50	3.16	ALD1
418	H1	H2O	270	9.152	24.455	19.953	.50	.00	ALD1
419	H2	H2O	270	8.246	25.663	19.834	.50	.00	ALD1
420	OH2	H2O	33	17.343	28.002	10.924	1.00	15.05	ALT0
421	H1	H2O	33	16.803	27.228	11.085	1.00	.00	ALT0
422	H2	H2O	33	16.718	28.727	10.895	1.00	.00	ALT0
423	OH2	H2O	35	18.011	24.017	12.522	1.00	5.19	ALT0
424	H1	H2O	35	18.155	24.271	13.434	1.00	.00	ALT0
425	H2	H2O	35	17.841	23.075	12.559	1.00	.00	ALT0
426	OH2	H2O	39	3.432	15.315	2.784	1.00	6.84	ALT0
427	H1	H2O	39	4.382	15.203	2.757	1.00	.00	ALT0
428	H2	H2O	39	3.144	15.133	1.890	1.00	.00	ALT0
429	OH2	H2O	42	17.738	31.889	10.446	1.00	12.85	ALT0

430	H1	H2O	42	18.383	32.536	10.731	1.00	.00	ALTO
431	H2	H2O	42	17.449	32.202	9.589	1.00	.00	ALTO
432	OH2	H2O	44	3.269	19.336	2.708	1.00	13.97	ALTO
433	H1	H2O	44	3.732	19.834	3.382	1.00	.00	ALTO
434	H2	H2O	44	3.668	18.467	2.737	1.00	.00	ALTO
435	OH2	H2O	45	3.507	7.758	2.946	1.00	7.51	ALTO
436	H1	H2O	45	3.206	7.684	2.041	1.00	.00	ALTO
437	H2	H2O	45	3.142	8.587	3.254	1.00	.00	ALTO
438	OH2	H2O	46	6.344	25.167	3.147	1.00	11.09	ALTO
439	H1	H2O	46	5.529	24.792	2.813	1.00	.00	ALTO
440	H2	H2O	46	6.859	24.409	3.425	1.00	.00	ALTO
441	OH2	H2O	47	5.937	27.917	2.484	1.00	14.22	ALTO
442	H1	H2O	47	6.673	28.447	2.789	1.00	.00	ALTO
443	H2	H2O	47	6.011	27.098	2.972	1.00	.00	ALTO
444	OH2	H2O	48	14.328	24.853	1.771	1.00	15.35	ALTO
445	H1	H2O	48	14.545	25.588	2.345	1.00	.00	ALTO
446	H2	H2O	48	13.373	24.791	1.811	1.00	.00	ALTO
447	OH2	H2O	49	15.276	4.662	3.025	1.00	10.35	ALTO
448	H1	H2O	49	15.896	4.770	3.745	1.00	.00	ALTO
449	H2	H2O	49	14.940	3.771	3.130	1.00	.00	ALTO
450	OH2	H2O	51	6.659	12.398	4.209	1.00	4.50	ALTO
451	H1	H2O	51	5.882	11.864	4.375	1.00	.00	ALTO
452	H2	H2O	51	7.206	12.275	4.985	1.00	.00	ALTO
453	OH2	H2O	52	16.941	8.563	21.205	1.00	17.56	ALTO
454	H1	H2O	52	16.516	7.825	20.768	1.00	.00	ALTO
455	H2	H2O	52	17.423	9.010	20.509	1.00	.00	ALTO
456	OH2	H2O	53	4.624	11.196	18.440	1.00	14.89	ALTO
457	H1	H2O	53	5.465	11.454	18.064	1.00	.00	ALTO
458	H2	H2O	53	4.164	12.023	18.589	1.00	.00	ALTO
459	OH2	H2O	55	11.151	6.882	5.376	1.00	10.80	ALTO
460	H1	H2O	55	11.984	6.746	5.829	1.00	.00	ALTO
461	H2	H2O	55	11.128	7.821	5.194	1.00	.00	ALTO
462	OH2	H2O	56	10.550	26.856	4.423	1.00	14.67	ALTO
463	H1	H2O	56	11.316	26.863	4.997	1.00	.00	ALTO
464	H2	H2O	56	10.140	26.005	4.578	1.00	.00	ALTO
465	OH2	H2O	57	6.465	17.203	5.444	1.00	7.05	ALTO
466	H1	H2O	57	6.095	17.499	4.613	1.00	.00	ALTO
467	H2	H2O	57	7.348	17.573	5.459	1.00	.00	ALTO
468	OH2	H2O	58	5.993	22.089	4.344	1.00	10.32	ALTO
469	H1	H2O	58	6.142	21.158	4.176	1.00	.00	ALTO
470	H2	H2O	58	6.263	22.214	5.254	1.00	.00	ALTO
471	OH2	H2O	59	9.410	24.073	4.533	1.00	6.19	ALTO
472	H1	H2O	59	10.337	23.856	4.633	1.00	.00	ALTO
473	H2	H2O	59	8.951	23.392	5.024	1.00	.00	ALTO
474	OH2	H2O	62	15.657	1.437	6.586	1.00	4.84	ALTO
475	H1	H2O	62	16.377	2.026	6.361	1.00	.00	ALTO
476	H2	H2O	62	15.292	1.171	5.741	1.00	.00	ALTO
477	OH2	H2O	63	1.216	21.569	2.359	1.00	29.06	ALTO
478	H1	H2O	63	1.855	20.905	2.619	1.00	.00	ALTO
479	H2	H2O	63	1.425	21.759	1.445	1.00	.00	ALTO
480	OH2	H2O	64	6.552	19.480	14.345	1.00	19.71	ALTO
481	H1	H2O	64	6.204	18.611	14.542	1.00	.00	ALTO
482	H2	H2O	64	6.844	19.422	13.435	1.00	.00	ALTO
483	OH2	H2O	65	4.231	13.734	6.877	1.00	12.95	ALTO
484	H1	H2O	65	4.817	13.021	7.132	1.00	.00	ALTO

485	H2	H2O	65	4.710	14.207	6.197	1.00	.00	ALT0
486	OH2	H2O	66	3.907	18.043	11.118	1.00	29.27	ALT0
487	H1	H2O	66	3.840	18.980	11.303	1.00	.00	ALT0
488	H2	H2O	66	4.278	17.998	10.237	1.00	.00	ALT0
489	OH2	H2O	67	9.639	6.311	8.007	1.00	17.04	ALT0
490	H1	H2O	67	9.792	5.449	7.621	1.00	.00	ALT0
491	H2	H2O	67	10.503	6.723	8.020	1.00	.00	ALT0
492	OH2	H2O	68	12.816	28.811	7.958	1.00	10.29	ALT0
493	H1	H2O	68	11.931	28.739	8.317	1.00	.00	ALT0
494	H2	H2O	68	13.164	29.616	8.341	1.00	.00	ALT0
495	OH2	H2O	69	13.390	2.096	8.112	1.00	39.14	ALT0
496	H1	H2O	69	13.878	1.788	7.349	1.00	.00	ALT0
497	H2	H2O	69	12.825	1.360	8.349	1.00	.00	ALT0
498	OH2	H2O	71	1.206	29.563	7.923	1.00	14.94	ALT0
499	H1	H2O	71	.874	30.303	7.414	1.00	.00	ALT0
500	H2	H2O	71	1.267	29.894	8.819	1.00	.00	ALT0
501	OH2	H2O	72	15.371	15.905	9.948	1.00	6.51	ALT0
502	H1	H2O	72	15.175	16.674	10.484	1.00	.00	ALT0
503	H2	H2O	72	14.581	15.767	9.425	1.00	.00	ALT0
504	OH2	H2O	74	10.699	4.986	10.300	1.00	6.08	ALT0
505	H1	H2O	74	10.281	5.332	11.088	1.00	.00	ALT0
506	H2	H2O	74	10.388	5.552	9.594	1.00	.00	ALT0
507	OH2	H2O	76	14.586	7.421	13.797	1.00	15.11	ALT0
508	H1	H2O	76	13.721	7.287	13.409	1.00	.00	ALT0
509	H2	H2O	76	14.924	6.537	13.940	1.00	.00	ALT0
510	OH2	H2O	78	15.178	30.322	8.217	1.00	29.09	ALT0
511	H1	H2O	78	15.256	30.909	7.466	1.00	.00	ALT0
512	H2	H2O	78	14.597	29.624	7.914	1.00	.00	ALT0
513	OH2	H2O	80	.423	1.190	7.745	1.00	9.35	ALT0
514	H1	H2O	80	.862	1.998	7.477	1.00	.00	ALT0
515	H2	H2O	80	-.367	1.158	7.206	1.00	.00	ALT0
516	OH2	H2O	81	17.191	18.598	7.457	1.00	16.40	ALT0
517	H1	H2O	81	16.933	18.822	8.352	1.00	.00	ALT0
518	H2	H2O	81	16.486	18.943	6.909	1.00	.00	ALT0
519	OH2	H2O	86	19.544	17.601	8.464	1.00	19.18	ALT0
520	H1	H2O	86	18.881	18.283	8.568	1.00	.00	ALT0
521	H2	H2O	86	19.358	16.977	9.166	1.00	.00	ALT0
522	OH2	H2O	87	18.148	17.072	3.229	1.00	7.65	ALT0
523	H1	H2O	87	17.352	17.604	3.221	1.00	.00	ALT0
524	H2	H2O	87	18.321	16.883	2.306	1.00	.00	ALT0
525	OH2	H2O	89	14.410	16.754	1.675	1.00	14.52	ALT0
526	H1	H2O	89	13.586	16.273	1.598	1.00	.00	ALT0
527	H2	H2O	89	14.903	16.519	.890	1.00	.00	ALT0
528	OH2	H2O	90	11.294	6.819	2.113	1.00	10.44	ALT0
529	H1	H2O	90	10.583	6.193	1.977	1.00	.00	ALT0
530	H2	H2O	90	11.411	6.846	3.063	1.00	.00	ALT0
531	XCL	XCL	93	6.932	8.095	.414	1.00	26.49	ALT0
532	OH2	H2O	96	5.764	25.135	-.672	1.00	20.43	ALT0
533	H1	H2O	96	5.382	24.259	-.609	1.00	.00	ALT0
534	H2	H2O	96	5.881	25.277	-1.611	1.00	.00	ALT0
535	OH2	H2O	98	9.962	8.704	-3.323	1.00	43.67	ALT0
536	H1	H2O	98	10.550	8.421	-2.623	1.00	.00	ALT0
537	H2	H2O	98	10.100	8.065	-4.022	1.00	.00	ALT0
538	OH2	H2O	100	22.191	10.681	-1.183	1.00	9.85	ALT0
539	H1	H2O	100	22.175	10.337	-2.076	1.00	.00	ALT0

540	H2	H2O	100	21.302	11.003	-1.033	1.00	.00	ALT0
541	OH2	H2O	101	4.017	7.773	.254	1.00	11.96	ALT0
542	H1	H2O	101	4.927	7.486	.188	1.00	.00	ALT0
543	H2	H2O	101	3.627	7.524	-.585	1.00	.00	ALT0
544	OH2	H2O	103	16.993	24.575	3.433	1.00	20.03	ALT0
545	H1	H2O	103	16.280	25.123	3.105	1.00	.00	ALT0
546	H2	H2O	103	16.555	23.883	3.929	1.00	.00	ALT0
547	OH2	H2O	105	14.373	26.828	8.982	1.00	16.60	ALT0
548	H1	H2O	105	13.842	26.452	9.684	1.00	.00	ALT0
549	H2	H2O	105	14.109	27.747	8.947	1.00	.00	ALT0
550	OH2	H2O	106	15.999	13.643	11.448	1.00	5.54	ALT0
551	H1	H2O	106	16.610	13.344	10.775	1.00	.00	ALT0
552	H2	H2O	106	15.387	14.213	10.982	1.00	.00	ALT0
553	OH2	H2O	110	10.690	23.554	2.025	1.00	36.11	ALT0
554	H1	H2O	110	10.873	23.806	2.930	1.00	.00	ALT0
555	H2	H2O	110	10.542	22.609	2.066	1.00	.00	ALT0
556	OH2	H2O	113	11.011	18.901	-.419	1.00	4.09	ALT0
557	H1	H2O	113	10.999	18.218	-1.090	1.00	.00	ALT0
558	H2	H2O	113	10.090	19.132	-.304	1.00	.00	ALT0
559	OH2	H2O	114	5.621	14.857	4.322	1.00	11.58	ALT0
560	H1	H2O	114	6.241	15.561	4.509	1.00	.00	ALT0
561	H2	H2O	114	6.077	14.294	3.696	1.00	.00	ALT0
562	OH2	H2O	115	15.988	23.842	6.221	1.00	37.96	ALT0
563	H1	H2O	115	15.328	24.277	6.762	1.00	.00	ALT0
564	H2	H2O	115	15.983	22.934	6.525	1.00	.00	ALT0
565	OH2	H2O	116	17.671	14.294	3.850	1.00	7.20	ALT0
566	H1	H2O	116	17.281	15.138	4.076	1.00	.00	ALT0
567	H2	H2O	116	17.157	13.652	4.340	1.00	.00	ALT0
568	OH2	H2O	118	20.890	20.745	-.746	1.00	20.12	ALT0
569	H1	H2O	118	21.227	19.932	-1.122	1.00	.00	ALT0
570	H2	H2O	118	20.564	21.239	-1.498	1.00	.00	ALT0
571	OH2	H2O	152	4.678	21.487	14.924	1.00	7.21	ALT0
572	H1	H2O	152	3.754	21.338	14.730	1.00	.00	ALT0
573	H2	H2O	152	5.092	20.635	14.790	1.00	.00	ALT0
574	OH2	H2O	154	13.863	19.046	.083	1.00	8.78	ALT0
575	H1	H2O	154	13.749	18.224	.560	1.00	.00	ALT0
576	H2	H2O	154	14.202	19.660	.735	1.00	.00	ALT0
577	OH2	H2O	158	15.064	16.402	5.815	1.00	19.02	ALT0
578	H1	H2O	158	15.108	17.357	5.864	1.00	.00	ALT0
579	H2	H2O	158	14.139	16.213	5.657	1.00	.00	ALT0
580	OH2	H2O	160	3.237	10.815	15.923	1.00	20.84	ALT0
581	H1	H2O	160	3.407	11.062	16.832	1.00	.00	ALT0
582	H2	H2O	160	3.864	11.327	15.413	1.00	.00	ALT0
583	OH2	H2O	162	.383	6.427	14.015	1.00	31.15	ALT0
584	H1	H2O	162	-.477	6.093	14.270	1.00	.00	ALT0
585	H2	H2O	162	.911	5.642	13.865	1.00	.00	ALT0
586	OH2	H2O	164	9.081	12.994	19.180	1.00	25.06	ALT0
587	H1	H2O	164	8.869	13.735	19.749	1.00	.00	ALT0
588	H2	H2O	164	8.581	13.155	18.380	1.00	.00	ALT0
589	OH2	H2O	166	3.263	10.371	9.658	1.00	13.56	ALT0
590	H1	H2O	166	4.143	10.036	9.486	1.00	.00	ALT0
591	H2	H2O	166	2.706	9.592	9.649	1.00	.00	ALT0
592	OH2	H2O	167	4.636	21.198	10.906	1.00	9.41	ALT0
593	H1	H2O	167	5.490	21.559	11.146	1.00	.00	ALT0
594	H2	H2O	167	4.262	20.897	11.734	1.00	.00	ALT0

595	OH2	H2O	168	18.203	6.453	17.266	1.00	31.75	ALT0
596	H1	H2O	168	17.851	6.977	16.546	1.00	.00	ALT0
597	H2	H2O	168	17.738	6.769	18.041	1.00	.00	ALT0
598	OH2	H2O	280	5.137	18.309	8.429	.50	2.00	ALD1
599	H1	H2O	280	4.533	17.634	8.741	.50	.00	ALD1
600	H2	H2O	280	5.818	17.826	7.961	.50	.00	ALD1
601	OH2	H2O	330	14.134	8.926	20.295	.50	4.70	ALD2
602	H1	H2O	330	13.863	9.613	20.904	.50	.00	ALD2
603	H2	H2O	330	15.082	8.865	20.407	.50	.00	ALD2
604	OH2	H2O	237	17.424	10.105	12.759	.50	5.71	ALD1
605	H1	H2O	237	17.716	10.834	12.212	.50	.00	ALD1
606	H2	H2O	237	18.229	9.757	13.145	.50	.00	ALD1
607	OH2	H2O	337	19.438	26.616	9.320	.50	21.79	ALD2
608	H1	H2O	337	19.180	27.290	9.950	.50	.00	ALD2
609	H2	H2O	337	19.214	25.790	9.748	.50	.00	ALD2
610	OH2	H2O	338	.353	9.067	12.722	.50	11.50	ALD2
611	H1	H2O	338	.939	9.273	13.451	.50	.00	ALD2
612	H2	H2O	338	.634	9.647	12.015	.50	.00	ALD2
613	OH2	H2O	366	2.671	16.521	5.798	.50	4.14	ALD2
614	H1	H2O	366	2.125	15.735	5.821	.50	.00	ALD2
615	H2	H2O	366	2.429	17.004	6.589	.50	.00	ALD2
616	OH2	H2O	266	3.257	18.419	6.262	.50	2.00	ALD1
617	H1	H2O	266	2.761	18.281	7.069	.50	.00	ALD1
618	H2	H2O	266	4.163	18.224	6.501	.50	.00	ALD1
619	OH2	H2O	367	4.945	19.506	4.979	.50	2.00	ALD2
620	H1	H2O	367	5.719	19.287	5.499	.50	.00	ALD2
621	H2	H2O	367	4.249	18.959	5.343	.50	.00	ALD2
622	OH2	H2O	263	6.098	8.113	4.518	.50	3.57	ALD1
623	H1	H2O	263	6.563	7.745	3.768	.50	.00	ALD1
624	H2	H2O	263	5.680	7.359	4.933	.50	.00	ALD1
625	OH2	H2O	363	6.769	9.228	4.425	.50	2.00	ALD2
626	H1	H2O	363	7.612	9.357	3.990	.50	.00	ALD2
627	H2	H2O	363	6.415	8.430	4.032	.50	.00	ALD2
628	OH2	H2O	256	15.701	19.094	2.907	.50	34.15	ALD1
629	H1	H2O	256	15.389	18.215	2.690	.50	.00	ALD1
630	H2	H2O	256	14.954	19.511	3.340	.50	.00	ALD1
631	OH2	H2O	356	15.800	19.614	4.718	.50	19.55	ALD2
632	H1	H2O	356	14.881	19.795	4.521	.50	.00	ALD2
633	H2	H2O	356	16.190	20.476	4.862	.50	.00	ALD2
634	OH2	H2O	257	14.820	19.952	6.846	.50	2.00	ALD1
635	H1	H2O	257	14.701	20.574	6.128	.50	.00	ALD1
636	H2	H2O	257	14.172	20.216	7.500	.50	.00	ALD1
637	OH2	H2O	230	14.701	9.012	19.457	.50	2.00	ALD1
638	H1	H2O	230	13.905	8.895	19.976	.50	.00	ALD1
639	H2	H2O	230	14.388	9.109	18.557	.50	.00	ALD1
640	OH2	H2O	368	2.790	20.277	7.389	.50	9.67	ALD2
641	H1	H2O	368	2.225	19.756	7.961	.50	.00	ALD2
642	H2	H2O	368	3.676	20.110	7.710	.50	.00	ALD2
643	OH2	H2O	277	5.442	15.238	12.094	.50	2.00	ALD1
644	H1	H2O	277	4.498	15.393	12.127	.50	.00	ALD1
645	H2	H2O	277	5.781	15.940	11.538	.50	.00	ALD1
646	OH2	H2O	377	6.123	15.557	12.667	.50	17.75	ALD2
647	H1	H2O	377	6.252	14.741	12.184	.50	.00	ALD2
648	H2	H2O	377	5.188	15.744	12.578	.50	.00	ALD2
649	OH2	H2O	279	5.600	15.213	8.960	.50	2.00	ALD1

```
650   H1   H2O   279      5.159   15.990    9.304    .50    .00   ALD1
651   H2   H2O   279      5.249   15.108    8.076    .50    .00   ALD1
652   OH2  H2O   331     14.915    9.198   17.253    .50   9.00   ALD2
653   H1   H2O   331     15.196    9.457   16.375    .50    .00   ALD2
654   H2   H2O   331     14.028    9.550   17.333    .50    .00   ALD2
655   OH2  H2O   370      9.731    9.273    1.723    .50  12.46   ALD2
656   H1   H2O   370      9.440    9.336     .813    .50    .00   ALD2
657   H2   H2O   370     10.361    9.987    1.816    .50    .00   ALD2
658   O3'  GUA     4     12.996   23.691    9.015    .50   6.62   ALT2
659   P    CYT     5     14.493   23.266    9.469    .50   6.56   ALT2
660   O1P  CYT     5     15.033   22.324    8.468    .50   6.59   ALT2
661   O2P  CYT     5     15.218   24.514    9.773    .50   6.55   ALT2
662   O5'  CYT     5     14.182   22.476   10.857    .50   6.50   ALT2
663   C5'  CYT     5     15.193   22.249   11.846    .50   6.41   ALT2
664   H5'  CYT     5     16.117   22.744   11.546    .50    .00   ALT2
665   H5"  CYT     5     14.863   22.669   12.796    .50    .00   ALT2
END
```

* The "ATOM" at the beginning of each line has been deleted and extra spaces removed for ease of reading. x,y,z are orthogonal Å coordinates. Q is the occupancy and B the temperature factor in units of $Å^2$ (note: B's for H's are not included or refined). The seqment ID (Seg. ID) identify wheather the atom is in the main structure, ALT0 with Q of 1.00, or one of several alternative conformations.

B. Orthogonal Coordinates (Å) of the AMHA molecule

	No.	Name	x	y	z
ATOM	1	O(1)	-1.870	3.638	2.149
ATOM	2	O(2)	1.315	-1.487	-0.223
ATOM	3	N(1)	2.390	1.285	6.293
ATOM	4	N(2)	3.248	3.956	3.119
ATOM	5	N(3)	2.343	7.473	-0.070
ATOM	6	C(1)	1.848	4.901	5.514
ATOM	7	C(2)	1.245	5.253	6.636
ATOM	8	C(3)	0.942	4.363	7.654
ATOM	9	C(4)	1.316	3.025	7.494
ATOM	10	C(4a)	1.967	2.608	6.342
ATOM	11	C(5a)	3.019	0.891	5.193
ATOM	12	C(5)	3.418	-0.524	5.096
ATOM	13	C(6)	4.038	-0.980	4.029
ATOM	14	C(7)	4.356	-0.181	2.940
ATOM	15	C(8a)	3.341	1.698	4.087
ATOM	16	C(8)	4.014	1.184	2.972
ATOM	17	C(9a)	2.248	3.512	5.284
ATOM	18	C(9)	2.920	3.096	4.156
ATOM	19	C(10)	2.311	4.700	2.351
ATOM	20	C(11)	0.939	4.450	2.481
ATOM	21	C(12)	0.043	5.182	1.690
ATOM	22	C(11)	0.526	6.161	0.826
ATOM	23	C(14)	1.899	6.414	0.722
ATOM	24	C(15)	2.779	5.675	1.509
ATOM	25	C(26)	-1.439	5.002	1.844
ATOM	26	H1	-1.530	3.402	3.244
ATOM	27	H2	2.017	5.551	4.836
ATOM	28	H3	1.013	6.158	6.749
ATOM	29	H4	0.496	4.657	8.440
ATOM	30	H5	1.123	2.386	8.179
ATOM	31	H6	3.227	-1.110	5.822
ATOM	32	H7	3.930	3.744	2.398
ATOM	33	H8	4.280	-1.897	3.999
ATOM	34	H9	4.774	-0.572	2.175
ATOM	35	H10	4.251	1.769	2.263
ATOM	36	H11	0.622	3.802	3.101
ATOM	37	H12	-0.093	6.652	0.282
ATOM	38	H13	3.712	5.848	1.461
ATOM	39	H14	1.129	-1.557	-1.356
ATOM	40	H15	1.560	-1.225	0.699
ATOM	41	H16	3.196	7.517	-0.259
ATOM	42	H17	2.003	7.467	-1.070
ATOM	43	H18	-1.745	5.596	2.532
ATOM	44	H19	-1.850	5.262	1.021

Bibliography

Chapter 1

1. F. Miescher, *Med. Chem. Unt.* **4**, 441-460 (1871).
2. I. Rosenfield, E. Ziff, *DNA for Beginners* (Writers and Readers Publishing, Inc, 1983).
3. E. Chargaff, *Experientia* **6**, 201-209 (1950).
4. J. D. Watson, F. H. Crick, *Nature* **171**, 737-738 (1953).
5. A. Rich, G. J. Quigley, A. H. J. Wang, *Biomolecular Stereodynamics* (Adenine Press, Guilderland, N.Y., USA, 1981), vol. I and II, pp. 35-52.
6. W. Saenger, *Principles of Nucleic Acid Structure* (Springer-Verlag New York Inc., Berlin, 1984).
7. L. Stryer, *Biochemistry* (W.H. Freeman and Company. New York. Fourth ed., 1995).
8. D. Rawn, *Biochemistry* (Neil Patterson Publishers, 1989).
9. A. H. J. Wang, G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel, A. Rich, *Nature* **282**, 680-686 (1979).
10. A. H. J. Wang, G. J. Quigley, F. J. Kolpak, G. van der Marel. A. Rich. J. H. van Boom, *Science* **211**, 171-176 (1981).
11. R. V. Gessner, C. A. Frederick, G. J. Quigley, A. Rich, A. H. Wang, *J Biol Chem* **264**, 7921-35 (1989).
12. S. Fujii, A. H. J. Wang, G. J. Quigley, H. Westerink, G. van der Marel, J. H. van Boom, A. Rich, *Biopolymers* **24**, 243-250 (1985).
13. A. Rich, A. Wang, A. Nordhiem, *Nucleic Acid Research* , 11-35 (1983).
14. S. Fujii, A. H. J. Wang, G. van der Marel, J. H. van Boom, A. Rich, *Nuc. Acids Res.* **10**, 7879-7892 (1982).

15. A. H. J. Wang, R. V. Gessner, G. A. van der Marel, J. H. van Boom, A. Rich, *P.N.A.S.* **82**, 3611-3615 (1985).

16. E. M. Lafer, A. Moller, A. Nordheim, B. D. Stollar, A. Rich, *Proc. Natl. Acad. Sci. USA* **78**, 3546-3550 (1981).

17. A. Moller, A. Nordheim, S. A. Kozlowski, D. Patel, A. Rich, *Biochemistry* **23**, 54-62 (1984).

18. R. V. Gessner, G. J. Quigley, A. H. J. Wang, G. A. van der Marel, J. H. van Boom, A. Rich, *Biochemistry* **24**, 237-240 (1985).

19. A. Rich, A. Nordheim, A. Wang, *Annual Review of Biochemistry* **53**. 791-846 (1984).

20. A. Razin, A. D. Riggs, *Scinece* **210**, 604-610 (1980).

21. M. A. Sirover, L. A. Loeb, *Biochemical and Biophysical Research Communications* **70**, 812 (1976).

22. T. Spiro, *Nucleic Acid-Metal Ion Interaction* (John Wiley & Sons, Inc. 1980).

23. A. S. Mildvan, L. A. Loeb, *CRC Critical Review in Biochemistry* **6**. 219-244 (1979).

24. T. Kornberg, A. Kornberg, *The Enzymes* **10**, 173 (1974).

25. L. A. Loeb, *The Enzymes* **10**, 173 (1974).

26. D. L. Sloan, L. A. Loeb, A. S. Mildvan, R. J. Feldmann, *J. Biol. Chem.* **250**, 8913 (1975).

27. J. P. Slater, I. Tamir, L. A. Loeb, A. S. Mildvan, *J. Biol. Chem.* **247**. 2784 (1972).

28. L. A. Leob, *Pharm, Ther. A* **2**, 117 (1977).

29. F. M. Pohl, T. M. Jovin, *J. Mol. Biol.* **67**, 375-396 (1972).

30. T. J. Thamann, R. C. Lord, A. Wang, A. Rich, *Nucleic Acids Res.* **9**. 5443-5457 (1981).

31. Y.-G. Gao, M. Sriram, A. Wang, *Nucleic Acids Res* **21**, 4093-4101 (1993).

32. P. S. Ho, C. A. Frederick, D. Saal, A. H. J. Wang, A. Rich, *J. Biomol. Struct. Dyn.* **4**, 521-534 (1987).

33. S. R. Holbrook, J. L. Sussman, R. W. Warrant, G. M. Church, S. H. Kim, *Nucl. Acids Res.* **4**, 2811-2820 (1977).

34. G. J. Quigley, M. M. Teeter, A. Rich, *Proc. Nat. Acad. Sci. USA* **75**. 64-68 (1978).

35. M. M. Teeter, G. J. Quigley, A. Rich, *Nucleic Acid-Metal Ion Interactions* (John Wiley and Sons, New York, 1980) pp. 145-177.

36. B. C. Baguley, *Anticancer Drug Des* **6**, 1-35 (1991).

37. T. D. Sakore, B. S. Reddy, H. M. Sobell, *J. Mol. Biol.* **135**, 763-786 (1979).

38. G. J. Quigley, A. H. J. Wang, G. Ughetto, G. van der Marel, J. H. van Boom, A. Rich, *P.N.A.S.* **77**, 7204-7208 (1980).

39. A. H. Wang, G. Ughetto, G. J. Quigley, A. Rich, *Biochemistry* **26**, 1152-63 (1987).

40. M. H. Moore, W. N. Hunter, B. L. d'Estaintot, O. Kennard, *J Mol Biol* **206**, 693-705 (1989).

41. C. M. Nunn, L. Van Meervelt, S. D. Zhang, M. H. Moore, O. Kennard, *J Mol Biol* **222**, 167-77 (1991).

42. G. J. Quigley, G. Ughetto, G. A. van der Marel, J. H. van Boom, A. H. J. Wang, A. Rich, *Science* **232**, 1255-1258 (1986).

43. A. H. J. Wang, G. Ughetto, G. J. Quigley, A. Rich, *J. Biomol. Struct. Dyn.* **4**, 319-342 (1986).

44. M. Cirilli, F. Bachechi, G. Ughetto, F. P. Colonna, M. L. Capobianco, *J Mol Biol* **230**, 878-89 (1993).

45. L. D. Williams, M. Egli, G. Qi, P. Bash, G. A. van der Marel, J. H. van Boom, A. Rich, C. A. Frederick, *Proc Natl Acad Sci U S A* **87**, 2225-9 (1990).

46. M. Egli, L. D. Williams, C. A. Frederick, A. Rich, *Biochemistry* **30**, 1364-72 (1991).

47. Q. Gao, L. D. Williams, M. Egli, D. Rabinovich, S. L. Chen, G. J. Quigley, A. Rich, *Proc Natl Acad Sci U S A* **88**, 2422-2426 (1991).

48. J. Markovits, Y. Pommier, M. R. Mattern, C. Esnault, B. P. Roques, J. B. Le Pecq, K. W. Kohn, *Cancer Res* **46**, 5821-6 (1986).

49. L. Larue, M. Quesne, J. Paoletti, *Biochem Pharmacol* **36**, 3563-9 (1987).

Chapter 2

1. L. Sir Bragg, *Sci. Amer.* **219**, 58-70 (1968).

2. M. M. Harding. *Chem. Brit.* **4**, 548-553 (1968).

3. J. P. Glusker, K. N. Trueblood, *Crystal Structure Analysis: a primer, (2nd edn.* 1985).

4. M. F. C. Ladd, R. A. Palmer, *Structure Determination by X-ray Crystallography* (1985).

5. T. L. Blundell, L. N. Johnson, *Protein Crystallography* (Academic press, 1976).

6. A. T. Brunger, *X-PLOR - A System for Crystallography and NMR* (The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, 1992).

7. Enraf-Nonius, *Molen-Structure Determination System* (1990).

8. R. Boistelle, J. P. Astier, *J. of Crystal Growth* **90**, 14-30 (1988).

9. A. McPHERSON, *Eur. J. Biochem.* **189**, 1-23 (1990).

10. A. Ducruix, R. Giege, *Crystallization of Nucleic Acids and Proteins* (Oxford University Press, New York, 1992).
11. A. McPHERSON, *Methods in Enzymology* (1985), vol. 114.
12. A. H. J. Wang, M. K. Teng, *J. Cryst. Growth* **90**, 295-310 (1988).
13. A. K. Aggarwal, *Methods:A companion to methods in enzymology* (1990), vol. 1.
14. M. J. CoX, P. C. Weber, *J. of Crystal Growth* **90**, 318-324 (1988).
15. P. C. Weber, *Methods: A Companion to Methods in Enzymology* **1**, 31-37 (1990).
16. D. J. Haas, M. Rossmann, *Acta Cryst.* **B26**, 998-1040 (1970).
17. G. Petsko, *J. Mol. Biol.* **96**, 381-392 (1975).
18. H. Hope, *Acta Cryst.* **B44**, 22-26 (1988).
19. H. Hartmann, F. Parak, W. Steigemann, G. A. Petsko, D. Ringe Ponzi, H. Frauenfelder, *Proc. Natl. Acad. Sci. USA* **79**, 4967-4971 (1982).
20. H. Hope, F. Frolow, K. von Bohlen, I. Makowski, C. Kratky, Y. Halfon, H. Danz, P. Webster, K. S. Bartels, H. G. Wittmann, A. Yonath, *Acta Cryst.* **B45**, 190-199 (1989).
21. W. Watt, A. Tulinsky, R. P. Swenson, K. D. Watenpaugh, *J. Mol. Biol.* **218**, 195-208 (1991).
22. T. Earnest, *Proteins: Structure, Function and Genetic* **10**, 171-187 (1991).
23. R. Sutton, *J. Chem. Soc. Faraday Trans* **87**, 101-105 (1991).
24. D. W. Rodgers, *Structure* **2**, 1135-1140 (1994).
25. Enraf-Nonius, *The CAD4 Manaul* (1988).
26. C. Klein, *The Xpis Manaul* (X-ray Research G.m.b.H., Germany, 1993).
27. N. Collaborative Computational Project, *The CCP4 Suit : Programs for Protein Crystallography* (Daresbury Laboratary, UK, 1996).

28. G. Sheldrick, *SHELXL-93* (Institut fuer Anorg. Chemie, Goettingen, Germany, 1993).

29. D. Gewirth, Z. Otwinowski, W. Minor, *The HKL Manaul* (Department of Molecular Biophysics and Biochemistry, Yale University, Edition 4. 1995).

30. D. Gewirth, Z. Otwinowski, *The Scalepack Manaul* (Department of Molecular Biophysics and Biochemistry, Yale University, 1995).

31. Baylor College of Medicine, *Chain Manaul* (Version 5.0, 1992).

Chapter 3

1. W. Saenger, *Principles of Nucleic Acid Structure,* (Springer-Verlag New York Inc., 1984).

2. G. J. Quigley, M. M. Teeter, A. Rich, *Proc. Nat. Acad. Sci. USA* **75**. 64-68 (1978).

3. R. V. Gessner, G. J. Quigley, A. H. J. Wang, G. A. van der Marel. J. H. van Boom, A. Rich, *Biochemistry* **24**, 237-240 (1985).

4. C. A. Frederick, M. Coll, D. M. G. A. Van, B. J. H. Van, A. H. J. Wang, *Biochemistry* **27**, 8350-8361 (1988).

5. Y.-C. Jean, Y.-G. Gao, A. Wang, *Biochemistry* **32**, 381-388 (1993).

6. Y.-G. Gao, M. Sriram, A. Wang, *Nucleic Acids Res* **21**. 4093-4101 (1993).

7. T. F. Kagawa, B. H. Geierstanger, A. H. Wang, P. S. Ho, *J. Biol. Chem.* **266**, 20175-20184 (1991).

8. R. G. Brennan, E. Westhof, M. Sundaralingam, *J. Biomol. Struct. Dyn.* **3**, 649-666 (1986).

9. P. S. Ho, C. A. Frederick, D. Saal, A. H. J. Wang, A. Rich, *J. Biomol. Struct. Dyn.* **4**, 521-534 (1987).

10. M. A. Young, B. Jayaram, D. L. Beveridge, *J. Am. Chem. Soc.* **119**, 56-69 (1997).

11. D. Gewirth, Z. Otwinowski, W. Minor, *The HKL Manaul* (Department of Molecular Biophysics and Biochemistry, Yale University, Edition 4, 1995).

12. D. Gewirth, Z. Otwinowski, *The Scalepack Manaul* (Department of Molecular Biophysics and Biochemistry, Yale University, 1995).

13. R. V. Gessner, C. A. Frederick, G. J. Quigley, A. Rich, A. H. Wang. *J Biol Chem* **264**, 7921-35 (1989).

14. M. Egli, L. D. Williams, Qi Gao, A. Rich, *Biochemistry* **30**, 11388-11402 (1991).

15. D. Bancroft, L. D. Williams, A. Rich, M. Egli, *Biochemistry* **33**, 1073-1086 (1994).

16. Shun-le Chen, A Dissertation for the Degree of Doctor of Philosophy (CUNY, 1993)

17. J. L. Fulton, D. M. Pfund, S. L. Wallen, M. Newville, E. A. Stern, Yanjun Ma, *J. Chem. Phys.* **105**, 2161-2166 (1996).

Chapter 4

1. T. L. Su, T. C. Chou, J. Y. Kim, J. T. Huang, G. Ciszewska, W. Y. Ren, G. M. Otter, F. M. Sirotnak, K. A. Watanabe, *J Med Chem* **38**, 3226-35 (1995).

2. T.-L. Su, J.-T. Huang, K. A. Watanabe, G. M. Otter, F. M. Sirotnak, X. -B. Kong, T.-C. Chou, *Proc. Am. Assoc. Cancer Pes.*, 2241 (1993).

3. T. L. Chou, F. Leteurte, T.-L. Su, K. A. Watanabe, X. B. Kong, W. T. Beck, Y. Prommier, *Proc. Am. Assoc. Cancer Res.* , 368 (1994).

4. B. F. Cain, G. J. Atwell, *Eur J Cancer* **10**, 539-49 (1974).

5. B. F. Cain, G. J. Atwell, W. A. Denny, *J. Med. Chem.* **18**, 1110-1117 (1975).

6. B. C. Baguley, W. A. Denny, G. J. Atwell, B. F. Cain, *J Med Chem* **24**, 170-7 (1981).

7. F. Hudecz, J. Kajtar, M. Szekerke, *Nucleic Acids Res* **9**, 6959-73 (1981).

8. B. Marshall, R. K. Ralph, *Advances in Cancer Research* (Academic Presss, Inc, 1985), vol. 44.

9. K. X. Chen, N. Gresh, B. Pullman, *Nucleic Acids Res* **16**, 3061-73 (1988).

10. R. M. Wadkins, D. E. Graves, *Biochemistry* **30**, 4277-83 (1991).

11. B. Granzen, D. E. Graves, B. C. Baguley, M. K. Danks, W. T. Beck, *Oncol Res* **4**, 489-96 (1992).

12. D. P. Figgitt, W. A. Denny, S. A. Gamage, R. K. Ralph, *Anticancer Drug Des* **9**, 199-208 (1994).

13. E. M. Nelson, K. M. Tewey, L. F. Liu, *Proc Natl Acad Sci U S A* **81**, 1361-5 (1984).

14. Y. Pommier, R. E. Schwartz, L. A. Zwelling, K. W. Kohn, *Biochemistry* **24**, 6406-10 (1985).

15. E. Schneider, Y.-H. Hsiang, L. F. Liu, *Adv. Pharmacol.* **21**, 149-181 (1990).

16. M. J. Robinson, N. Osheroff, *Biochemistry* **29**, 2511-5 (1990).

17. M. J. Robinson, N. Osheroff, *Biochemistry* **30**, 1807-13 (1991).

18. J. Cummings, J. F. Smyth, *Ann Oncol* **4**, 533-43 (1993).

19. M. J. Waring, *Eur J Cancer* **12**, 995-1001 (1976).

20. W. R. Wilson, B. C. Baguley, L. P. Wakelin, M. J. Waring, *Mol Pharmacol* **20**, 404-14 (1981).

21. C. Bailly, W. Denny, L. Mellor, L. Wakelin, M. Waring, *Biochemistry* **31**, 3514-3524 (1992).

22. I. Chourpa, H. Morjani, J. F. Riou, M. Manfait, *FEBS Lett* **397**, 61-4 (1996).

23. A. Maxwell, *J Antimicrob Chemother* **30**, 409-14 (1992).

24. B. S. Glisson, A. M. Killary, P. Merta, W. E. Ross, J. Siciliano, M. J. Siciliano, *Cancer Chemother Pharmacol* **31**, 131-8 (1992).

25. D. Gewirth, Z. Otwinowski, W. Minor, *The HKL Manaul* (Department of Molecular Biophysics and Biochemistry, Yale University, Edition 4. 1995).

26 D. Gewirth, Z. Otwinowski, *The Scalepack Manaul* (Department of Molecular Biophysics and Biochemistry, Yale University, 1995).

27. G. J. Quigley, A. H. Wang, G. Ughetto, G. van der Marel, J. H. van Boom, A. Rich, *Proc Natl Acad Sci U S A* **77**, 7204-8 (1980).

28. Z. H. L. Abraham, Shirley D. Cutbush, Reiko Kuroda, Stephen Neidle, R. Morrin Acheson, Grahame N. Taylor, *Chem. Soc. Perkin Trans II* . 461-465 (1985).

29. J. M. Karle, R. L. Cysyk, I. L. Karle, *Acta Cryst.* **B36**, 3012-3016 (1980).

Chapter 5

1. H. Frauenfelder, G. A. Petsko, D. Tsernoglou, *Nature* **280**, 558-563 (1979).

2. S. H. Northrup, M. R. Pear, J. A. McCammon, M. Karplus, T. Takano, *Nature* **287**, 659-660 (1980).

3. B. Mao, M. R. Pear, J. A. McCammon, *Biopolymers* **21**, 1979-1989 (1982).

4. M. Karplus, J. A. McCammon, *Ann. Re. Biochem.* **53**, 263-300 (1983).

5. J. L. Smith, W. A. Hendrickson, R. B. Honzatko, S. Sheriff, *Biochemistry* **25**, 5018-5027 (1986).

6. T. Ichiye, M. Karplus, *Proteins* **2**, 236-259 (1987).

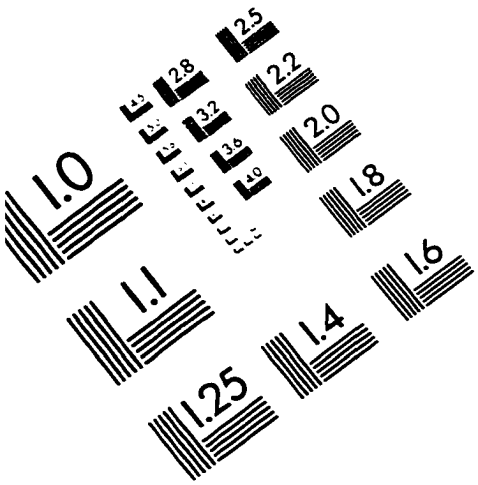
7. J. Kuriyan, M. Karplus, G. A. Petsko, *Proteins* **2**, 1-12 (1987).

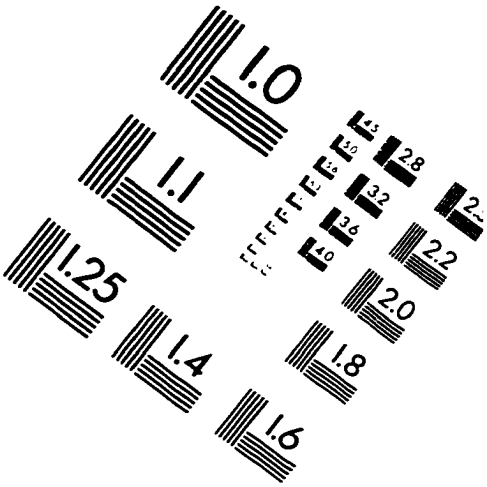
8. J. Kuriyan, W. I. Weis, *Proc. Natl. Acad. Sci. USA* **88** (1991).

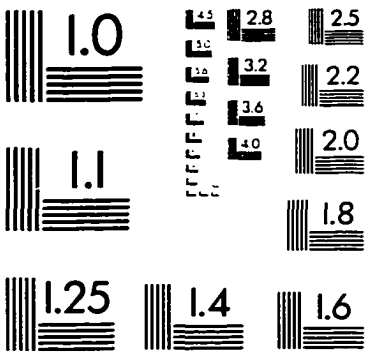
9. F. T. Burling, W. I. Weis, K. M. Flaherty, A. T. Brunger, *Science* **271**, 72-77 (1996).

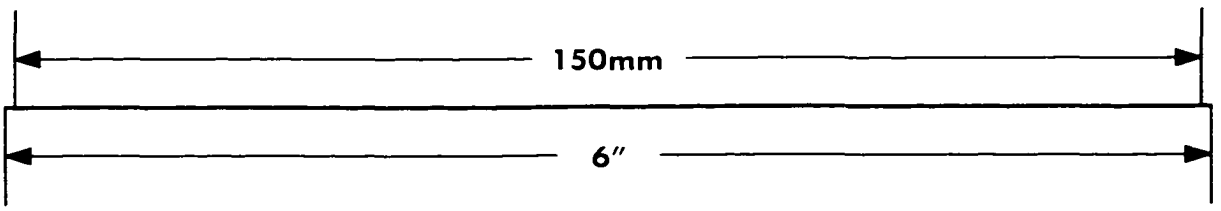
10. John Kuriyan, Klara Osapay, Stephen K Burley, Axel T. Brunger, Wayne A. Hendrickson, Martin Kplus, *Proteins* **10**, 340-358 (1991)

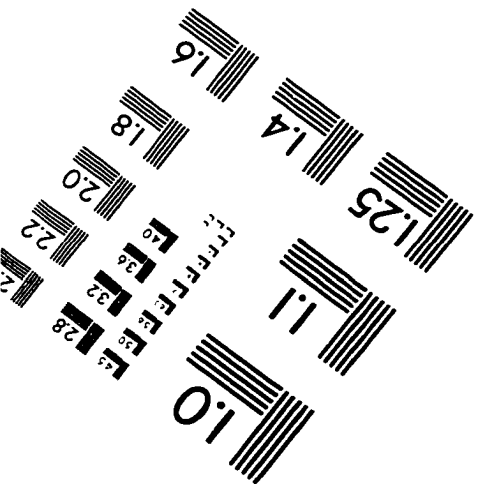
IMAGE EVALUATION
TEST TARGET (QA-3)

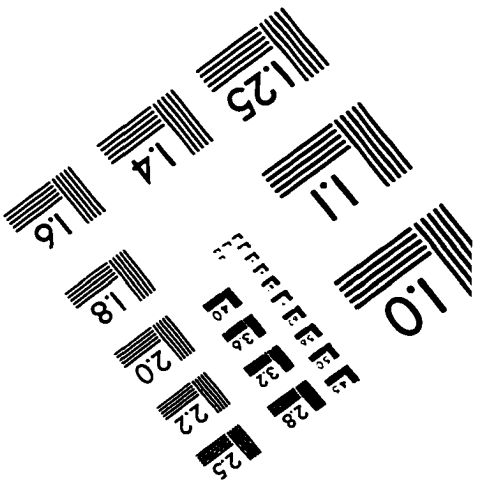












APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989