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.

U·M·I

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

Order Number 9405512

Understanding DNA structure and packing in crystals

Chen, Shun-Le, Ph.D.

City University of New York, 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

A

UNDERSTANDING DNA STRUCTURE AND PACKING IN CRYSTALS

by

SHUN-LE CHEN

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

1993

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.

May 12, 1993
Date

Chair of Examining Committee

5/17/93
Date

Executive Officer

Supervisory Committee

The City University of New York

Abstract

Understanding DNA Structure and Packing in Crystals

by

Shun-Le Chen

Adviser: Professor Gary Joseph Quigley

Three structures of DNA oligomers and DNA-drug complex have been examined, using X-ray diffraction method.

The high resolution low temperature structure of Z-DNA, $d(CGCGCG)_2$ shows a dual conformation. At GpC steps of the oligomer, the phosphate group can assume either a ZI or ZII conformation. The relative probabilities of the two conformations are determined to be 75/25 percent for phosphate P3 and 85/15 percent for phosphate P9. The counterion, Mg(II), in the crystal structure shows a strong conformational and crystallographic packing dependency. Some Mg(II) are partially occupied. The magnesium-water complex dimmer has been determined to be 75 percent occupancy. The structure of DNA and solvent shows highly anisotropic thermal motion.

The structure of copper(II)-Z-DNA complex shows a strong base-specific interaction. Not only do the guanine bases interact with

copper(II) ion, but also the adenine bases, which are in a particular crystallographic environment, can interact with the copper(II) ion strongly. The geometry of copper-solvent complexes can be octahedral or trigonal bipyramidal. All of copper(II) ions are partially occupied.

Ditercalinium bis-intercalates into $d(CGCG)_2$ at the two CpG steps. The intercalation causes helical unwinding, widening of the minor groove, and a 15° kink on the DNA oligomer. These structural variations might be recognized by the uvrABC proteins, and be the trigger of DNA repair machinery.

To my wife, Ming-Ying, my daughter, Yang, and my son, Wei

List of Tables

List of Figures

TABLE OF CONTENTS

CHAPTER I INTRODUCTION

Nucleic Acids constitute the library of genetic information for all living organisms. They do not only carry and transform the genetic messages from generation to generation, but also play a regulatory role in many biological processes concerned with utilization of genetic information. In the process of studies of nucleic acids, it becomes clear that the biological functions of nucleic acids must be understood in terms of their structures. The double-helical model of DNA, proposed by Watson and Crick in 1953, suggested the structural basis for its biological role, but this insight into nucleic acid structure seems to have generated as many questions as it has provided answers. The concept of structure-function relationship of nucleic acids remains a major concern within the scientific community.

In this study, the crystal structures of several DNA oligomers and DNA complexes has been explored in detail.

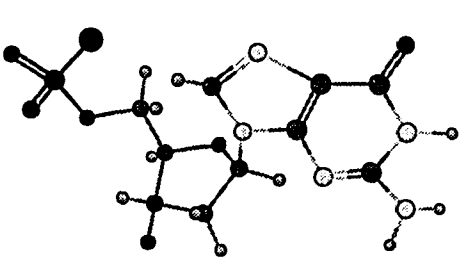
1.1 Structure and Functions of DNA

DNA is the molecule of heredity. DNA is a subgroup of nucleic acids, that were discovered by Friedrich Miescher in 1869 [1] and then identified to fall into two categories, DNA (deoxyribonucleic acid) and RNA (ribonucleic acid) in 1931[2]. The genetic role of DNA

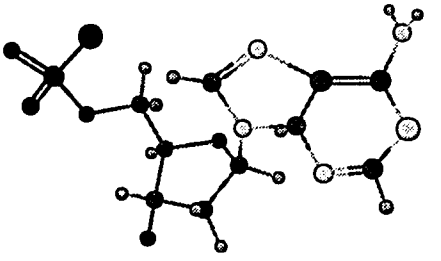
was proved by Olswald Avery in his study on the type transform of *Pneumococcus* in 1944[3]. Avery successfully demonstrated that the purified DNA of *S Pneumococci* transforms live *R Pneumococci* into *S Pneumococci.* Then in the year of 1952 Alfred Hershey and Martha Chase successfully showed that the DNA of the T2 virus/phage not its protein content enters and infects *E. coli* cells [4].

Now it is well understood that the deoxyribonucleic acid, DNA, is the molecule of heredity. The genes of all cells, prokaryotic and eucaryotic organisms, and many viruses are made of DNA. (Some viruses use RNA, ribonucleic acid, as their genetic material.)

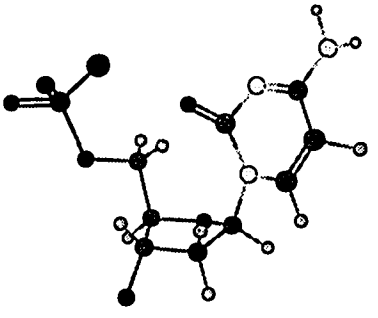
Primary structure. DNA, deoxyribonucleic acid, is a very long, threadlike macromolecule made up of a large number of deoxyribonucleotides, each composed of a base, a sugar, and a phosphate group. The bases of DNA molecules carry the genetic information, whereas their sugar and phosphate groups perform a structural role. There are four bases, two purines, adenine and guanine, and two pyrimidines, cytosine and thymine, that constitute DNA. The fundamental units of DNA molecules are nucleotides, that are made of bases, deoxyribose, and phosphate group (Fig. 1.1).



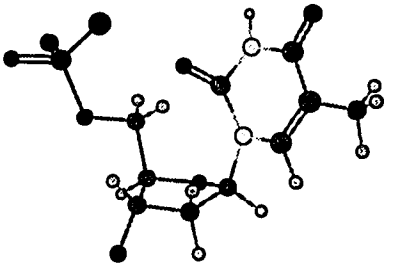
dGMP



dAMP



dTMP



dCMP

Fig. 1.1. The four deoxynucleotids

As shown in Figure 1.1, the bases connect to deoxyribose at the 1' carbon, and the nucleotides will polymerized via phosphate group at 3' hydroxyl group and 5' carbon. The direction of a polynucleotide is defined from 5' to 3' end. (IUPAC-IUB notation) Fig. 1.2.

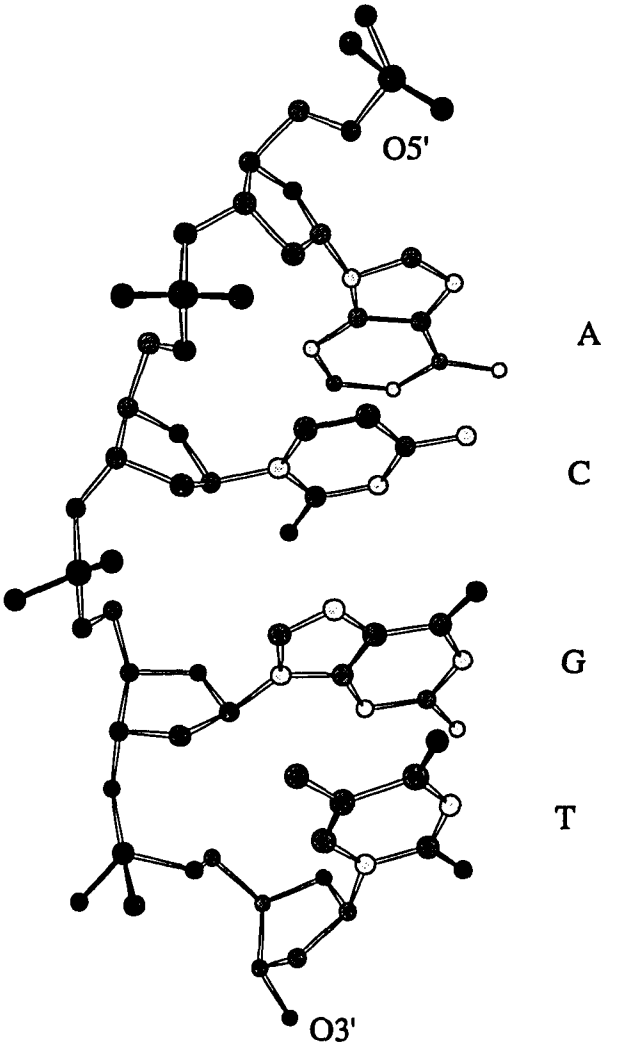


Fig. 1.2. Segment of a polydeoxyribonucleotide

(Hydrogens are not shown)

The length of a single DNA molecule can vary from nanometers with a few thousands base pairs to centimeters with millions of base pairs. (Tab. 1.1) The ratios (A)/(T) and (C)/(G) are usually close to one, as are the overall ratio (purine)/(pyrimidine), while the ratios (G)/(A) and (C)/(T) vary for different organisms but are constant for different tissues within each organism [5].

Table 1.1. Sizes of DNA molecules

Organism	Base pairs (thousands, or kb)	Length (μm)
Viruses		
Polyoma (SV40)	5.1	1.7
λ phage	48.6	17
T2 phage	166	56
Vaccinia	190	65
Bacteria		
Mycoplasma	760	260
E. coli	4000	1360
Eucaryotes		
Yeast	13500	4600
Drosophila	165000	56000
Human	2900000	990000

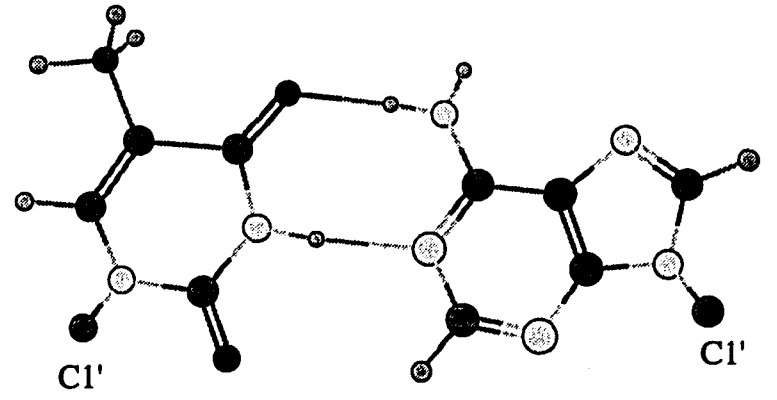
After A. Kornberg. DNA replication [7]

Secondary structure - a double helical structure. The general structure of DNA was believed to be a double helix, as proposed by James Watson and Francis Crick in 1953 [6] (Fig. 1.3). Gathering previous knowledge on DNA structure which were scattered and incomplete, Watson and Crick introduced the concept of a 'base pair' in

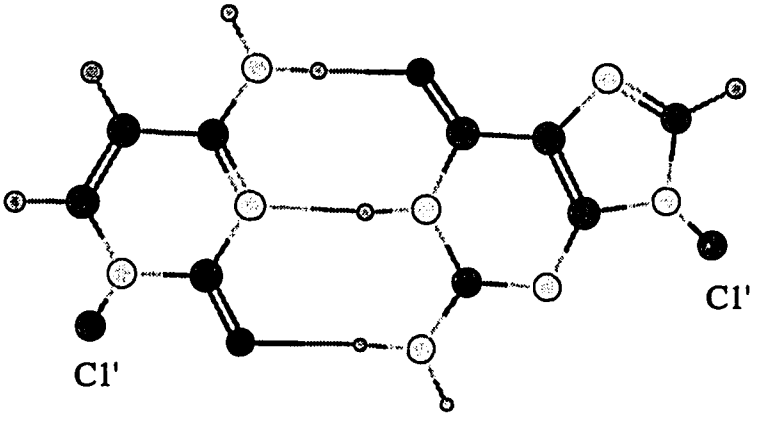
their famous double helix model. They described a DNA molecule as a double strand right-hand helix. the two strands are anti- parallel to each other.

Fig. 1.3. Skeletal model of double-helical DNA

The base pairs were held together by specific hydrogen bonds between purines and pyrimidines, i.e. G-C and T-A. Those base pairs later were called 'Watson-Crick base pair' (Fig. 1.4).



A-T



C-G

Fig. 1.4. Watson-Crick base-pairs

The concept of the base pair immediately leads to the mechanism of DNA replication and transcription. Therefor some of the question of "how does DNA function?" could be answered with satisfaction. Most previous knowledge supported the Watson - Crick double helix model indirectly. There was no direct experimental evidence at that time, except the clues suggested by the results of fiber diffraction.

The double helical structure has several remarkable features.

1. Two helical polynucleotide chains are coiled around a common axis to form a right hand double helix. The chains run in opposite directions.

2. The purine and pyrimidine bases are on the inside of helix, whereas the phosphate and deoxyribose unit are on the outside. The planes of the bases are perpendicular to the helix axis. The planes of the sugars are nearly at right angle to those of the bases.

3. The diameter of the helix is approximately 20 Å. Adjacent bases are separated by 3.4 Å along the helix axis and related by a rotation of 36° about the axis. Hence, the helical structure repeats after ten residues on each chain; that is, at an interval of 34 Å.

4. The two chains are held together by hydrogen bonds between pairs of bases. Adenine is always paired with thymine. Guanine is always paired with cytosine.

5. The sequence of bases along a polynucleotide chain is not restricted in any way. DNAs of identical base composition can differ entirely in sequence. The precise sequence of bases

carries the genetic information.

6. Stacking interactions between the flat aromatic surfaces of the bases stabilize the helical structure against repulsion forces of the negative charged phosphate groups.

7. Torsion angles within backbone, deoxyribose and phosphate group can vary in a wide range. Thus a DNA molecule can induce various conformation.

8. Antiparallel double helices can also form between a DNA chain and a RNA chain. These are called DNA-RNA hybrids.

The most important feature of the double helical structure for DNA is the concept of complimentarity [6]. Based on it, many functions of DNA could be explained.

In the cells, or organisms, DNA structures are highly organized in various structures. In terms of structure, the primary structure refers to the length and the base sequence; the secondary structure refers to the conformation of the helix; the higher order structure refers to the packing of DNA in a cell.

The function of DNA is structurally dependent. The major processes that DNA involves are **replication** (multiplication and

inheritance of genetic information), **transcription** (copying information for processing), and **recombination** (gene exchange between participating DNAs). Without knowledge of DNA structure, the mechanism of those processes can not be understood clearly.

During the processes of replication and transcription, a DNA double helix dissociates (or partially dissociates) into two separated polydeoxynucleotides as the templates, then the new synthesized polydeoxynucleotides (or polynucleotides), that are complimentary to the templates, combine with the templates to form new double helices.

The biological functions of DNA are highly structure dependent. The processes that DNA molecules involved are all associate with enzymes, DNA binding proteins, and/or other physical/chemical reagents [7-9]. In general, an enzyme is a structure specific protein. A particular enzyme can only recognize a specific structure to induce a particular process.

For instance, deoxyribonulease I (Dnase I) binds to double strand DNA and cuts DNA molecules into short segments regardless of its base sequence [10]. EcoRI endonuclease recognizes the section of -GAATTC- in a DNA molecule and cut the DNA of both end at the sites between G and A [11]. DNA Polymerase I initiates replication of E. coli DNA at a unique site [12,13]. RecA protein induces the site specific recombination of λ phage DNA and E. coli DNA [14-17].

The **genetic-on-off** switch region of a lysogenic strain of *E. coli* DNA has three operators, OR1, OR2, and OR3. A certain dose of UV light irradiation can stop the normal growth of the *E. coli* and initiate the production of bacteriophage. There are two repressors, **repressor** and **Cro** involved in this process for phages λ and 434. The mechanism of this process is believed to be that the regulatory protein **repressor** binds to OR1 and OR2, thereby turning off the synthesis of **Cro**. With UV irradiation **Cro** binds to OR3, thereby turning on the synthesis of **Cro**. The structures of the complexes 434 **repressor**-operators and 434 **Cro**-operators have been examined by X-ray crystallographic method. In both cases, the proteins form a dimer that interacts with operators distorting the DNA duplexes. The distortion of DNA conformation induce surfaces that are complimentary to the surfaces of proteins. It is shown that the **repressor** has a weaker interaction with OR3 than the **Cro** does. This can be explained in terms of DNA structure. The sequences of the switch region of the DNA are listed as follows.

OR1	5'	^{1}**TATCACCGCCAGTGGTA**^{17}	3'
	3'	**ATAGTGGCGGTCACCAT**	5'

OR3	5'	^{1}**TATCACCGCAAGGGATA**^{17}	3'
	3'	**ATAGTGGCGTTCCCTAT**	5'

The recognition regions of both operators are located at the base pairs 1 to 6 and 12 to 17 (the boldface regions). The OR1 provides two recognition regions that are almost identical. These two regions can induce complimentary surfaces to **repressor** and **Cro**. While OR3, for the different sequences of two recognition region, can not. That explains the different affinities of OR3 to repressor and **Cro** [18].

Anti-cancer drugs attack DNA molecules of cancerous cells forming various drug-DNA complexes. The interactions of intra-strand cross linking, inter-strand cross linking, and intercalation have been found to block DNA replication effectively or induce DNA repair [19-21].

Theoretical studies need accurate structures of DNA. Scientific theory has its ultimate goal the understanding and prediction of natural phenomena. While biological macromolecules are very complicated subjects, theoretical study of the conformation and structure of biological macromolecules has its unique importance for reasons discussed above.

Molecular simulations, molecular modeling (i. e. static) as well as dynamics [22,23] are rapidly growing fields that have a significant impact on many areas of biological macromolecule research. Biological macromolecules are inherently dynamic systems, and it was the

demonstration that calculations based on simple physical models could give new and testable insights into their internal motions that have convinced many biochemists that biophysical theory has become relevant. To perform a molecular modeling or simulation, a set of functional form and parameters of all potential energy equations and an initial molecular structure with initial velocities of every atoms are required. The functional potential energy equations are highly approximate and can only be tested by how well they predict a variety of experimentally observable results with detailed molecular structure. The initial structure can be obtained from X-ray crystallographic or NMR structure data [24,25], or by model-building [26]. Structures determined by X-ray crystallographic methods have served not only as initial models but also for experimental verification of molecular modeling and simulations [23].

Before these methods can be safely used in a predictive manner, their behavior for well understood systems must be well characterized. In most cases, a well understood system is based on well defined structure determined at high resolution by X-ray crystallography.

On the other hand molecular dynamics simulation is also used for structure refinement in X-ray crystallography [27].

X-ray crystallography is the most effective and practical technique for structure determination. There are

various techniques by which the DNA structures can be studied. Crystallographic and spectroscopic methods are the most effective techniques. Ultraviolet (UV), circular dichroism (CD), linear dichroism (LD), infrared (IR), laser raman, fluorescence, and nuclear magnetic resonance (NMR) are frequently employed to study DNA structures. NMR can produce some inter-atomic distances information of DNA molecule by J-coupling and nuclear overhauser effects (NOE), experiments. Based on such information, three dimensional structure can be predicted with certainty [28]. Crystallography is the only way to get an unambiguous three dimensional structure for a DNA molecule, and the X-ray single crystal diffraction is the most convenient and practical method. In this study the crystal structure of DNA is our concern.

1.2 Approaches of X-ray Crystallography

By the 1950s, X-ray crystallography was accepted as a powerful tools for determining the three dimensional structure of small molecules. In this section we will only review the approaches that are related to the structure of DNA. The principle and practical methods of X-ray crystallography will be reviewed later.

The first X-ray diffraction pattern of DNA was obtained by William Astbury in 1938 [29]. At that time there were no single crystals

of DNA, thus, Astbury's experiment was based on quasi-crystalline 'fiber'. The fiber was made of purified native DNA, a mixture of molecules with various length and sequence. From the X-ray diffraction pattern Astbury concluded that in a DNA 'fiber' the bases were parallel to each other and perpendicular to the fiber axis, and the separation between bases were about 3.34 Å. A photograph of typical fiber diffraction is shown in Fig. 1.5.

Fig. 1.5. Diffraction pattern of DNA fiber

(Reproduced by permission of Dr. R. Langridge, Computer Graphics Laboratory, University of California, San Francisco)

In 1952, Furberg [30] published the structure of cytidine, solved by single crystal X-ray diffraction. Subsquently more results of DNA fiber diffraction and single crystal diffraction of isolated bases and nucleosides were published [30-32]. Based on those results several

models of DNA structure were proposed [30,33]. The most remarkable achievement was made by James Watson and Francis Crick [6]. By Watson's model, the fiber diffraction patterns can be interpreted meaningfully. The questions about how the nucleic acids functioning can be answered. There was no direct experimental evidence at that time. The model of double helical structure was still a proposal to be confirmed.

Natural DNA molecules are flexible due to their extraordinary elongated shape and dynamic nature. Therefore, no single crystal grown from natural DNA molecules have been reported. Based on further fiber diffraction experiments, several families of DNA, with slightly different helical pitches and helical symmetry, have been recognized [34]. They were named A- [35], B- [36], B'- [37], C- [38], C'-, C"-, D- [39], E-, and T-DNA [40]. Among them the A-conformation and B-conformation occur most frequently. Contemporarily, many single nucleosides and nucleotides had been crystallized successfully, and their three dimensional structures solved. These results have been reviewed by Saenger [41]. A few structures of nucleotide dimer had also been reported [42-48].

Before the crystal structure of a DNA was determined experimentally, the double helical structure of a nucleic acid was observed in yeast $tRNA^{Phe}$ at near atomic resolution [49]. This is the first direct visualization of double helical structure in a larger than

dimeric nucleic acid molecule. The molecule of $tRNA^{Phe}$ consists of 76 residues that can arrange in a cloverleaf secondary structure. The three dimensional structure of $tRNA^{Phe}$ was determined to be folded into a L shape tertiary structure (Fig. 1.6) by X-ray diffraction of single crystal. In the structure, there are four regions, containing four to seven Watson-Crick base-pairs, arranged into double helices.

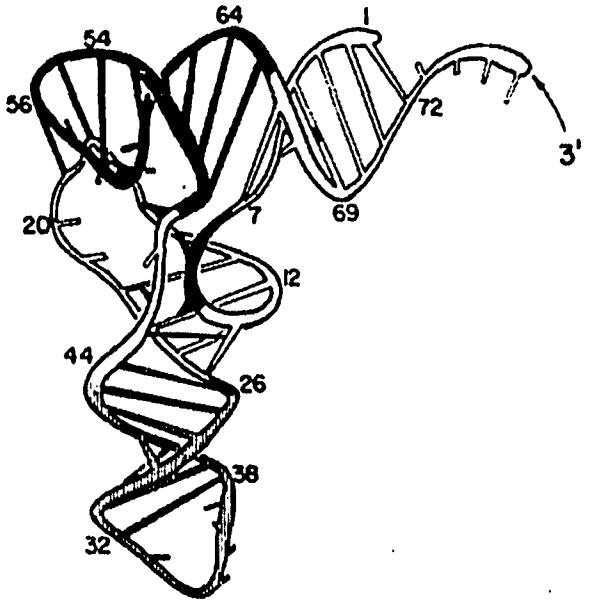


Fig. 1.6. The three dimensional structure of $tRNA^{Phe}$

(from [223])

The structure of yeast $tRNA^{Phe}$ provides a rich display of structural capabilities of nucleic acids. Besides the typical Watson-Crick base-pairs there are many unusual base-base interactions in tRNA structures [41]. The tRNA structures and the technique developed for

crystallization and structure solution provided valuable knowledge and insight for subsequent DNA structural studies.

In order to obtain the exact three dimensional structure of DNA double helix, short DNA oligomer have been synthesized in various laboratories. An oligomer with a few base pairs seemed more likely to have a well defined structure and geometric shape, which facilitate growing single crystal. Various DNA oligomers with different length and sequences have been tried by various research groups.

In 1979, A. Wang, et al. published the first single crystal structure of DNA oligomer, the hexamer $d(CGCGCG)_2$, at 0.9 Å resolution [50]. This was the very first structure of large DNA molecule, six base pairs, determined by X-ray diffraction. Surprisingly, it was a left hand double helical structure, instead of the right hand helix proposed by Watson and Crick. It was named Z-DNA for the zigzag pattern of its ribose-phosphate backbone. One year later, Wing, et al. reported the first right hand helical structure of a B-DNA, dodecamer $d(CGCGAATTCGCG)_2$ [51]. Following their discoveries of the structures of Z-DNA and B-DNA, those two groups and others reported their results of A-DNA structures diffraction experiments of single crystals with different sequences in the following years. The sequences of DNA oligomers were: $d(GGTATACC)_2$ [52], $d(GGCCGGCC)_2$ [53], and $d(CCGG)_2$ [54]. Since then more DNA oligomer structures have been reported.

The three dimensional structures of DNA oligomers revealed several important rules of DNA structure in crystals. They can be summarized as follows:

1. DNA structure is sequence dependent. For instance, an alternative purine-pyrimidine sequence facilities the formation of the Z-conformation [55]; homo-guanosine/homo-cytidine sequence usually form A-DNA [53,56].

2. Chemical modification of bases facilitates certain conformations. For instance, the methylation or bromination at C5 of cytidine can stabilize the Z-conformation [55,57-59]; de-methylation of thymine (i. e. conversion of the thymine to uracil) also favors the Z-conformation [60].

3. Miss-matched base pairs, such as G-T, G-U, A-C, G-A or C-T, usually do not influence the structure significantly [61-68].

4. Solvent molecules, usually water, play important roles in the stabilization of crystal structures. Generally, in DNA crystals the water contents can exceed 50 percent by weight. Within the crystal lattice there are often large solvent channels, additional ion or drugs can invade the lattice without significantly disturbing the DNA structure [69]. This makes

the application of multi-isomorphous replacement possible in solving the DNA structures.

5. The ionic strength of the solution can affect the conformation. For instance, at the low ionic strength an oligomer of alternating purine-pyrimidine tends to the B-form; at high ionic strength it tends to the Z-form [70,71].

6. Counterions and/or polyamines are necessary to stabilize the crystal structures for all conformations, since the crystal and therefor each unit cell must be neutral [50].

7. Certain metal ion complexes, such as cobalt hexamine, stabilize certain conformation strongly [72,73].

All these rules are not rigid. A specific DNA can disobey any of these rules. In the following section we will discuss the structure features of the three major groups of DNA, i.e. A-, B-, and Z-DNA.

1.3 The Structure of DNA

Definition of terms in the structure of DNA. In this study we use the terms recommended by IUPAC-IUB, International Union of

Pure and Applied Chemistry (IUPAC) and International Union of Biochemistry (IUB). The torsion angles in nucleotide are shown in Fig. 1. 7 and Table 1.2.

Table. 1.2 The definition of torsion angles in nucleotides

Torsion angle	Atoms involved
α	(n-1)O3'-P-O5'-C5'
β	P-O5'-C5'-C4'
γ	O5'-C5'-C4'-C3'
δ	C5'-C4'-C3'-O3'
ε	C4'-C3'-O3'-P
ξ	C3'-O3'-P-O5'(n+1)
χ	O4'-C1'-N9-C2 or O4'-C1'-N1-C4
υ0	C4'-O4'-C1'-C2'
υ1	O4'-C1'-C2'-C3'
υ2	C1'-C2'-C3'-C4'
υ3	C2'-C3'-C4'-O4'
υ4	C3'-C4'-O4'-C1'

* Atoms designed (n-1) or (n+1) belong to adjacent units.

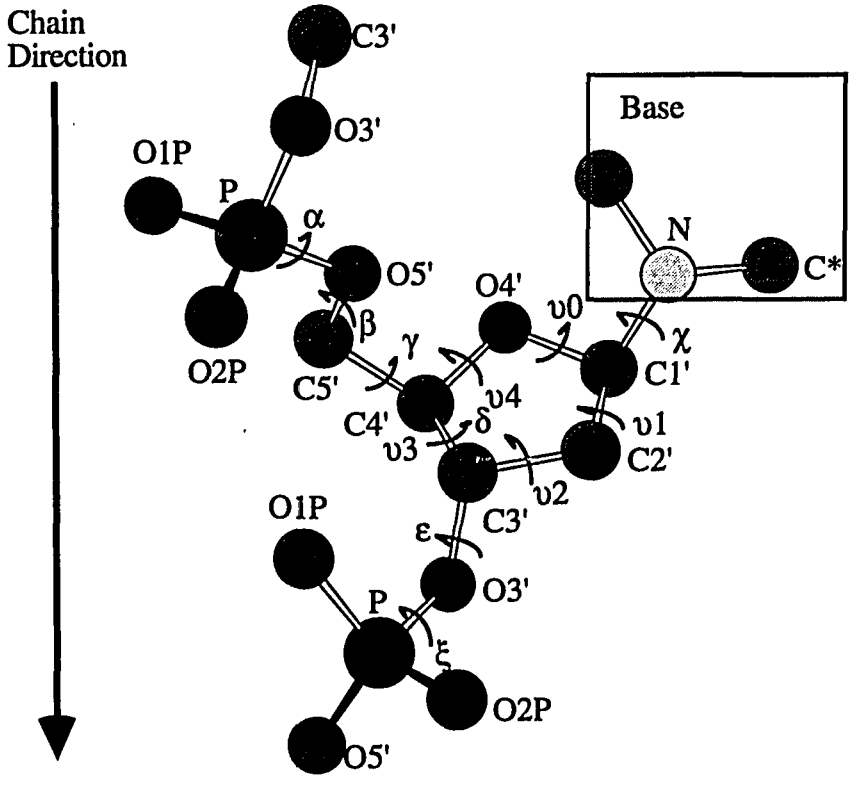


Fig 1.7. Defination of torsion angles in nucleotide

(Hydrogens are not shown)

Where N is N1 in pyrimidine and N9 in purine; C* is C2 inpyrimidine and C4 in purine.

DNA double helices fall into three major conformational groups, A type, B type, and Z type by their structure features. The A family has only one member, A-DNA, while the B family consists B-, B'-, C-, C'-, C"-, D-, E-, and T-DNA [41]. Both A and B family are of right hand double helical structures as opposite to the left hand double helices of Z family. The major differences among these three families are listed in Tab. 1.3.

Table 1.3. Comparison of structural characteristics of A-, B-, Z-DNA

Structure type	Pitch Å	Helical symmetry	Axial rise per repeat (h)Å	Turn angle per repeat (t)	Minor groove width Å	Major groove width Å	Minor groove depth Å	Major groove depth Å
A	28.2	11_1	2.56	32.7°	11.0	2.7	2.8	13.5
B	33.8	10_1	3.38	36.0°	5.7	11.7	7.5	8.5
Z*	44.6	$6_5=6_{-1}$	2x3.70	-60.0°	2.7	N/A	9.0	N/A

* two nucleotides per repeat.

The A family has approximate P_{11} symmetry, B family P_{10}, and Z family P_{-6}. Both A- and B- families are found in natural DNA and synthesized oligomers; Z family are only found in high ionic strength conditions.

Generally, the planes of base pair are not exactly perpendicular to the helical axis. They tilt at different angles according to their structure type. The definitions of helical parameters are given in Fig. 1.8. Within the base pair the propeller twist angle, can be up to 10°. (Fig. 1.8). The typical bond length of hydrogen bonds between the paired bases is about 2.8 Å. The geometries of typical A-T and C-G base pairs are shown in Figure 1.4.

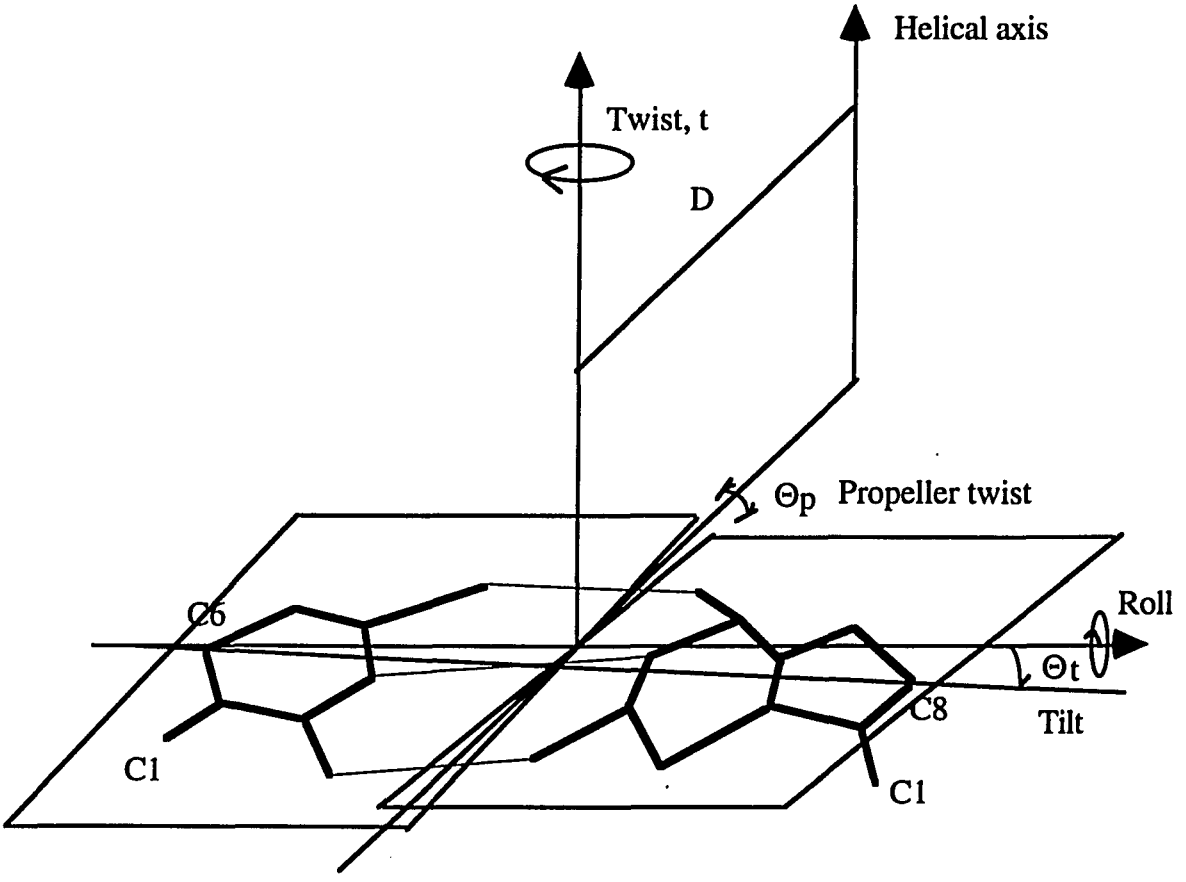


Fig. 1.8. The definition of helical parameters.

For the favorable energy, the furanose sugar always puckers into non-planar geometry. They usually fall into one of two classes of conformations N, around C3'-endo, or S, around C2'-endo. (Fig. 1.9) The geometry about the glycol bond can be anti or syn. All the parameters mentioned above are of average values. For each specific structures the parameters vary [221, 222].

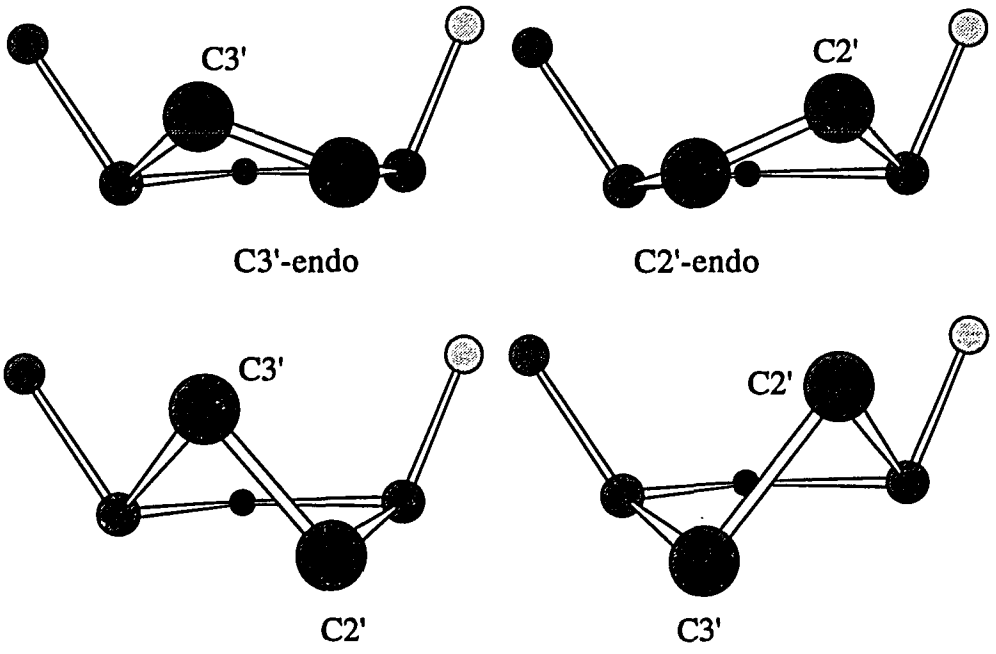


Fig. 1.9. Conformations of sugar pucker

1.3.1 The Structure of A-DNA

The sequences of homo-guanosine/homo-cytidine tend to form A-DNA [53,56]. All reported single crystal structures of A-DNA have -CC- and -GG- sequences. [52,53,56,62,63,65,74-85] The typical geometric parameters of A-DNA are summarized in table 1.4.

The DNA oligomers have been crystallized in dehydrating conditions [86]. The major features that A-DNA structures include specific sugar pucker and uniform base conformation. The sugar pucker

is restricted to C3'-endo giving a P-P distances of about 6 Å. A strong correlation of torsion angles α, γ suggests that rotation moves the bases to avoid the purine-purine clashes. The values of the torsion angle δ are in a narrow range of 89-91° [74]. There appears to be less sequence dependence than that in B-DNA. Figure 1.10 shows a typical A-DNA structure.

Table 1.4. Average helix parameters in A-DNA type oligomer and in A-DNA

Oligo-nucleotide	Twist (t) per residue	Rise per residue (Å)	Base tilt (°)	Dis-placement D (Å)	Major groove Width (Å)	Minor groove width (Å)
d(GGTATACC)	32.2(1)	2.87(1)	13.5	4.0	10.2	6.3
d(CCGG)	33.9(2)	2.85(3)	14.0	3.7	10.6	4.8
d(GGCCGGCC)	32.6(3)	3.03(4)	12.0	3.6	9.6	7.9
A-DNA	32.7	2.56	20	4.5	10.9	3.7

From the X-ray crystallographic structures listed above, the major structure features of A-DNA can be summarized as:

1. The number of residues within one full turn is eleven, and the base-pairs twist about 33° per residue. The pitch is about 28-29 Å.
2. The displacement of bases from the helical axis, D, is between 3.6 Å to 4.5 Å (Fig. 1.9).

3. The sugar pucker is C3'-endo conformation. Base-pairs tilt about 10-20°.
4. The width of major groove is much larger than that of the minor groove.

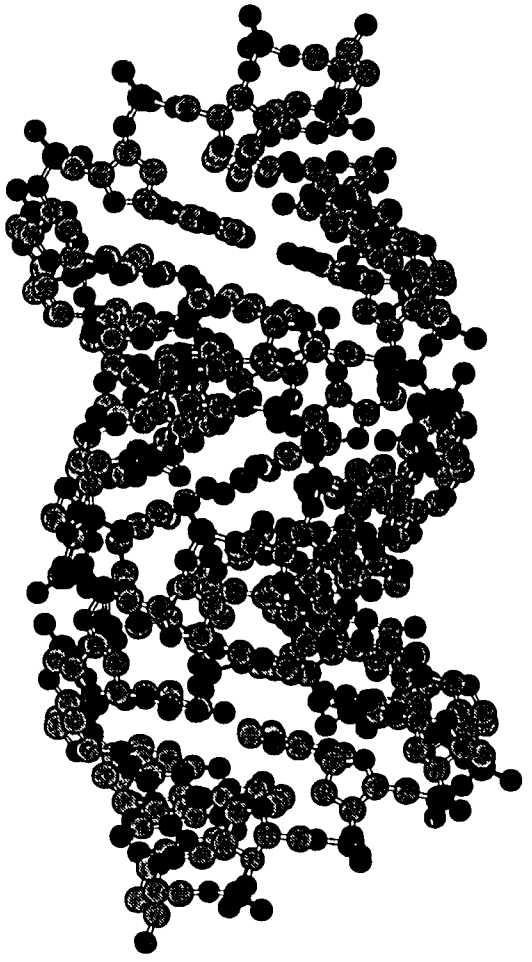


Fig. 1.10. The structure of A-DNA

(generated by HyperChem)

1.3.2 Structure of B-type DNA

While the A-DNA is the only member of A family, there are more members in the B family including B-DNA, C-DNA, D-DNA and others. One of the significant features of B-type DNA is the sugar pucker of C2'-endo (or the related C3'-exo). Thus the distances between adjacent phosphate groups are about 7 Å. The χ angle that defines the rotation about the glycosyl link C1'-N is in the range of -90°-120°. In B-DNA, there are approximately ten base-pairs per turn with an axial rise per nucleotide of 3.3 to 3.4 Å, resulting in a small negative tilt (-6°) of bases. The helical symmetry is about P_{10}.

Natural DNA and synthetic oligomers with various sequences show a B-type conformation. A right-hand double helix with a per base-pair rise of 3.4 Å is identified as a B-type structure. B-, B'-, C-, D-, E-, are considered members of B family with different twist angles, and helical symmetries distinct from each others. Many crystal structures of B-DNA have been published [64,-68,78,87-115]. The typical structure of B-DNA can be summarized as follows.

1. The double helix of B-DNA exhibits a certain flexibility, which can be seen from the bend of the dodecamer d(CGCGAATTCGCG)$_2$. In the crystal the overall helix is not straight but rather is bent by 19° corresponding to a radius of curvature of 112 Å (Fig. 1.11). According to R. Wing, this

distortion seems induced by the crystal packing force rather than being an intrinsic structural property of the complex [51]. From studies of DNA duplex in solution, only 0.25 to 0.5 Kcal/mole dodecamer would be sufficient to bring the observed curvature [116-119]. Bending of DNA could be important for its packing and induce higher order structure.

2. The sugar puckering and conformation of nucleotide are not constant over the whole double helix. They show a strong sequence dependence. The sugar puckering while always in the C2'-endo class varies between C3'-exo and C4'-endo.

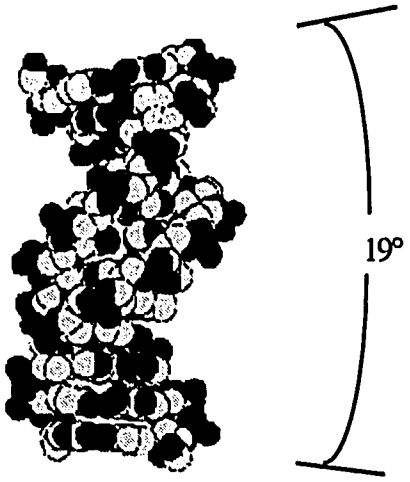


Fig. 1.11 Kink of B-DNA (after[51])

3. The base-pairs are located on the helical axis.

4. The base-pairs tilt by a small angle about -6°.

5. The major and minor grooves are more nearly equal in size than the same sites in A-DNA.

6. The overall conformations of B-type oligomers can vary in a wide range depending on the sequences. The oligomers with different sequences can induce typical B-DNA, or C-DNA, or D-DNA, and other various conformation. In different conformation the helical symmetries and detail structural features are quiet different. Figure 1.12 shows the structure of typical B-DNA.

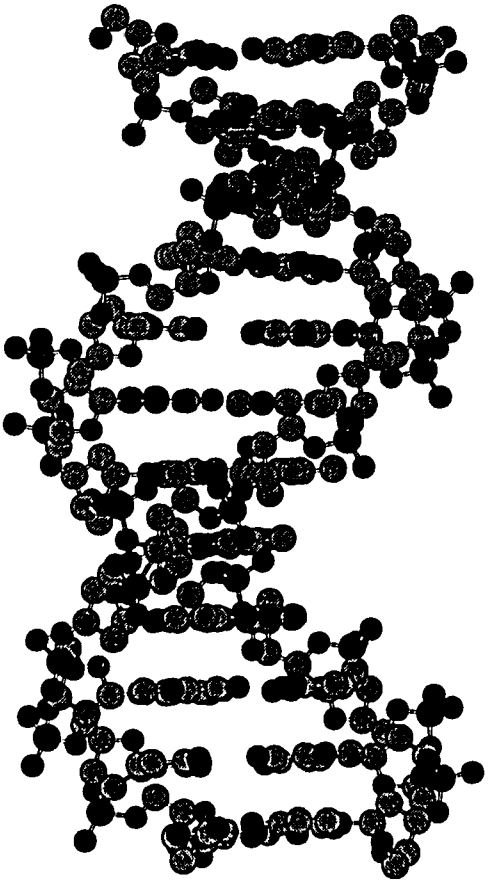


Fig. 1.12. Structure of typical B-DNA

(generated by HyperChem)

1.3.3 Structure of Z-DNA

Z-DNA was discovered by X-ray crystallography in 1979 [50]. It was noted earlier that these DNA sequences gave unusual CD and NMR spectroscopic signals which are now associated with this novel left-handed double helix [70,71]. It was not recognized that the left-handed double helix existed, until Wang, et al. completed their single crystal diffraction experiments on the oligomer $d(CGCGCG)_2$ in 1979. Since then many structures of Z-DNA have been published [55,57,59,61,64,120-132].

The helical symmetry of the Z-DNA double helix is P_{-6} (two base pairs per repeat). The two strands are anti-parallel to each other. The rise per base-pair is about 3.7 Å, and the twist angles per base-pair are different for the CpG steps and GpC steps. The twist angle of CpG step is -45°; while the twist angle of GpC step is -15°. That makes the repeat unit in Z-DNA to contain two consecutive base pair. The twist angle of per repeat unit is -60°. All base-pairs show typical Watson-Crick type pairing but with syn-purine and anti-pyrimidine. The sugars are puckered in C3'-endo (syn-purine) and C2'-endo (anti-pyrimidine) pairwise. This makes the phosphates in the sequences d(CpG) and d(GpC) nonequivalent. The radii of alternative phosphate are not equal, 6.2 Å for the phosphorus in sequence -CpG- and 7.6 Å in sequence -GpC-. Consequently, the phosphorus atoms in a strand form a zig-zag

pattern, which gives the left-hand double helix its name Z-DNA (Fig. 1.13).

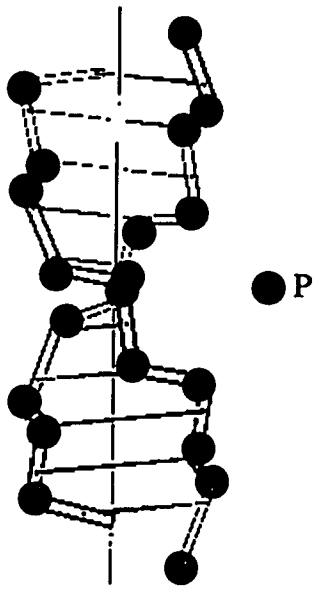


Fig. 1.13. Zig-zag course of line connecting phosphate groups in Z-DNA (after Saenger)

The phosphate groups that link the GC steps can induce two different conformations. They are named ZI and ZII. Figure 1.14 show the base stacking of CG and GC step in Z-DNA.

The base-pairs in Z-DNA are shifted toward the periphery, compared to B-DNA. Thus the major groove is filled with cytosine C5 and guanine N7, C8 atoms. In the overall structure only the minor groove persists. The minor groove is deep, extended to the helical axis.

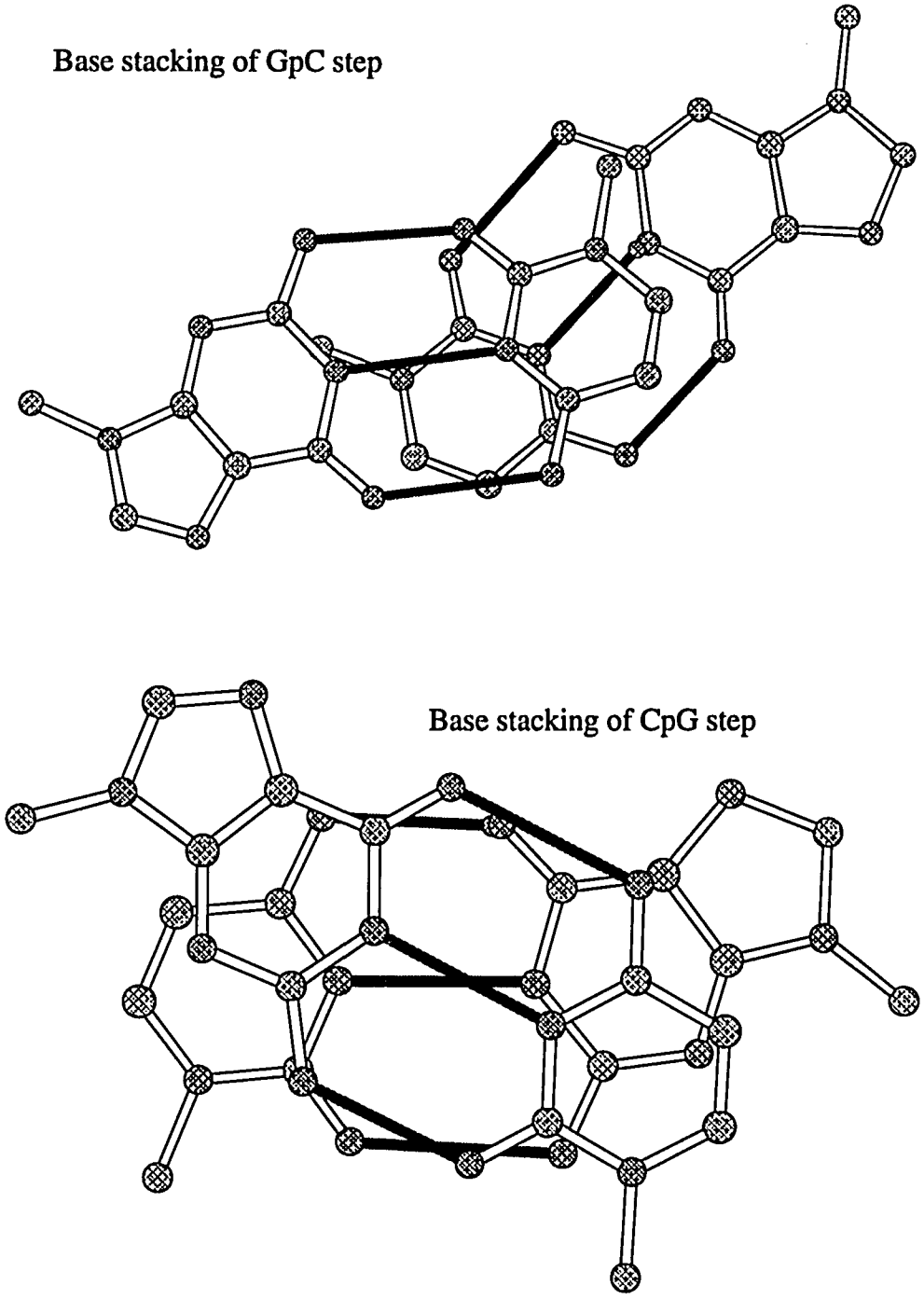


Fig. 1.14. Base stucking in Z-DNA

The different helical twist angles at CpG setp and GpC step are shown.

The structure features of Z-DNA are summarized in Table 1.5. Figure 1.15 shows the overall structure of typical Z-DNA.

Table 1.5. Helical parameters of Z-DNA

	Helix		
	Z_I	Z_{II}	Z_F
Bases per turn	12	12	12
Pitch height (Å)	44.6	44.6	43.5
Rotation per nucleotide repeat (°)	-60	-60	-60
Rise per nucleotide repeat	7.43	7.43	7.43
Base tilt (°)	-7	-7	-5
Radius (Å) of phosphate			
d(CpG)	6.3	6.1	
d(GpC)	7.3	8.0	

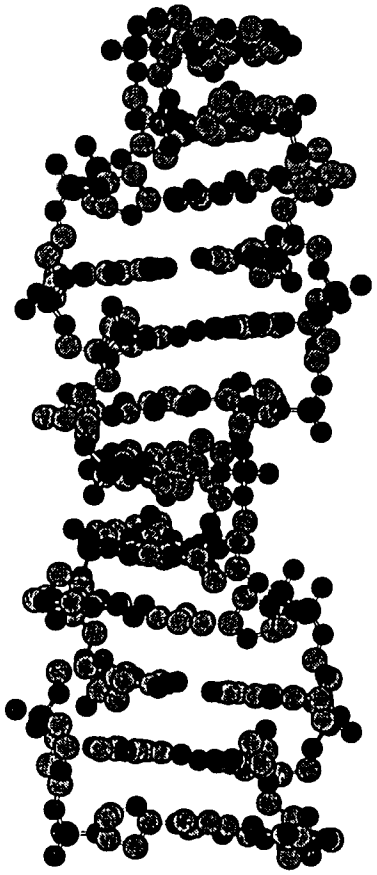


Fig. 1.15. Overall structure of Z-DNA

(generated by HyperChem)

1.4 The Interaction Between Metal Ions and DNA

DNA is a polyanion and as such must be surrounded by sufficient cations to neutralize its negative charge. DNA molecules are always in the form of DNA-cation complexes. There are extensive studies on metal-DNA interaction, since metal ions play important rules in both biological processes and DNA structure. In the procedures of replication and transcription, that is catalyzed by RNA Polymerase, manganese (II) increases the affinity of RNA polymerases to DNA template up to 100 fold [133]. The mechanisms of metal ion in such processes have been reviewed. [134,-137]. In solution the ionic strength is important to the B-Z equilibrium. Multivalent ions, such as Mg(II), $Co(NH_3)_6^{+3}$ are very effective at inducing Z form. In crystals, the counterions or their complexes neutralize the repulsion between the negative charges of DNA molecules and stabilize the structure. The presence of metal ions is a basic requirement to form DNA single crystal [50,139]. On the other hand, metal ion can also have a destabilizing effect on DNA double helical structure. For instance, copper (II) destabilizes DNA double helix and promotes transition into random coil state. [138] The effects of metal ions on DNA structure according to the nature of the metal ion and the conformation of DNA molecules. Metal ions are present in all living bodies at certain levels (Table 1.6).

Table 1.6. Metal contents in human body

Cation	Total concentration in body (g/70kg)	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		

(After Saenger)

The interaction between metal ions and DNA involves mainly charge-charge interaction with some coordination bonding. By the early 80's several review papers summarizing the research on interactions between metal ions and bases, nucleosides, nucleotides, and polynucleotides had been published [136,150,-154]. The four potential binding sites are of so called "hard" ligands [155]. Therefore, only certain types of metal ions are able to bind, i.e. alkali, alkaline earth, and transition metal ions. Metal ions can bind to DNA molecules in four different regions. They are the phosphate oxygen atoms, the ribose hydroxyls and ring oxygen, the heterocycle nitrogen atoms, and the exocyclic base keto groups (Fig. 1.16).

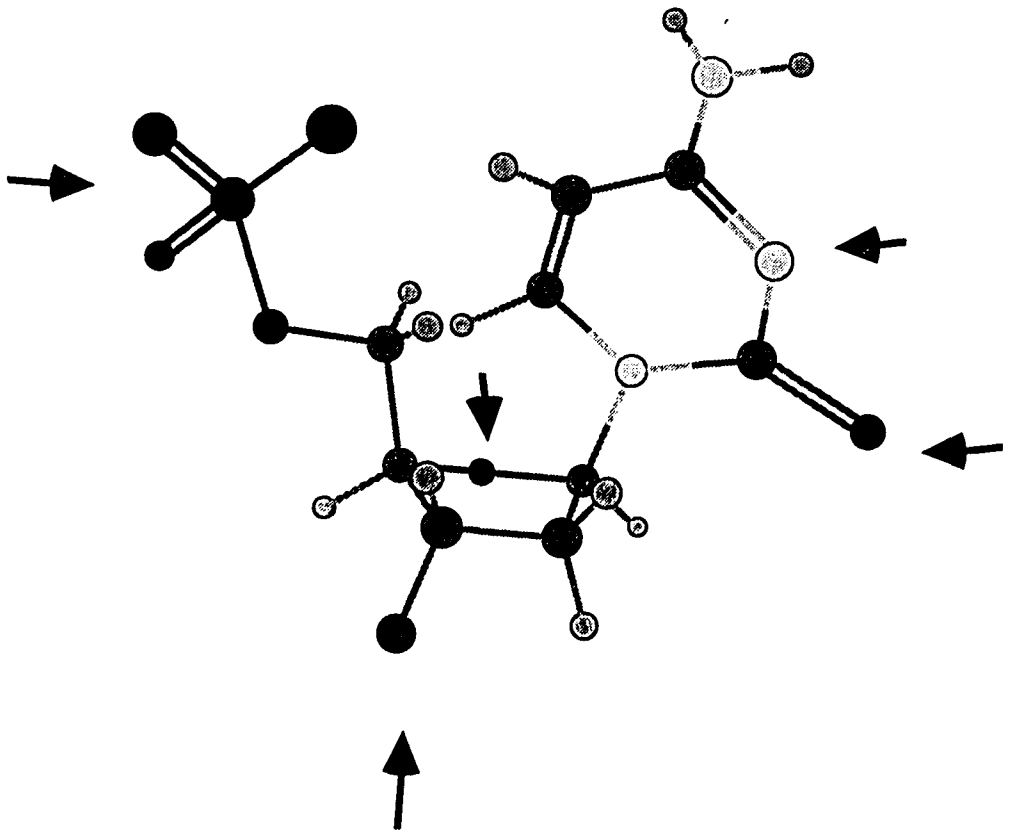


Fig. 1.16. Potential binding sites of metal ion.

The four potential binding sites have various affinities to different metal ions. Since alkali metals and alkali earth metals are "soft", i.e. they interact with DNA mainly by charge-charge interaction, they can bind to all four types of sites. The transition metals, which are more likely to form coordination bonds with DNA, prefer base keto oxygens, nitrogens, and phosphate oxygens. Since we are concerned with the interactions between metal ions and polynucleotides, the purine N9 and pyrimidine N1 sites are not considered.

The alkali metals can bind to all four types of binding sites, usually in the hexahydrate form with octahedral geometry. The alkali metal-water cationic complexes are located in the solvent channel in crystal structure [140] (Fig. 1.17).

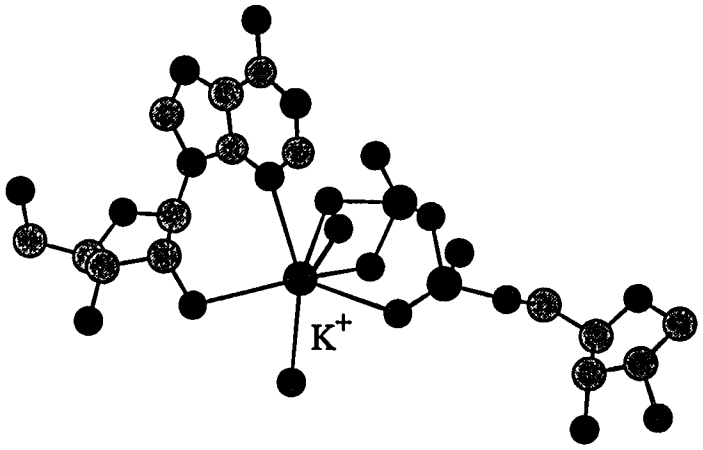


Fig. 1.17. Binding of potassium ions

(after Saenger, graphics is generated by HyperChem)

The divalent alkali earth cations, especially magnesium, bind to DNA more tightly than the alkali metals do, because of their higher charge and smaller radius. The binding sites are less restricted than the monovalent alkali metals. In the crystal structure of $d(CGCGCG)_2$, which is crystallized under conditions of high magnesium(II)

concentration (120 mM), an unusual magnesium(II) dimer complex is reported to be observed [73] (Fig. 1.18).

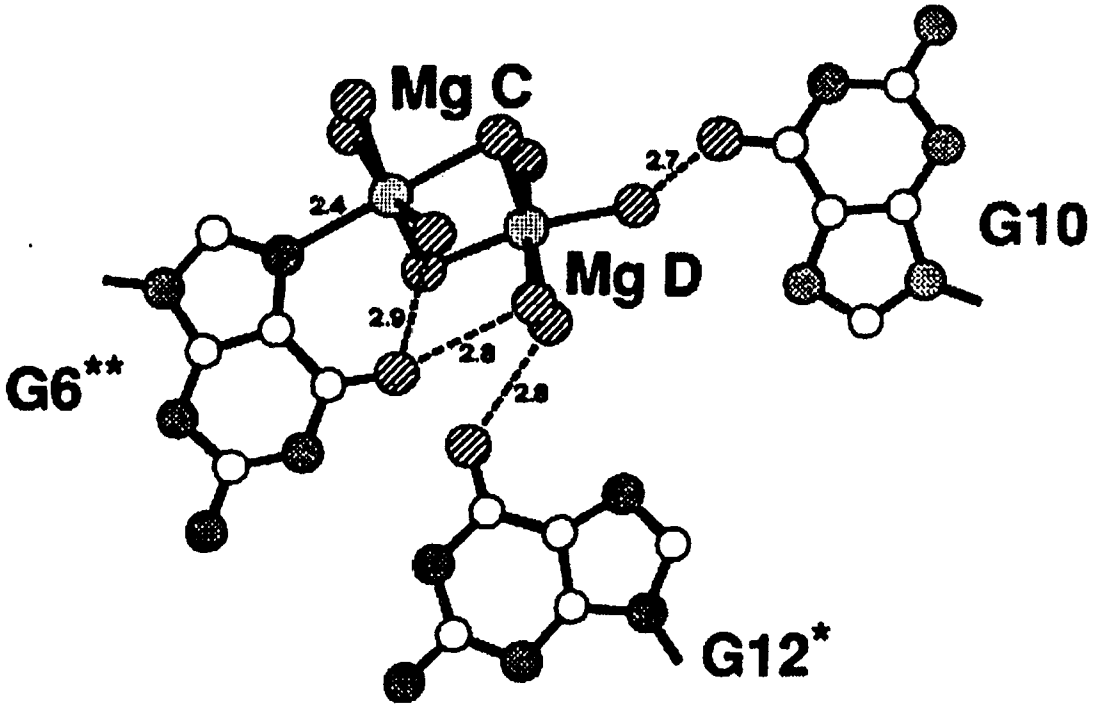


Fig. 1.18. Magnesium Binding sites in Z-DNA (from [73])

Transition metals show various behaviors. Their preferred binding sites are purine N7, phosphate oxygen and pyrimidine N3. Cadmium(II) can bind to purine N7, pyrimidine N3, phosphate oxygen, ribose O2', and O3'. (Figure 1.19) [141]. The geometry of Cd(II) is usually octahedral. Cobalt(III), nickel(II), copper(II), zinc(II), and platinum(II) also usually bind to purine N7, pyrimidine N3, and phosphate oxygen atoms.

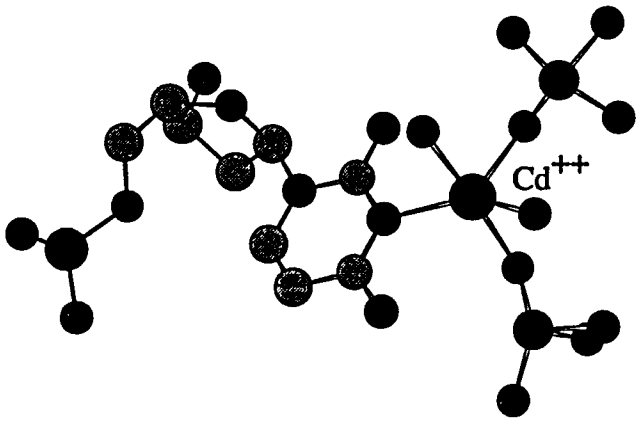


Fig. 1.19. Cadmium binding sites (after Orbell)

The geometries of the transition metal ions are: Co(III) octahedral and tetrahedral; Ni(II) octahedral; copper(II) octahedral, square plane, square pyramid, and trigonal bipyramid; Zn(II) tetrahedral; Pt(II) square plane [41]. Under strong alkaline conditions Ag(I) can directly bond to pyrimidine O4 and deprotonate N3 [142] (Figure 1.20).

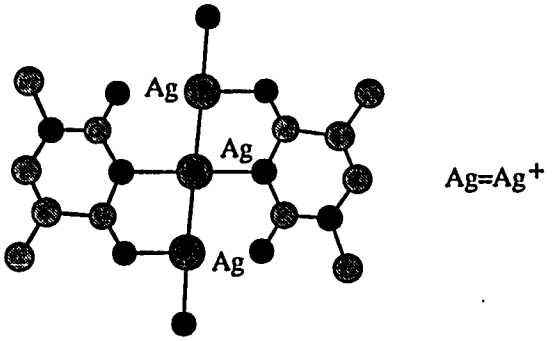


Fig. 1.20. Coordinated Ag(I) 1-methylthymine complex (after [142])

Platinum(II) and its complexes can interact with DNA by direct binding [143] and intercalation [144]. The interactions between platinum and DNA have been thoroughly examined by many research groups. We will not cite all the literature here. The antitumor activities of cis-dichlorodiamineplatinum (II) have made platinum(II) compounds a "hot topic" and advanced the research on metal-nucleic acids complexes [145-149].

1.5 DNA-Drug Complexes

In clinical anticancer chemotherapy, DNA is a target of antibiotics and anticancer drugs. The drug-DNA interaction have been found to have various effects on DNA repair and inhibition of DNA replication and transcription. The drug induced inhibition of replication and transcription have found wide clinical use [19,20,156,157]. The drug induced DNA repair is a rather new aspect [158].

The interactions between drugs and DNA can be classified into four groups: (1) inter strand cross link, (2) intercalation, (3) chemical modification, and (4) groove binders. In this study we focused on the structure characteristics of intercalated DNA-drug complexes.

Some early approaches to structures of DNA-intercalator complexes were provided by fiber diffraction experiments on DNA-

actinomycin complexes [159] and DNA-platinum derivatives [144,160]. From those observations the basic structures of intercalated DNA-drug complexes were ascertained. The fundamental features of those complexes are as follows [159,161-164]:

1. The intercalators have a large hydrophobic planar portion in their molecular structures. When the drug molecule intercalates into DNA, this hydrophobic portion inserts between the stacked base pairs. The separations between base pair planes and drug plane are about 3.4 Å, which is of the same distances between the normal base pairs, the thickness of an aromatic group.

2. The nearest-neighbors of intercalated base pairs are excluded from additional intercalation, so that every second (next neighbor) intercalation site remains empty when intercalation occurs.

3. Intercalation causes DNA unwinding of about 10-20°.

4. Some intercalators are sequence specific.

The crystal structure of DNA-daunomycin complexes [113,165-167], of DNA-triostin A complex [114,168,169], of DNA-echinomtcin [115], and of DNA-ditercalinium complex [21] were reported in the

1980s. Other drug and dye complexes of DNA have also been crystallized and the structures were determined [170]. The detailed structures obtained by X-ray single crystal diffraction have confirmed the results of fiber diffraction and have led to a better understanding of the DNA-drug complexes. Like base-base stacking in DNA, the base-drug stacking in the complexes have shown hydrogen bonding and other interactions. There are not only local structural changes, such as unwinding and kinking at the intercalation sites, but also longer range distortions have been observed [113,171].

Ditercalinium, 7H-pyridocarbazole dimer, is a synthetic anticancer drug first reported in 1978 [172]. Ditercalinium binds to DNA as a bis-intercalator and triggers a DNA repair process [173-175]. The theoretical and NMR studies on the interaction of oligo-DNA, $d(CGCG)_2$, and ditercalinium had resulted in a three dimensional structure model [176,177], which provides us helpful clues in the complete determination of the molecular structure of the complex. In 1990, L. Williams successfully crystallized DNA-Ditercalinium complex, which were examined in this study.

1.6 Solvent Interactions

Water molecules play an important role in stabilizing the structure of DNA both in solution and crystals. In early studies of fiber

structures, Gosling reported the observations of conformation change of DNA structures with changes of the relative humidity [32]. By examining the changes of diffraction pattern, he pointed out that the B-DNA form, which is stable at high humidity (approx. 92% relative humidity), changes to the A-DNA form, when the relative humidity drops to about 75%, and continued drying led to gradual loss of the double helical diffraction pattern. Gosling and many later researchers believed that the water molecules must be associated most tightly with backbone phosphate oxygen atoms, and less tightly with polar N and O atoms of sugar rings and base-pair edges [95]. Single crystal diffraction experiments confirmed the existence of water molecules in the crystals, but disagreed with the prediction of the binding sites of water molecules with DNA [50,87]. The results of single crystal diffraction experiments show that the most well defined water molecules are found to be associated with polar N and O atoms of base-pair edges and sugar rings. At room temperature there are no visible water molecules that associate with the phosphate oxygen atoms in B-DNA. However, at low temperatures (16° and 280° K) some of the water molecules are found to be associated with phosphate oxygen atoms in B-DNA [92].

The hydration shells of DNA may be described as two discrete layers, the primary and secondary layer of water [178]. The primary layer consists of water molecules immediately adjacent to the DNA, which are in a different chemical state than liquid water [179,180]. In the secondary shell, the water characteristics are essentially those of liquid

water. It was proposed that in the primary shell 10 to 12 water molecules per nucleotide associate with base-pairs and 8 to 10 associate with backbone phosphate group. That makes the total number of water molecules to be 18 to 23 per base-pair.

Crystal studies show that the interaction between the DNA molecule and water is mainly though hydrogen bonding. The DNA/RNA molecule has a number of potential hydrogen bond acceptors as well as several hydrogen bond donors. The phosphate oxygen atoms (O1P, O2P), furanose hydroxyl groups (O2', O3', O5') and furanose ring oxygen O4', cyclic keto groups (O6 of pyrimidine and O2, O4 of purine), and cyclic nitrogen (N7, N3 of purines, N1 of adenine, and N3 of cytosine) atoms are all hydrogen bond acceptors; while the amino groups (N1 of guanine, N2, N6 of purine, N4 of cytosine, and N3 of thymine) are the hydrogen bond donor (Fig. 1.21). Among these donor-acceptor pairs, the N4(C)-O6, N1(G)-N3, N2(G)-O2, N6(A)-O4, and N3(T)-N1 form the hydrogen bonds in Watson-Crick base-pairs.

In the structures of DNA we can classify the hydrogen bonds into two groups, intra-molecular hydrogen bonds (the hydrogen bonds formed within the DNA molecules), and inter-molecular hydrogen bonds (the hydrogen bonds formed between DNA molecule and water molecules). The intra-molecular hydrogen bonds were already discussed in section 1.4. Specifically these hydrogen bonds, the Watson-Crick hydrogen bonds are: cytosine O2(acceptor) - guanine N2(donor),

cytosine N3 (acceptor) - guanine N1 (donor), cytosine N4 (donor) - guanine O6 (acceptor), thymine N3 (donor) - adenine N1 (acceptor), and thymine O4 (acceptor) - adenine N6 (donor).

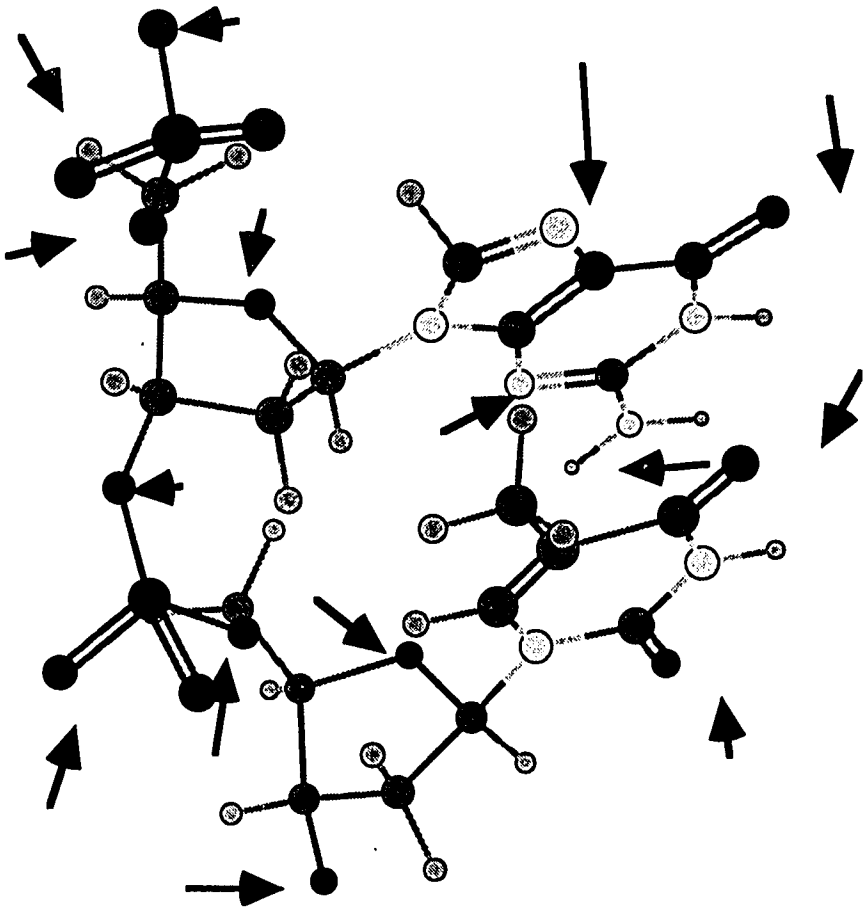


Fig. 1.21. Potential binding sites of water

Generally, water molecules occupy the inter-molecule channels and grooves. Not all the water molecules can be determined by crystallographic methods because of their high mobility. In the first crystal structure of DNA, Z-form of $d(CGCGCG)_2$, 62 water molecules were located with well defined positions [50]. In the structure of B-form

of d(CGCGAATTCGCG)$_2$ 83 water molecules were determined [51]. The water molecules with undefined positions are called free water.

The free waters cause a continuous background in the X-ray diffraction pattern, which is called the solvent mask. The solvent mask is an important factor that limits the precision of the crystal structures. Some water molecules form complexes with metal cations via coordination bonds. For alkali metal cations, water molecules usually form a octahedral six-coordinated complexes (Fig. 1.18). These complexes usually form bridges between phosphate oxygens (O1P, O2P), sugar hydroxyl (O2'), and base (N3) [140]. For the alkali earth metal cations, for example Mg(II), a dimer of Mg-water complex, that directly coordinates to N7 of guanine, had been observed besides the simple octahedral complex (Fig. 1.19). A second category is the water chains spanning the grooves. These water molecules form intra-molecule water bridges [92]. They are (Fig. 1.22): the phosphate oxygen-water-N2 bridge, the O6-water-O6 bridge, the N2-water-water-phosphate oxygen bridge, and the N4-water-water-N4 bridge

A third category consists of pentagonal water networks that are formed by the water-water hydrogen bonds [181] (Fig. 1.23).

Most of the well defined water molecules belong to the primary hydration shell. Together with counter ions the primary hydration shell stabilizes the structures of DNA both in crystals and in solution. The water molecules within secondary hydration shell are highly mobile and

frequently can not be observed by X-ray diffraction.

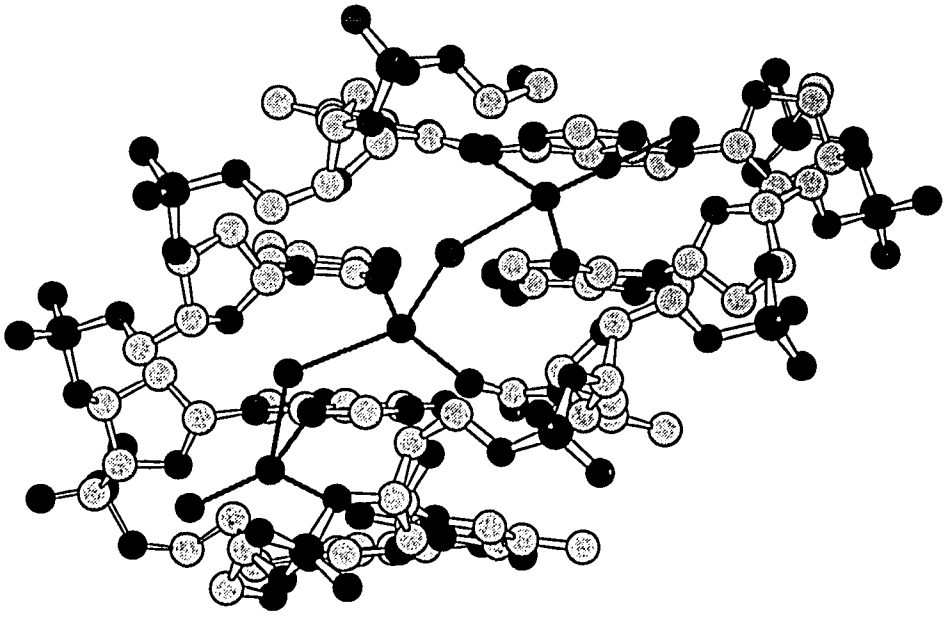


Fig. 1.22. Water structure in minor groove (after [87])

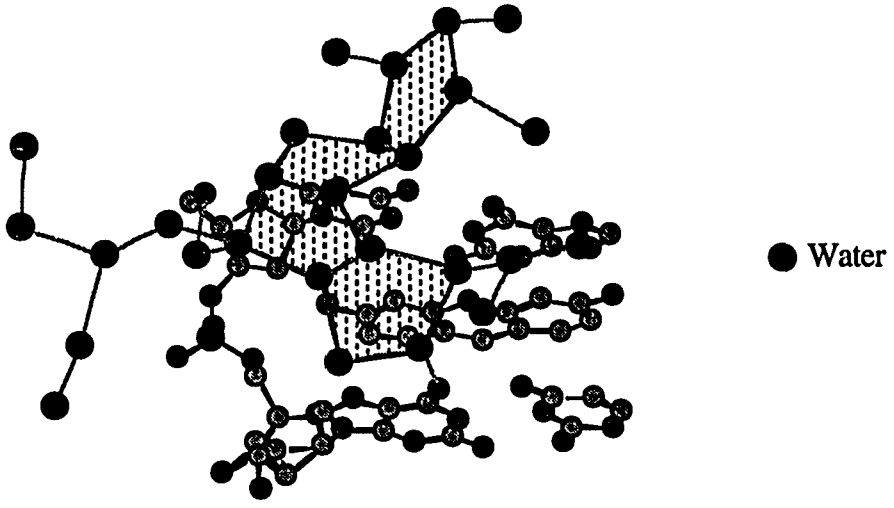


Fig. 1.23. The water pentagon structure

(after S. Neidle)

CHAPTER II METHODOLOGY

X-ray crystallography is the fundamental technique for DNA structure studies. Usually, once a quality single crystal has been obtained one can solve the structure, but one must first obtain high quality single crystals. Due to the difficulties of growing such single crystals, the technique of crystallizing the DNA molecules becomes a critical factor.

2.1 Crystallization of DNA Oligomer

The driving force for macromolecules, such as DNA oligomers and proteins, to crystallize is the same as for all thermodynamic systems --- the minimization of free energy. The general principles that govern the crystallization of small molecules also hold for the macro-molecular systems. In general, the degree of supersaturation, the pH, the temperature, and the speed of crystallization are major factors. In practice, the crystallization of macromolecules is much more complex and difficult than of small molecules because of their geometric shapes, flexibilities, and multi-conformational natures. The first crystallized protein, pepsin, was reported in 1934 [182], while the first DNA single crystal, Z-conformation of $d(CGCGCG)_2$, was reported forty-five years later in 1979 [50].

The factors that caused difficulties in crystallizing DNA were the

lack of pure oligomers and hence the lack of conformational homogeneity in solution. The high flexibility and multi-conformational nature of DNA in solution makes such solutions multiphase equilibrium systems that hardly meet the requirement of homogeneity. Prior to any other consideration, choosing species which form a homogeneous solution is the first priority in the crystallization of DNA oligomer, as well as protein crystallization.

The strategy used to induce crystallization of DNA oligomer is to bring the system toward a limited degree of supersaturation. In reaching the supersaturation condition, the homogeneity of solution must not be broken. Owing to their shapes, polyvalent character, and general physical properties, the behavior of macromolecule in solution is complicated. This complexity can lead to multi-minimum of solubility dependent upon the concentration and type of the electrolyte, the concentration of DNA molecules, the pH, the temperature and other influences [60,183]. Thus, a given DNA oligomer can often be crystallized in different unit cells.

The general procedure is to bring the system to low supersaturation state slowly, then to maintain such a condition for a certain time period. There are a number of techniques employed successfully to induce the growth of DNA single crystals. All of them are adopted from protein crystallization. The following techniques have been reviewed by McPherson [184].

1. Batch method. This is a classical technique described for early enzyme crystallization. The protein is dissolved at low ionic strength to give a solution of high concentration. The precipitating agent (salt or organic solvent) is then gradually added to bring the solution to a state of slight supersaturation. After a certain time period, without disturbing, crystals appear. Usually, a relative large amount of material is required. This method has been successfully employed to produce large single crystals of lysozyme, a small globular protein, and many others.

2. Equilibrium dialysis. This technique was introduced by Theorell [185], and was used in the crystallization of hemoglobin [186]. By use of a dialysis membrane the dialysis solution and mother liquor containing macromolecule and salts can be separated. The small ions and solvent molecules can diffuse through the membrane but the macromolecules will not. The ionic strength and pH of a macromolecule solution enclosed in such a membrane is adjusted by equilibration against a macromolecule free solution. By the early 1970's, dialysis method had been modified to work at micro-liter scale. [187,188]

3. Vapor diffusion. This is currently the most widely used technique and it has various versions. Regardless of variations in details the principle is the same. McPherson [184] and Wang [183] have reviewed the technique. The principle of the vapor diffusion method is

that two separate solutions, a working solution and a reservoir, equilibrate with each other via vapor diffusion in a sealed container to bring the working solution to a low supersaturation state. To stabilize the overall condition the volume of the reservoir is much larger than that of the working solution. In practice, both standing drop and hanging drop methods are being used successfully for various material. Normally, 10 to 40 micro-liter drop will produce single crystals of suitable size.

The working solution consists of DNA molecules, buffer, counterions, and low concentration precipitating agent. The concentrations of DNA and precipitating agent are carefully set up to make the system slightly undersaturated. The reservoir is usually a water solution of precipitating agent, the concentration of which will guarantee the working solution will be a the low supersaturation state when the equilibrium state is established.

For crystallization of DNA oligomers 2-methyl-2,4-pentanediol (MPD), polyethylene glycol 400 (PEG 400) as well as higher molecule weight PEG, and isopropanol (iP) are used most commonly as precipitating agents.

2.2 Diffraction of X-rays

X-ray diffraction is a widely used experimental technique in structure determinations. The following sections give a very brief discussions. More details can be found in various books [189-191].

A collimated beam of monochromatic X-ray are scattered in all directions by the electron complement of every atom in an object, when the beam is directed through an object. If the object is a single crystal, which is a highly ordered three dimensional periodical lattice, the scattered X-ray beam will form a specific diffraction pattern (Fig. 2.1). The smallest repeat unit of the three dimensional lattice is defined as the unit cell. The diffraction pattern is determined by the three dimensional lattice of the object, i.e. the symmetry and distribution of the electron complement. Different crystals generate different diffraction patterns. The diffraction patterns can be recorded by film, by photoelectron multiplier, and or by area detectors. The bright spots in a diffraction pattern are called reflections. The reflections are distinct by their positions and intensities.

The intensity of a reflection, **I**(hkl), can be simply expressed as the product of a structure factor, **F**(hkl), and its complex conjugate **F***(hkl).

$$\mathbf{I}(hkl)=\mathrm{Scale}\mathbf{F}(hkl)\mathbf{F}^{*}(hkl)$$

A structure factor, **F**(hkl), is defined as a Fourier transform of the distribution of the electron density in the crystal.

$$\mathbf{F}(hkl)=\Sigma f_i e^{i\,2\pi(hx_i+ky_i+lz_i)}$$

Here **I**(hkl) is the intensity of a reflection with Miller indices (hkl); **F**(hkl) is the structure factor of the reflection; "Scale" is the scale factor determined by geometry of the crystal and data collection factors; x_i, y_i, z_i, are the fractional coordinates of atom i; f_i is the scattering factor of atom i; the summation is over the whole unit cell in three dimensions; and **F*** is the complex conjugate of **F**.

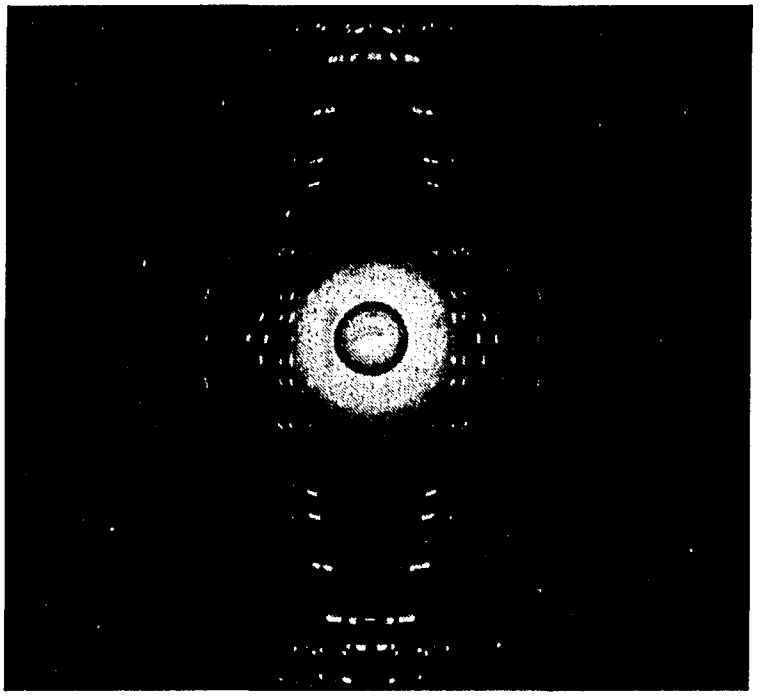


Fig. 2.1. X-ray diffraction pattern of Z-DNA

In principle, one can perform a reverse Fourier transform on the diffraction pattern to solve the molecular structure that constitutes the single crystal. Practically, there is no simple straight forward way to do this. Experimentally, one can only record the intensities of reflections. The phases of reflections are completely lost. In order to perform the reverse Fourier transform one must has a phase to every reflection. This is the well known phase problem. Various approaches have been employed in solving the phase problem. They are Direct Methods, Patterson Function, Multi-Isomorphous Replacement, Anomalous Scattering, and Molecular Replacement. With the assigned phases and intensity data of sufficient resolution, one can solve the structures with certainty. Due to imperfections in the crystal, thermal vibration of the lattice, the lack of precision of measuring scattering factors, the errors of phase assignment, and other experimental errors, structure refinement is always necessary.

In today's scientific community, structure determination of small molecules is generally no longer a problem. By using modern X-ray diffractometer, electronic computer, and commercial software, X-ray crystallographers solve structure of small molecules routinely, usually spending a few days to solve one structure. For the biological macro-molecules, the structure determination is still a difficult research project, usually, such a project can include material preparation, such as synthesis or separation from natural products, molecular modification, crystallization, data collection, phase determination, model building and

rebuilding, molecular dynamics simulation, and structure refinement. Every step require creative work.

2.3 The Phasing Technique

There are several techniques to solve the phase problem in structural determination of DNA, or more generally macromolecules. They are Multiple-Isomorphous Replacement (MIR), Anomalous Scattering (AS), Molecular Replacement (MR), and the closely related Rigid Body Search (RBS) and Patterson methods. Since the overall structures of DNA double helices show less variation than other macromolecule, the Rigid Body Search methods are frequently effective in primary phasing. Rigid Body searching can be performed in both real space or Patterson vector space. These phasing techniques are briefly discussed in the following sections.

2.3.1 Multi-Isomorphous Replacement

Heavy atoms with large scattering factors can make a significant contribution to the structure factors. In general, by introducing heavy atoms into a macromolecular crystal, one can make more than one heavy atom derivative of a macromolecule. Some heavy atom derivatives will not disturb the native crystal structure too much. In this case the phases of the observed reflections can be determined by the

Multiple Isomorphous Replacement methods. The structure factors of native structure and its heavy atom derivative are shown in Fig. 2.2.

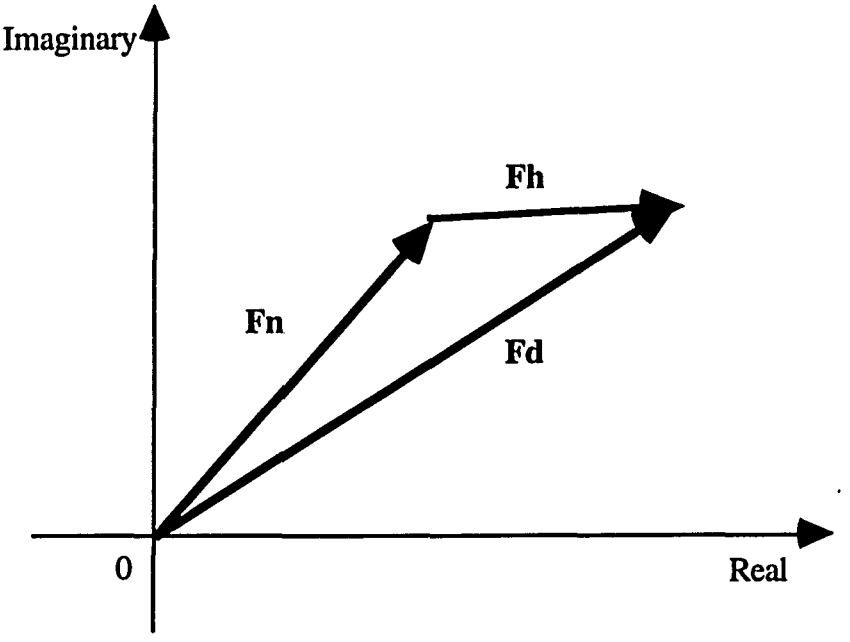


Fig. 2.2. A vector diagram of structure factors **Fd=Fn+Fh.** Where **Fn** is the vector component from native structure, **Fh** is the component of heavy atom, **Fd** is the synthetic structure factor.

The position of heavy atoms in unit cell can frequently be determined easily by the Patterson method or direct methods. After the positions of heavy atoms have been determined, the structure factors, **Fh**, of heavy atoms can be obtained by direct summation or Fourier transform. Combining the calculated structure factor of heavy atoms,

Fh, and the observed amplitudes of structure factors of the native structure, $|F_n|$, the heavy atom derivatives, $|F_d|$, the phases of F_n, and F_d can be determined with two fold ambiguity (Fig. 2.3.).

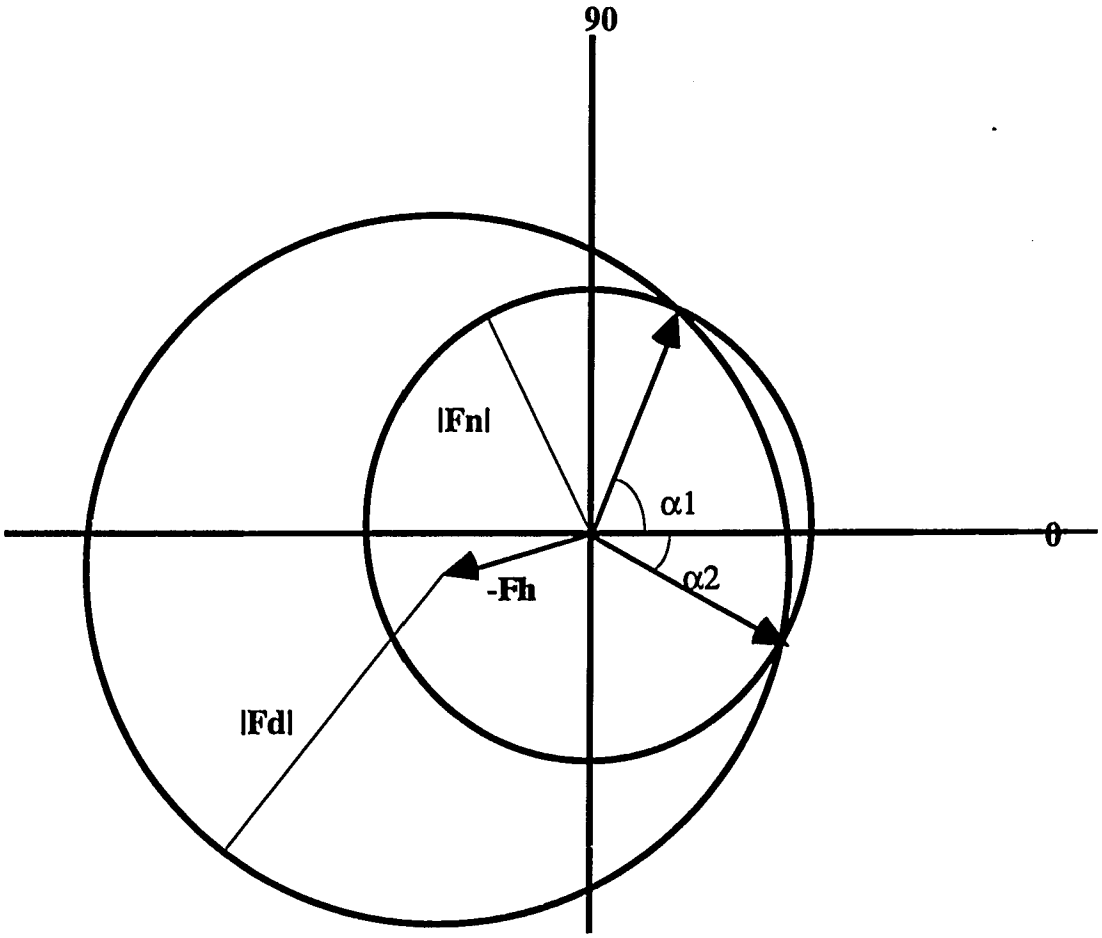


Fig. 2.3a. Phase determination using one heavy atom derivative.

Where **|Fn|** and **|Fd|** are structure factor of amplitudes of native structure and structure factor of heavy atom derivative, respectively. **α1** and **α2** are two possible phase angles. **-Fh** is the negative of structure factor of heavy atom.

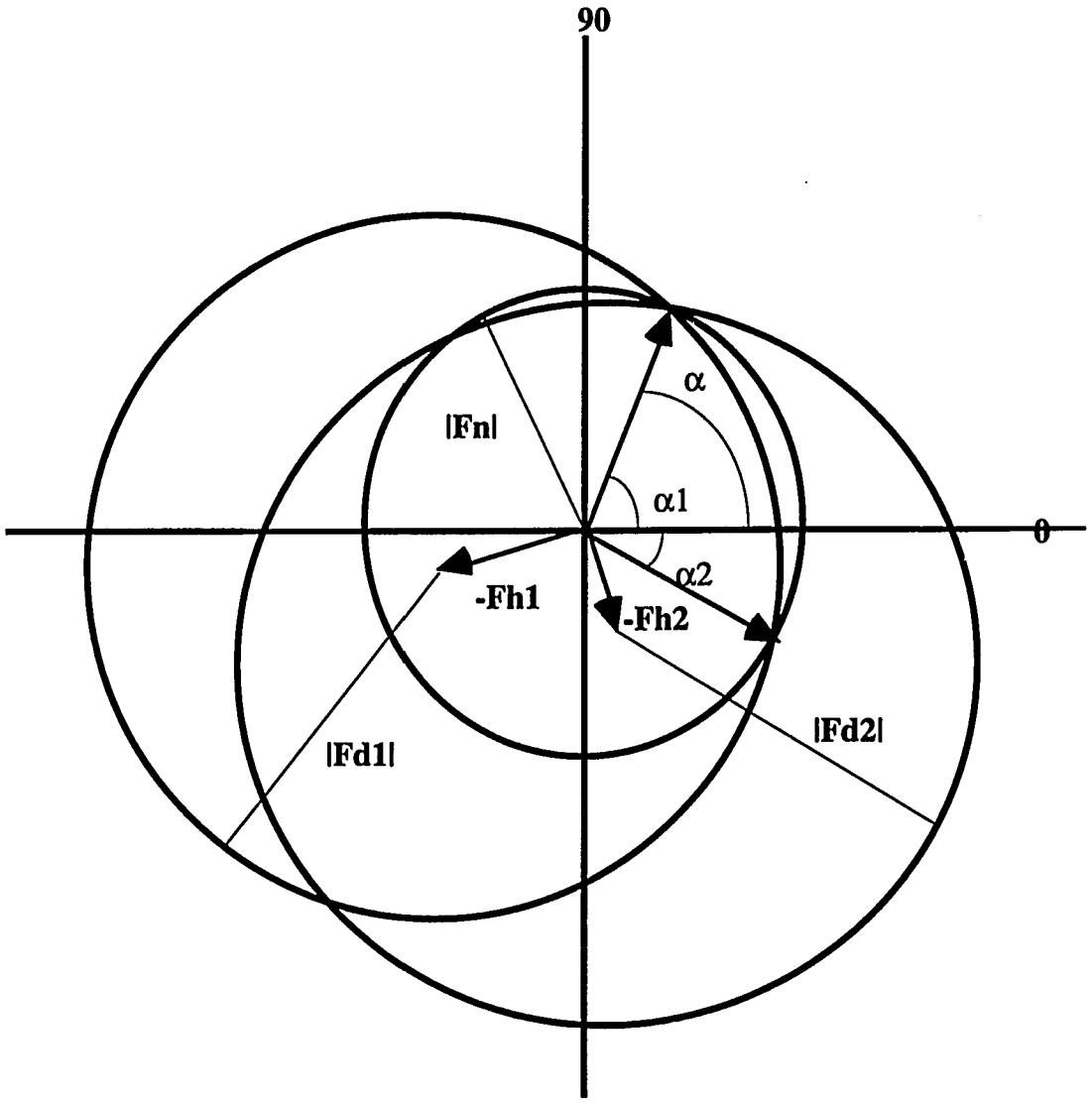


Fig. 2.3b. Phase determination using two or more heavy atom derivatives.

Where α is the determined phase.

In order to determine the phases of structure factors, at least two different isomorphous derivatives are needed in addition to the native structure. In the crystals of DNA oligomers there are always inter molecule solvent channels. Heavy atoms can easily position themselves in these inter molecule channels without significant distortion on the native structure.

The commonly used heavy atoms include strontium, barium, cobalt, mercury, uranium, platinum, and others. For DNA molecules 5Br or 5I (bromine or iodine coordinates to C-5 position of Cytosine, or Uridine bases) are commonly used synthetic analogs which can often be crystallized isomorphously.

2.3.2 Anomalous Scattering

Multiple isomorphous derivatives of DNA crystal are not always available. When lacking multiple isomorphous derivatives, the use of anomalous scattering is an alternative technique to solving the phase problem. Some heavy atoms exhibit strong anomalous scattering when the incident X-ray beam is near the characteristic absorption wavelengths of the atoms. When anomalous scattering occurs, the pairwise reflections with Friedel related indices, say reflection **F**(hkl) and **F**($\bar{h}\bar{k}\bar{l}$), show a significant differences in intensities. This feature can also be used in phase determination. The structure factors with anomalous scattering are shown in fig. 2.4.

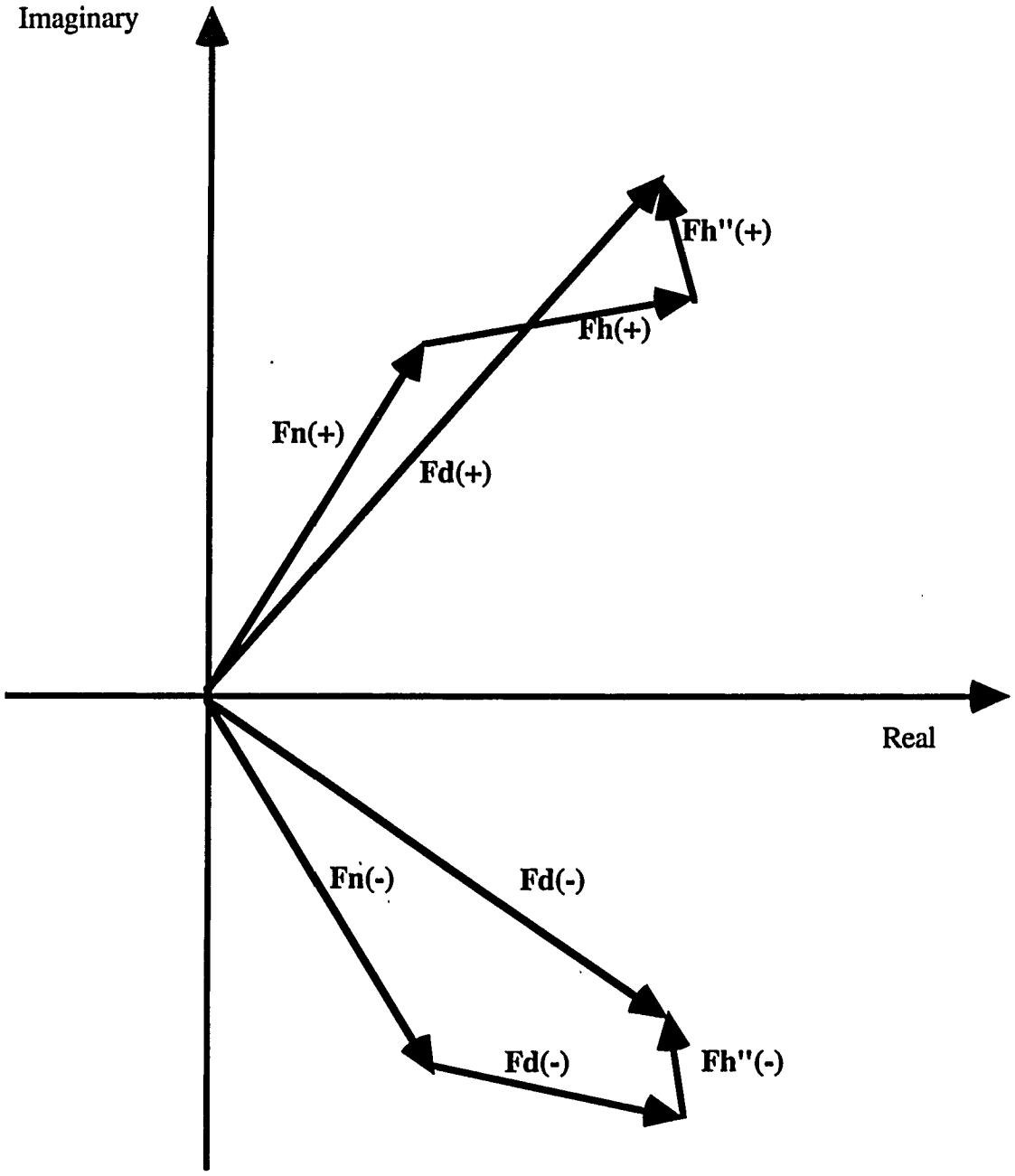


Fig. 2.4. The structure factors with anomalous scattering

Where Fh"(+) and Fh"(-) are the anomalous scattering; Fd(+) and Fd(-) are the structure factors of Friedel's pair.

Once the positions of heavy atoms have been located, the anomalous scattering can be used to calculate the vectors of anomalous scattering, Fh". Using the structure factors (Fh) and the vector of anomalous scattering (Fh"), a procedure similar to that employed in MIR methods can be performed to determine the phases (Fig. 2.5). The use of anomalous scattering allows phase determination by single heavy atom derivative.

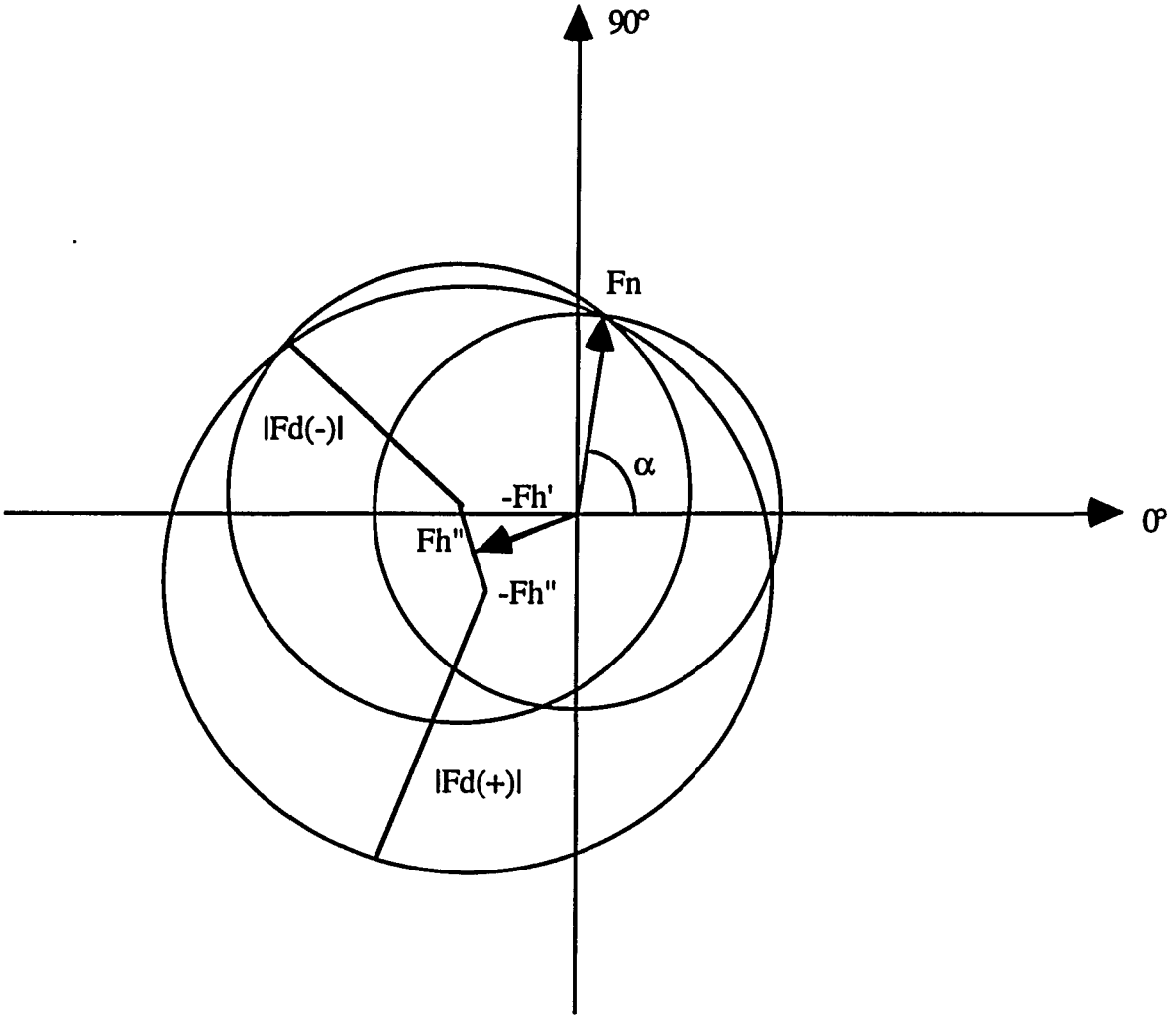


Fig. 2.5. Phase determination by isomorphous replacement with anomalous scattering

2.3.3 Rigid Body Searching

In most cases of structural determination of DNA oligomer, the rigid body searching technique is extraordinarily useful. The DNA double helices are quit defined structures. The variations in detail frequently do not change the overall shape. One can often take a pre-established model, which could be a section of a known structure or a computer generated model, to perform translation and rotation searching. The range to be searched may be the entire unit cell or only a portion of it according to the symmetry and space group.

The search can be conducted in real space as well as in Patterson space defined by the Patterson function

$$P(uvw)=\frac{1}{V}\sum_{h}.\sum_{k}\sum_{l}|F|^2 e^{-i\,2\pi(hu+kv+lw)}$$

where $u=x_i-x_j$, $v=y_i-y_j$, and $w=z_i-z_j$. (i, j, k are from 1 to the total number of atoms, N); V is the volume of unit cell.

In the Patterson function the structure factor is replaced by the square of structure factor. By using the Patterson function, the phase problem can be avoided. The densities in a Patterson map are interpreted by the different terms of the coordinates, (u, v, w).

Applying a rotation and translation search on a model can locate a position where the calculated structure factors (rigid body search) or densities of Patterson function (Patterson search) fit the observed ones. When the calculated and observed data reach a fair agreement, the search process is complete.

The structure of ditercalinium - $d(CGCG)_2$ complex has been solved by the rigid body searching method [21]. Since the models employed in the searching procedure is usually only a rough approximation of the structure, searching should start at fairly low resolution. Normally 5 to 6 Å resolution is a good starting point.

Summarizing the phasing techniques, each method mentioned above can only give approximate sets of phases. As a result of the approximate nature of the phasing techniques and the unavoidable experimental and calculation errors, structure refinement is always necessary to achieve the final structure.

2.4 Structure Refinement

Using observed data and calculated phases, one can generate an electron density map by means of Fourier transformation, using the relationship

$$\rho(xyz)=\frac{1}{V}\sum_{h}\sum_{k}\sum_{l}F(hkl)e^{i\,2\pi(xh+yk+zl)},$$

where ρ is electron density at point (xyz), $F(hkl)=|F_{obs}(hkl)|e^{i\,\phi_{calc}(hkl)}$. From these maps a structure model can be established.

The atomic coordinates based on such models are not very accurate, partially because the maps are of insufficient resolution to show the constituent atoms individually and partially because the phases used to calculate the electron densities are not very accurate. In order to understand much of the chemistry of DNA molecules and how they function more accurate structural parameters are required.

Usually, the resolution of X-ray diffraction of DNA single crystals is in the range of 1.2 Å to 2.5 Å, the best previously reported diffraction data being 0.9 Å resolution [50]. Therefore, the refinement is essential. The traditional nonlinear least square refinement, that have been used to refine small structures successfully, are not sufficient to refine the molecular structures of macromolecules.

Mathematically, the number of independent observations must at least equal the number of parameters to be determined. In the structure determination every atom needs four parameters to define its geometric positions and isotropic thermal motion. Thus to determine the structure

of a molecule consisting of N atoms, at least 4N unique reflections are needed. In practice, to minimize experimental errors by the least square technique the number of unique reflections should greatly exceed the number of parameters. The total number of available reflections for a certain single crystal, N_t, is given by the equation

$$N_t=(4.2)*V/nd_{min}^3 \qquad 2.1$$

where V stands for the volume of the unit cell, n is a constant depends on the symmetry, d_{min} is the resolution limit [191].

For the structure determination of DNA oligomer, the ratio of the number of observations to number of parameters (coordinates, thermal factors, etc.) can be larger or less than unity. Therefor, the conventional least square refinement will not normally work for DNA structures. The following techniques developed for protein structure refinement are used in DNA structure refinement.

• PROLSQ, Protein Least Square, created by Wayne A. Hendrickson and John Konnert. This approach introduces geometric constraints (bond length, bond angle, planarity, chirality, and some torsion angles) to increase the ratio of over determination.

• RSR, Real Space Refinement, developed by R. Diamond. This approach treats some portion of the molecule, such as cyclic bases, with

fixed bond lengths and bond angles and adjustable rotation angle around the bonds to increase the ratio of over determination.

• CORELS, Constrained-Restrained Least-Squares, developed by Joel L. Sussman. This approach combines geometric constraints and energy restraints to improve the ratio of over determination.

• X-PLOR, developed by Axel T. Brünger. In X-PLOR, two approaches have been used. First an empirical energy term combining all constraints/restraints, electrostatic interactions, and agreement between observed data and calculated parameters, etc. is used to improve the ratio of over determination. Second molecular dynamic simulation is introduced to overcome the problem of multiple minimum by a simulated annealing procedure (SA). A simulated annealing applies a fast heating and slow cooling procedure on the structure, to allow the release of internal stress of the model. Some details will be given in the following chapters.

The fundamental difference of these refinement techniques with the conventional methods is the introduction of constraints and restraints to reduce the total number of independent parameters or effectively increase the number of " observations". Since the bond lengths, bond angles, and some torsion/dihedral angels are pretty rigid, application of constraints and restraints on DNA structures is quite acceptable. By reducing the total number of independent parameters, the ratio of

observations to the number of degrees of freedom increases. That makes the refinement practically available. We will discuss some of the techniques in more detail later.

2.5 Modeling of DNA Oligomer

Since the derived electron density maps rarely provide an accurate and clear image at atomic resolution, constructing an atomic model is an essential step. The molecular modeling in DNA structure determination consists of two major parts, model building and fitting the model into the electron density map. The initial model can be taken from known structures or generated by computer. There are several model building software programs in use. The commonly used programs in our laboratory include Macro-Model [192], and Helix [193]. These programs can generate A-, B-, Z-form double helixes or single strands. Some times the initial models could be part of the structures. FRODO/CHAIN [194,195], Macro-Model or Dock [196] software packages can be used to modify the initial models. The fitting and further modification is usually conducted at graphic terminals with the program FRODO/CHAIN.

Among the three molecular fragments of a nucleotide, the bases (cytosine, guanine, thymine, adenine) and the phosphate are quite rigid;

the deoxyribose sugar and connections are the variable part. In fitting the model into an electron density map, the local adjustment usually takes place in the sugar puckering and the torsion angles α, β, γ, δ, ε, ζ, and χ (Fig. 1.7).

CHAPTER III OUTLINE THE PROBLEMS

In the context of the preceding survey of the current knowledge of DNA structure, some important questions can be addressed as follows.

• Close examination of reported crystal structures of oligo-DNAs reveal many phenomena that have not been interpreted with confidence and satisfaction, such as the unusual high thermal motion of phosphate group P3 in the structure of $d(CGCGCG)_2$ [73]. Is there any possibilities of obtaining more reliable results?

• Molecular dynamics simulation is a useful method, and crystal structures can be employed as initial models for molecular dynamics studies. What are the dynamically behaviors of DNA molecules in crystals?

• Crystal fields are periodic environments governed by particular symmetry factors that are different from the environments in solution. Specific oligo-DNA molecules form characteristic packing patterns in crystals. What are the internal relationships that induce the specific packing pattern?

• Solvent water plays an extremely important role in determining the three dimensional structure and, therefore, the biological functions of DNA. Is it possible to get a more defined and unambiguous understanding of solvent structure in DNA crystals?

• Metal ions interact with DNA in diverse ways. Each metal species binds to DNA at various sites with different interactions. Copper, although it is present at low concentrations in living systems, does effect many important biological processes. What are the structural and functional behaviors of copper ions with DNA?

In order gain further insight into some of these issues, we focus here on the detailed examination of crystal structures and crystal packing of a few interesting oligo-DNA and DNA-drug complexes. Three subjects are chosen in our study.

1. Oligo-DNA, $d(CGCGCG)_2$, in the Z-form at low temperature.
2. Oligo-DNA, $d(m5CGUAm5CG)_2$-Cu(II) complex , in the Z-form.
3. DNA-drug complex, $d(CGCG)_2$-ditercalinium.

In general, the higher the resolution of the crystallographic diffraction data the more details that can be seen in the electron density map. Therefore, the more accurately that the structure can be determined. A series of electron density maps, generated from this

work by truncating the diffraction data at different resolutions, demonstrates the relationship of the structural accuracy and the resolution of the diffraction data (Figure 3.1).

The resolution range of DNA structures reported in the literature range from 0.9 Å to 2.8 Å. The highest resolution, reported by Wang, et al. [50] is 0.9 Å for the Z-DNA of the oligomer $d(CGCGCG)_2$. In practice, the resolution limit is determined mainly by the quality of the single crystals. The instrumentation limit is determined by the maximum 2θ angle and the wavelength of the X-ray used. in our laboratory the maximum 2θ angle is 70° and the wavelength is 1.54 Å(Cu Kα). Therefore the instrumentation limit is 0.82 Å determined by the Bragg equation

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

or in more practical crystallographic definition

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

Yielding

$$d_{min}=\lambda/2\sin\theta_{max}$$

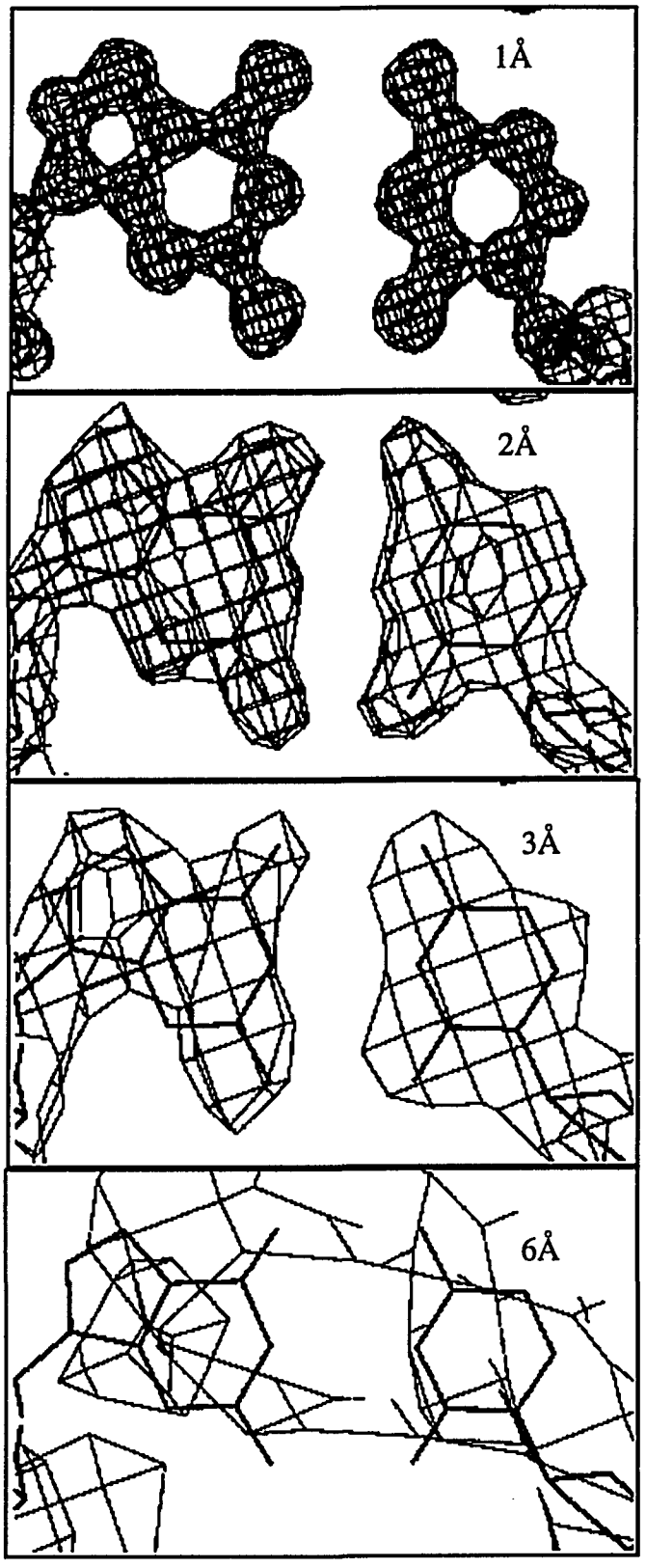


Fig. 3.1. Electron density maps at different resolutions.

A single crystal always has some imperfection. Growth imperfection, lattice disorder, lattice thermal motion, and multiple conformations, etc. all tend to limit the perfection of a single crystal. For a normal, imperfect single crystal the practical limit of resolution is difined by the intensity of the diffracted X-rays dropping to undetectable level with increasing 2θ angle.

The purpose of this study is to provide a detailed structure of DNA and its solvent environments in a single crystal. Thus choosing a suitable oligomer and carefully determining the experimental conditions are essential. Upon examining all published structures of DNA and the experimental conditions used, we decided to take the oligomer $d(CGCGCG)_2$ in the Z-form as our subject. This oligomer has been crystallized under various conditions [122], and the crystals have been reported to diffract to 0.9 Å resolution. Appropriate concentrations of Mg(II), spermine, Mg(II) and spermine, and $Co(NH_3)_6^{+3}$ are sufficient to induce this oligomer into the Z-conformation. In order to simplify the counterion and solvent environments, we used only Mg(II) ion in these experiments for crystallization.

There are several reasons for conducting low temperature data collection on single crystals of biological macromolecules. First, single crystals of a bio-macromolecule degrade with time, due to damage from exposure to high intensity X-rays. Usually, a complete diffraction data set can only be obtained by using several crystals. The inter-crystal

variation is taken into account in the final results by combining the diffraction data into one data set. Unfortunately this does not eliminate small but significant variations such as crystal to crystal variations, solvent structure, disorder, and effective pH. Setting the single crystals at very low temperature can significantly reduce the radiation damage. A complete data collection can often be carried out using one crystal without serious radiation damage. Thus inter-crystal variation will be eliminated. Second, at room temperature, the thermal vibration of the lattice can cause uncertainty in the position of atoms in the structure. In some cases, multi-conformational feature of a structure can not be identified from of the high degree of thermal motion. All these effects lead to less accurate final results. Performance of diffraction experiment at very low temperature reduces or eliminates these difficulties significantly. The structural distortion caused by high degree thermal vibration and natural decay will be reduced significantly. The low temperature diffraction experiment described were conducted using liquid nitrogen cooling to a temperature of 150° K.

In the sections below we will address the problems in three subjects as follows.

Part 1. High resolution low temperature study of Z-DNA, $d(CGCGCG)_2$.

The oligo-DNA, $d(CGCGCG)_2$, has been studied in great detail. It

has been crystallized under various conditions and with different counterions including Mg(II), spermine, MG(II)+spermine, Co(III)-hexamine, and Cu(II). It has been considered the most well defined structure of DNA. The initial work of Wang et al. [50] reported that the structure had been refined to 0.9Å resolution. There are, however, some structural features that are still not clear. For example, in the structure of $d(CGCGCG)_2$-Mg(II) crystal, an unusual dimer of $Mg_2(OH_2)_9$ and high thermal factor group, phosphate-3, have been reported. In addition there is not a detailed understanding of the ion and water structure in these crystals.

The first experiment of our study is designed to study the crystal structure of $d(CGCGCG)_2$ at 150°K (liquid nitrogen temperature).

The purposes of this portion of the study are:

- To determine the crystal structure of $d(CGCGCG)_2$ at 150° K.
- To examine the counterion and solvent structure.
- To examine the correlation of DNA conformation and counterion/solvent structure.
- To examine the thermal motion behavior at 150° K.
- To examine the relationship of crystal packing and molecular structure.

Part 2. Study of Cu(II)-d(mCpGpUpApmCpG)$_2$ crystal at Room temperature

Copper is found in living systems with significant concentrations. (see table 1.6). Copper(II) cation is found to bind to the bases of DNA with higher affinity than that of any other divalent cations [138]. Earlier crystal structures of isolated nucleotide and base derivatives have shown that the N-7 atoms of the purine bases, both guanine and adenine, are susceptible to coordinated binding to copper(II) cations [154,197,198]. One important effect of copper on DNA is its role in the copper(II)-catalyzed oxidative cleavage of DNA [199]. In the studies on direct DNA damage, the adenine residues were found to be resistant to damage induced by copper(II) and hydrogen peroxide [200,201]. This raises the question as to whether copper(II) binds specifically to guanine bases or more generally to purine bases in duplex DNA.

In the structure of Z-DNA, d(CGCGCG)$_2$, copper(II) cations have been found to bind to all six guanine bases [202].

The purposes of this portion of experiment are:

- To determine the crystal structure of the Cu(II)-d(m5CGUAm5CG)$_2$ crystal.
- To examine the base specificity of Cu(II) interaction.
- To examine the isomorphism of the copper derivatives.

Part 3. Determination of the crystal structure of $d(CGCG)_2$-ditercalinium complex

Ditercalinium(Fig. 3.2), 7H-pyridocarbazolium dimer, has a very high affinity for DNA with a binding constant of 10^8 [203]. It has drawn a lot of attention as a potential anticancer drug. Theoretical and experimental studies shows that the rigid linker, bipiperidine, prevents ditercalinium from self-stacking, and facilitates it as a bis-intercalator of DNA [176,177]. A predicted geometry of $d(CGCG)_2$-ditercalinium complex has been made based on 2-D NMR experiment and theoretical analysis [176]. The structure features are as follows:

- The DNA is a right hand double helix.
- Ditercalinium bis-intercalates into the double helix at the two C-G steps.
- The linker of ditercalinium is located in the major groove.
- The stacking of base-base and base-drug are about 3.4 Å.

The questions we will examine are:

- How accurate is the predicted structure of the $d(CGCG)_2$-Ditercalinium complex?

- Can we use a model that is built based on the predicted structure

to solve the crystal structure of d(CGCG)$_2$-ditercalinium complex? If yes, how do we build the model and how to use the model in a molecular replacement procedure to solve the structure?

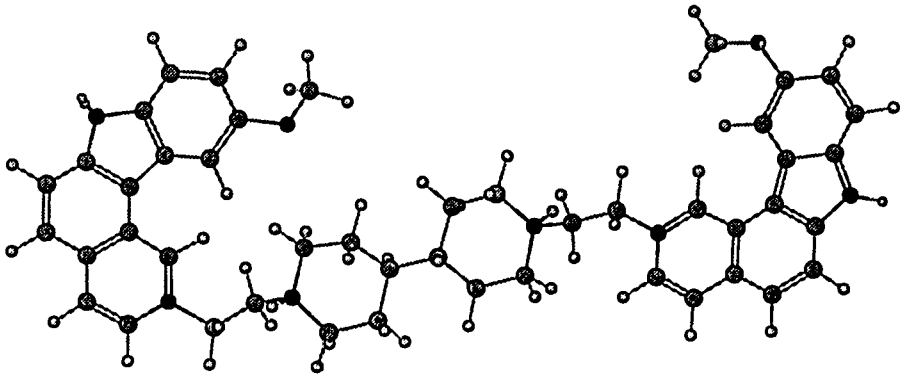


Fig. 3.2 Ditercalinium

In summary, this study will focus on the following:

- The structure of d(CGCGCG)$_2$ crystal at liquid nitrogen temperature.
- The structural features of d(m5CGUAm5CG)$_2$, containing Cu(II).
- The structure of d(CGCG)$_2$-Ditercalinium complex and the relationship of its structure and function.

CHAPTER IV EXPERIMENTS AND STRUCTURE DETERMINATION

4.1 Materials and Equipment

4.1.1 DNA Oligomer and Chemicals

1. Oligomer $d(CGCGCG)_2$

The DNA oligomer was synthesized by Mr. Stephen Bobin of the RCMI sequence and synthesis facility, Hunter College, CUNY using phosphodiester chemistry. The DNA is purified, following deprotection, by HPLC on a DuPont Zorbax Bio Series oligonucleotide column, using a NaCl gradient in a solvent system containing 20% acetonitrile and 80% 0.02 M sodium phosphate. Fractions from the major peak were collected, concentrated, desalted, and lyophilized. The purification was performed in N. R. Kallenbach's laboratory, New York University, with the assistance of Dr. M. Lou and Dr. Q. Guo.

2. $d(m5CGUAm5CG)_2$-Cu complex

The synthetic DNA oligomer and crystallization of Cu(II)-$d(m5CGUAm5CG)_2$ was carried out in collaboration with Professor P. S. Ho's group, Department of Biochemistry, Oregon State University.

3. Ditercalinium-d(CGCG)$_2$ complex

The synthesis of the DNA oligomer and the crystallization of ditercalinium-d(CGCG)$_2$ was carried out in collaboration with Dr. L. Williams and colleagues, Department of Biology, Massachusetts Institute of Technology.

MPD, 2-Methyl-2,4-Pentondiol, is from Aldrich. All other chemicals are from Sigma.

4.1.2 Equipment and Software

Crystallization. The crystallization set up used in this laboratory is assembled from a Pyrex spot plates (#7220, 9 wells, Dow Corning), plastic petri dishes (#3488-B28), and plastic sandwich box (4x4x1 in. Tristate Plastic). See Fig. 4.1.

X-ray diffractometer. Enraf Nonius CAD4 diffractometer system with fine focus GX21 rotating anode X-ray generator and RF586 high performance constant potential regulator with the maximum power of 4 KW [204].

Low temperature device. Enraf Nonius RF558 liquid nitrogen system. This system provides constant nitrogen flow to cool crystal to a desired low temperature. The lowest temperature provided by this system is about -150° C (123° K) [205].

Computer and software. All computation and graphic modeling were carry out on the Hunter College Chemistry VAX cluster, Evans-Southern's PS390 graphic station and SiliconGraphics' INDIGO ELAN graphics work station. IBM PCs and Apple Macintosh were used in word processing and in some data analysis. The software employed include: CAD4 (Nonius Enraf) for diffractometer control and data collection; SDP and Molen (Nonius Enraf) for data reduction; XPLOR (A. Brünger) for structure solution and refinement; FRODO and CHAIN for modeling and graphics; AMPAC, molecular orbital (quantum) calculation; Chem-3D, model examination; and programs written by G. J. Quigley and Shun-Le Chen, data scaling, peak searching, and transformation.

4.2 Crystallization

4.2.1 General Procedures

In this study all crystallization was done using the vapor diffusion technique with the sitting drop method. This technique was developed for crystallization of tRNA [206] and successfully used to crystallize many DNA oligomers. For different samples various conditions, in terms of pH, buffer concentrations, counterions, precipitating agent, anti-cancer drug, and temperature, were used. The precipitating agents

used included 2-MPD, PEG 400, and isopropanol. All experiments were set up on silicon coated spot plates (Dow Corning Pyrex glass, nine wells catalog #7724). The spot plates with sitting drops and reservoir (30 ml solution of precipitating agent in water) were sealed in clear plastic sandwich boxes (Tristate plastic, 4x4x1 1/8 in, catalog #9052) Fig. 4.1. The boxes are sealed with high vacuum silicon grease. The sealed environment allows for equilibration of volatile components between the sitting drops and the reservoir. Boxes containing crystallization solutions were stored either at room temperature or low temperature (4° C, cold room). The crystallization set ups were monitored regularly using an M8 stereo microscopy (Wild Heerbrugg Co.).

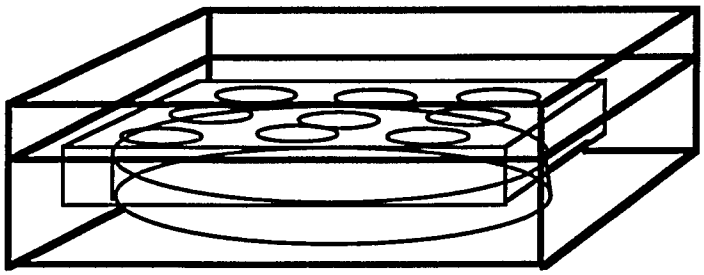


Fig. 4.1. Crystallization device

When the single crystals had grown to suitable size (0.3-0.5 mm. each side), they were mounted into X-ray diffraction capillaries (Charles Supper Co.) individually with tiny amount of mother liquor (Fig. 4.2). Unmounted crystals are stored in the original spot plates in a cold room. After one day of stabilization in the capillary, the crystal was mounted on the X-ray diffractometer for data collection.

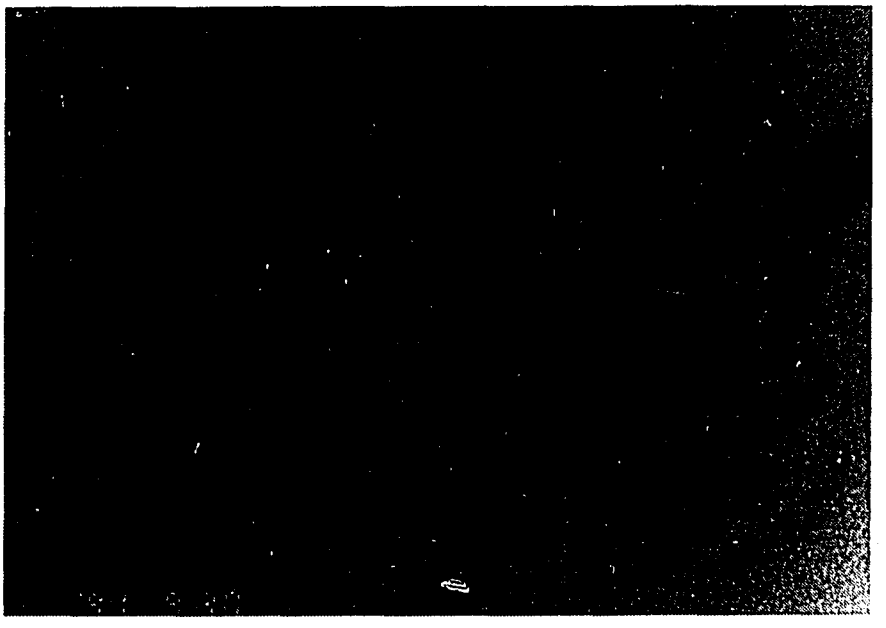


Fig. 4.2. A mounted crystal (Holliday junction)
(1:20)

4.2.2 Crystallization d(CpGpCpGpCpG)$_2$

The DNA crystallization was carry out by the vapor diffusion method as described above. The setup conditions were: DNA 55 mg/ml (5 mM), $MgCl_2$ 120 mM, MPD 1.5% (v/v), and sodium cacodylate buffer (PH=6.8) 10 mM. The reservoir is 10% (v/v) MPD in water. 60

μl of crystallization solution was used per depression. Plates were sealed in the sandwich box with 30 ml reservoir and stored at 4°C for approximately two weeks.

After two weeks, the single crystals had stabilized with the dimension of about 0.5 X 0.4 X 1.0 mm^3 in quasi-hexagonal rodlike shape. Fig. 4.3.

Fig. 4.3. Crystals of Z-DNA

Another set of crystallization were carry out to grow large size single crystals for neutron diffraction. The consideration was increasing the volume of solution (the quantity of DNA) and reducing the initial

concentration both in working solution and reservoir. The reason to do so were limiting the possible number of crystal nuclei and reducing the driving speed to supersaturation. For this experiment, 180 µl of working solution with DNA concentration of 30 mg/ml per depression was used .

The concentration of reservoir was 5% (v/v) of MPD. After a month in one depression two single crystals were obtained with dimension of 0.9x0.8x1.5 mm^3 and 0.9x0.8x0.8 mm^3. These two crystals were taken to Brookhaven National Laboratory. for a test data collection. Using the high flux neutron beam source, the observed reflections on the first crystal showed broad peaks due to the damage due to deuterium-hydrogen exchange. The second crystal behaved well.

4.2.3 Crystallization of Cu(II) Substrated d(m5CGUAm5CG)$_2$

Crystals of d(m5CGUAm5CG)$_2$ as Z-DNA were also grown by the vapor diffusion method. The setup conditions were: 4.0 mM (4.8 mg/ml) oligomer, 28 mM sodium cacodylate buffer at pH 7.0, 15 mM magnesium chloride, and 8.5% 2-MPD. Cu(II) (chloride) was added to the mother liquor gradually over 48 hours after the drops were set up. The final molar ratio of Cu(II) to DNA was 3 to 2. The crystals are yellowish and rod like.

4.2.4 Crystallization of Complex of Ditercalinium-d(CGCG)$_2$

Crystals of Ditercalinium-d(CGCG)$_2$ complex were growing at room temperature using the sitting drop vapor diffusion method. The setup conditions are: 16.8 mM cacodylate buffer, pH 6.0, 14 mM ammonium acetate, 0.3 mM spermine tetrahydrochloride, 0.8 mM magnesium chloride, 6% 2-MPD, 0.7 mM DNA (single strand concentration), and 0.2 mM ditercalinium. The sitting drops were equilibrated against a reservoir of 30% 2-MPD. Within one month reddish single crystals of suitable size for X-ray diffraction were obtained.

4.3 Crystal Mounting

The crystals are mounted in X-ray capillaries (Charles Supper Co.). The capillaries are sealed with Five minute Epoxy (Sigma Co), then set on goniometer head.

When mounting a crystal into a capillary, removing excessive mother liquor is very important. The excess mother liquor around the crystal will degrade the quality of diffraction data since it contributes to background scattering and to absorption. Furthermore, excess mother liquor may cause unexpected shift of crystal along the capillary, that will make the data collection impossible.

For the low temperature experiments special care must be taken when mounting a crystal. During the procedure of flash cooling, excess mother liquor will seriously damage the crystal. As much as possible of the excess mother liquor is carefully removed. Actually, getting the crystal dry is the key factor in performing a successful low temperature experiment. If there is too much mother liquor around the crystal, the crystal will be destroyed due to the freezing of the mother liquor. The phase transition of mother liquor builds up internal stress inside the crystal. This stress causes the development of cracks inside the crystal, when the capillary containing crystal is cooled to very low temperature. Visually, two kinds of cracks, micro or visible, might occur during the cooling procedure. The visible cracks can be seen through a microscope; the micro-cracks will turn the crystal opaque.

The capillary containing the crystal is mounted on the goniometer head, being sure that the crystal is well centered in the X-ray beam and at the center of motion of goniostat. For the low temperature experiments a special procedure is used. First, the capillary is mounted on the goniometer head, then the capillary is aligned vertically under the end of thermal insulation pipe, where the low temperature stream of nitrogen exits. The distance between crystal and the end of the pipe is set to 6.0 mm. Prior to turning on the low temperature stream, the crystal is carefully pre-centered. Being sure that the crystal is well centered in the X-ray beam and at the center of motion of goniostat and that the position of center of the low temperature stream is aligned with

the crystal. Once the crystal is mounted properly, the low temperature stream is turned on, but the stream is initially deflected from the crystal with a piece of paper. When the temperature has reached the desired point, typically 150° K, and is running steadily, the paper is removed, and the crystal is flash cooled to the desired temperature. The system is allowed to stabilize for 10 minutes, then a minor adjustment to the centering is made to correct for the small shift of the crystal due to the temperature change.

4.4 Data Collection and Reduction

Data collection include the following parts: (1) Crystal orientation and Unit cell determination, (2) Space group (symmetry) determination, and (3) Intensity data collection. In this study the crystal orientation and unit cell determination are completed by auto-indexing procedure of the CAD4 system.

The experimental determined scattering vector $\underline{s}$ corresponds to a given diffraction spot of the crystal is related to the crystal orientation, diffraction angle, and unit cell size and orientation by

$$\underline{s}=h\underline{a}^*+k\underline{b}^*+l\underline{c}^*$$

where h, k, l are the integral Miller indices of the particular diffraction spot and $\underline{a}^*$, $\underline{b}^*$, $\underline{c}^*$ are the so called reciprocal lattice vector which relate to the unit cell vectors $\underline{a}$, $\underline{b}$, $\underline{c}$ by the relationship

$\underline{a}\bullet\underline{a}^*=1 \qquad \underline{a}\bullet\underline{b}^*=0 \qquad \underline{a}\bullet\underline{c}^*=0$

$\underline{b}\bullet\underline{a}^*=0 \qquad \underline{b}\bullet\underline{b}^*=1 \qquad \underline{b}\bullet\underline{c}^*=0$

$\underline{c}\bullet\underline{c}^*=1$

A set of n observed diffraction spots yields n^2 three dimensional linear equations with 3n integer coefficients and a common unknown unit cell base consisting of the unit cell distances, angles and reference orientation (9 parameters). The auto-indexing procedure seeks out a consistent set of these parameters with the smallest values of the Miller indices. Then the orientation of the unit cell and the unit cell parameters, dimensions (a, b, c), and angles (α, β, γ) can be determined. Once the unit cell has been determined, the space group of the unit cell can be determined. Experimentally, finding the symmetry of the diffraction pattern and the conditions of systematic absence can determine the space group. Knowing the unit cell and space group, one can decide on the range of data collection, in term of Miller indices. Data collection was accomplished with our CAD4 system using the automatic data collection procedure. In addition to the main data set, one or two sets of "psi" scan data were collected for absorption correction (will be described in following sections). Usually, a complete data collection for a DNA crystal requires several days or longer. In order to monitor the variation of the instrument system and the decay of crystal, several reflections should be set up as intensity standards and orientation references. These reflections will also be used in the

following data reduction as standards and references.

Data reduction includes the following:

1. The intensity correction for Lorenz, polarization, absorption, and degradation effects. Lorenz and polarization are corrections dependent on the physics and geometry of data collection. The absorption correction dependent on the size, shape, and chemical composition of the crystal and the wavelength of the X-ray used. The degradation correction is based on the decay of intensities of standard reflections.

2. If multiple crystals are used, every set of data should be collected with a reasonable overlap to another data set. After all segments of data have been collected, the multiple data set will be scaled to a common scale by a least square minimization procedure using the data that overlapped between segments. The scaled data segments are combined together to form a complete data set.

4.4.1 The Crystal of $d(CGCGCG)_2$ at Liquid Nitrogen Temperature

The data collection was performed on our Enraf-Nonius rotating anode X-ray diffractometer. The power setting were: 40 KV and 90 mA. The unit cell parameters were determined to be: a=17.85±0.02 Å,

b=31.10±0.02 Å, c=44.18±0.02 Å, α=β=γ=90.0±0.05°. The data collection was completed on one single crystal using shells from high resolution to low resolution. The data collection shells are shown in table 4.1.

Table 4.1. The scheme of data collection shells
(a=17.85 Å, b=31.10 Å, c=44.18 Å, α=β=γ=90.0°;
Space group $P2_12_12_1$)

Shell	d (Å)	1/d	$(1/d)^3$	θ (°)	Vol.	Unique Refs.
1	1.71	0.58	0.20	26.80	0.84	2737
2	1.36	0.74	0.40	34.61	1.68	5474
3	1.19	0.84	0.60	40.56	2.51	8211
4	1.08	0.93	0.80	45.70	3.35	10948
5	1.00	1.00	1.00	50.43	4.19	13685
6	0.94	1.06	1.20	55.01	5.03	16422
7	0.89	1.12	1.40	59.59	5.86	19159
8	0.82	1.22	1.60	69.95	6.70	21896

According to equation 2.1 the number of reflections in a shell is proportional to $(1/d)^3$. The shells are divided in term of uniform steps in $(1/d)^3$. Thus the numbers of reflections in every shell are approximately the same. The data were collected from a total of five shells, ranging from 1.0 Å to infinity. 14,878 reflections are collected. For the collected data, based on intensity **I**, 13,297 are stronger than

1.0xσ(I), (σ(I) is based on counting statistics) 12,592 are stronger than 2.0xσ(I), 11,583 are stronger than 3.0xσ(I). The inter-shell overlaps are 0.5° in 2θ angle. The total number of unique reflections is 12,498.

Since the shape of the crystal is usually not a sphere, the absorption effects in different directions are not uniform. An absorption correction is necessary. A set of several reflections collected by a special method, a so-called "psi" scan, is used for this purpose. Reflections that meet the condition ($\chi \approx \pm 90°$) are used to carry out a psi scan. The psi scan collects intensity data of one or several chosen reflections at different positions (usually with 10° steps in the range of 0° to 360° of the ϕ angle). The intensities of the different orientations on a particular reflection vary according to the absorption from different directions of the crystal. The psi scan was carried out both before and after the data collection, using reflection (0 0 4), (0 0 12), and (0 0 24). These data are used for the absorption correction.

After the completion of data collection, the crystal did not show noticeable damage based on examination of the intensity of the standard reflections. Therefore, an expanded data collection was carry out in shells to a resolution of 0.94 Å and then 0.85 Å. The data collection on the shell of resolution 0.94 Å was completed. Unfortunately, the shell of resolution 0.85 Å was not completed due to the failure in the diffractometer. The data in the shell of 0.85 Å resolution that was collected consists about 2/3 of the total number of reflections on the

shell.

Determination of Space Group.

According to the previously determined unit cell parameters, the crystal belongs to the orthorhombic system. By examining the intensities of the special zones, (h,0,0), (0,k,0), and (0,0,l) the appearance of significant intensity for even number reflections only indicates that the symmetry is $P2_12_12_1$ (Space group #19). (see table 4.2, 4.3, 4.4) .

Table 4.2. List of reflections in zone (h,0,0)

h	k	l	Fo
1	0	0	0.61
2	0	0	46.19
3	0	0	2.73
4	0	0	27.30
5	0	0	-0.79
6	0	0	22.91
7	0	0	1.19
8	0	0	17.97
9	0	0	1.72
10	0	0	28.54

Table 4.3. List of reflections in zone (0,k,0)

h	k	l	Fo
0	1	0	-0.50
0	2	0	11.59
0	3	0	0.63
0	4	0	24.72
0	5	0	1.01
0	6	0	13.91
0	7	0	1,79
0	8	0	53.49
0	9	0	0.70
0	10	0	2.40
0	11	0	2.11
0	12	0	34.12
0	13	0	1.64
0	14	0	38.12
0	15	0	2.31
0	16	0	-2.41
0	17	0	1.93
0	18	0	15.13

Table 4.4. List of reflections in zone (0,0,l)

h	k	l	Fo
0	0	1	0.70
0	0	2	8.25
0	0	3	1.11
0	0	4	16.63
0	0	5	0.69
0	0	6	11.57
0	0	7	-0.29
0	0	8	3.36
0	0	9	1.51
0	0	10	1.32
0	0	11	-2.78
0	0	12	300.89
0	0	13	-1.35
0	0	14	11.97
0	0	15	2.17
0	0	16	29.28
0	0	17	0.68
0	0	18	5.70

Data reduction and scaling.

The shell-wise collected data sets have been reduced by anisotropic intensity correction and absorption correction, using sub-programs of Molen, Decay, Psiscan, and Psicalc individually. The reduced data sets, then, have been converted into Xplor readable format and scaled together by least square method. The target function is:

$$\Sigma\Sigma(||F(h_j)_i|-|F(h_j)_k||)^2=0 \qquad (4.1)$$

where the h_j are the indices of the reflections and i, k (i not equal to k) are the number of shell. The scaling is necessary because of degradation effects.

The programs used in these tasks are the modified Molen sub-program Ftable (modified by the author), and the least square scaling program, Scale, written by the author.

4.4.2 Copper-d(m5CGUAm5CG)$_2$ Complex

One single crystal with dimensions of 0.45x0.25x0.2 mm^3 was mounted in glass capillary. Data to 1.3 Å were collected on a Enraf-Nonius rotating anode X-ray diffractometer (CAD4 system) in $\omega-\theta$ scan mode with Cu Kα radiation. The crystal was determined to be in the orthorhombic $P2_12_12_1$ space group, with unit cell parameters of

a=17.59 Å, b=30.58 Å, and c=44.52 Å. A total 4357 reflections were collected on this crystal.

The data were reduced by the DATA procedure of SDP software (Enraf-Nonius Co.), Structure Determination Program. The intensities were corrected for Lorenz, polarization and absorption effects. Systematic absences and 2σ cutoff criteria gave a total of 2587 reflections for use in the structure determination. Details of the methods used are as described above with the exception of the low temperatures technique which were not applied.

4.4.3 Ditercalinium-d(CGCG)$_2$ Complex

The space group and unit cell parameters were determined from precession photograph and CAD4 auto index procedure. The DNA-Ditercalinium complex crystallized in tetragonal space group $P4_12_12$ with the unit cell dimensions a=b=26.88 Å, c=82.60 Å. The data to 1.7 Å were collected on the Enraf-Nonius rotating anode X-ray diffractometer and a Rigaku AFC5 rotating anode diffractometer in the θ scan mode with Cu Kα radiation. Intensities were corrected for Lorentz, polarization, and absorption effects. Two sets of data were scaled together and reduced by the sub-procedure DATA of SDP (the older version of Molen) software. After systematic absence removal and 3σ(I) cutoff, a total of 2211 unique reflections were used to solve the structure. Among the data, 542 reflections are between 2.0 Å and 1.7 Å

4.5 Structure Determination

Before discussing the structure determination, we will introduce a few definitions.

1. **R-factor.** This is a measurement of the agreement between the calculated and observed structure factors. The lower the R value is the better the agreement.

R-factor is defined by equation (4.2).

$$R=\Sigma||Fo(h_j)|-|Fc(h_j)||/\Sigma|Fo(h_j)| \qquad (4.2)$$

2. **B factor.** This is a measurement of the thermal motion (vibration) of atoms in the structure. The thermal factor B is defined by equation (4.3)

$$B=8\pi^2\bar{u}^2 \qquad (4.3)$$

Where $\bar{u}$ is the RMS displacement of the atom.

3. **Electron density maps.** The electron density is defined in section 2.4. In this study several types of electron density

maps are employed. These are 2Fo-Fc, Fo-Fc, and Fc-Fo maps. These map are defined as follows.

• 2Fo-Fc map.

$$\rho(x,y,z)=\frac{1}{V}\sum_{h}\sum_{k}\sum_{l}(2Fo(hkl)-Fc(hkl))e^{-i2\pi(hx+ky+lz)}$$

$$=\frac{1}{V}\sum_{h}\sum_{k}\sum_{l}(2|Fo(hkl)|-|Fc(hkl)|)e^{-i\phi_c(hkl)}e^{-i2\pi(hx+ky+lz)}$$

• Fo-Fc map

$$\rho(x,y,z)=\frac{1}{V}\sum_{h}\sum_{k}\sum_{l}(|Fo(hkl)|-|Fc(hkl)|)e^{-i\phi_c(hkl)}e^{-i2\pi(hx+ky+lz)}$$

• Fc-Fo map.

$$\rho(x,y,z)=\frac{1}{V}\sum_{h}\sum_{k}\sum_{l}(|Fc(hkl)|-|Fo(hkl)|)e^{-i\phi_c(hkl)}e^{-i2\pi(hx+ky+lz)}$$

4.5.1 The Structure of $d(CGCGCG)_2$ at Low Temperature

Structure solution and refinement.

The basic technique used to solve the structure were molecular replacement and differential Fourier transform. The room temperature structure of $d(CGCGCG)_2$ [122] was taken as the initial model. The R-

factor at the beginning was 24%. Further refinement did not reduce the R-factor significantly. This suggested that the system seems to be trapped in a local minimum. Adjustment on the initial model must be made. At this point a Simulated Annealing [27] was performed. First the model was set to 2000° K, and the slow cooling method was used to minimize the empirical energy. At such a high temperature the thermal vibration energy of atoms allowed the system to overcome the energy barrier of local minimum. Then the velocity scaling method [27], including the E_X term, the "energy" defined by the difference between the model structure and the calculated structure from X-ray data, was used to slowly cool the system to 150° K (the temperature of experiments) to allow the empirical energy to reach its minimum. After simulated annealing the R-factor dropped to 19.0%.

At this point, the B values of atoms of the DNA double helix were about 3 to 4 with a few exceptions. The B value of counterions and water molecules varied from 3 to 28. The unexpectedly high B values suggested incorrect positions or occupancies of corresponding atoms. The initial electron density maps, including 2Fo-Fc, Fo-Fc, and Fc-Fo maps, were generated by running differential Fourier transforms. The three maps were used together to eliminate possible miss-interpretations that might arise through the using of difference map only. It is known that a difference Fourier transform frequently introduces unexpected "noise" into the difference electron density map. Examination of the three maps, revealed several significant disagreements between the

the model structure and the electron densities. They are the regions around Mg-15, phosphate 3, and phosphate 9. See figures 4.4 and 4.5.

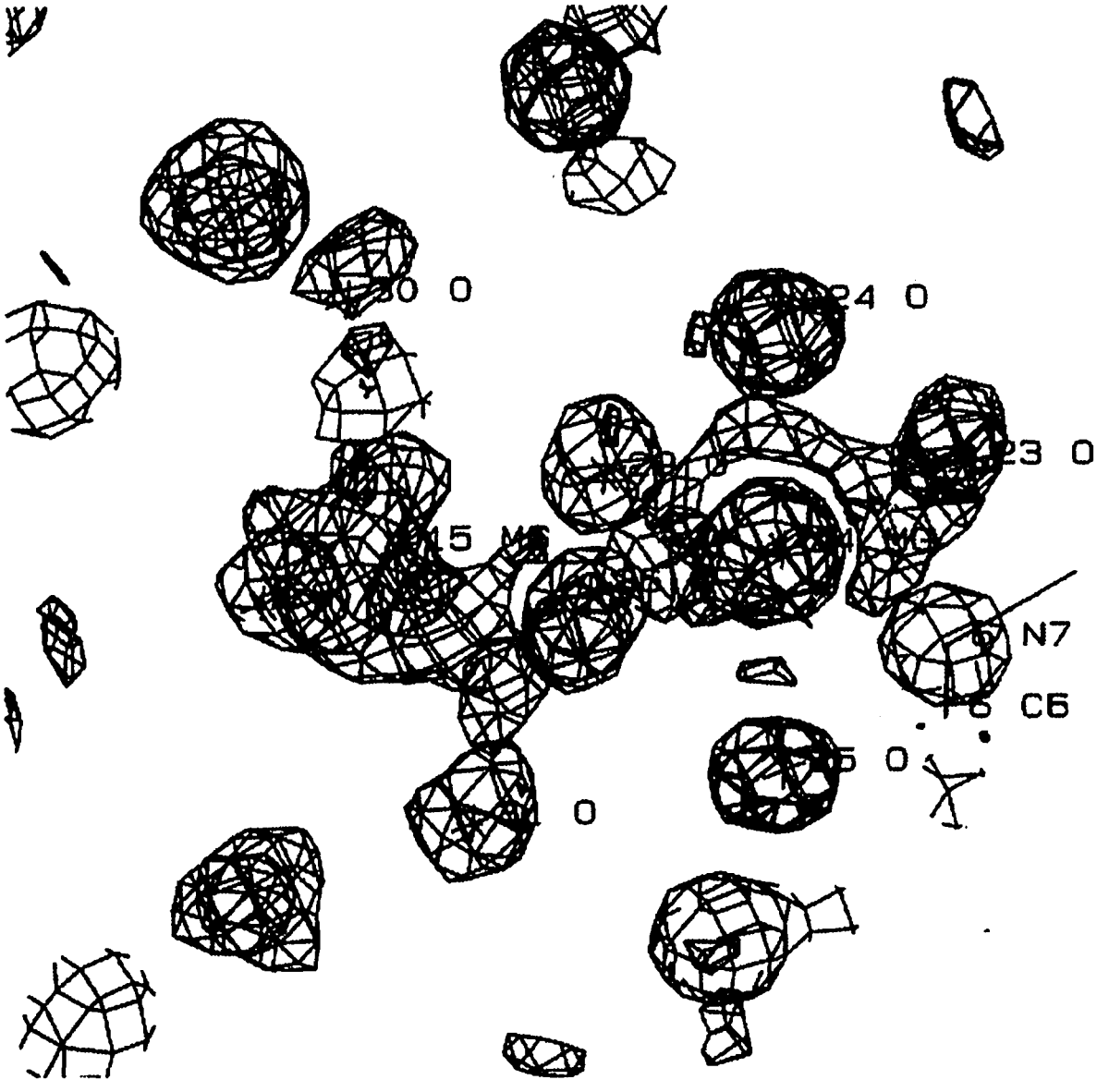


Fig. 4.4. The electron density maps around site Mg-15

— 2Fo-Fc — Fc-Fo — Fo-Fc

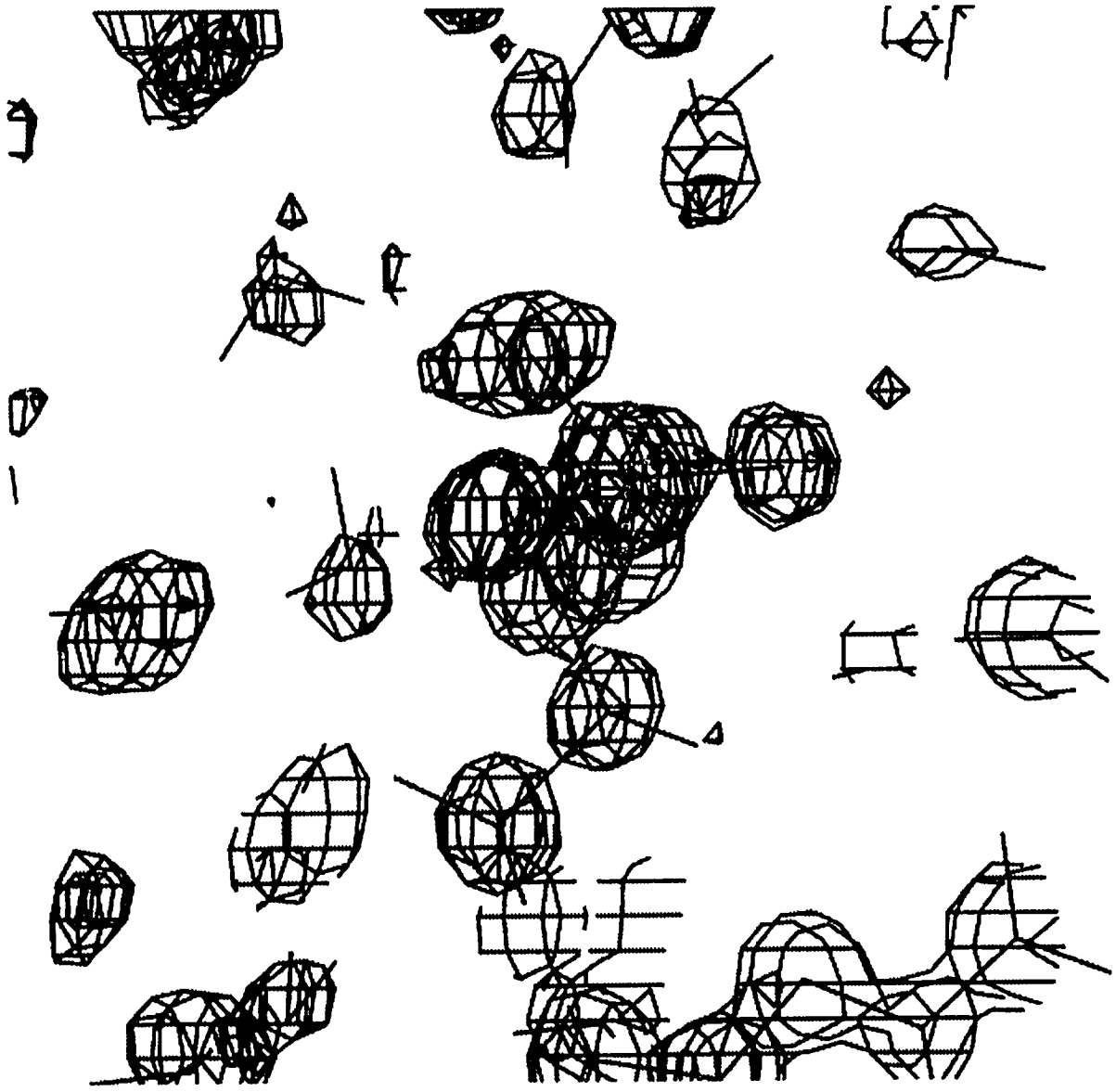


Fig. 4.5 Electron density map of P3

— 2Fo-Fc — Fo-Fc — Fc-Fo

At the site of MG-15 a large circular negative density surrounds atom MG-15. The formation of negative density is explained below. Suppose the thermal vibration is isotropic and experimental error can be neglected, the electron density, $\rho(r)$, of an atom (or ion) can be described by equation (4.4),

$$\rho(r)=Q*Z*(4\pi/B')^{3/2}*e^{-4(\pi r)^2/B'} \qquad (4.4)$$

where Z is the total number of electrons; r is the radius from the center of atom; B'=B+B", with B the isotropic thermal factor and B" the corresponding to Gaussian representing inherent electron distribution (i.e. transform of f(s), the atomic scattering factor), and Q is the occupancy.

It is clear that the distribution of electron density will vary as the B' value changes. The bigger the B' value the broader and lower the distribution of electron density. Some idealized distributions of electron density are shown in Figure 4.6.

Suppose there is an atom (or ion) with lower occupancy and smaller B' value, the corresponding electron density would be similar to curve 2 (Q=0.3, B'=4). If this atom is incorrectly assigned full occupancy, the refinement procedure will force the B' value to increase. Therefor a partial occupied atom will seem like a full occupied atom with high B' value. The electron density will be similar to curve 1

(Q=1, B'=20). In the differential map, (Fo-Fc) map, there will be a circular negative density around the atom. See Fig. 4.6, curve 3.

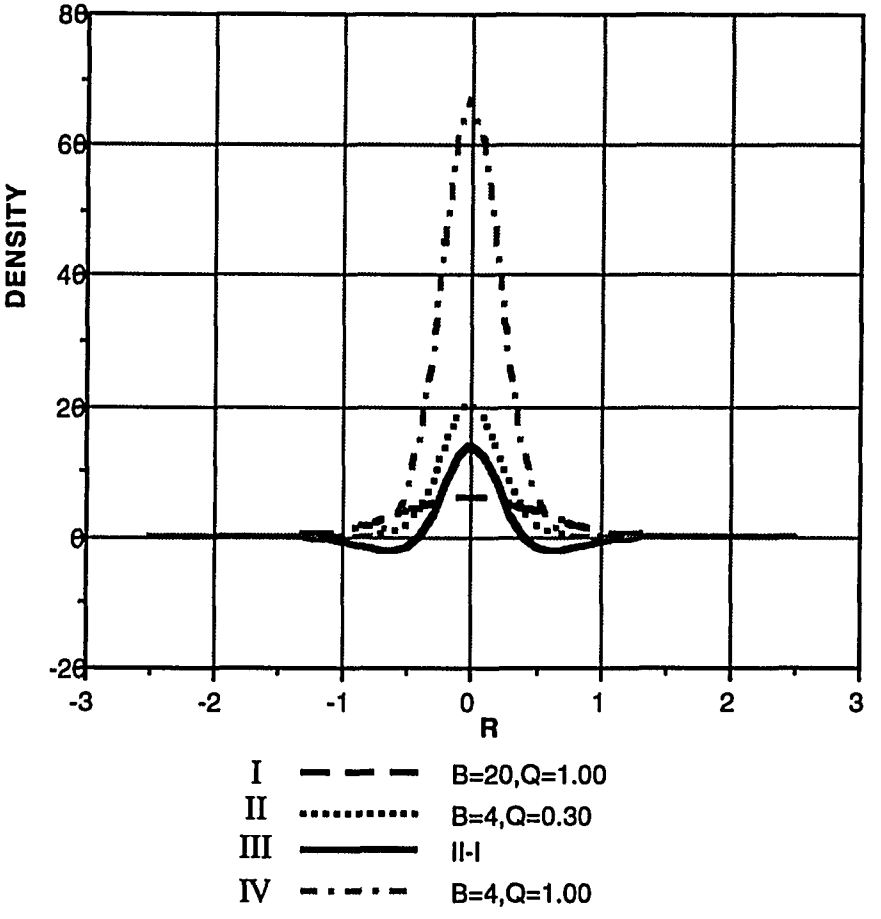


Fig. 4.6. The scheme of formation of circular negative density around partial occupied site

Based on the analysis above, we believe that the site of Mg-15 should be interpreted as partially occupied. Thus we assign Mg-15 and associated water molecules a lower occupancy. Also, some of the coordinated water molecules were assigned partial occupancies. The

occupancies of partial occupied atoms (ion) were also be refined at the final refinement as described later.

At the site of phosphate 3, two extra strong peaks are located at the positions that suggested that Phosphate 3 has dual conformations. A similar feature was observed at Phosphate 9. An alternative conformation has been assigned to atoms P3, O1P3, O2P3, P9, O1P9, and O2P9 to fit the electron density maps. (Fig. 4.5) At the final refinement these partial occupancies are also refined to give more accurate values.

A group of electron density peaks has been observed in the solvent channel near phosphate 3, the geometry of this group of electron density can fit a typical octahedron of a $Mg(H_2O)_6$ complex. See Fig. 4.7. Thus an additional partially occupied Mg(II)-water complex, Mg-17 and coordinated water molecules, has been added to the structure to fit the electron density maps. In addition, the charge balance is also considered. Before the Mg-17 was add to the structure, there was a total of 7 positive charges in this structure. That was insufficient to balance the 10 negative charges of the DNA. An additional magnesium(II) improved the charge unbalanced situation. The occupancies of Mg-17 and associated water molecules are initially set to 30%, then refined to 25% later.

The additional partially occupied Mg(II), Mg-17, and coordinated

water molecules have been added to fit the electron densities (Fig. 4.7).

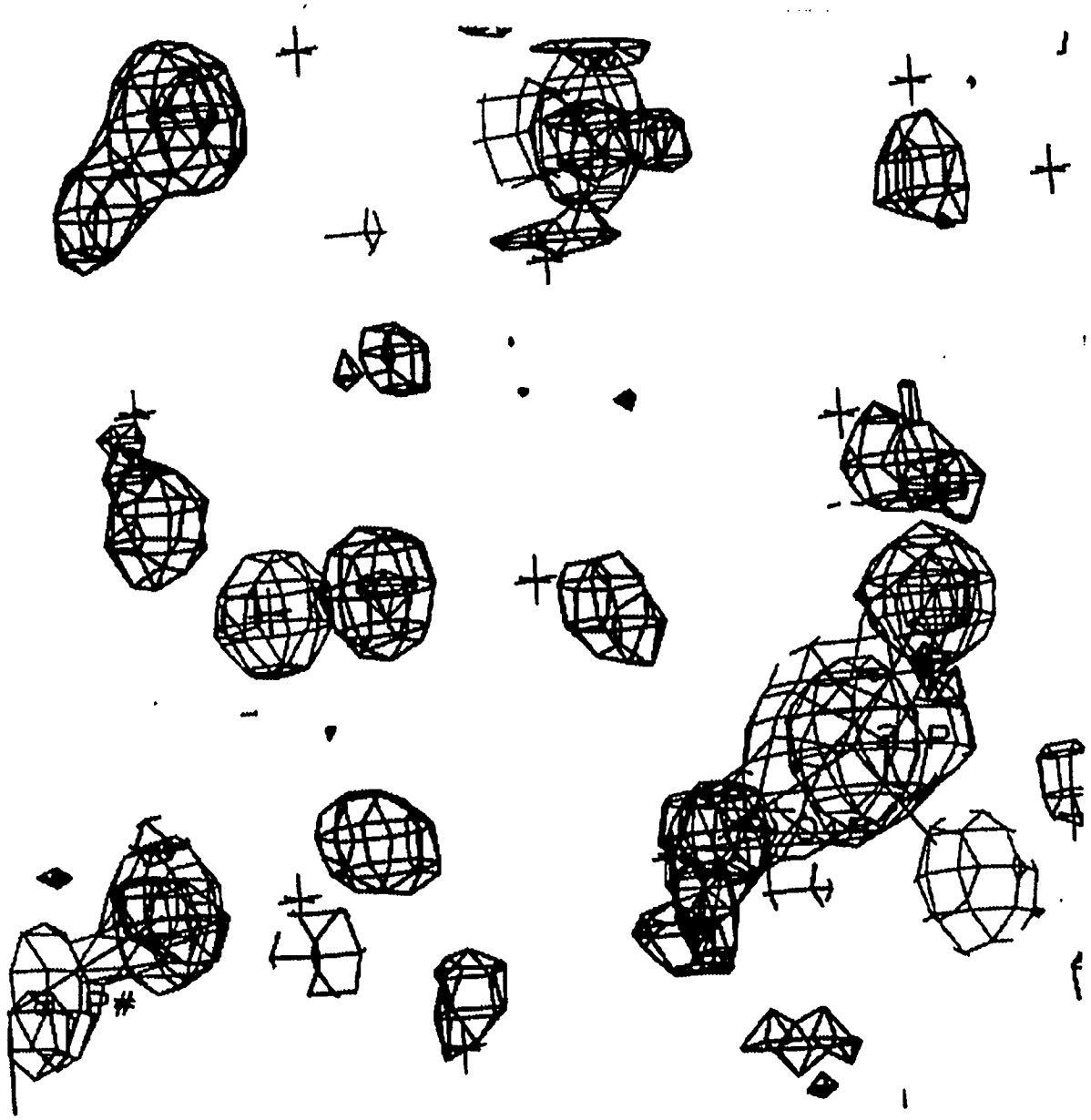
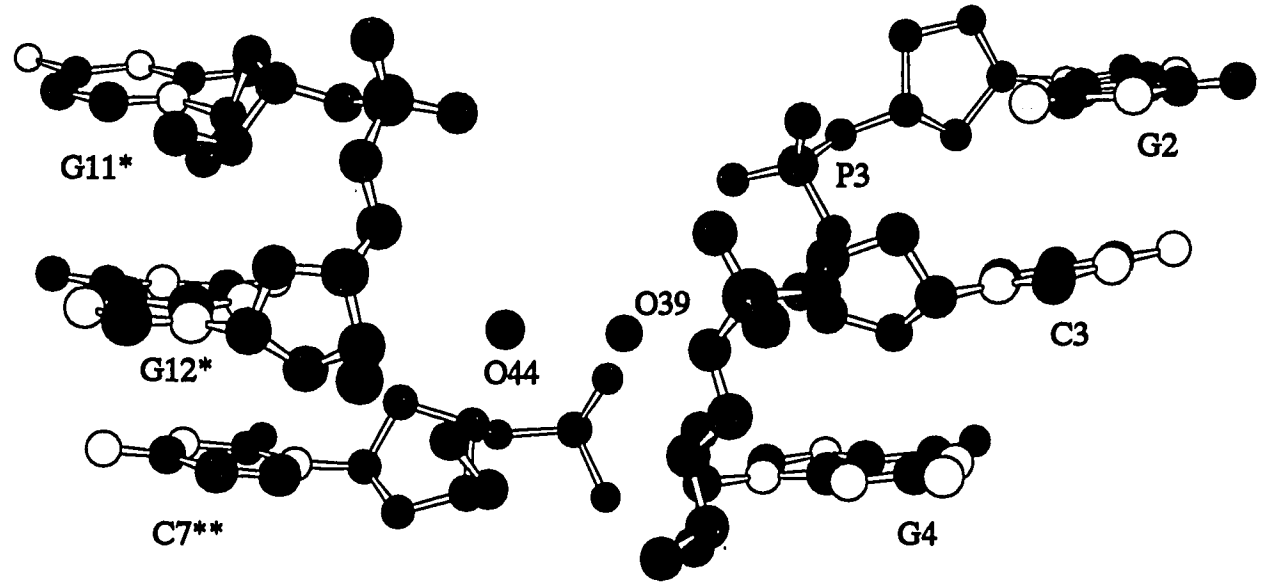


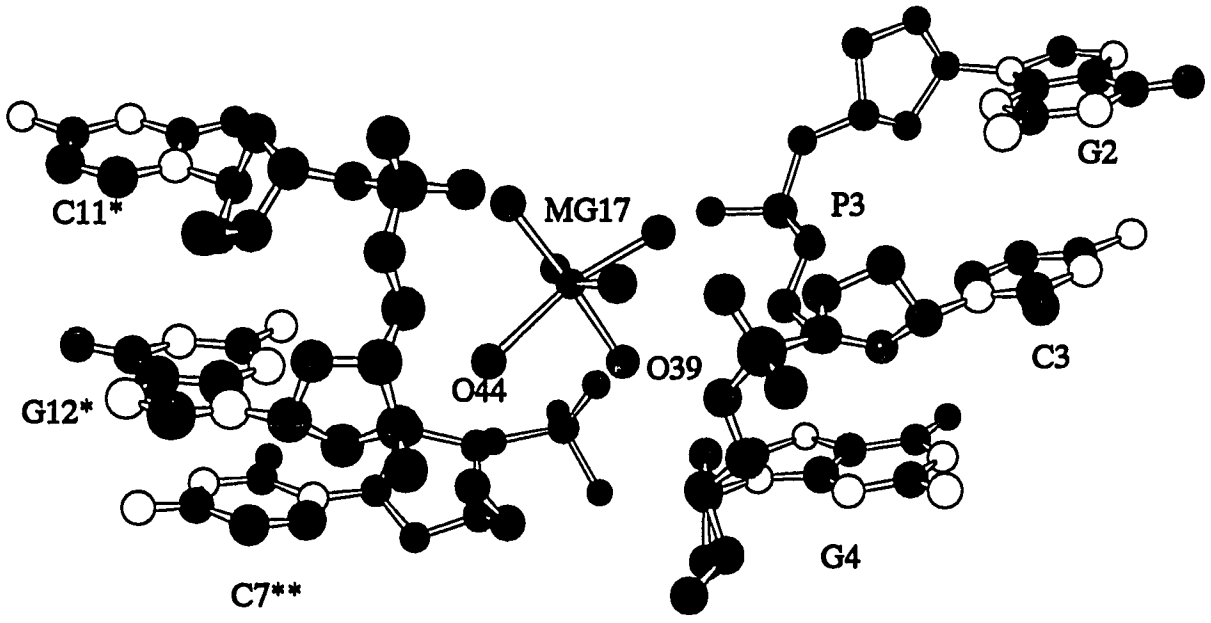
Fig. 4. 7. difference map of the region of Mg-17

— 2Fo-Fc — Fo-Fc — Fc-Fo

The structure of Mg-17 and its' neighborhood are shown in Fig. 4.8.



a. Mg-17 absent (75% probability)



b. Mg-17 presents (25% probability)

Fig. 4.8. The structure of Mg17 and vicinity

The structure shows significant anisotropic thermal motion. Fig. 4.9 shows a section of the DNA in this structure which shows a significant anisotropic thermal motion.

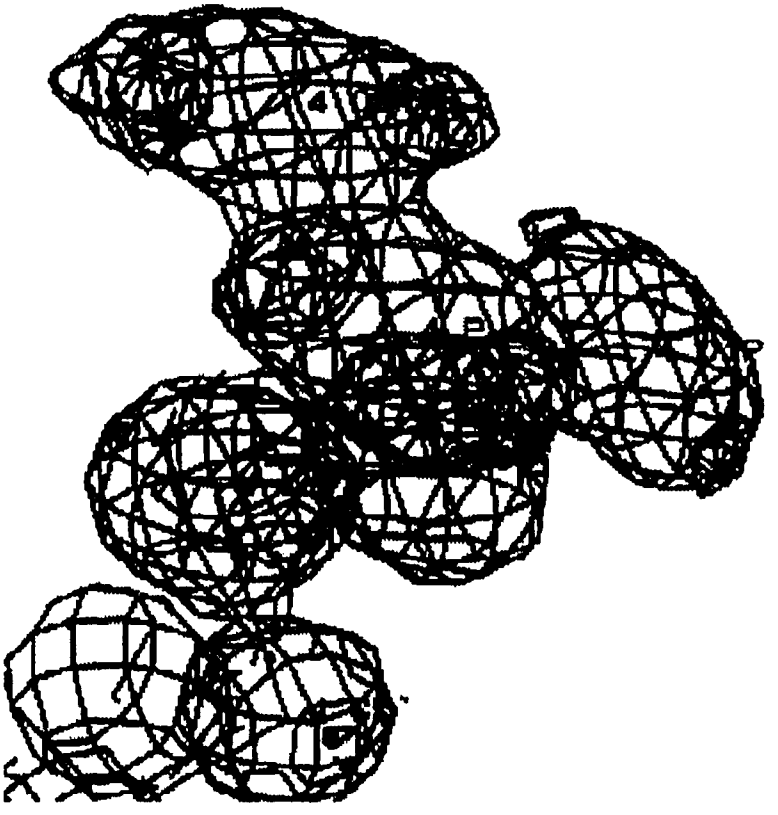


Fig. 4.9 Electron densities around section of phosphate group, P-4, anisotropic thermal motion is clear.

—— 2Fo-Fc —— Fo-Fc

The refinement procedure used in this study was designed to treat the thermal motion of every individual atoms/ions isotropically. The electron density maps show strongly anisotropic features for the thermal motion of most atoms in this structure. At this stage we applied an

anisotropic overall B scaling on the observed structure factors. By using the rescaled structure factors, further refinement reduces the R-factor to 14.8%.

The final structure including five Mg(II)-water complexes is shown in Fig. 4.10.

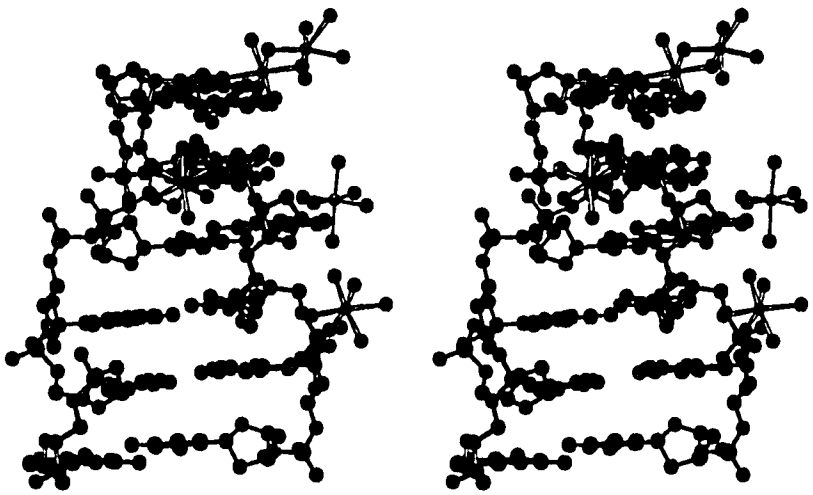


Fig. 4.10 Stereo view of Z-DNA, d(CGCGCG)$_2$ with 5 magnesium(II)

4.5.2 The Structure of d(m5CGUAm5CG)$_2$-Cu(II) Complex

The structure was solved by molecular replacement, taking the native structure of d(m5CGUAm5CG)$_2$ [132] as the initial model. The initial model was refined on 2 Å in the absence of solvent molecules using Konnert-Hendrickson nonlinear least square constrained refinement method, PROLSQ [207]. The current working version of PROLSQ in this laboratory was modified to be specific for nucleic acids by G. J. Quigley [193]. The copper(II) ions were identified in a (2Fo-Fc) map as intense positive electron densities located approximately 2 Å from the N7 nitrogen of the purine bases. On further refinement, four partially occupied crystallographically unique copper(II) complexes (with various numbers of water ligands) and 71 well defined free water molecules were located using 1.3 Å resolution data (Fig. 4.11).

The final refinement was completed by using XPLOR software with empirical energy minimization parameters as described previously. The final R-factor was 20.9% to 1.3 Å resolution.

The occupancy of copper(II) in the structure had been verified collaboratively at Prof. Ho's laboratory. The verifying analysis include Atomic Absorption spectrometer and UV spectrometric absorption determination of re-dissolved crystal. These results support the partial occupied Copper structure. [208]

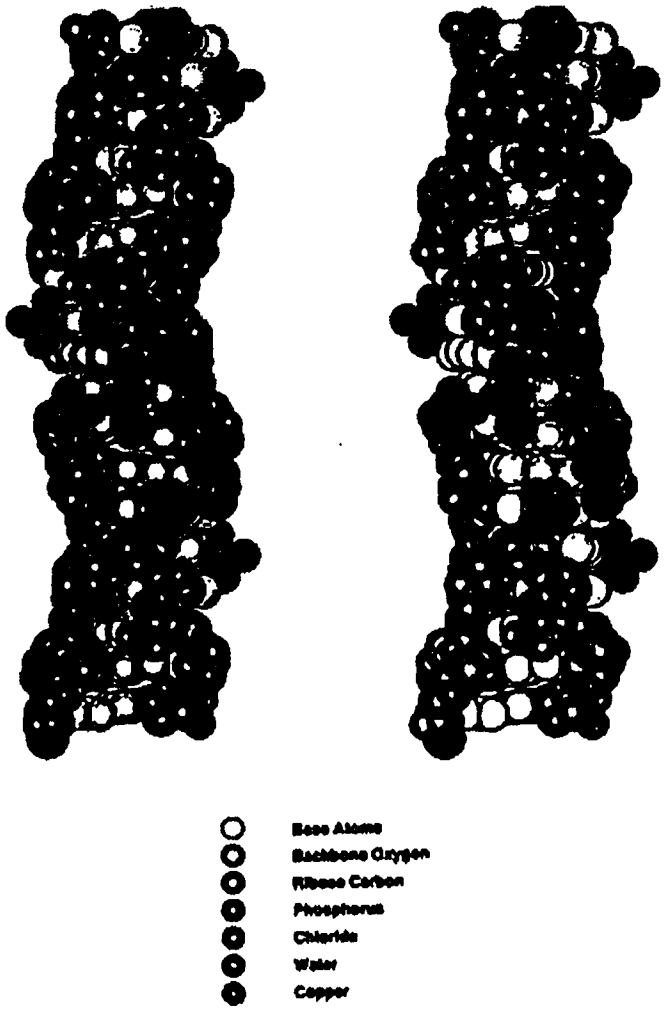
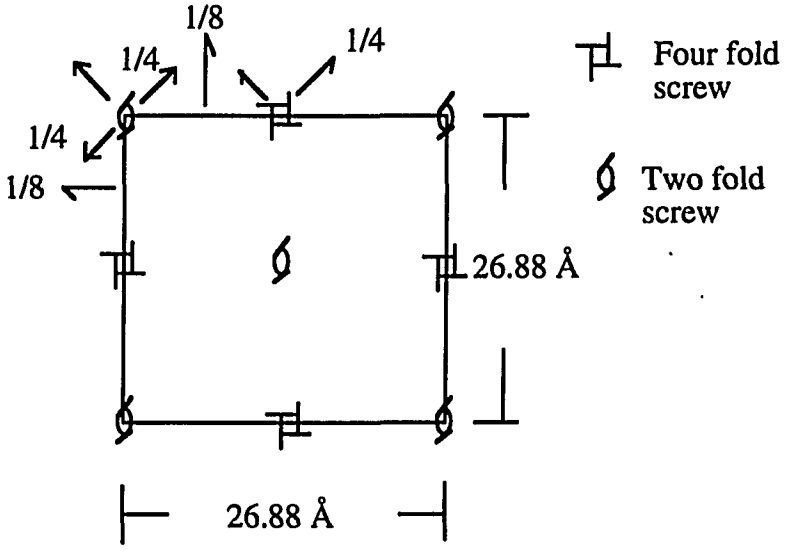


Fig. 4.11. Space filling structure of Cu-d(m5CGUAm5CG)$_2$

4.5.3 Ditercalinium-d(CGCG)$_2$ Complex

The Structure determination of Ditercalinium-d(CGCG)$_2$ was a challenge to us because of the limited data. There are no heavy atom derivatives available. We only have the diffraction data to 1.7 Å resolution, the parameters of unit cell, the symmetry of unit cell, and a predicted model (without coordinates).

The unit cell was determined to be 26.88x26.88x82.60 Å^3 with $P4_12_12$ symmetry (space group #92). The symmetry is shown in Fig. 4.12



Not all the symmetry operators are shown

Fig. 4.12. The unit cell and symmetry of space group $P4_12_12$ (#92, after international tables for x-ray crystallography)

Taking the CpG intercalation step of the $d(CGATCG)_2$-adriamycin complex [209] as a starting point, several models of the ditercalinium-$d(CGCG)_2$ were constructed. Based on the assumption that the drug and DNA first forms a stable complex (known from NMR etc.), and that this then crystallizes, four models seemed plausible as shown schematically in Fig. 4.13.

These consisted of two of each having the linker in either the minor or major grooves. The two models of each consisted of having the two chromophores intercalated between the two terminal base pairs. (Fig 4.13. B) This requires opening two intercalation sites but both are at ends and both chromophores are stacked between DNA bases removing them, to a large degree, from the aqueous environment. The second pair of possibilities involves one chromophore intercalated between the middle two base pairs with the remaining chromophore stacked on one end. (Fig. 4.13. C) While, this only requires opening a single intercalation site, it is internal and therefore more restricted. In addition, the remaining chromophore is highly exposed to the solvent, however, a terminal base pair is buried. The NMR results suggested that the former one, i. e. the bis-intercalation model to be the most likely structure of this complex. Taking the major groove binding bis-intercalation structure, the model is roughly a cylinder of 20 Å diameter and 20.4 Å long. (Fig. 4.13 C. left). According to the symmetry and the size of unit cell, there should be eight molecules in a unit cell, and one molecule in an asymmetric unit.

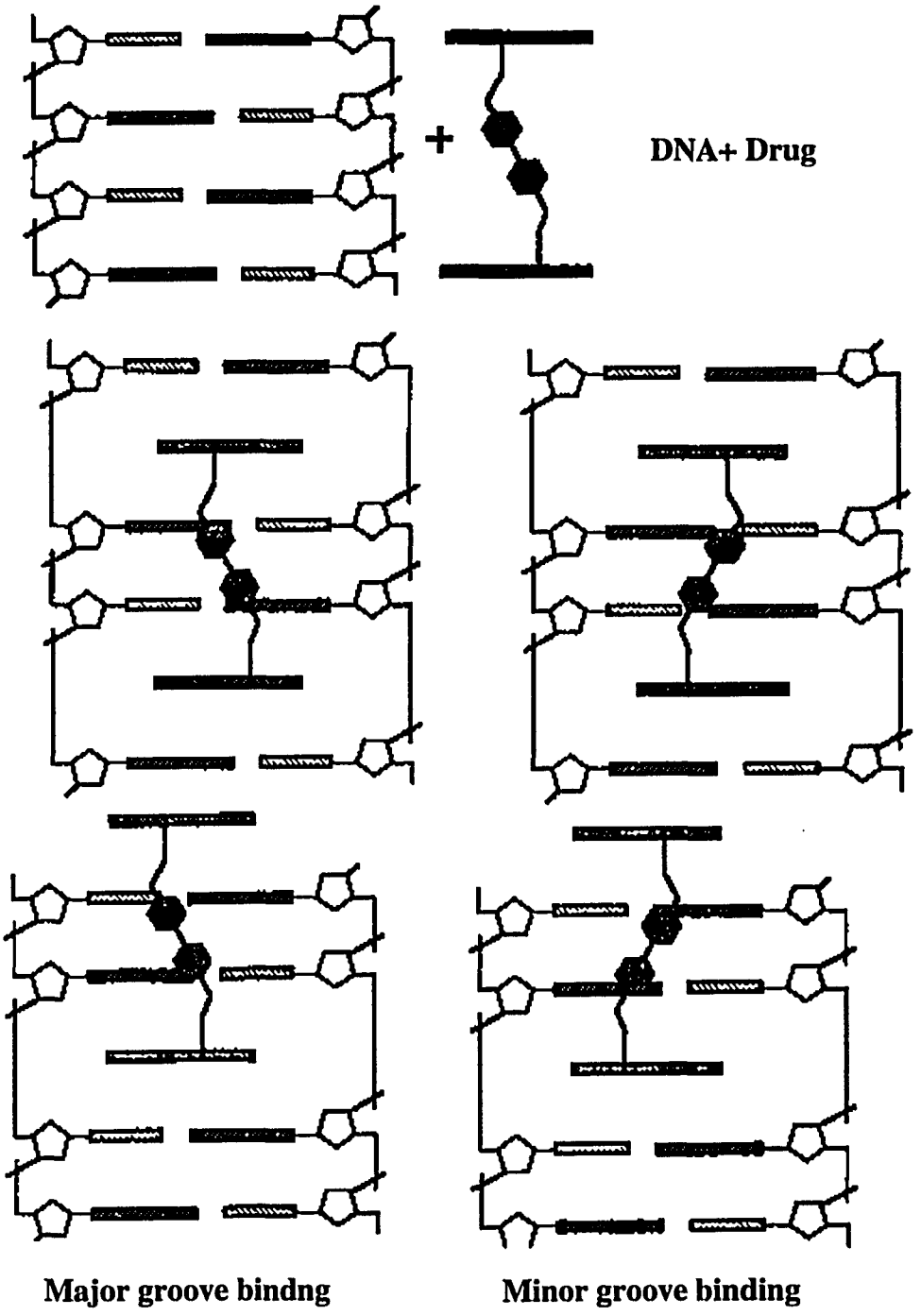


Fig. 4.13. Models of ditercalinium-d(CGCG)$_2$ complex

In order to solve the structure, an initial Patterson map was produced. The Patterson map indicated a strong stacking at a 3.44 Å interval along the c axis (Fig. 4.14).

Such a pattern strongly suggested that there are 24 planar groups, either base pairs or chromophores of ditercalinium, which would place four DNA-ditercalinium complexes stacking along the c axis. Therefore, the complexes would occupy the general positions with the asymmetric unit consisting one DNA duplex and one ditercalinium molecule.

Because no heavy atom derivative is available in this work, the molecular replacement method was employed. Taking the initial model, based on the predicted structure, and considering the symmetric operation of the space group ($P4_12_12$) and the size and shape of the model, the center of the molecule can only be located in two regions. They are position I (around x=0, y=0, and z=1/8,) and position II (around x=0, y=0.5, and z undefined but only 0-1/4 unique). If the molecule sits at position I, the search range in x and y direction should not exceed 1/4. The search range in z direction is very small by the limitation of the two perpendicular two-fold symmetry operators. If the molecule sits at position II, the search ranges in x, y, and z directions should not excess a quarter of the dimensions of the unit cell, because the two-fold screw axises do not restricts the translation along z axis.

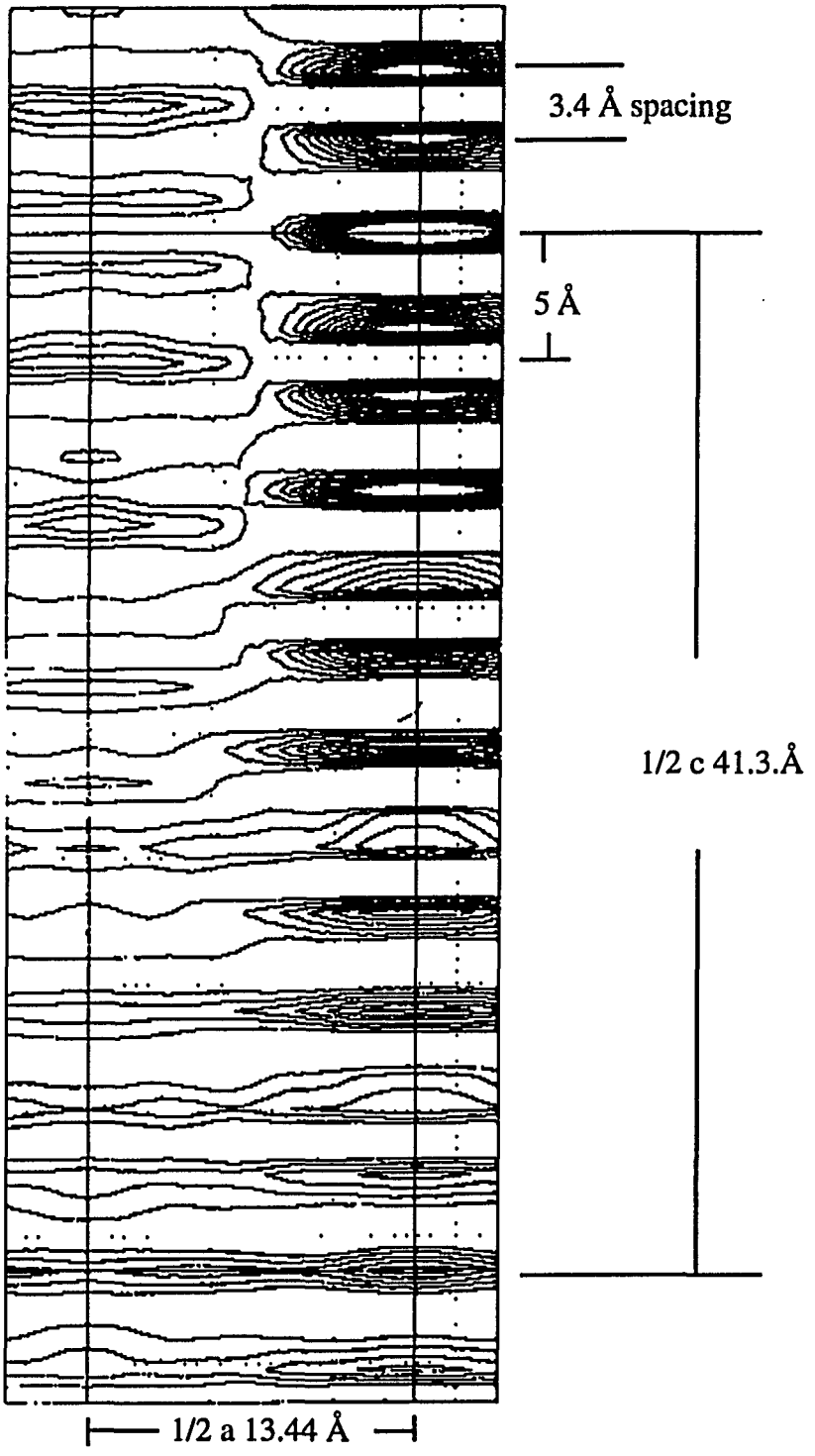


Fig. 4.14. Patterson map of ditercalinium-DNA complex 2Å resolution.

Initial attempts to solve the structure using a series of search techniques were completely unsuccessful. Close examination of the data revealed that the intensity distribution was highly anisotropic. An anisotropic rescaling program was written by G. J. Quigley for recombining the various data sets giving a substantial improvement in agreement between data sets. This was subsequently reconfirmed using the then newly acquired version of XPLOR which also contained an anisotropic rescaling option.

Continued searching both with rotation function and R search techniques continued to fail until it was recognized that the search models were tending to deviate from the expected stacking direction using the limited 4-5 Å data typically used for these searches. To avoid this problem, limited number of strong reflections that corresponding to the 3.4 Å stacking, wewe included in the R search. With the inclusion of this data the primary candidate model (that closest to the NMR structure) showed a clear minimum. This model was subsequently refined first by rigid body techniques (XPLOR) and then releasing the rigid body constraints and gradually adding the higher resolution data.

The structure was finally refined to R factor equal to 20 per cent at 1.7 Å resolution.

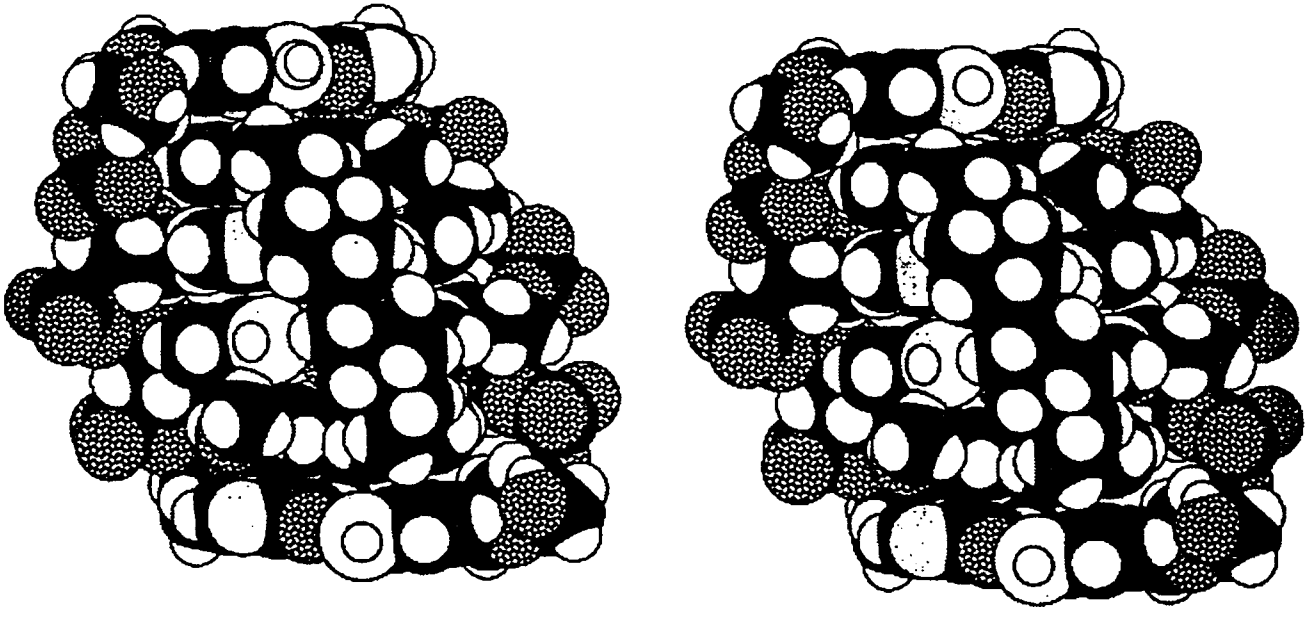


Fig. 4.15. Space filling structure of ditercalinium-d(CGCG)$_2$

CHAPTER V RESULTS AND DISCUSS

5.1 Z-DNA d(CGCGCG)$_2$ at Low Temperature

5.1.1 Structure Feature of d(CGCGCG)$_2$ at Liquid Nitrogen Temperature

The overall shape of the oligomer is almost a perfect cylindrical rod. The oligomer forms a left-hand double helix with a wide "surface" major groove and a very deep narrow minor groove. The six base pairs form exactly a half turn of helix, and two symmetry related molecules form a complete turn. The stacking between those two molecules is so similar to that of GpC stacking within the molecule that it makes the two adjacent molecules look like a continuous double helix, with two phosphate group missing. See Fig. 5. 1 and Fig. 1.14.

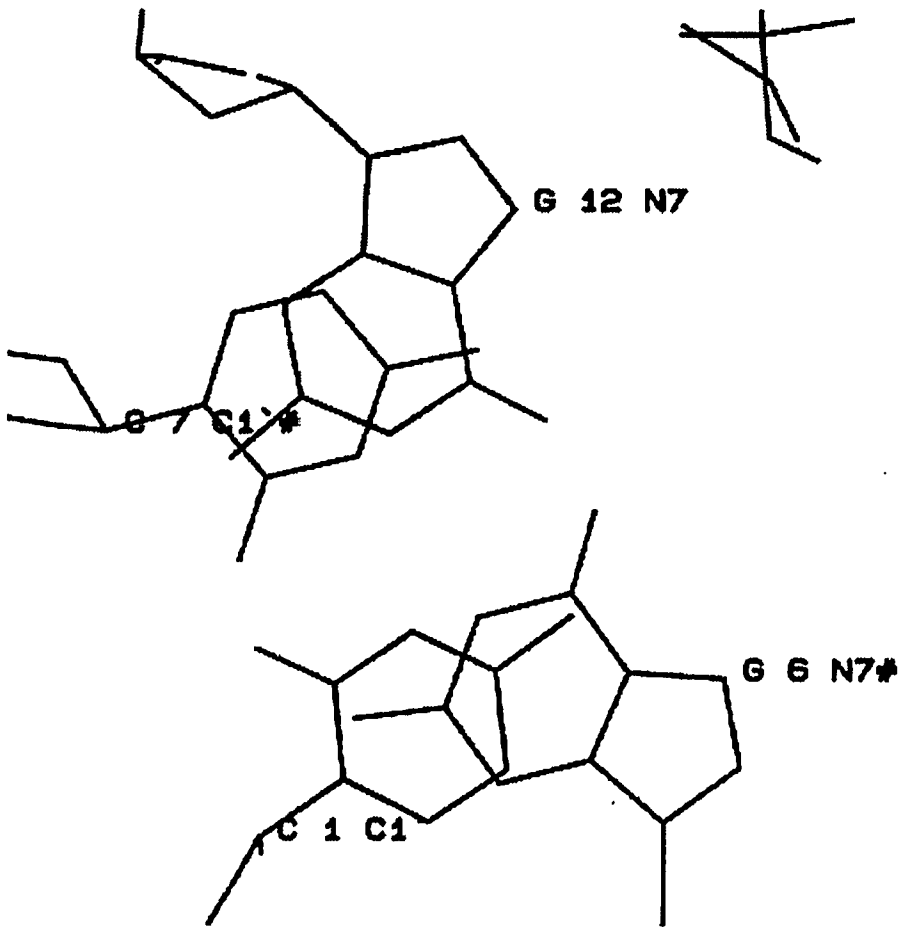


Fig. 5. 1. Intermolecule base stacking in Z-DNA crystal
(G6*-C7* and C1-G12)

The general structure of d(CpGpCpGpCpG)2 at 150° K is similar to the structure at room temperature of the same sequence. Fig. 5.2 shows the superposition of the two structures.

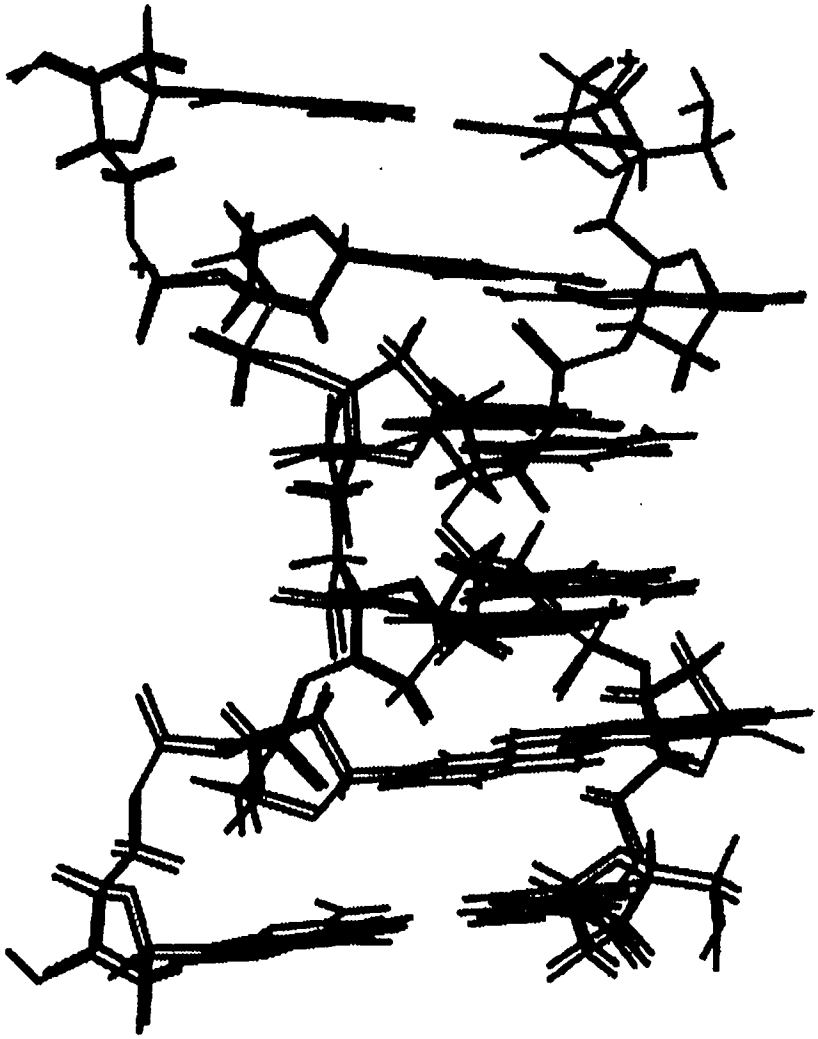


Fig. 5.2. Superposition of the structures of Z-DNA at low and room temperature

——— room T. ——— low T.

The backbone GpC sugar-phosphate chains are mainly in the ZI conformation with 25% of phosphate 3 and 15% of phosphate 9 in the ZII conformation. It is interesting that the two multi-conformation phosphate groups are related by a pseudo two fold rotation symmetric axis. See Fig. 5.3.

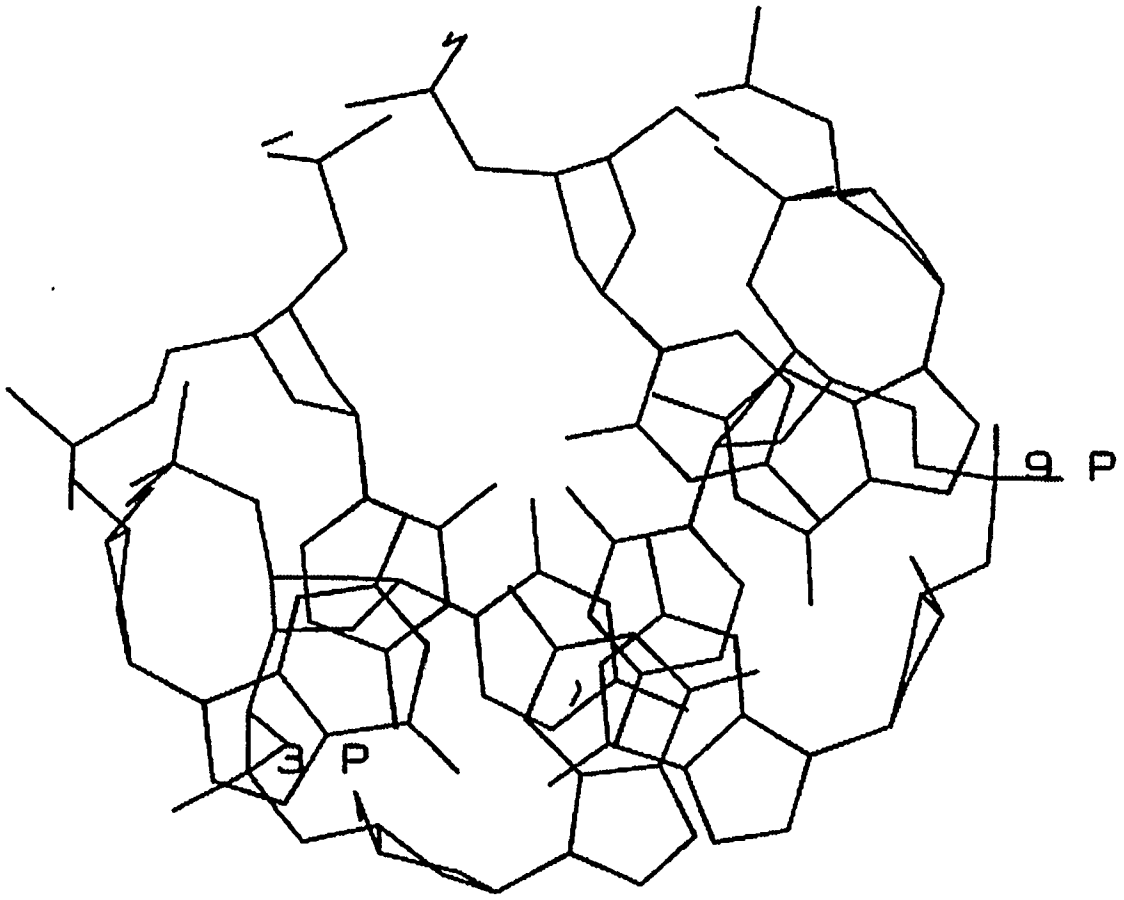


Fig. 5.3. Phosphate 3 and 9 symmetric relation

Due to the existence of alternative conformations, the solvent shell around each of the two conformations in these regions must be different. One partially occupied water molecule, OH2 112, has been determined associated with ZI phosphate 3 and no additional water molecule can be identify in the region of phosphate 9. The different occupancies of the ZI and ZII conformation for these two phosphate groups (75/25 for P3 and 85/15 for P9) may be explained as the result of the effect of crystal packing. Figure 5.4. shows the sliced packing pattern by base pair. It is clear that the environments of P3 and P9 are different. The phosphate group of P3 is located in a widely open solvent channel, while the phosphate group of P9 is in a narrow solvent channel.

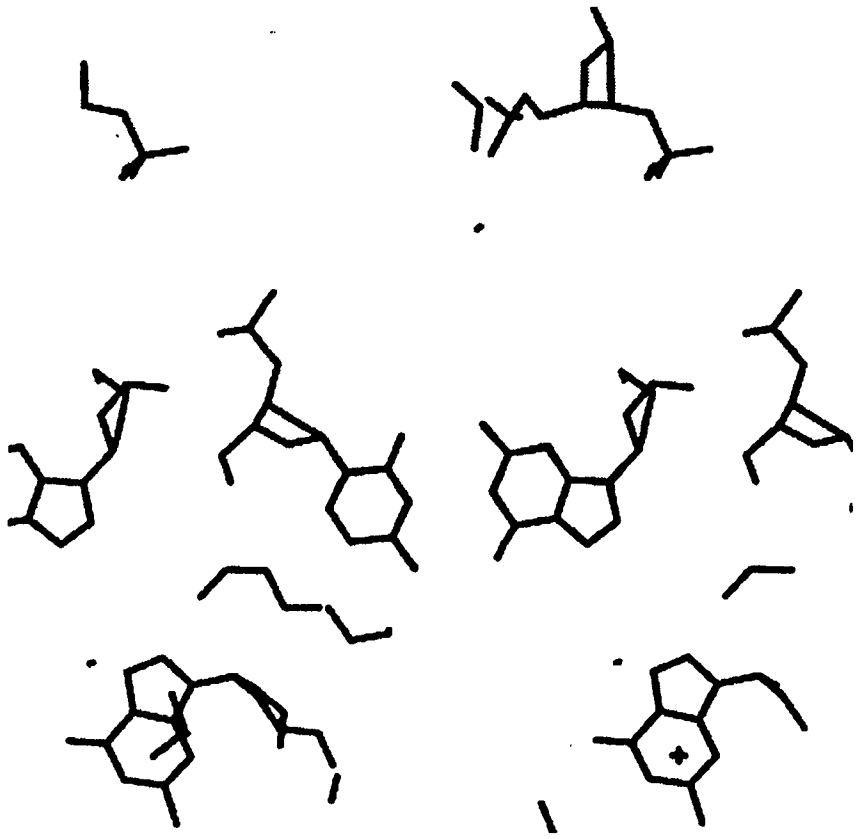


Fig. 5.4A. Packing of C1-G12

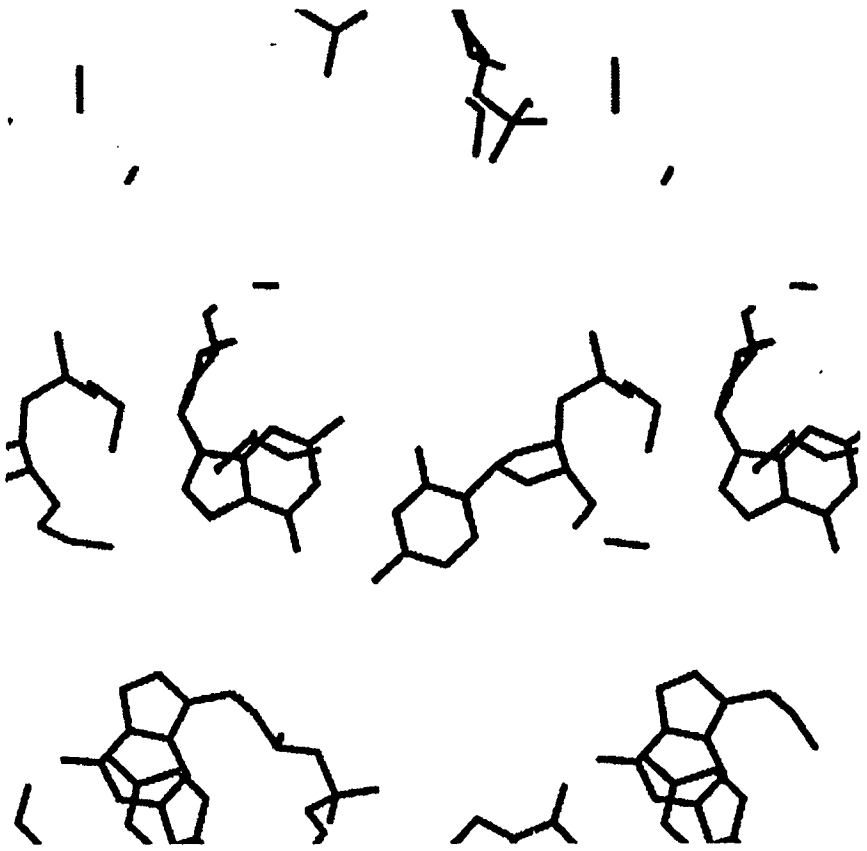


Fig. 5.4B. Packing of G2-C11

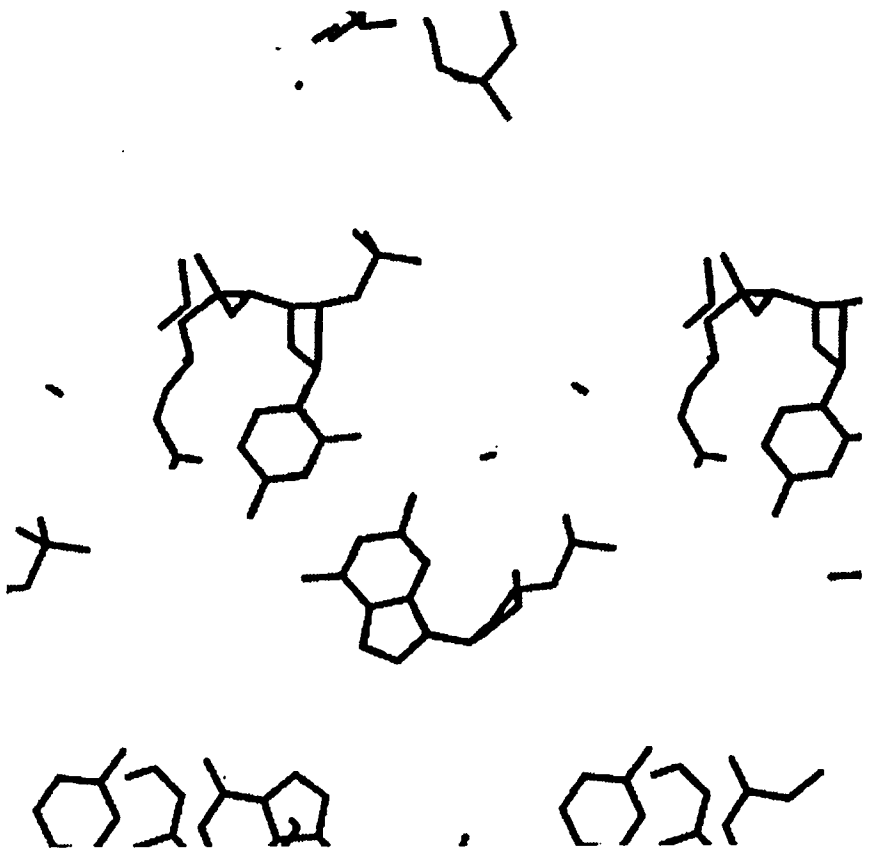


Fig. 5.4C. Packing of C3-G10

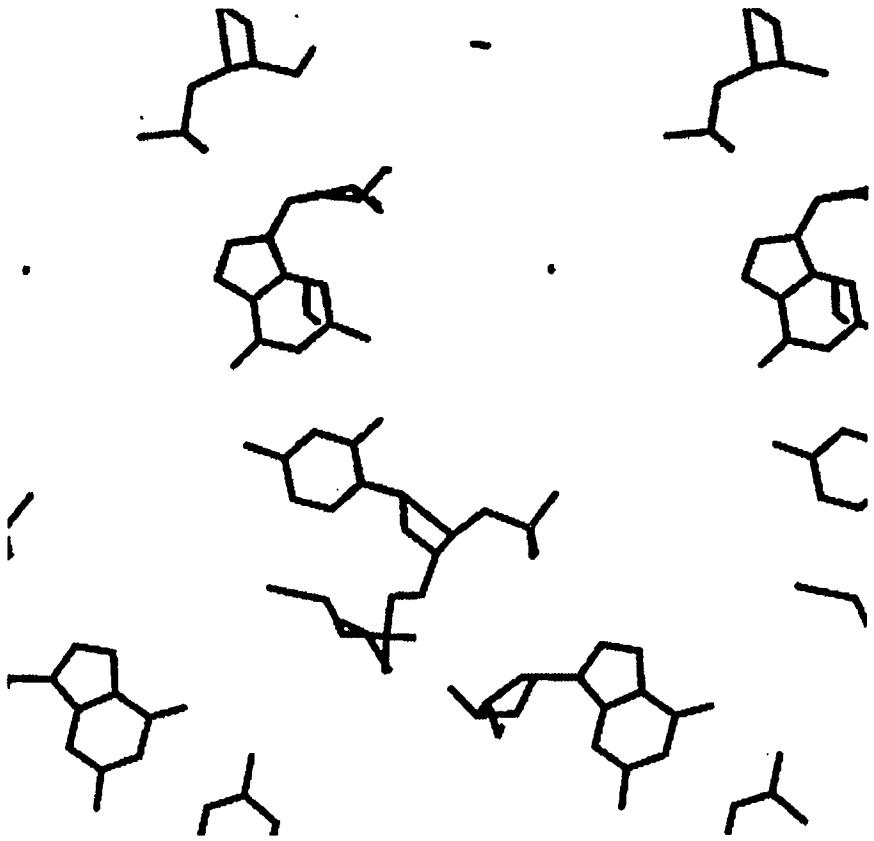


Fig. 5.4D. Packing of G4-C9

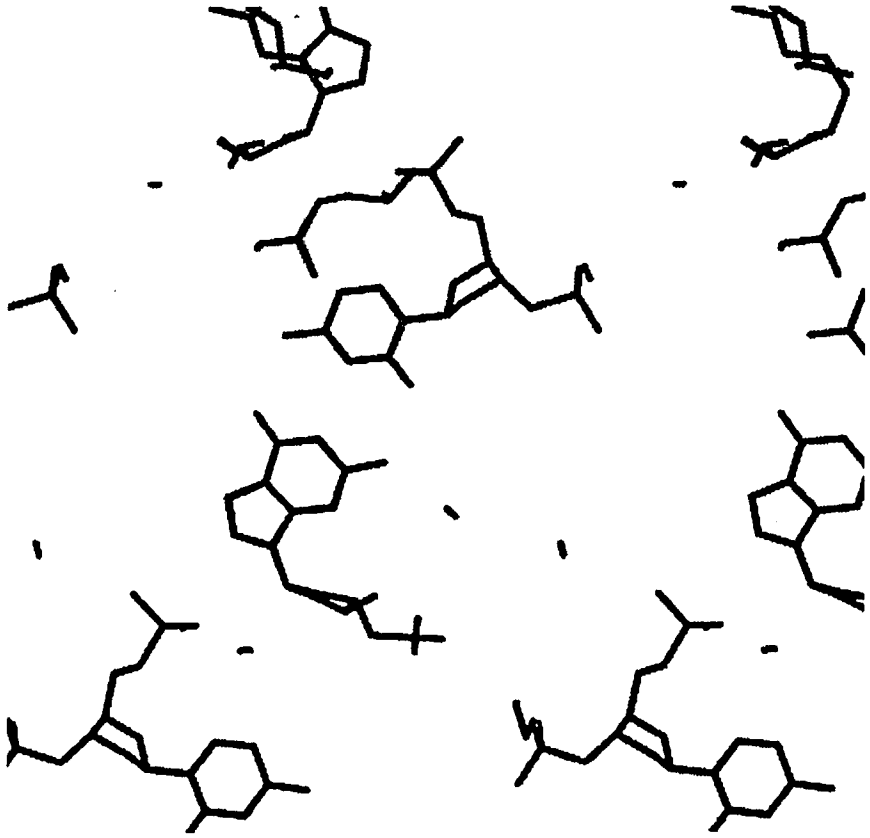


Fig. 5.4E. Packing of C5-G8

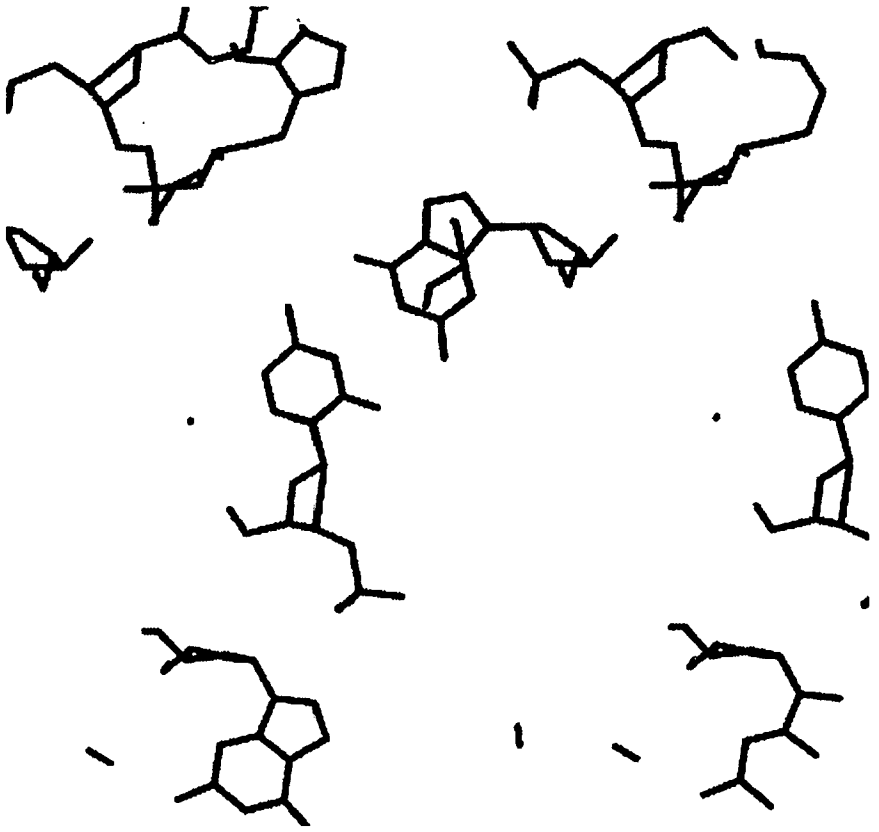


Fig. 5.4F. Packing of G6-C7

5.1.2 The Counterion and Solvent Structure.

A total of five unique counterion, Mg(II), sites have been identified in the structure. These divalent magnesium are named MG-13, MG-14, MG-15, MG-16, and MG-17 respectively. According to their interaction with DNA and occupancies, the magnesium(II) cations can be classified into groups.

By occupancies of magnesium(II) there are two groups.

Group I: fully occupied magnesium, MG-13, MG-14, and MG-16;

Group II: partially occupied magnesium, MG-15 and MG-17.

By the interaction to DNA, there are also two different groups.

Group A: directly binding to DNA via N7-MG coordinate bond, MG-14.

Group B: not directly binding to DNA, MG-13, MG-15, MG-16, and MG-17.

The five magnesium(II)-water complex are shown in figure 5.5.

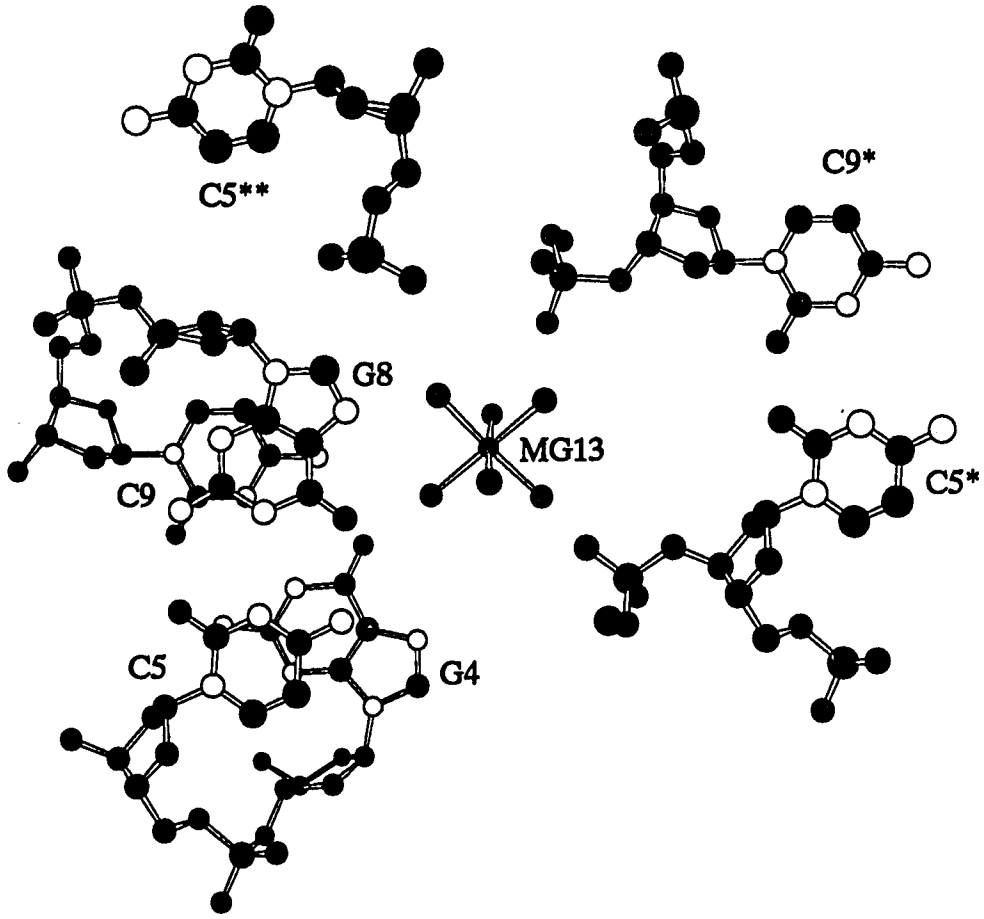


Fig 5.5 A. Mg13 and vicinity

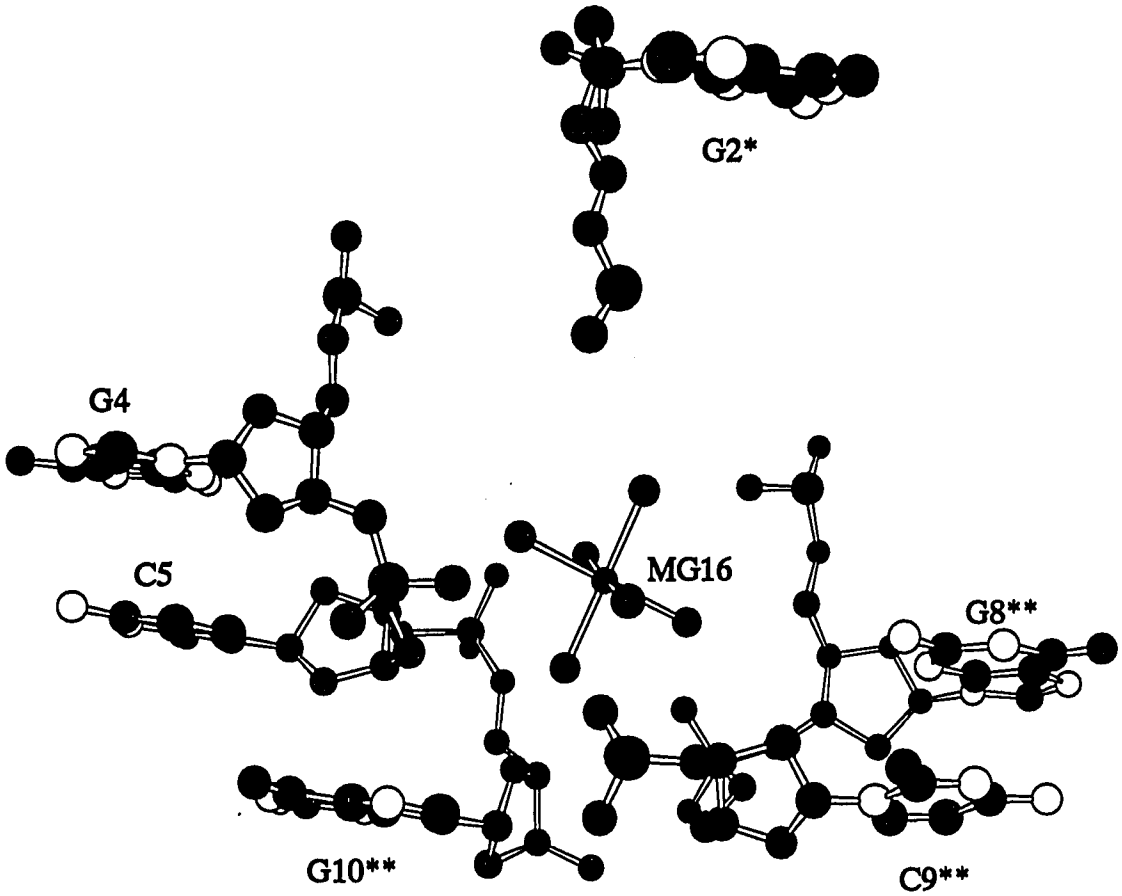
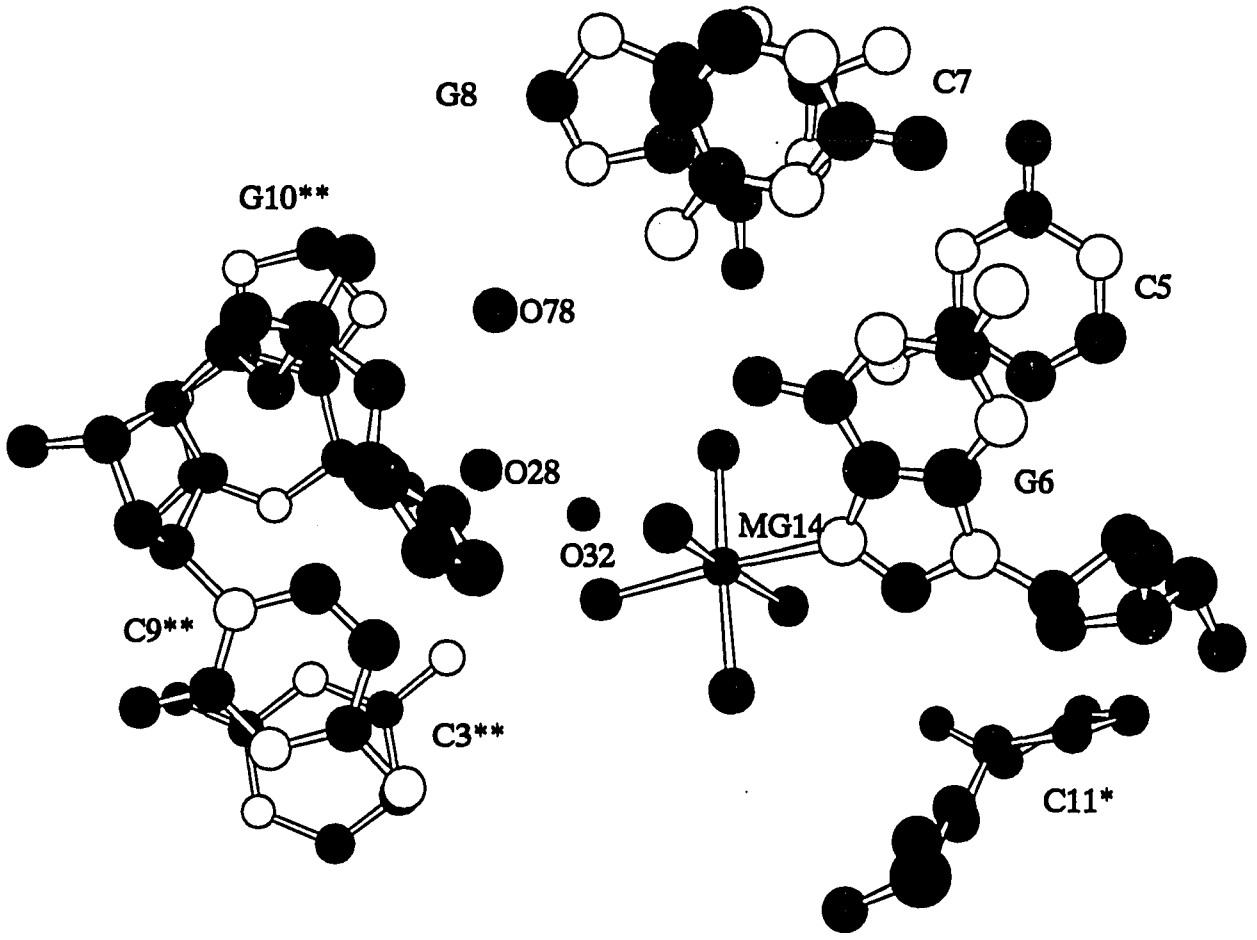
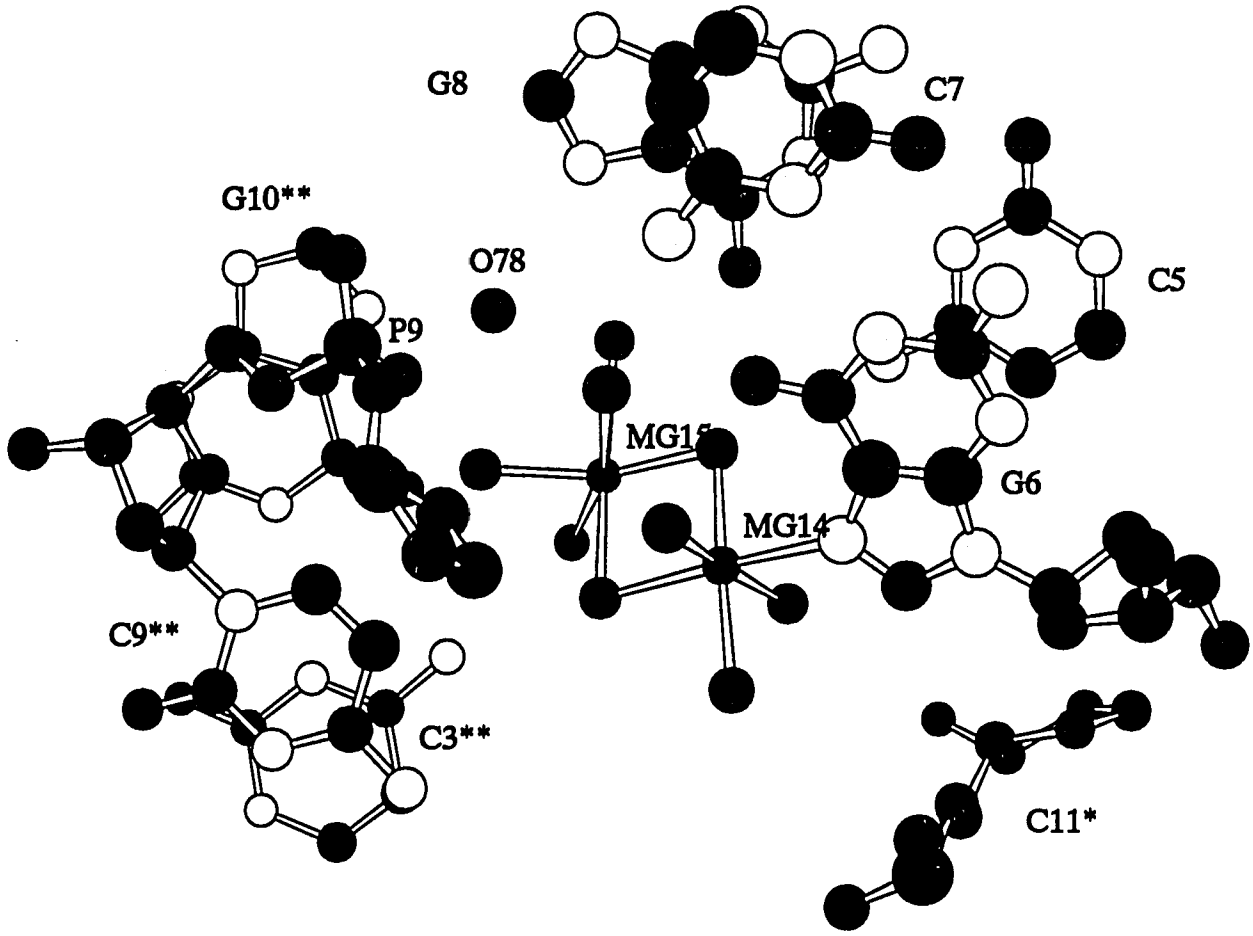


Fig 5.5 B. Mg16 and vicinity



Mg-15 absent (75% probability)

Fig. 5.5Ca. Mg-14, Mg-15 and vicinity-I



b. Mg-15 present (25% probability)

Fig. 5.5Cb. Mg14, Mg-15 and vicinity-II

All magnesium cations form nearly perfect octahedral geometry with six coordinated ligands. MG-14 has five water ligands and a nitrogen atom ligand, N7-G6. The others have six apparent water ligands each. The sites of MG-14 and MG-15 have been interpreted as a $Mg_2(H_2O)_9$ dimmer bonded to N7-G6. [210] See Fig. 1.19. At the lower temperature this appears to be only partially true. MG-14 can be interpreted as a full occupied cation site, while the MG-15 is partially occupied with occupancy 25 percent at most. The interpretation as a fully occupied MG-14 and partially occupied MG-15 is clearly better from the electron density map. After changing the occupancy of MG-14, the R-factor reduced significantly.

A total of 94 water molecules have been determined per asymmetric unit. 27 water molecules are associated with magnesium cations, 56 water molecules are directly interacting with bases or backbones. 7 water molecules show clearly dual occupied sites. 12 water molecules are partial occupied.

The water molecule occupancy in the vicinity of MG-14 and MG-15 shows large variations. Due to the partial occupied Mg-15, the water molecules (OH2-29 and OH2-31), that only coordinate to Mg-14 have the same occupancy with central atom (Mg-15). The water molecules (OH2-27 and OH2-30), that coordinate to both Mg-14 and Mg-15, are fully occupied. The water molecule, OH2-32, is fully occupied with two alternative positions. When Mg-15 is present, OH2-32 (25% occupancy)

coordinates to Mg-15. When Mg-15 is absent, OH2-32 occupies another position, that near to the empty site of Mg-15. There is another dual-position water molecule, OH2-78. Its' position also depends on the presence of Mg-15. The water molecules in the vicinities of phosphate 3 and phosphate 9 also show a variation in occupancy, position, and orientation. The four water molecules (OH2-39, OH2-40, OH2-41, and OH2-42) associate with MG-17 are partial occupied. Water molecule OH2-43 is fully occupied. Water molecule OH2-44 is fully occupied with two alternative positions, depending on the presence of Mg-17.

5.2 Copper Binding d(m5CGUAm5CG)$_2$

5.2.1 The Structure of Copper Binding d(m5CGUAm5CG)$_2$

The self-complimentary DNA oligomer d(m5CGUAm5CG)$_2$ was crystallized in the Z-form. Copper(II) ions were introduced into the crystal by soaking the crystal by copper(II) chloride solution. In the overall structure (Fig. 5.9) a total of four copper(II) complexes per DNA molecule were found. The copper cations were bound to N7 nitrogen of purines and occurred in the intermolecular channels. The space group and unit cell parameters of this and the native d(m5CGUAm5CG)$_2$ structure [132] are similar, and both are isomorphous with most other hexamer duplexes crystallized as Z-DNA. See Table 5.1.

Table 5.1 Comparison of unit cell of native and copper binding d(m^5CGUAm^5CG)$_2$

Space Group $P2_12_12_1$

	a	b	c
Native* DNA	17.82	30.44	44.52
DNA-Cu(II)	17.59	30.58	44.52

* after Zhou and Ho

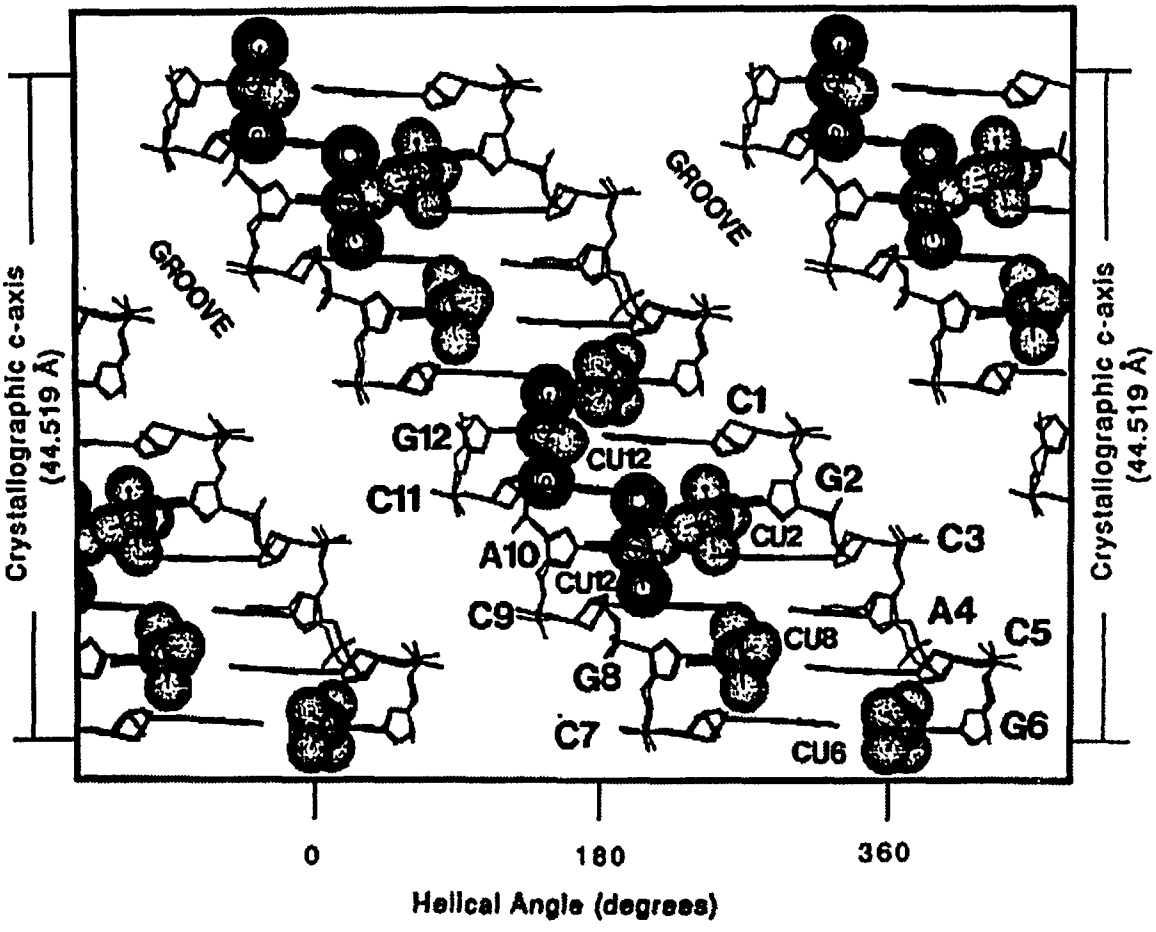


Fig. 5.6. Cylindrical plot of Cu-d(CGCGCG)$_2$ and Cu-d(m5CGUAm5CG)$_2$

In the structure copper(II) shows a strong base specific binding to guanine. The copper(II)s are coordinated to the N7 nitrogen of G2, G6, G8, and G12. Figure 5.6 is a superposition of the cylindrical projections of the copper(II) soaked versus native structures of $d(m5CGUAm5CG)_2$. The cylindrical projections were generated by projecting the atomic coordinates of the model from the helical axis onto a cylindrical surface wrapped around the DNA. When this surface is unwound, we get a plot of the atomic coordinates of the structure with the helical axis along the vertical axis of the plot, versus the helical angle along the horizontal axis. The model for the copper(II) chloride-soaked structure is shown with light bonds, while that of the native structure is shown with heavy bonds. The residues are numbered C1 to G6 in the 5' to 3' direction along one strand, and C7 to G12 in the 5' to 3' direction along the complementary strand. The copper(II) complexes are shown as van der Waals spheres, and are labeled according to the purine base to which they are bound (e.g., the Cu2 complex is bound to the guanine base at position G2). From this comparison, it is evident that the DNA structures of the two models nearly overlap. This is consistent with the low overall root mean square (RMS) deviation between the two DNA structures of 0.358 Å. There are, however, some details that differ between the models, that are directly or indirectly related to specific binding of copper (II) to the DNA bases and their subsequent effects on the bases and backbone of the hexamer duplex. The conformation of the ribose sugar ring of guanine residue G12 has inverted from a 3'-endo in the native structure to a 2'-endo pucker in

the presence of copper(II). This results in a rotation about the C5' carbon-O5' oxygen bond, and the subsequent displacement of the phosphodiester linking adenine A4 and cytosine C5. The minor groove is thus narrower at this point of the hexamer than in the native structure. Additional minor rotations of the phosphodiester linkage between residues U9 and A10, C5 and G6, and A4 and U3 result in slight narrowing of the minor groove crevice at these points in the copper(II) chloride soaked crystal structure.

5.2.2 Copper(II) Content of the d(m5CGUAm5CG)$_2$ Crystal

The total content of copper(lI) in this crystal structure was estimated to be 1.3 copper(II) ions per hexamer duplex by numerical integration of the electron densities of the cations in the crystallographic (2Fo-Fc) map. This result agrees with the result of 1.1 copper(II) ions per hexamer duplex collaboratively verified by spectroscopic measurements. [202] These two values are identical within the limits of accuracy in the molar extinction coefficient value for determining the DNA content. Thus-the numerical integration method for estimating the copper(II) content in the DNA crystal, and subsequently the distribution of copper, appears to be valid.

5.2.3 Distribution and Coordination Geometries of Copper(II) Complexes at the Guanine Bases

The structure of the $d(CGCGCG)_2$ crystal soaked with copper(II) chloride [202] showed that the micro environment at each potential metal binding site, as defined by the packing of the hexamer duplexes in the crystal, affects both the affinities and the chemical properties of the complexes at the various sites in the DNA. To illustrate the similarities of the micro environments between the copper(II) chloride-soaked $d(CGCGCG)_2$ and $d(m5CGUAm5CG)_2$ crystals, the distribution of copper(II) and the coordination geometries of the copper complexes at the four guanine residues that are common to both structures (residues G2, G6, G8 and G12), are compared in Table 5.3 and Table 5.4 respectively. A more detailed description of these copper(II) complexes is given by Kagawa et al., 1991. For this comparison, we have defined the relative partial occupancy of copper at each site as the percent of copper(II) bound to that site versus the total quantity of copper bound to the hexamer (Table 5.2). This allows us to compare the affinities of each potential binding site between crystals having different absolute quantities of copper(II) bound (1.1 copper ions for this crystal versus 1.9 copper ions for the $d(CGCGCG)_2$ sequence.

In both the $d(CGCGCG)_2$ and the $d(m5CGUAm5CG)_2$ crystals, the Cu6 and Cu8 complexes account for 33% and 12 to 14% of the total copper(II) ions, respectively. The coordination geometry of the Cu6 copper complex is octahedral, while that of the Cu8 complex is a trigonal bipyramidal in both crystal structures (Fig. 5.7).

Table 5.2. quantities of Cu(II) in crystal

Copper site	d(CGUACG) % Cu(II)[1]	occupancy (%)[2]	d(CGCGCG) % Cu(II)[1]	occupancy (%)[2]
Cu2 at G2	17.9	16.3	6.7	3.5
Cu4 at G4/A4	1.8	1.6	19.9	10.5
Cu6 at G6	36.3	33.0	62.6	32.9
Cu8 at G8	14.9	13.5	23.0	12.1
Cu10 at G10/A10	NA	NA	39.0	20.5
Cu12 at G12	40.9	37.2	39.0	20.5
sum	110.0 1.1 Cu(II)/per hexamer	100.0	190.0 1.9 Cu(II)/per hexamer	100.0

The %Cu(II) and occupancies were determined from the (2Fo-Fc) electron density maps.

[1] The %Cu(II) at each site represents the integrated intensity of the electron density in the region of the site related to 27, the number of electrons for 100% occupied site.

[2] The relative occupancy is the %Cu(II) at the site normalized against the total copper in the crystal.

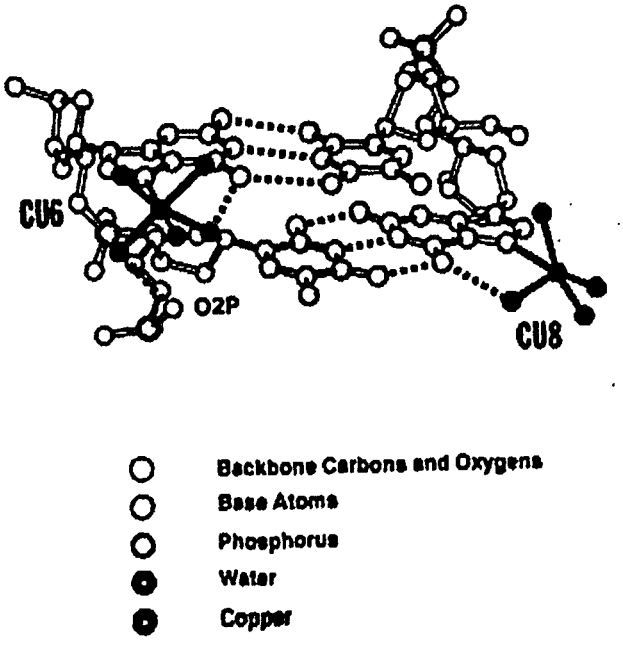


Fig. 5.7. Two geometries of Cu(II)-water complex

The atoms of the copper(II)-water complexes are shown as dark shaded bonded spheres. The octahedral coordination geometry of the Cu6-complex bound to N7-G6. The copper cation Cu8 coordinates to five ligands, four waters and N7-G8.

In these complexes, the N7 nitrogen of the guanine base and a water molecule act as the two axial ligands, while three to four waters form the equatorial plane of the coordination complex. The relative partial occupancy of the Cu2 copper complex is higher in this structure (16%) versus the $d(CpGpCpGpCpG)_2$ structure (4%). The lower

binding affinity for guanine residue G2 in the $d(CGCGCG)_2$ crystal results from a greater distortion of the coordination bond relative to the plane of the guanine base, and fewer stabilizing interactions between this copper complex and the DNA backbone and with other complexes, relative to the $d(m5CGUAm5CG)_2$ structure. These factors result in a more regular coordination geometry for the Cu2 complex in this structure (distorted octahedron) than in the $d(CGCGCG)_2$ structure. The N7 nitrogen to copper bond distance, and the trigonal bipyramid geometry of the Cu2 complex in this crystal is distorted by the proximity of a negatively charged phosphate group from a neighboring hexamer duplex. This site, therefore, is slightly more open in the $d(m5CGUAm5CG)_2$, but is still affected by the crystal packing that is characteristic of this site in Z-DNA crystals.

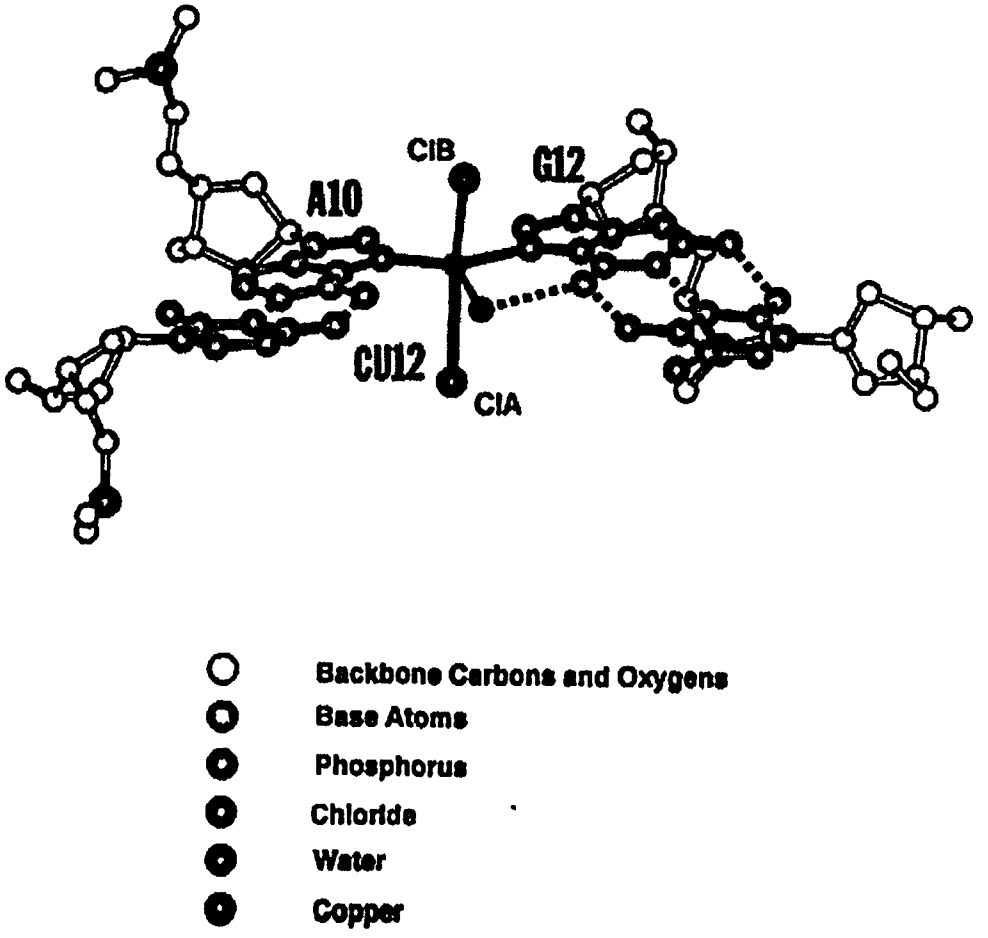


Fig. 5.8. The geometry of CU12 with A10and G12

Finally, residue G12 in the $d(CGCGCG)_2$ crystal is in close proximity to guanine residue G10 of an adjacent symmetry related hexamer (Fig. 5.8). The metal centers of these complexes are separated by 1.55Å, and the guanine base planes are displaced by approximately 1Å along the z-axis. The two potential copper(II) binding sites, therefore, must be alternatively partially occupied. The relative partial occupancies of the two sites are 21% each, accounting for 42% of the total copper in the asymmetric unit. In the $d(m^5CGUAm^5CG)_2$

structure, guanine residue G12 is packed adjacent to the adenine residue A10 of a neighboring duplex (Fig. 5.8.A). Rather than forming two mutually exclusive sites, however, the two purine bases form a single copper(II) binding site, with a relative partial occupancy of 37%. The G12 residue and its partner, therefore, account for approximately 40% of the copper(II) present in both crystals. The properties of this site will be discussed in a later section in the context of the affinity of copper(II) binding to adenine bases in the crystal.

With the exception of the Cu2 site, the distribution and coordination geometries of copper(II) complexes at analogous guanine residues appear to be similar in the two crystals. This indicates that the C5 methyl group of the cytosines, and the replacement of the two internal guanine bases in the hexamer does not adversely affect the ability of the metal to bind to accessible sites in the isomorphous crystals. A comparison of the copper(II) affinities at the adenine A4 and A10 residues with the analogous G4 and G10 residues in the $d(CGCGCG)_2$ crystal, therefore, allows us to draw some valid conclusions concerning the susceptibility of adenine bases to copper(II) interaction in duplex DNA.

5.2.4 Copper(II) Binding to the Open Adenine A4 Site

The Cu4 binding site at residue G4 in the $d(CGCGCG)_2$ crystal structure is located in an open solvent channel, and is characterized by

its accessibility and the lack of available stabilizing interactions. Copper binding at this open G4 site would best model copper binding to a guanine residue of double stranded DNA in solution. The binding affinity for this site (a relative partial occupancy of only 11%) was lower than the other accessible copper binding sites in the crystal; nevertheless, the presence of this complex is clear (Fig. 5.9. B).

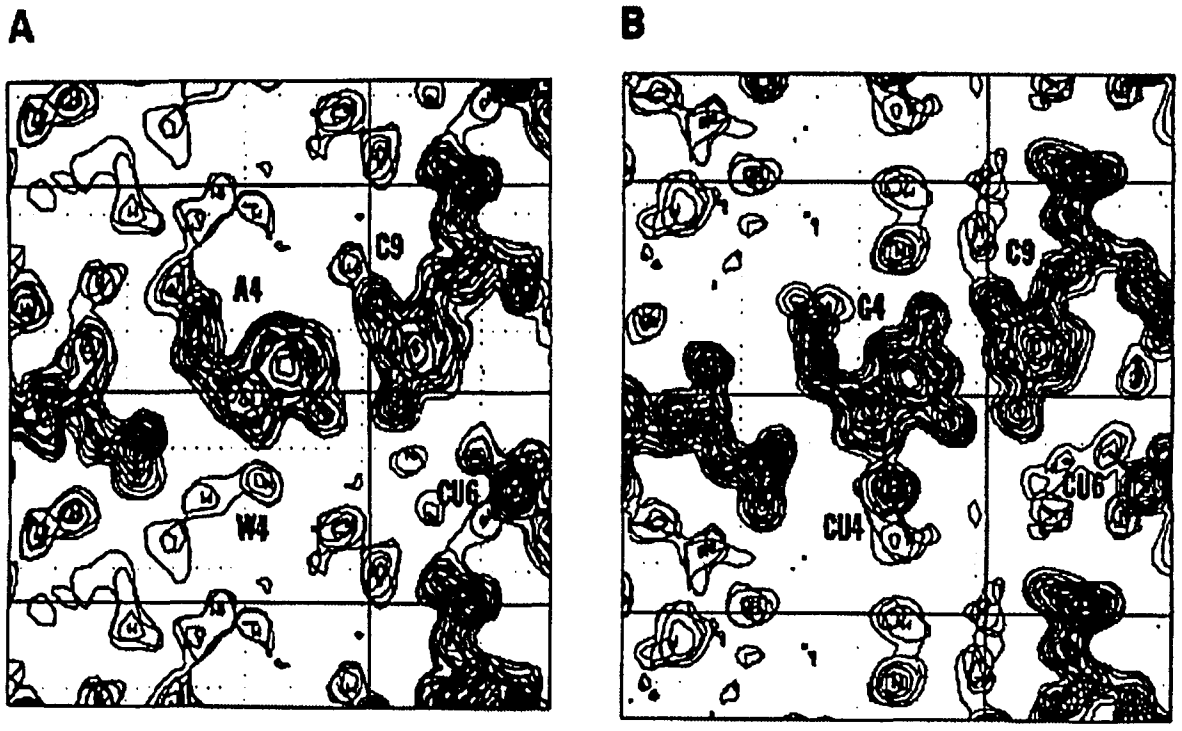


Fig. 5.9. Electron densities of A4-C9 and G4-C9

The center of the electron density is 2.34 Å from the N7 nitrogen of guanine residue G4. The coordination geometry at this site could not be assigned, but the presence of an axial water ligand suggests that this copper complex is either a trigonal pyramid or an octahedral complex.

In the present d(m5CGUAm5CG)$_2$ crystal structure, the potential copper binding site at adenine residue A4, as with the G4 site in the d(CGCGCG)$_2$ crystal, is located in an open solvent channel, again mimicking a free solution environment. We would therefore predict, by analogy, that the relative partial occupancy of copper(II) complexes at this site would be about 10%. In fact, there is a well defined set of contours observed in the electron density map located near the N7 nitrogen of adenine residue A4 (Fig. 5.9. A).

The integrated electron density of these contours, however, corresponds to a relative partial occupancy of only 1.6%, identical to the electron density of a fully occupied water molecule. The center of this density is positioned 2.75 Å from the N7 nitrogen, which is significantly longer than the copper to nitrogen distance observed for any copper complex in either DNA duplex structure (Cu to N7 distances are in the range from 1.92 Å to 2.57 Å; and compatible with a water site. Table 5.3).

Table 5.3. The geometry of copper complexes

		d(m5CGUAm5CG)$_2$		d(CGCGCG)$_2$	
Cu(II) site	ligand	average distance Å	geometry	average distance Å	geometry
Cu2 at G	N7	2.30	distorted octahedron	2.57	three ligands
	water	2.11		2.15	
Cu4 at G	N7		no copper	2.34	two ligands
	water			2.25	
W4 at A	N7	2.75			
Cu6 at G	N7	2.15	octahedron	2.17	octahedron
	water	2.04		2.20	
Cu8 at G	N7	2.38	trigonal bipyramid	2.31	trigonal bipyramid
	water	2.10		2.42	
Cu10 at G	N7			2.39	one ligand
Cu10 at A	N7	1.82	(see Cu12)		
Cu12 at G	N7	2.12	trigonal bipyramid	2.38	one ligand
	Cl	2.90			
	water	2.30			

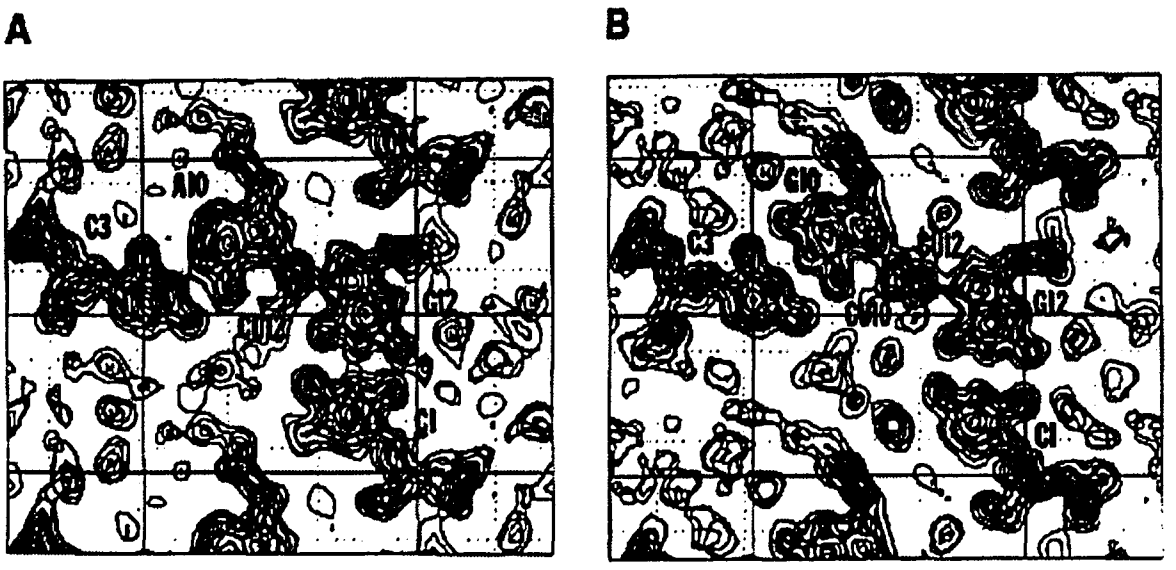


Fig. 5.10. Electron densities of A10-G12 and G10-G12

This electron density is also situated approximately ~37° out of the plane of the purine base (Fig. 5.10). A similar distortion in the N7 coordination bond was observed for the Cu2 copper complex in the d(CGCGCG)$_2$ crystal structure, however, this distortion was due to the presence of a nearby phosphate group. Adenine residue A4 is solvent accessible, and is not in close proximity to any negatively charged phosphate groups. Since there is no a priori reason for a copper complex at the A4 site to be distorted in either its bond lengths or geometry, or to exhibit a low relative partial occupancy, this set of electron density was assigned as a water molecule. The absence of copper binding at this open A4 site, therefore, suggests that copper(II) complexes would not bind to adenine bases of duplex DNA in solution.

5.2.5 Copper Binding to the Adenine A10 Residue

Adenine residue A10 of d(m5CGUAm5CG)$_2$ is adjacent to the G12 residue of a neighboring hexamer duplex. This is analogous to the G10 position at the interface between the major groove surfaces of two neighboring hexamer duplexes in the d(CGCGCG)$_2$ crystal structure. In this latter structure, the two guanine bases define mutually exclusive binding sites with a unique copper(II) complex bound to each purine base. The electron densities associated with the water ligands of the two complexes could not be assigned, therefore, it was not possible to unambiguously assign the geometries of these complexes.

In the present d(m5CGUAm5CG)$_2$ structure, the adenine A10 residue and guanine G12 residue of the adjacent hexamer appear to share a single copper(II) complex (Fig. 5.10A). It may be argued that the complex is bound only to the guanine base, and that the adenine is only fortuitously placed so that it appears to be bound. The structure at this region, however, argues against this simple model. If the copper complex were bound only to the guanine base, it would either be distorted or the adjacent hexamer containing the adenine A10 residue would necessarily be distorted. The copper(II) to N7 nitrogen distance of the adenine base is 1.92 Å, which is much less than the sum of the van der Waals radii of the two atoms. The high partial occupancy of this complex (41%; Table 5.2) indicates that any major distortion to the DNA would be observed in the crystal structure. In addition, the N7 to

N7 nitrogen distances of the adenine A10 and guanine G12 purine bases (3.88 Å) is 0.16 Å shorter than that of the native d(m5CGUAm5CG)$_2$ structure. Thus, the adjacent A10 and G12 bases act in concert to define a single shared copper(II) binding site.

The coordination geometry of this copper complex is defined by the position of the two purine bases. A water molecule is located in the plane of the two bases, and displaced 2.30 Å from the metal center. This suggests either a trigonal planar or trigonal bipyramidal coordination geometry for this complex. Two sets of strong electron densities were located, one 2.69 Å above and the second 3.10 Å below the metal center, implying that the complex is a trigonal bipyramid (Fig. 5.8). The bond distances, however, are much longer than any copper(II) to water ligand distances in the DNA duplex structures. This, along with the relatively short metal to purine bases distances (Table 5.3), indicates a very strong Jahn-Teller distortion of the coordination complex. A numerical integration of the electron densities of the axial ligands gave higher total electron counts than would be predicted for water molecules (data not shown). These two ligands were therefore assigned as putative chloride ions. A more thorough analysis of the integrated electron densities and hydrogen bonding interactions between the axial ligands and atoms of the two adjacent duplexes (data not shown), however, indicated that the ligand 3.10 Å below the metal center is most likely a water molecule. The structure at the A10 adenine base in the copper(II) chloride soaked d(m5CGUAm5CG)$_2$ hexamer, therefore,

appears to bind copper(II) complexes in an axially distorted trigonal bipyramid coordination geometry. An adjacent guanine base appears to be necessary to facilitate binding at this purine base.

5.3 The Structure of Ditercalinium-d(CGCG)$_2$ Complex

In the x-ray structure of ditercalinium bound to d(CGCG)$_2$, the ditercalinium molecule bis-intercalates at the two CpG steps of the DNA fragment (Figures 5.11 and 5.12). The DNA retains an underwound, right-handed double-helical conformation with the linker of ditercalinium in the major groove as first shown by NMR [177,212-214]. Although both the ditercalinium molecule and d(CGCG)$_2$ have the potential to adopt two-fold symmetry, the complex lacks symmetry in the x-ray structure. The conformation and DNA-interactions of one half of the ditercalinium molecule are different from the other half.

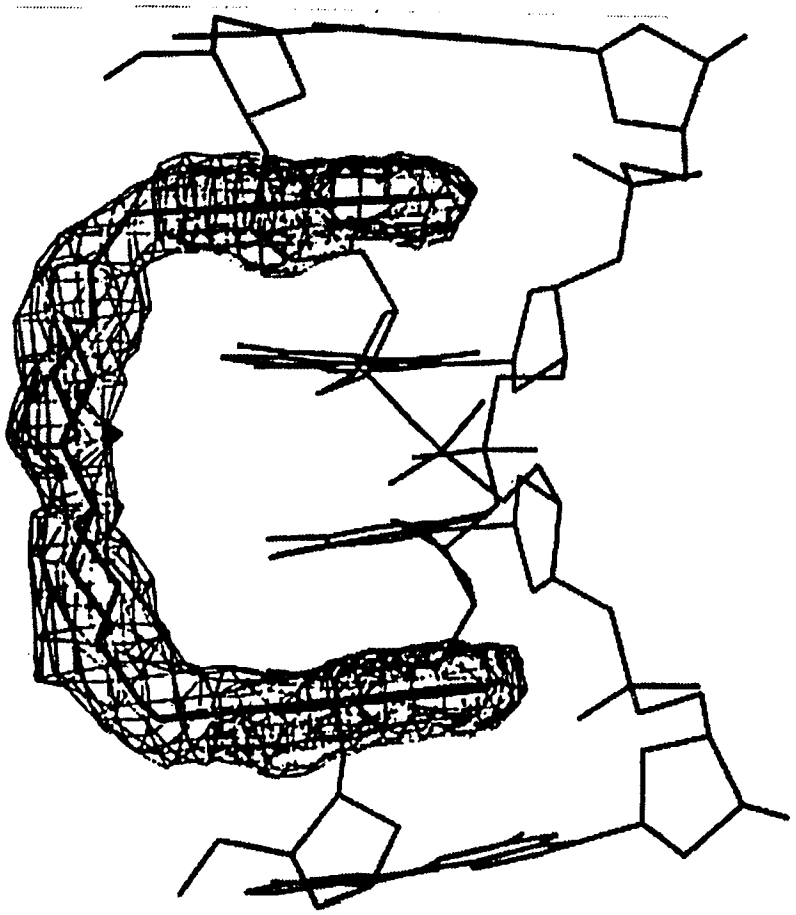


Fig. 5.11. The stick drawing of the d(CGCG)$_2$-ditercalinium complex.

The 2Fo-Fc electron densities surrounding the bis-intercalated drug includes.

The DNA in the complex with ditercalinium is kinked. This abrupt bend of around 15° in the helical axis towards the minor groove is clearly observable in Figure 5.11. The kink appears to arise from a combination of several factors. First, torsional constraints of the linker prevent co-planarity of the two chromophores in the complex. Second, the large surface areas of the intercalated chromophores induce each base pair to be co-planar with the adjacent chromophore. Third, on the major groove side of the complex, the linker maintains an axial distance between the two chromophores (10.4 Å) which slightly exceeds the expected rise of four base pairs of B-DNA (10.2 Å). If the linker connecting the two bis-intercalated chromophores were flexible, we would expect the axial rise of the chromophores to be the same as the axial rise of four base pairs of B-DNA. The combined effects of the torsional constraints, the excess length of the rigid linker and the large surface areas of the intercalated chromophores are shown schematically in Figure 5.12.

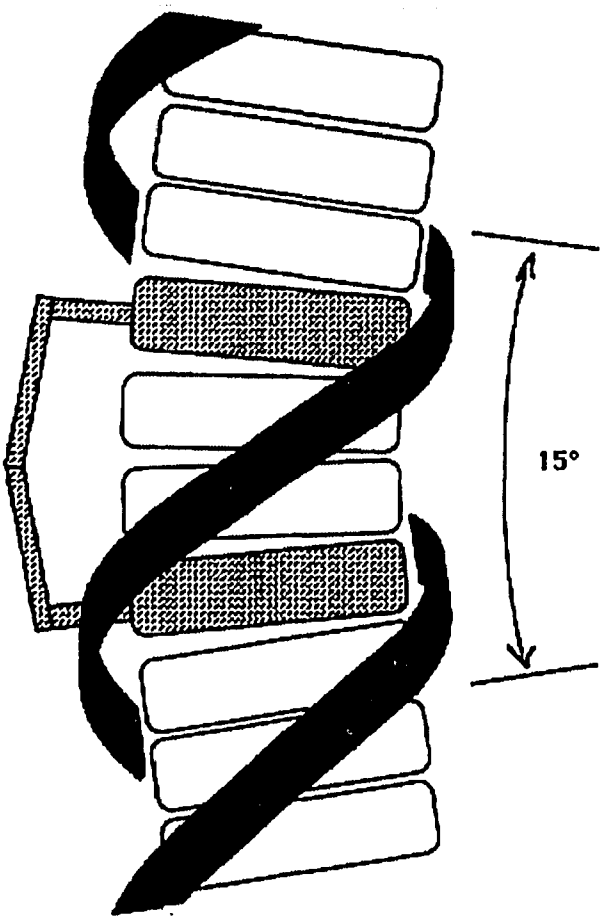


Fig. 5.12. Schematic diagram of DNA-ditercalinium complex. The kink of DNA induced by the bis-intercalation is shown.

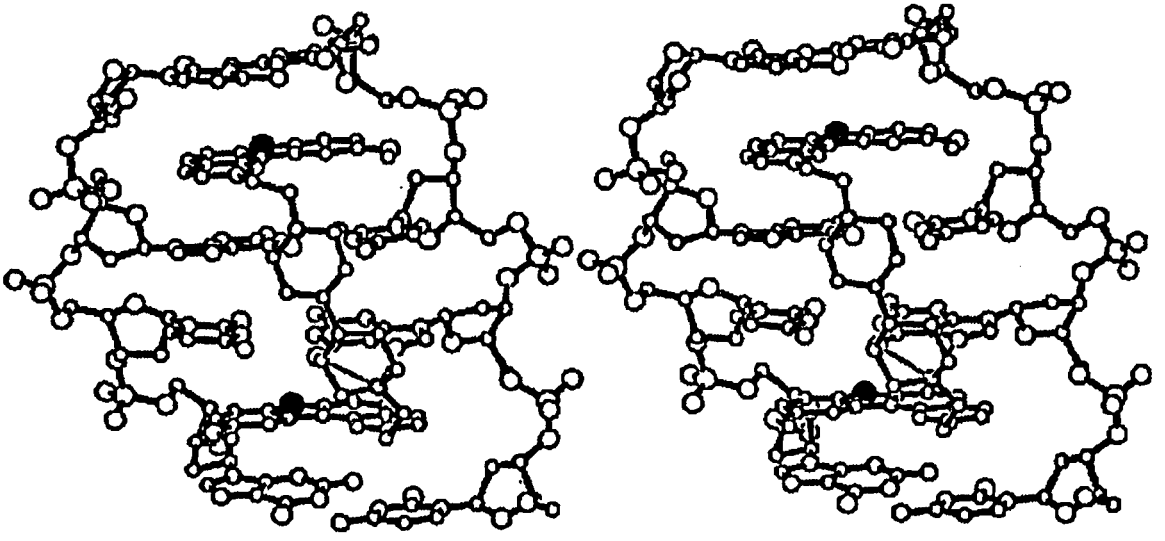


Fig. 5.13. Stereo view of DNA-ditercalinium complex

The DNA in the complex with ditercalinium is severely underwound. The helical twist is 22° at the first step, 30° at the second step, and 20° at the third step of the DNA tetramer. Compared to B-DNA which averages a helical twist of 36° per step, the total extent of DNA unwinding is 36° in the fragment of DNA complexes with ditercalinium. The helical unwinding is observable in the axial views of successive base pair-chromophore steps and base pair-base pair steps (Figure 5.14).

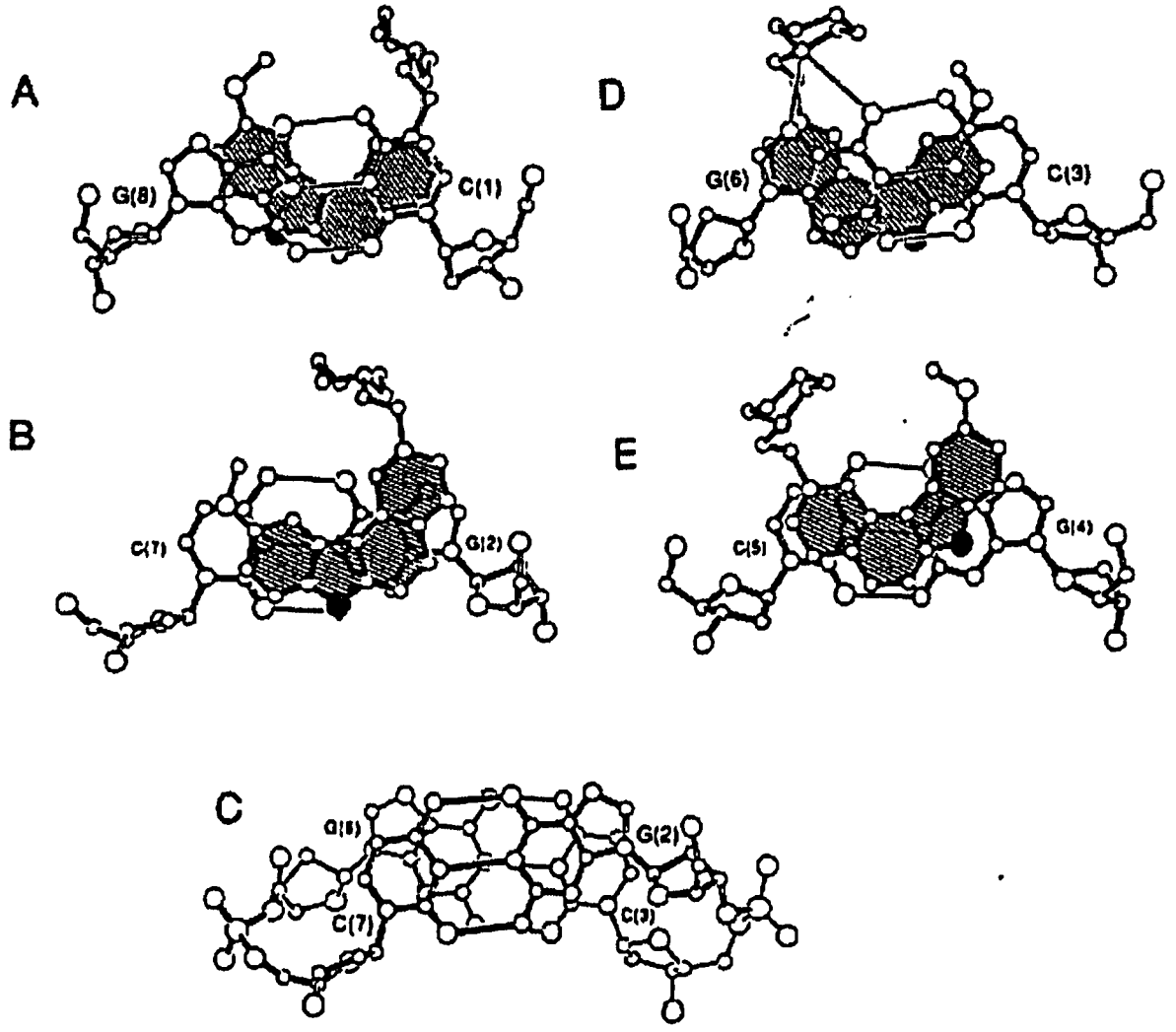


Fig. 5.14. Stacking of bases and ditercalinium

The kinking and unwinding of the DNA is accompanied by broadening of both grooves. For example in the ditercalinium complex the P4 to P8 distance (minus 5.8 Å, the van der Waals radii of the phosphate groups) is 10.5 Å while the width of the minor groove of B-DNA, as indicated by distances between analogous phosphates, is 5.7 Å. Similarly, in the

ditercalinium complex the P2 to P6 distance (minus the van der Waals radii of the phosphate groups) is 14.0 Å while the width of the major groove of B-DNA, as indicated by distances between analogous phosphates, is 11.7 Å.

The linker of ditercalinium is located in the major groove. Although certain conformational details of one of the pyridyl rings remains unclear in the electron density maps, it is clear that the conformation and DNA-interactions of one half of the linker are different from the other. The defined half of the linker (marked with primes in Figures 5.12 and 5.13) is located closer to the floor of the major groove than the "disordered" half. The proton of the piperidinic nitrogen (N 17') of the defined ring is directed towards the floor of the major groove (Figures 5.13 and 5.14D). This nitrogen forms what appears to be a bifurcated hydrogen bond, donating a proton simultaneously to both N7 (2.95 Å) and O6 (3.44 Å) of residue G(6). The hydrogen bond pulls the end of the linker towards the floor of the major groove and swivels the chromophore approximately around the DNA helical axis such that the Ol0'-methyl, which is located on the other end of this chromophore, protrudes out into the major groove (Figures 5.14D and 5.14E).

Although the conformation of the second pyridyl ring may be disordered (two possible rotamers of the piperidine ring fit the density), it is clear that the piperidinic nitrogen (N17) of this ring does not form

hydrogen bonds to the DNA. At the current stage of refinement, this nitrogen is 4.81 Å from the N7 and 3.94 Å from the O6 of residue G(2). In one conformation the proton on N17 is directed out, away from the floor of the major groove. In comparison with the hydrogen-bonded end of the linker, the absence of a hydrogen bond from N17 to the G(2)-C(7) base pair accompanies an increase in the distance between the linker and the base pair. This enables the O10-methyl, unlike the O10'-methyl (described above), to stack on an adjacent cytosine base (Figure 5.14B).

Ditercalinium bis-intercalates with the rigid linker in the major groove. The linker runs diagonally across the major groove such that the drug takes on the appearance of a backwards Z when viewed from the major groove (Figures 5.12 and 5.14A). The linker runs counter to the twist of the DNA helix. The long axis of each of the chromophores is oriented nearly parallel to those of the flanking base pairs. The indole N7-H and N7'-H groups of the chromophores of ditercalinium are directed out towards the minor groove and are juxtaposed near the N2 atoms of guanines when viewing down the helical axis (Figure 5.13). The N7 and N7' atoms are flush with the floor of the minor groove and do not protrude into the groove. The view into the minor groove reveals the close mimicry of DNA base pairs by the drug molecule (Figure 5.14B). The N7 and N7' atoms of ditercalinium appear to simulate N2 atoms of guanines.

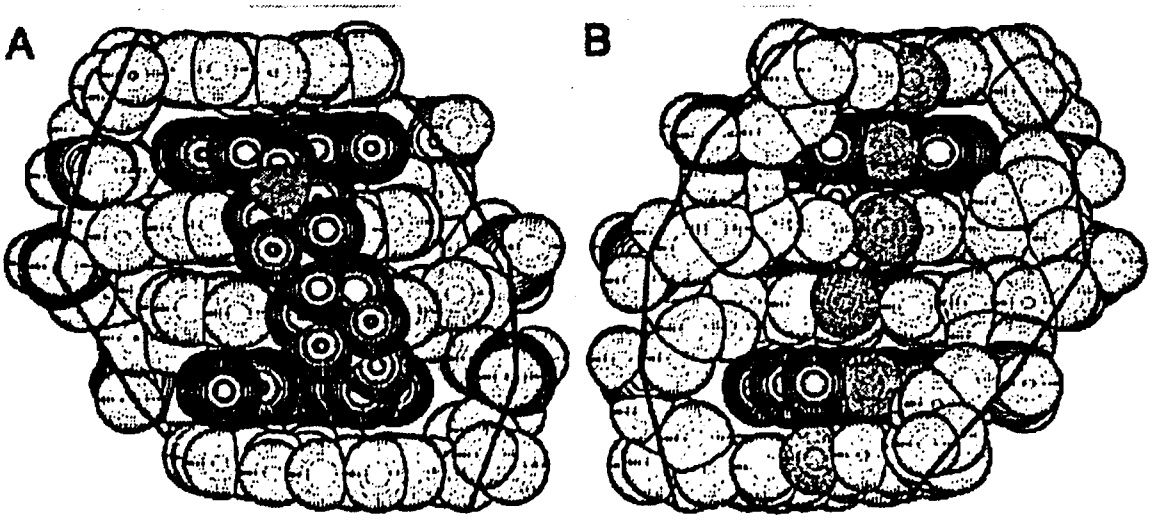


Fig. 5.15. Two views of DNA-ditercalinium complex form major and minor groove

In the first step viewing roughly along the helical axis, (Figure 5.14A), the cytosine base of C(1) is stacked directly on the pyridyl ring of the ditercalinium chromophore while the guanosine base of G(8) is partially stacked on the indole and benzyl rings. In the second step, the guanine base of residue G(2) is nearly completely stacked on the pyridyl and benzyl rings while the cytosine is only partially stacked on the O10-methyl group (Figure 5.14B). In the third step, the axial view (Figure 5.14C) indicates good stacking of the C(3)-G(6) base pair with the G(2)-C(7) base pair. In the fourth step, the guanine base of G(6) is nearly completely stacked on the chromophore while the complement cytosine base of C(3) is less well stacked (Figure 5.14D). The O10'-methyl of

ditercalinium is unstacked and protrudes out into the major groove. In the fifth step (Figure 5.14E), the cytosine base of C(5) is nearly completely stacked on the chromophore although not as well as C(1), the other terminal cytosine (Figure 5.14A). The guanine base of G(4) is less well stacked than the complementary cytosine. In the van der Waals representation, the gaps between the ditercalinium chromophores and the terminal C-G base pairs (Figure 5.15) indicate that the shapes of ditercalinium and the intercalation cavity within the DNA fragment are not closely matched.

5.4 Discussion

5.4.1 d(CGCGCG)$_2$ at Low Temperature

The overall structure of d(CGCGCG)$_2$ and its counterion/solvent environment in crystal show a strong correlation with the crystallographic packing.

The packing dependence of interaction between DNA and magnesium cations. If we do not consider the crystallographic packing, say for an isolated DNA oligomer, d(CGCGCG)$_2$, residue G6 and residue G12 are chemical equivalent. The nitrogen atoms, N7-G6 and N7-G12, would have the same potential to coordinate to metal cations. In the crystal, these two residues are not equivalent (Fig. 5.16). In the crystal only N7-G6 is found to coordinate to a magnesium (II)

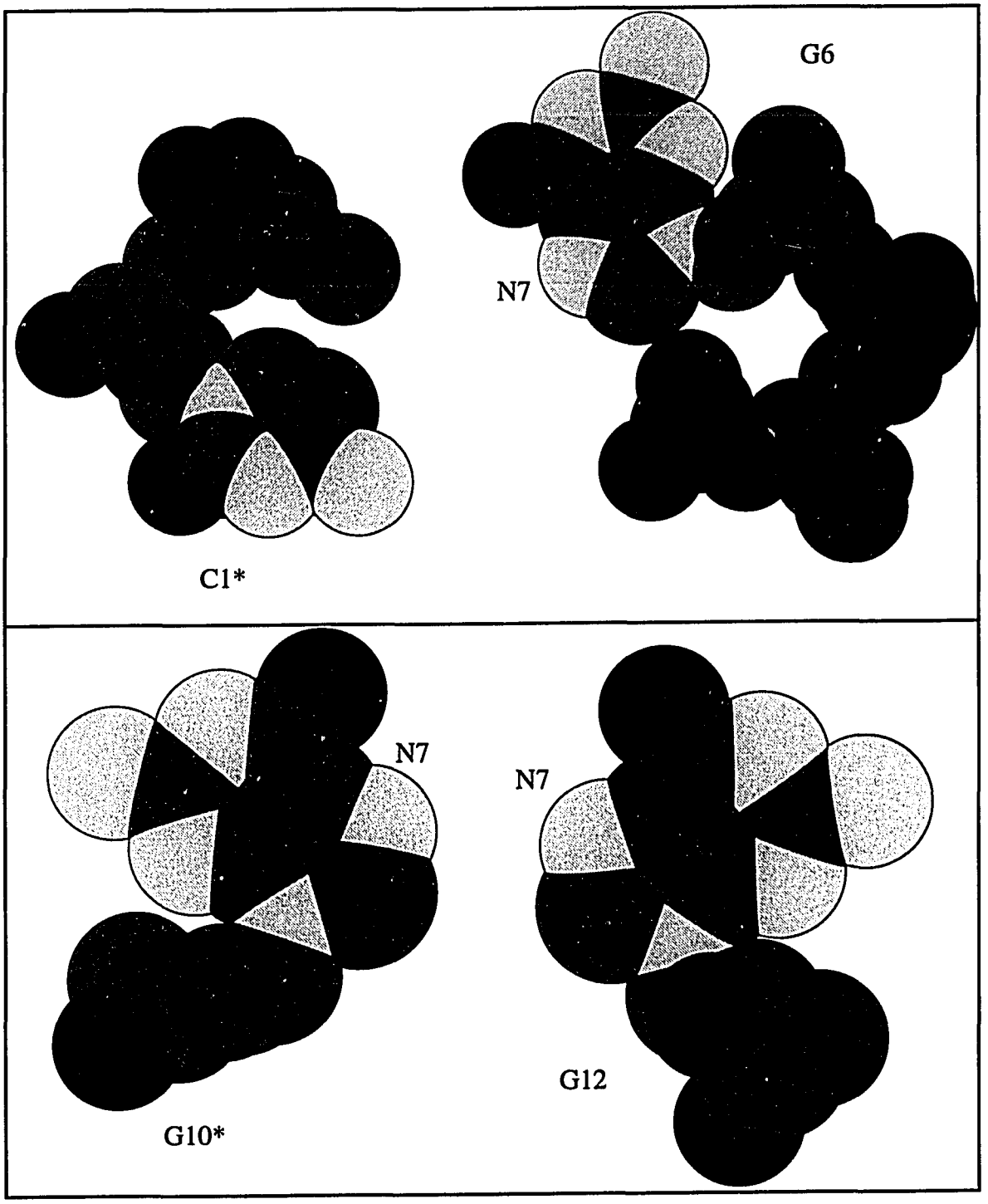


Fig. 5.16. The packing patern around N7-G6 and N7-G12.

cation. The two nitrogen atoms N7-G6 and N7-G12 are quit different, because of the effect of crystallographic symmetry. Figure 5.16 shows the different environments of these two nitrogen atoms in crytsal. The solvent chanel between the G6 and the symmetry related C1 is wide open. This allows a magnesium cation to coordinate to the N7 atom. The solvent chanel between the G12 and the symmetrey related G10 is narrow. This makes the binding of a magnesium cation to N7-G12 unfavorable.

The correlation of conformation and crystallographic packing. There are two dual conformation sites in this structure, phosphate 3 and phosphate 9. Phosphate 3 is the linker of residues G2 and C3; while phosphate 9 is the linker of residues G8 and C9. The percentage of each conformation in this structure have been determined quantitatively to be 75 percent ZI and 25 percent ZII at phosphate 3; 85 percent ZI and 15 percent ZII at phosphate 9. That makes the DNA duplex in a ZI/ZII hybrid conformation. From an isolated DNA duplex, it seems hard to explain why this would happen. Examining the entire packing pattern, these dual conformation sites are located near to the interfaces of two symmetry related DNA duplexes. In the interface regions between two consecutive base-pairs there are no phosphate groups. The absence of phosphate groups might produce a slightly different electrical and van de Waals field and different water structure than the regions in a continuous strand. The variation in field may

induce the ZI-ZII transition.

Gessner reported a relatively high B value on phosphate 3. [122] They interpreted this phenomenon as " disorder " of the two phosphate groups. The similar phenomenon had been reported by Wang and Quigley as alternative conformations, qualitatively. The observed hybrid conformation can be considered as the evidence of dynamics transition between ZI and ZII conformation in crystal at normal conditions. In this study, the temperature was set at 150° K. At such a low temperature the thermal motion of lattice is small (B=3-4 for atoms of DNA duplex. while B=8-12 at room temperature. That correspond to a RMS displacement of 0.19 to 0.23 Å and 0.32 to 0.39 Å, respectively). The results of this study suggests the structure freezes out in two different conformation with different probabilities.

Furthermore, one might propose that in the dynamic system of a Z-DNA crystal there is an equilibrium of different conformations. These equilibrium could be investigated by carefully designed experiments. The experimental results of crystal dynamics would be helpful in the study of molecular dynamics.

The correlation between back bone conformation and counterion structure. Previous studies have reported that MG-14 and MG-15 together with coordinated water molecules form a $Mg_2(H_2O)_9$-N7 complex. This complex includes the N7 of G-6 via a coordinate

bond. [122]

In this study, it is clear that MG-15 is of much lower occupancy than MG-14 is. Quantitative determination has shown that the occupancy of MG-15 is about 25%; and two water molecules, H_2O-31 and H_2O-32 (Fig. 5.6), that coordinate to MG-15, are of the same occupancies with MG-15. Some other water molecules around MG-15 exhibit either highly anisotropic motion or have multi-occupied sites. The features above strongly suggest a different interpretation of this region. (Figure 4.4)

The structure has been refined with these two interpretations separately. Both have similar final R values of 14.6 percent. In the following paragraphs, we will discuss the two possible interpretations of this region.

Model I. A full occupied MG-14 and partially occupied MG-15.
Model II. Both Mg-14 and Mg-15 are partially occupied.

These two models result same final R-factor of 14.6 percent. The associated water molecules and the water molecules in the nearest neighbor show similar structure.

The water molecules that coordinate to Mg-14 are OH2-24, OH2 -25, OH2-26, OH2-27, and OH2-30. The water molecules that

coordinate to Mg-15 are OH2-27, OH2-28, OH2-29, OH2-30, OH2-31, and OH2-32. Among them, OH2-27 and OH2-30 are shared by the two magnesium(II). With the Model II of two alternatively partially occupied magnesium(II), the final B factor of Mg-14 is unusually low (2.18). This B-factor is lower than the B-factors of the DNA molecule. The final different electron density map (2Fo-Fc) shows a additional strong positive density on the site of Mg-14. This feature suggests that the partially occupied Mg-14 may not be a correct interpretation. As in the final report we take the Model I, i.e. a fully occupied Mg-14 and a partially occupied Mg-15. See Fig. 5.5. D.

What is the exact structure of this region may be determined by neutron diffraction. In X-ray diffraction hydrogen atom can not be seen. Therefore the orientation of a water molecule can not be determined experimentally. With neutron diffraction, hydrogen atoms can be determined, then the orientation of water molecule can be determined. With determined water molecules the structure of this region and the entire molecule can be determined without ambiguity.

Correlation of DNA conformation and counterion /solvent structure. The occupancies of two conformations of DNA are 0.25 and 0.75 at phosphate 3. The occupancies of MG-15 and MG-17 are 0.25. This is hardly a coincidence. In order to examine the correlation, we plot an asymmetry unit and its symmetric related vicinity of radius 35 Å. In Figure 5.4, we plot the crystal packing

pattern slice by slice. The slices are divided along the Z-axis. Every slice include one base pair in the asymmetric unit and six other base pairs belonging to symmetry related molecules.

In the sliced packing pattern, it is clear that the five Mg(II)-water complexes can be divided into groups. In term of chemical structure, there are two groups of Mg(II)-water complexes.

group I: MG-14, which coordinates to N7-G6 with five coordinated water molecules.

group II: MG-13, MG-15, MG-16, MG-17: which coordinate to six water molecules and locate in the solvent channels between DNA molecules.

In term of crystal packing environment, the group II can be divided into three subgroup:

subgroup I: MG-13, which is located between G8-C5 and G4-C9 and the groves of symmetry related DNA molecules.

subgroup II: MG-15, which is located between base-pairs G2-C11 and G12-C1.

subgroup III: MG-16, MG-17, which are located between backbone and grove of symmetry related DNA molecules.

In term of occupancies, there are two groups of Mg(II)-water

complexes.

fully occupied: MG-14, MG-13, and MG-16.
partially occupied: MG-15 and MG-17.

It is interesting to notice that MG-15 and MG-17 are near to the dual-conformation phosphate groups P-3 and P-9* (* indicates the atom is on a symmetry related molecule) respectively. The unambiguous correlation between DNA backbone conformation and counterion structure suggest that - the alternative conformation of DNA backbone and the partial occupied Mg(II) complexes are not independent. In other word, they are two manifestation of one structure feature.

The solvent structure. The major feature of solvent structure also shows a multiple conformations. That is a logical consequence of the multiple conformations of the DNA and counterions. Many water molecules are determined to be dual positions with partial occupancies.

The other significant feature of solvent structure is high anisotropic thermal motion of some water molecules. This can be seen clearly on the electron density map. In the 2Fo-Fc map, a water molecule with isotropic thermal motion will appear as sphere shape density; while a water molecule with highly anisotropic thermal motion will appear as an elliptical ball shape density (football like). Figure 5.8 shows some obviously anisotropic water molecule. Water molecule that

is near the partially occupied Mg(II) complexes show multi-positional and anisotropic nature. Water molecules, OH2-31, OH2-44 and OH2-20 are typical examples.

In some cases, it is difficult to distinguish between positionally disordered and anisotropic water molecules. In principle, if a density can be expressed as three dimensional simple Gaussian functions, it can be interpreted as anisotropic water; otherwise a multi-positional partial occupied water are a better interpretation.

To obtain a better agreement between the structure and electron density, Holbrook has suggested the use of an anisotropic temperature factors for the DNA with isotropic temperature factors for solvent molecules. Using this approach he obtained a better R-value on Daunomycin-DNA complex. [120,211] In our work, it appears that the solvent molecules are more anisotropic than the DNA.

5.4.2 d(m5CGUAm5CG)$_2$-Cu(II) Complex

In this study, we present the Z-DNA crystal structure of d(m5CGUAm5CG)$_2$ soaked with copper(II) chloride at 1.3 Å resolution. The binding of copper(II) to the corresponding guanine bases in this crystal versus the Z-DNA crystal of the d(CGCGCG)$_2$ sequence were nearly identical in terms of the distribution of metal ions

along the hexamer duplex, and the coordination geometries of the metal complexes. This suggests that the microenvironment defined by the crystal packing of the two hexamer sequences are nearly identical, and that analogous potential copper (II) binding sites can be compared between these crystals. The differences in copper(II) binding to the purine bases that differ between the two structures (G4 and G10 for $d(CGCGCG)_2$ versus A4 and A10 for $d(m5CGUAm5CG)_2$) are thus intrinsic to the properties of the bases in the DNA duplex, and not an artifact of the microenvironment caused by crystal packing.

The adenine base at the A4 position sits in an open solvent channel in the crystal, and directly mimics the environment one would expect for an adenine base of a DNA duplex in aqueous solution. A copper(II) complex was not observed to be bound at this site, suggesting that adenine bases of duplex DNA in dilute solution would not be susceptible to copper(II) modification. This is in contrast to predictions from the crystal structures of copper(II) binding to isolated adenine nucleosides and derivatives [198], but is consistent with the observed sequence specificity for copper(II) induced oxidative DNA damage [200,201]. The sensitivity of thymine and guanine residues located 5' to guanine residues, and the resistance of adenine residues to oxidative damage can therefore, at least in part, be explained in terms of preferential copper binding to the guanine bases.

Adenine residue A10, on the other hand, does show significant

copper(II) binding. The binding of this complex, however, is facilitated by, or requires, concomitant binding of the metal to a guanine base. In this type of binding, where two purines act in concert to define the equatorial plane of the copper(II) complex, the metal to adenine N7 bond is as strong, if not stronger than that to the guanine N7 (as evidence by the 0.20 Å shorter Cu to adenine N7 bond length). Thus, adenine bases are susceptible to copper(II) modification, but may require additional, if not specific, DNA-DNA interactions to facilitate or induce binding. This suggests that, although copper(II) modification and the subsequent cleavage induced by the metal is sequence specific in dilute solution, this may not be the case where specific DNA-DNA interactions may occur. In the densely packed environment of the nuclear matrix in cells, where the DNA concentration and the environment may be more crystalline, we may expect that copper(II) binding, and the chemistry associated with this binding, to show much less discrimination between the purine bases.

5.4.3 Ditercalinium-d(CGCG)$_2$ Complex

Ditercalinium is a synthetic bis-intercalator with high affinity for DNA and strong anti-tumor properties. We described the three-dimensional x-ray structure of a DNA-ditercalinium complex, which is a substrate of a DNA repair system. X-ray structures of ditercalinium and other substrates of DNA repair systems will be increasingly

important for understanding mechanisms of repair and for further drug design. Ditercalinium activates DNA repair processes in both prokaryotic and eucaryotic cells. In prokaryotes, the DNA complex of ditercalinium is incorrectly recognized by the uvrABC repair system as a covalent lesion on DNA. A diverse array of covalent damage to DNA is excised by the uvrABC repair system (reviewed in reference 1). UvrABC excises pyrimidine dimmers, 6-4 photoadducts, and chemical adducts of psoralen, cis-platinum, mitomycin C, benz(a)pyrene and other reactive chemicals. In general, the damage excised by uvrABC causes distortions of DNA. The diversity of the covalent damage excised by uvrABC suggests that the DNA distortions are recognized by uvrABC rather than intrinsic structural motifs of the adducts. However, not all DNA distortions are recognized. UvrABC does not repair mispaired bases, O6-methylguanine or 3-methyladenine. To account for uvrABC recognition of such a diverse array of DNA lesions, it has been proposed (1) that uvrABC recognizes DNA kinks and binds to the face of the DNA which does not contain the adduct.

The DNA in the complex with ditercalinium is kinked (by 15°) and severely unwound (by 36°) with exceptionally wide major and minor grooves. If uvrABC does recognize kinked DNA and bind to the undamaged face, the repair system might recognize the DNA kink induced by ditercalinium and bind to the minor groove of the complex where the shape and hydrogen bonding pattern mimics that of normal (undamaged) DNA. The two chromophores of a ditercalinium molecule

substitute for two DNA base pairs with the N7 and N7' atoms of ditercalinium replacing N2 atoms of guanines. The loss of activity of certain 7-substituted analogies of ditercalinium could then result from loss of uvrABC binding within the minor groove where the complexes would differ significantly from the minor groove of DNA, both in shape and in arrangement of hydrogen bonding functionalities.

Upon carefully examination of the structure of the drug-DNA complex, one can find that there are three unusual stacking regions. The stacking between the center base pairs, i.e. the GpC step, and the stacking between the end base pairs and the drug show a little unparalleled between the involved chromophores. On the other hand, the symmetry related molecules show a perfect stacking.

A series of NMR studies of DNA-ditercalinium complexes [177,214] allow detailed comparison of those results with the x-ray structure described here. A detailed comparison of a three-dimensional x-ray structure to an analogous NMR structure is impossible as the NMR data was not used to quantify inter-atomic distances. In general, NMR spectroscopy and x-ray crystallography are in many ways complementary techniques for the analysis of biological conformation. NMR has the advantage of making measurements in solution and can detect kinetic events on a biologically interesting time scale. X-ray crystallography has the disadvantage of requiring a crystal, which is frequently impossible to grow and sometimes yields a conformation

which differs from that predominating in solution. However fine structural details, obtainable by x-ray but inaccessible by NMR may be necessary to understand macromolecular interactions.

Many of the characteristics of the x-ray structure of ditercalinium bound to $d(CGCG)_2$ were predicted from the ^1H and ^{31}P NMR data. Ditercalinium was designed with a rigid linker to prevent intramolecular interactions (self-stacking between chromophores) that would compete with DNA binding and decrease DNA affinity. The NMR experiments suggested that ditercalinium bis-intercalates in right-handed $d(CGCG)_2$ with the bases in the anticonformation and the linker of ditercalinium in the major groove. Additional important features such as the kink in the helical axis, the details of the stacking interactions, especially the differences between the two chromophores, and the hydrogen bonds between the linker and the major groove were observed crystallographically but were not apparent from the NMR experiments. It is interesting to note that, a molecular mechanics study [176] of ditercalinium bound to $d(CGCG)_2$ did anticipate many aspects of the major groove hydrogen bonding scheme observed in the x-ray structure.

The positively charged nitrogens of ditercalinium are located near the floors of the major groove of the DNA. Favorable charge interactions within the grooves of DNA appear to be an important factor in stability and conformation of complexes. The floors of the

grooves are regions of DNA with the greatest electronegative potential [215]. X-ray crystallographic studies suggest that mono- and polyamines uniformly bind to DNA so that the positive charges reside near the floors of the major or minor grooves. Positive charge is located near the floors of the grooves in the three dimensional crystal structures of DNA complexed with (a) intercalators which place positive charge in the major [216,217] or in the minor groove [167,209,218], (b) minor groove binding compounds [219], and (c) spermine in the major groove of A-DNA [220] and in the major groove of DNA-anthracycline complexes [218]. Thus placement of charged groups may be a primary consideration for designing agents that bind to DNA.

We would like to know which features of DNA-ditercalinium complexes activate DNA repair systems. The present structure leads to a number of suggestions but no firm answers. Could it be that the fixed 15° kink in the helical axis of the DNA-ditercalinium complex is similar to helical distortions resulting from covalent modifications? This is possible but additional factors must be considered including other helical distortions and, in the minor groove, the unusual mimicry of base pairs by the chromophores of ditercalinium. We would hope to obtain answers to these questions by solving the crystal structures of DNA complexes of ditercalinium derivatives that do not induce DNA repair responses.

CHAPTER VI CONCLUSIONS

This study consists of three different structures of DNA oligomers and complexes. The results of this study suggest following conclusion:

The structure of $d(CGCGCG)_2$ at liquid nitrogen temperature has two conformation.

The multiple conformation sites of $d(CGCGCG)_2$ located at phosphate 3 and phosphate 9. The two conformations have different occupancies. The major conformation, ZI, have 75 percent and 85 percent occupancies for P3 and P9 respectively. Correspondingly, the minor conformation have 25 and 15 percent occupancies for P3 and P9 respectively. The atoms involved in alternative conformations are phosphors and O1P, O2P in both sites.

Taking the result of this study, together with the room temperature structure [122], we concluded that the crystal structure of $d(CGCGCG)_2$ is of a dynamically equilibration system of ZI/ZII. At liquid nitrogen temperature, the system is "frozen" at two conformations, which can not be observed at room temperature. Instead unusual high thermal factors at the multiple conformation sites, P3 and P9, were observed. This suggests further studies on the dynamics of Z-

DNA crystal structure, under different temperature, could be interesting. Such result as building up a well defined system which can be used to verify molecular dynamic algorithm.

The counterion-DNA interaction is highly dependent on the crystallographic packing and DNA conformation.

The counterions, magnesium(II) in $d(CGCGCG)_2$ and copper(II) in $d(m5CGUAm5CG)_2$, interact to DNA differently at different sites. Magnesium(II) is a rather "hard" ion, which always form a six ligands octahedral structure. Copper(II) is a "soft" ion, which allows various coordinate number and different geometries.

Because of the particular packing in the crystal, atom N7-G6 and N7-G12 are crystallographically different. Atom N7-G6 faces a wider channel, which allows MG-14 coordinates to N7-G6. Atom N7-G12 faces a narrow channel; there is no magnesium(II) coordinates to it. Atom N7-A4 and N7-A10 are crystallographically different. N7-A4 face a wider channel. While N7-A10 faces a narrow channel. Thus CU-12 forms a bridge between N7-A10 and N7-G12 with a trigonal bipyramidal geometry. We would like to say that in this study copper(II) interacts to DNA differently governed by the combined effect of crystallographic packing and base specification.

Because of the multiple conformation feature of the crystal

structure of $d(CGCGCG)_2$, the magnesium(II) shows some conformation dependency. MG-17 has been found near the multiple conformation site, P3, with the same occupancy with the ZII conformation. We conclude the MG-17 only associates to ZII conformation. This makes a total of five different magnesium(II) in an asymmetry unit.

Some water molecules in crystal structure show anisotropic thermal motion.

By examination the electron density maps, we found that the thermal motion of some water molecule are highly anisotropic. For instance, OH2-18 and OH2-20, which coordinate to MG-13 and OH2-78, which locate at the solvent channel, are of anisotropic thermal motion. The anisotropic nature of thermal motion can be a reason for higher R value at high resolution. Thus an anisotropic refinement is necessary to obtain more accurate result. Furthermore, we recognize that water molecule are more anisotropic than the DNA molecule, in general. Holbrook has refined the structure of B-DNA with anisotropic DNA and isotropic water molecules. We would do further refinement with anisotropic B factors for water molecule and the DNA molecule, or anisotropic water and isotropic DNA.

It is possible to growth large Z-DNA crystals which are suitable to conduct neutron diffraction experiments. Neutron

diffraction can determine the positions of hydrogen atoms in a crystal structure. The key factor to perform a neutron diffraction experiment is large sized single crystals. In this study we have successfully grown Z-DNA crystals with the size of 1.5x0.9x0.9 mm^3, which have been taken to Brookhaven National Laboratory to run a trial experiment at the high flux neutron source. The results are quite satisfactory.

The structure variation in d(CGCG)$_2$-ditercalinium complex might be related to drug induced DNA repair. The DNA in the complex containing ditercalinium kinked 15° and severally unwounded by 36° with exceptionally wide major groove and minor groove due to the bis-intercalation of ditercalinium. The structure variation can be the key factor that can be recognized by uvrABC in prokaryotes, or by mitochondria DNA repair system in eucaryotes.

The structure predicted based on NMR results can be taken as the initial model to solve crystal structure. In this study a model of d(CGCG)$_2$-ditercalinium complex has been built up. This model has successfully served as initial model in the procedure of structure determination. This might be a particular case. However, one can always takes advantage of other structure information obtained from other techniques to help solve crystal structures, especially when there is a lack the isomorphous derivatives or a good initial model.

In summary, the results of this study on the three different DNA/DNA complexes have provided some clear information on the crystal structures and the explanation on their biological function. While the results have explained some questions which were not clear before, they have also raised more questions which remain to be answered. Further studies are expected.

Future goals

Our future studies include the following projects.

• **Neutron diffraction experiment of Z-DNA.** We have obtained relatively large and good quality single crystals of Z-DNA, which behavior well at the High Flux neutron beam. We expected that larger single crystal will be grown and neutron diffraction will be done in the near future. With the neutron diffraction data the hydrogen atoms will be determined experimentally. The complete structure information will be obtain by this study.

• **Anisotropic refinement of low temperature structure of Z-DNA.** Our results on Z-DNA structure indicate that the thermal motion of the structure are significantly anisotropic. With low

temperature technique, a sufficiently large data set can be collected. That makes the anisotropic refinement of Z-DNA structure possible.

• **The dynamic features study of Z-DNA crystal.** The structural studies of Z-DNA at different temperatures will allow a detail investigation on the dynamic features of crystal structure. Such studies will be very useful both in structure analysis and theoretical study.

• **The studies on modification of ditercalinium.** This study exposed the structure feature of DNA-ditercalinium complex. It seems that the linker of the drug is little longer than the space of the bis-intercalation site of the DNA It should be interesting to examine the interaction of DNA and a modified drug with shorter linker between the two chromophores.

Acknowledgment

This work was supported by the NIH research grants 1R01GM41359 and RR-03037. We especially thank Dr. N. R. Kallenbach, Dr. B. Schoengbern, Dr. L. Day, Dr. M Diem, Dr. M. Mezei, Dr. D. Goss, Dr. M. Lu, Dr. Q. Guo

Supplementary material available

Complete coordinates of oligonucleotide d(CGCGCG)$_2$-Mg5 with water are available.

References

1. Miescher, F. Med. Chem. Unt., **1871**, 4, 441.
2. Levene, P.A. and Bass, L.W. Nucleic Acids (Chemical Catalogue Co., New York, 1931).
3. Avery, O.T., MacLeod, C.M. and MacCarthy, M. J. Exp. Med., **1944**, 79, 137.
4. Hershey, A.D. and Chase, M. J. Gen. Physiol, **1952**, 36, 39.
5. Chargaff, E. Experientia, **1950**, 6, 201.
6. Watson, J.D. and Crick, F.H. Nature, **1953**, 171, 737.
7. Kornberg, A. DNA Replication (Freeman, San Francisco, 1980).
8. Friedberg, E.C. DNA Repair (Freeman, San Francisco, 1985).
9. Stryer, L. Biochemistry (W. H. Freeman and Co., New York, 1988).
10. Suck, D. and Oefner, C. Nature, **1986**, 321, 620.
11. McClarin, J.A., Frederic, C.A., Wang, B.-C., *et al.* Science, **1986**, 234, 1526.
12. Prescott, D.M. and Kuemple, P.L. Proc. Nat. Acad. Sci. USA, **1972**, 69, 2842.
13. Bird, R.E., Lourarn, J., Martuscelli, J. and Caro, L. J. Mol. Biol., **1972**, 70, 549.
14. Mathews, C.K. and van Holde, K.E. Biochemistry 1-883 (The Benjamin/ Cummings Publishing Comany, Inc., New York, 1990).
15. Holliday, R. Genet. Res., **1964**, 85, 283.
16. West, S.C., Cassuto, E. and Howard-Flanders, P. Proc. Natl. Acad.

Sci. USA, **1981**, 78, 2100.

17. Landy, A. Ann. Rev. Biochem., **1989**, 58, 913.
18. Branden, C. and Tooze, J. Introduction to Protein Structure 1-87 (Garland Publishiong, Inc., New York, Londen, 1991).
19. Gale, E.F., Cundiffe, E., Reynolds, P.E., Richmond, M.H. and Waring, M.J. Molecular Basis of Antibiotic Action 1-258 (John Wiley, London, 1981).
20. Corcoran, J.W. and Hahn, F.E. Antibiotics (Spring-Verlag, New York, 1975).
21. Gao, Q.I., Williams, L.D., Egli, M., *et al.* Proc. Natl. Acad. Sci. U. S. A., **1991**, 88, 2422.
22. Brooks, C.L.I., Karplus, M. and Pettitt, B.M. Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics 1-259 (John Wiley & Son, New York, 1988).
23. McCammon, A. and Harvey, S. Dynamics of Proteins and Nucleic Acids (Cambridge University Press, New York, 1987).
24. Brunger, A.T., Clore, G.M., Gronenborn, A.M. and Karplus, M. P.N.A.S., **1986**, 83, 3801.
25. Brunger, A.T., Campbell, R.L., Clore, G.M., *et al.* Science, **1987**, 235, 1049.
26. Karplus, M. and Pestko, G.A. Nature, **1990**, 347, 631.
27. Brunger, A.T. X-PLOR Manual, Version 2.1 (Yale University, 1990).
28. Altona, C. 1981. In Methods in Structural Molecular Biology, Davies, D.B., Saenger, W. & Danyluk, S.S. eds., Plenum Press.

London.

29. Astbury, W.T. and Bell, F.O. Nature, **1938**, 141, 747.

30. Furberg, S. Acta Chemica Scandinavica, **1952**, 6, 634.

31. Wilkins, M.H.F., Gosling, R.G. and Seeds, W.E. Nature, **1951**, 167, 759.

32. Franklin, R.E. and Gosling, R.G. Nature, **1953**, 171, 740.

33. Pauling, L. and Corey, R.B. Proc. Natl. Acad. Sci. USA, **1953**, 39, 84.

34. Leslie, A.G.W., Arnott, S., Chandrasekaran, R. and Ratliff., R.L. J. Mol. Biol., **1980**, 143, 49.

35. Fuller, W., Wilkins, M.H.F., Wilson, H.R. and Hamilton., L.D. J. Mol. Biol., **1965**, 12, 60.

36. Langridge, R., Marvin, D.A., Seeds, W.E., *et al.* J. Mol. Biol., **1960**, 2, 38.

37. Arnott, S. and Hukins, D.W.L. Biochem. Biophys. Res. Commun., **1972**, 47, 1504.

38. Marvin, D.A., Spencer, M., Wilkins, M.H.F. and Hamilton, L.D. J. Mol. Biol., **1961**, 3, 547.

39. Arnott, S. and Selsing, E. J. Mol. Biol., **1974**, 88, 509.

40. Mokul'skii, M.A., Kapitanova, K.A. and Mokul'skaya. J. Mol. Biol., **1972**, 6, 714.

41. Saenger, W. Principles of Nucleic Acid Structure (Springer-Verlag, Berlin, 1984).

42. Sussman, J.L., Seeman, N.A., Kim, S.-H. and Berman, H.M. J. Mol. Biol., **1972**, 66, 403.

43. Rosenberg, J.M., Seeman, N.C., Day, R.O. and Rich, A. J. Mol. Biol., **1976**, 104, 145.

44. Hingerty, B., Subramanian, E., Stellman, S.D., *et al.* Acta Cryst., **1976**, B32, 2998.

45. Seeman, N.C., Rosenberg, J.M., Suddath, F.L., Kim, J.J.P. and Rich, A. J. Mol. Biol., **1976**, 104, 109.

46. Suck, D., Manor, P.C. and Saenger, W. Acta Cryst., **1976**, B32, 1727.

47. Camerman, N., Fawcell, J.K. and Camerman, A. J. Mol. Biol., **1976**, 107, 601.

48. Viswamitra, M.A., Kennard, O., Jones, P.G., *et al.* Nature, **1978**, 273, 687.

49. Kim, S.-H. Prog. Nucl. Acid Res. Mol. Biol., **1976**, 17, 181.

50. Wang, A.H.J., Quigley, G.J., Kolpak, F.J., *et al.* Nature, **1979**, 282, 680.

51. Wing, R., Drew, H., Takano, T., *et al.* Nature, **1980**, 287, 755.

52. Shakked, Z., Rabinovich, D., Cruse, W.B.T., *et al.* Proc R Soc Lond B Biol Sci, **1981**, 213, 479.

53. Wang, A.H.J., Fujii, S., Boom, J.H.v. and Rich, A. P.N.A.S., **1982**, 79, 3968.

54. Conner, B.N., Takano, T., Tanaka, S., Itakura, K. and Dickerson, R.E. Nature, **1982**, 295, 294.

55. Wang, A.H.J., Hakoshima, T., Marel, G.v.d., Boom, J.H.v. and Rich, A. Cell, **1984**, 37, 321.

56. McCall, M., Brown, T. and Kennard, O. J. Mol. Biol., **1985**, 183,

385.

57. Fujii, S., Wang, A.H.-J., Marel, G.v.d., Boom, J.H.v. and Rich, A. Nuc. Acids Res., **1982**, 10, 7879.

58. Westhof, E., Prange, T., Chevrier, J.B. and Moras, D. Biochimie, **1985**, 67, 811.

59. Chevrier, B., Dock, A.C., Hartmann, J.B., *et al.* J. Mol. Biol., **1986**, 188, 707.

60. Ho, P.S., Kagawa, T.F., Tseng, K., Schroth, G.P. and Zhou, G. Science, **1991**, 254, 1003.

61. Ho, P.S., Frederick, C.A., Quigley, G.J., *et al.* EMBO Journal, **1985**, 4, 3617.

62. Kneale, G., Brown, T., Kennard, O. and Rabinovich, D. J. Mol. Biol., **1985**, 186, 805.

63. Brown, T., Kennard, O., Kneale, G. and Rabinovich, D. Nature, **1985**, 315, 604.

64. Brown, T., Hunter, W.N., Kneale, G. and Kennard, O. Proc. Natl. Acad. Sci. USA, **1986**, 83, 2402.

65. Hunter, W.N., Brown, T. and Kennard, O. J. Biomol. Struct. Dyn., **1986**, 4, 173.

66. Hunter, W.N., Kneale, G., Brown, T., Rabinovich, D. and Kennard, O. J. Mol. Biol., **1986**, 190, 605.

67. Prive, G.G., Heinemann, U., Chandrasegaran, S., *et al.* Science, **1987**, 238, 498.

68. Prive, G.G., Heinemann, U., Chandrasegaran, S., *et al.* 1988. In Structure & Expression, Sarma, R.H. & Sarma, M.H. eds., Adenine

Press. New York.

69. Wang, A.H.-J., Quigley, G.J. and Rich, A. Nuclic Acid Res., **1979**, 6, 3879.

70. Patel, D.J., Chanuel, L.L. and Pohl, F.M. Proc. Natl. Acad. Sci. USA, **1979**, 76, 2508.

71. Pohl, F.M. and Jovin, T.M. J. Mol. Biol., **1972**, 67, 375.

72. Aoki, K. Bull. Chem. Soc. Japan, **1975**, 48, 1260.

73. Gessner, R.V., Frederick, C.A., Quigley, G.J., Rich, A. and Wang, A.H.J. J. Biol. Chem., **1989**, 264, 7921.

74. Shakked, Z., Rabinovich, D., Kennard, O., *et al.* J. Mol. Biol., **1983**, 166, 183.

75. Kennard, O., Cruse, W.B.T., Nachman, J., *et al.* J. Biomol. Struct. Dyn., **1986**, 3, 623.

76. Rabinovich, D., Haran, T., Eisenstein, M. and Shakked, Z. J. Mol. Biol., **1988**, 200, 151.

77. Frederick, C.A., Quigley, G.J., Teng, M.-K., *et al. Molecular Structure of an A-DNA Decamer: d(ACCGGCCGGT)* 1-in Press (1988).

78. Dickerson, R.E., Drew, H.R., Conner, B.N., Kopka, M.L. and Pjura, P.E. Cold Spring Harbor Symp. Quant. Biol., **1982**, 47, 13.

79. Conner, B.N., Yoon, C., Dickerson, J.L. and Dickerson, R.E. J. Mol. Biol., **1984**, 174, 663.

80. Haran, T.E., Shakked, Z., Wang, A.H.J. and Rich, A. J. Biomol. Struct. Dyn., **1987**, 5, 199.

81. Heinemann, U., Lauble, H., Frank, R. and Bloecker, H. Biol Chem

Hoppe-Seyler, **1987**, 368, 1055.

82. Frederick, C.A., Saal, D., Marel, G.A.v.d., *et al.* Biopolymers, **1987**, 26, S145.

83. Jain, S., Zon, G. and Sundaralingam, M. J. Mol. Biol., **1987**, 197, 141.

84. McCall, M., Brown, T., Hunter, W.N. and Kennard, O. Nature, **1986**, 322, 661.

85. Dock-Bregeon, A.C., Chevrier, B., Podijarny, A., *et al.* Nature, **1988**, 335, 375.

86. Zimmerman, S.B. and Pheiffer, B.H. J. Mol. Biol., **1979**, 135, 1023.

87. Drew, H.R. and Dickerson, R.E. J. Mol. Biol., **1981**, 151, 535.

88. Dickerson, R.E. and Drew, H.R. J. Mol. Biol., **1981**, 149, 761.

89. Drew, H.R., Wing, R.M., Takano, T., *et al.* P.N.A.S., **1981**, 78, 2179.

90. Drew, H.R., Samson, S. and Dickerson, R.E. P.N.A.S., **1982**, 79, 4040.

91. Drew, H.R. and Dickerson, R.E. EMBO. J., **1982**, 1, 663.

92. Kopka, M.I., Fratini, A.V., Drew, H.R. and Dickerson, R.E. J. Mol. Biol., **1982**, 163, 129.

93. Dickerson, R.E., Kopka, M.L. and Pjura, P. P.N.A.S., **1983**, 80, 7099.

94. Dickerson, R.E. J. Mol. Biol., **1983**, 166, 419.

95. Kopka, M.L., Fratini, A.V., Drew, H.R. and Dickerson, R.E. J. Mol. Biol., **1983**, 163, 129.

96. Kopka, M.L., Pjura, P., Yoon, C., Goodsell, D. and Dickerson, R.E. 1985. in Structure and Motion: Membranes, Nucleic Acids and Proteins; Proceedings of the International Symposium on Structure and Dynamics of Membranes, Nucleic Acids and Proteins, Rome, Italy, Apr. 23-27, 1984, Clementi, E.E.A. eds., Adenine Press. Guilderland, N.Y., USA.

97. Goodsell, D.S. and Dickerson, R.E. Proc. Natl. Acad. Sci. USA, **1988**, 85, 6332.

98. Pjura, P.E., Grzeskowiak, K. and Dickerson, R.E. J. Mol. Biol., **1987**, 197, 257.

99. Teng, M.K., Usman, N., Frederick, C.A. and Wang, A.H.-J. Nucleic Acids Res, **1988**, 16, 2671.

100. Kennard, O. J. Biomol. Struct. Dyn., **1985**, 3, 205.

101. Hunter, W.N., Brown, T., Kneale, G., *et al.* J Biol Chem, **1987**, 262, 9962.

102. Corfield, W., Hunter, W.N., Brown, T., Robinson, R.E. and Kennard, O. Nuclic Acid Res., **1987**, 15, 7935.

103. Nelson, H.C.M., Finch, J.T., Luisi, B.F. and Klug, A. Nature, **1987**, 330, 221.

104. Coll, M., Solans, X., Altaba, M.F. and Subirana, J.A. J. Biomol. Struct. Dyn., **1987**, 4, 797.

105. Coll, M., Aymami, J., Marel, G.A.V.D., *et al.* Biochemistry, **1989**, 28, 310.

106. Yoon, C., Priv, G.G., Goodsell, D.S. and Dickerson, R.E. Proc. Natl. Acad. Sci. USA, **1988**, 85, 6332.

107. DiGabriele, A.D., Sanderson, M.R. and Steitz, T.A. Proc. Natl. Acad. Sci. USA, **1989**, 86, 1816.

108. Saper, M.A., Eldar, H., Mizuuchi, K., *et al.* J. Mol Biol., **1986**, 188, 111.

109. Joshua-Tor, L., Rabinovich, D., Frolow, F., Appella, E. and Sussman, J.L. Nature, **1988**, 334, 82.

110. Miller, M., Wlodawer, A., Appella, E. and Sussman, J.L. J. Mol. Biol., **1987**, 195, 967.

111. Miller, M., Wlodawer, A., Appella, E. and Sussman, J.L. Nature, **1988**, 334, 1931.

112. Cruse, W.B.T., Salisbury, S.A., Brown, T., *et al.* J. Mol. Biol., **1986**, 192, 891.

113. Quigley, G.J., Wang, A.H.J., Ughetto, G., *et al.* P.N.A.S., **1980**, 77, 7204.

114. Quigley, G.J., Ughetto, G., Marel, G.A.v.d., *et al.* Science, **1986**, 232, 1255.

115. Ughetto, G., Wang, A.H.-J., Quigley, G.J., *et al.* Nucleic Acids Research, **1985**, 13, 2305.

116. Jolly, D. and Eisenberg, H. Biopolymers, **1976**, 15, 61.

117. Frontali, C., Dore, E., Ferranto, A., *et al.* Biopolymers, **1979**, 18, 1353.

118. Schellman, J.A. Biopolymers, **1974**, 13, 217.

119. Miller, K.J., MacRea, J. and Pycoir, J.F. Biopolymers, **1980**, 19, 2067.

120. Holbrook, S.R., Wang, A.H.-J., Rich, A. and Kim, S.H. J. Mol.

Biol., **1986**, 187, 429.

121. Gessner, R.V., Quigley, G.J., Wang, A.H.J., *et al.* Biochemistry, **1985**, 24, 237.

122. Gessner, R. *High Resolution X-ray Diffraction Studies of Z-DNA* (Freien Universit t Berlin, 1989).

123. Ho, P.S., Frederick, C.A., Saal, D., Wang, A.H.J. and Rich, A. J. Biomol. Struct. Dyn., **1987**, 4, 521.

124. Fujii, S., Wang, A.H.-J., Quigley, G.J., *et al.* Biopolymers, **1985**, 24, 243.

125. Westhof, E., Dumas, P. and Moras, D. J. Mol. Biol., **1985**, 184, 119.

126. Wang, A.H.J., Gessner, R.V., Marel, G.A.v.d., Boom, J.H.v. and Rich, A. P.N.A.S., **1985**, 82, 3611.

127. Coll, M., Wang, A.H.J., Marel, G.A.v.d., Boom, J.H.v. and Rich, A. J. Biomol. Struct. Dyn., **1986**, 4, 157.

128. Drew, H., Takano, T., Tanaka, S., Itakura, K. and Dickerson, R.E. Nature, **1980**, 286, 567.

129. Crawford, J.L., Kolpak, F.J., Wang, A.H.J., *et al.* P.N.A.S., **1980**, 77, 4016.

130. Brennan, R.G., Westhof, E. and Sundaralingam, M. J. Biomol. Struct. Dyn., **1986**, 3, 649.

131. Chattopadhyaya, R., Grzeskowiak, K. and Dickerson, R.E. J. Mol. Biol., **1990**, 211, 189.

132. Zhou, G. and Ho, P.S. Biochem., **1990**, 29, 7229.

133. Kornberg, A. 1980. eds., DNA Replication, Freeman. San

Francisco.

134. Eichhorn, G.L., Butzow, J.J. and Shin, Y.A. J. Biosci., **1985**, 8, 527.

135. Sissoeff, I., Grisvard, J. and Guille, E. Prog. Biophys. Mol. Biol., **1976**, 31, 165.

136. Pezzano, H. and Podo, F. Chem. Rev., **1980**, 80, 365.

137. Mildvan, A.S. and Loeb, L.A. CRC Crit. Rev. Biochem., **1979**, 6, 219.

138. Eichhorn, G.L. and Shin, Y.A. J. Amer. Chem. Soc., **1968**, 90, 7323.

139. Quigley, G.J. 1982. in Molecular Structure and Biological Activity, North Holland. New York.

140. Adamika, D.A. and Saenger, W. Acta Crystallogr., **1980**, B36, 2585.

141. Clark, G.R. and Orbell, J.D. J. Chem.Soc. Chem. Commun., **1975**, 697.

142. Guay, F. and Beauchamp, A.L. J. Amer. Chem. Soc., **1979**, 101, 6260.

143. Wing, R.M., Pjura, P., Drew, H.R. and Dickerson, R.E. EMBO J., **1984**, 3, 1201.

144. Arnott, S., Bond, P.J. and Chandrasekaran, R. Nature, **1980**, 287, 561.

145. Rosenberg, B., Camp, L.V., Trosko, J.E. and Mansour, V.H. Nature, **1969**, 222, 385.

146. Harrison, R.C. and McAuliffe, C.A. Inorg. Respect. Biol.

Med.,**1978**, 1, 261.

147. Pestayko, A.W., Crooke, S.T. and Carter, S.K. Cisplatin: Current Status and new Developments. (Academic Press, New York, 1980).

148. Millard, M.M., Macquet, J.P. and Theophanides, T. J. Biochim. Biophys. Acta, **1975**, 402, 166.

149. Marcellis, A.T.M. and van Kralingen, C.G. J. Inorg. Biochem. **1980**, 13, 213.

150. Aoki, K. J. Cryst. Soc. Japan, **1981**, 23, 309.

151. Hodgson, D.J. Prog. Inorg. Chem., **1977**, 23, 211.

152. Spiro, T.G. Metal Ions in Biology. Nuclic Acid-metal Ion Interaction, 1980, John Wiley & sons, New York.

153. Gellert, R.W. and Bau, R. 1979. In Metal ions in biological systems, Sigel, H. eds., Dekker. New York.

154. Swaminathan, V. and Sundaralingam, M. Crit Rev Biochem, **1979**, 6, 245.

155. Ahrland, S., Chatt, J. and Davies, N.K. Quart. Rev. Chem.Soc, **1958**, 12, 265.

156. Waring, M.J. Annu. Rev. biochem., **1981**, 50, 159.

157. Berman, H.M. and Young, P.R. Annu. Rev. Biophys. Bioegn., **1981**, 10, 87.

158. Sancar, A. and Sancar, G.B. Annu. Rev. Biochem., **1988**, 57, 29.

159. Lerman, L.S. J. Mol. Biol., **1961**, 3, 18.

160. Bond, P.J., Langridge, R., Jennette, K.W. and Lippard, S.J. Proc. Nat. Acad. Sci. USA, **1975**, 72, 4825.

161. Fuller, W. and Waring, M. Ber. Bunsenges. Phys. Chem., **1964**,

68, 805.

162. Neville, D.M. and Davies, D.R. J. Mol. Biol., **1966**, 17,

163. Carins, J. Cold spring Harbor Symp. Quant. Biol., **1962**, 27, 311.

164. Waring, M.J. and Wakelin, L.P.G. Nature, **1974**, 252,

165. Wang, A.H.-J., Ughetto, G., Quigley, G.J. and Rich, A. Biochemistry, **1987**, 26, 1152.

166. Holbrook, S.R. and Kim, S.H. Int J Biol Macromol, **1979**, 1, 233.

167. Williams, L.D., Egli, M., Ughetto, G., *et al.* J. Mol. Biol., **1990c**, 215, 313.

168. Wang, A.H.J., Ughetto, G., Quigley, G.J., *et al.* Science, **1984**, 225, 1115.

169. Wang, A.H.-J., Ughetto, G., Quigley, G.J. and Rich, A. J. Biomol. Struct. Dyn., **1986**, 4, 319.

170. Kennard, O. and Hunter, W.N. Angew. Chem. Int. Ed. Engl., **1991**, 30, 1254.

171. Moore, M.H. and Hunter, W.N. J. Mol. Biol., **1989**, 206, 693.

172. Roques, B.P., Le Pecq, J.B., Pelaprat, D. and Le Guen, I. Franch Patent 72.23801, **1978**,

173. Lambert, B., Roques, B.P. and Le Pecq, J.B. Nucleic Acid Res., **1988**, 3, 1063.

174. Lambert, B., Jones, B.K., Rooques, B.P., Le Pecq, J.B. and Yeung, A.T. Proc. Natl. Sci. Acad. USA, **1989**, 86, 6557.

175. Lambert, B., Segal-Bendirdjian, E., Esnault, C., *et al.* Anticancer Drug Design, **1990**, 5, 43.

176. Maroun, R., Delepierre, M. and Roques, B. J. Biom. Struc. Dyna.,

1989, 7, 607.

177. Delbarre, A., Delepierre, M., Garbay, C., *et al.* Proc. Natl. Acad. Sci. USA, **1987**, 84, 2155.

178. Wolf, B. and Hanlon, S. Biochem., **1975**, 14, 1661.

179. Falk, M., Hartman, J.K.A. and Lord, R.C. J. Amer. Chem. Soc., **1962**, 84, 3843.

180. Falk, M., Hartman, J.K.A. and Lord, R.C. J. Amer. Chem. Soc., **1963**, 85, 387.

181. Neidle, S., Berman, H.M. and Shieh., H.S. Nature, **1980**, 288, 129.

182. Bernal, J.D. and Crowfoot, D.C. Nature, **1934**, 133, 794.

183. Wang, A. and Gao, Y.-G. Methods, **1990**, 1, 91.

184. McPherson, A. Preparation and Analysis of Protein Crystals 1-371 (R. E. Kreiger, Melbourne, Florida, USA, 1989).

185. Theorell, H. Biochem. J., **1932**, 252, 1.

186. Boyes-Watson, I., Davidson, E. and Perutz, M.F. Proc. Roy. Soc, Ser., **1947**, A191, 83.

187. Zeppezauer, E., Soderburg, B.O., Branden, C.I., Akenson, A. and Theorell, H. Acta Chem. Scand., **1976**, 21, 1099.

188. Zeppezauer, M. Adv. Emzymol, **1971**, 22, 253.

189. James, R.W. The Optical Princeples of the Diffraction of X-ray (Cornell University Press, Ithaca, New York, 1965).

190. Blundell, T.L. and Johnson, L.N. Protein Crystallography (Academic Press, New York, 1976).

191. Glusker, J.P. and Trueblood, K.N. Crystal Structure Analysis

(Oxford University Press, New York, Oxford, 1985).

192. Still, W. C. et al. MacroModel manual, 1989, Columbia University, New York

193. Quigley, G.J., Gehrke, L., Roth, D.A. and Auron, P.E. Nucleic Acids Research, **1984**, 12, 347.

194. Jones, T. A. in Computational Crystallography, 1982, Clarendon Press, Xoford.

195. CHAIN Manual, 1991, Baylor College of Medicine

196. Stodola, R. K., Manion, F. J., DOCK manual, 1985, The Fox Chase Cancer Center

197. Aoki, K., Clark, G.R. and Orbell, J.D. **1976**, 425, 369.

198. Marzilli, L.G., Kistenmacher, T.J. and Eichhorn, G.L. 1980. In Nucleic acid-metal ion interactions, Spiro, T.G. eds., John Wiley & Sons, Inc. New York.

199. Scheaaper, R.M., Koplitz, R.M., Tkeshelashvili, L.K. and Loedb, L.A. **1987**, 177, 179.

200. Yamamoto, K. and Kawanishi, S. J. Biol. Chem., **1989**, 264, 15435.

201. Sagripanti, J.-L. and Kraemer, K.H. J. Biol. Chem., **1989**, 264, 1729.

202. Kagawa, T., Geierstanger, B.H., Wang, A.H.-J. and Ho, P.S. J. Biol. Chem. **1991**, 266, 20175.

203. Delaprat, D., Delbarre, A., Oberlin, R., *et al.* J. Med. Chem., **1980**, 23, 1336.

204. Enraf Nonius, CAD4 manual, 1988, Enraf Nonius Co. New York

205. Eneaf Nonius, RF558 manual, 1988, Enraf Nonius Co. New York

206. Kim, S.H., Quigley, G.J., Suddath, F.L. and Rich, A. Proc. Nat. Acad. Sci. USA, **1971**, 68, 841.

207. Hendrickson, W.A. and Konnert, J. 1979. In Biomolecular Structure, Conformation, Function and Evolution, Srinvasan, R. eds., Pergamon Press. Oxford.

208. Gererstanger, B.H., Kagawa, T.F., Chen, S.-L., Quigley, G.J. and Ho, P.S. J. Biol.Chem., **1991**, 266, 20185.

209. Frederick, C.A., Williams, L.D., Ughetto, G., *et al.* Biochemistry, **1990**, 29, 2538.

210. Gessner, R.V., Frederick, C.A., Quigley, G.J., Wang, A.H.-J. and Rich, A. *The Molecular Structure of Z-DNA at High Resolution* (1990).

211. Holbrook, S.R., Wang, A.H.J., Rich, A. and Kim, S.H. J. Mol Biol, **1988**, 199, 349.

212. Delbarre, A., Delepierre, M., Igolen, J. and Le Pecq, B.P. Biochim. **1985**, 67, 823.

213. Delepierre, M., Igolen, J. and Roques, B.P. Biopolymers., **1988**, 27, 957.

214. Delepierre, M., Maroun, R., Garbay-Jaureguiberry, C., Igolen, J. and Roques, B.P. J. Biol. Chem., **1989**, 210, 211.

215. Lavery, R. and Pullman, B. J. Biomol. Struc. Dynam., **1985**, 2, 1021.

216. Williams, L.D., Egli, M., Gao, Q., *et al.* Proc. Natl. Acad. Sci. USA, **1990**, 87, 2225.

217. Egli, M., Gessner, R.V., Williams, L.D., *et al.* Proc Natl Acad Sci U S A, **1990e**, 87, 3235.

218. Williams, L.D., Frederick, C.A., Ughetto, G. and Rich, A. Nucl. Acid. R., **1990b**, 18, 5533.

219. Coll, M., Frederick, C.A., Wang, A.H.J. and Rich, A. P.N.A.S., **1987**, 84, 8385.

220. Jain, S., Zon, G. and Sundaralingam, M. Biochem., **1989**, 28, 2360.

221. Marzec, C. J., Day, L. A. J. Biomol. Struc. Dyn. **1993**, 10, 1091-1123

222. Marzec, C. J., Day, L. A. J. Biomol. Struc. Dyn. **1993**, 10, 1125-1154.

223. Quigley, G. J. and Rich, A. Science **1976**, 194, 796-806.