Order Number 9521277

X-ray crystallographic studies of small molecules and their interaction with DNA

He, Ru, Ph.D.

City University of New York, 1995

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

X-RAY CRYSTALLOGRAPHIC STUDIES OF SMALL MOLECULES AND THEIR INTERACTION WITH DNA

by

RU HE

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

1995

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.

1/20/95
Date

Chair of Examining Committee

1/24/95
Date

Executive Officer

Supervisory Committee

The City University of New York

Abstract

X-RAY CRYSTALLOGRAPHIC STUDIES OF SMALL MOLECULES AND THEIR INTERACTION WITH DNA

by
Ru He

Research Supervisor: Professor Gary Joseph Quigley.

X-ray crystallography and molecular mechanics were used to study the three-dimensional structure of three small molecules, one drug-DNA complex, and several metal ion-DNA complexes.

Structural features of the d(CGCGCGCGCGCG)-ditercalinium complex have been inferred by geometric and energetic analysis using molecular mechanics. A systematic comparison of the dodecamer model complex with the d(CGCG)-ditercalinium crystal structure and B-form DNA have been made. The driving force that stabilizes the complex is the interaction between the chromophores of ditercalinium and bases of DNA. Molecular features of the complex that are possibly responsible for the drug's unique mechanism, include the accessibility of the minor groove, the bend in the helical axis, the rigidity of linker, and the introduced asymmetry parameters.

X-ray crystallography was used to obtain the structure of three organic molecules. In (2s,3r,4s,5s)-3-(allyloxy)-4-[n-(benzyloxycarbonyl)-amino]-5,6-dihydroxy-2-[1-(allyoxy)naphth-3-yl]-2-hexanol structure, the β-

configuration was assigned to the compound because the thymine substituent at C-1' was elucidated to be cis to the substituent at C-4'. The S-methyl substituent at C-2' was determined to be trans to the thymine substituent at C-1' and cis to the 3'-hydroxyl which places this compound in the ribose series. In structure (2s,3r,4s,5s)-3-(allyloxy)-4-[n-(benzyloxycarbonyl)-amino]-5,6-dihydroxy-2-[1-(allyoxy)naphth-3-yl]-2-hexanol, the 5S configuration was determined instead of the 5R which is considered as the necessary conformation for activity. Full crystallographic details of the determinations of all three structures of are given in chapter 4.

Features of the cation-DNA interactions that stabilize the DNA crystal structure have been inferred from the x-ray crystallography structure of the hexanucleotide d(CGCGCG) at 1.3Å high-resolution (structure I). Cation binding sites and DNA solvation have been visualized in great detail. The crystal structure displays substantial solvent and solute order; the hydrated magnesium complexes exhibit specific contacts with DNA and there is only one conformation (ZI). This contrasts with the identical hexamer crystallized using a higher ionic strength mixed salt buffer (structure II). Structure II exhibits greater solvent and solute disorder, less well defined cation-DNA contacts and two conformations (ZI and ZII) of the Z-form nucleic acid.

To my father, my mother and my son

ACKNOWLEDGMENTS

I thank Professor G.J. Quigley for the opportunity to work under his tutelage in his laboratory, and for his generous support.

I thank Professors M. Mezei, D.J. Goss, K. Grohmann and L. Massa for their advice, help and encouragement. I am especially grateful for Professor P. Mills's constant unending guidance and care during my tenure at the graduate program and for her review of the earlier version of my thesis.

I thank the many people who have made contributions towards my attaining my degree, academic and otherwise. I thank my colleagues Drs. S.L. Chen, L.S. Huang, A.J.A. Soirat, Q. He and Z.P. Liu for their help and time. I give special thanks to O. Schonberg who had read and corrected the earlier version of this thesis. I thank my dear friends E. Bonner, P. Mendenhall and D. Tan for their emotional support and encouragement through these years.

I would like to thank Dr. J. L. Kim and Professor S.K. Burley of The Rockefeller University. Their advice, suggestions and help are greatly appreciated.

Special thanks also go to Mr. K. Tes for his generosity, kindness and love. Finally, I thank my parents, my sister and my son for their unconditional love and support.

Table of Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION AND BACKGROUND

1.1. THESIS PLAN

> *"We also now appreciate that molecular biology is not a trivial aspect of biological systems. It is at the heart of the matter. Almost all aspects of life are engineered at the molecular level, and without understanding molecules we can only have a very sketchy understanding of life itself. All approaches at a higher level are suspect until confirmed at the molecular level".*
>
> Francis Crick, What Mad Pursuit, 1988

In order to obtain an in-depth understanding of the biological activity of DNA, we study the three-dimensional structures of DNA. We seek to gain insight into the general principles of its structure stabilization and to clarify the origins of the structural effects on DNA caused by ligand. The oligonucleotide's local or global structure can be sensitive to the presence of ligand. The variability in structure depends significantly on the oligonucleotide sequence and on the ligand types. X-ray crystallographic methods have been used here to study the conformational effects of cations on oligonucleotide and molecular modeling has been used to study the structural perturbation caused by the anticancer drug ditercalinium. This study was facilitated by the methods of x-ray crystallography and molecular modeling that allow us to gain insight in the greatest possible detail and accuracy. We are addressing the following issues: drug intercalation into DNA, metal binding to DNA, and water molecules binding to DNA.

The primary focus: The primary focus of this work is to study the structure and function relationship of double-helical oligonucleotide by structure analysis.

Specifically we seek:

1) To define the binding sites and geometry of mono- and multivalent cations and their effects on the geometry of the oligonucleotide.

2) To characterize the structural features and distortions induced by active and inactive drug-DNA complexes.

3) To analyze the geometric and energetic aspects of the complexes in detail.

Specific aims. The first aim is the definition of the effect of mono- and multivalent cations binding on the geometry of the oligonucleotide. X-ray structure determination of metal binding to DNA gives direct evidence of cation-induced structural change and allows visualization of the ions. In order to optimize information on these two classes of interactions, two conditions were used in crystallization: a) K^+ as the only cation in the solution. Without the presence of multivalent cations, it gives an opportunity to see the monocation mode of binding in detail. b) Mg^{2+} as the only metal cation in the solution, in the absence of the monocation one should be able to find all of the Mg^{2+} in the structure.

The second aim is the characterization of the structural features of active and inactive drug-DNA complexes and, the identification of DNA distortion induced by the drug. For example, the biologically active ditercalinium-DNA complex has the character of kinking, this is proposed

to be a recognition signal for the enzyme activity. Therefore, by crystallizing and determining the structures of active and inactive drug-DNA complexes we can compare their structural features and gain insight into the mechanism of enzyme recognition specificity.

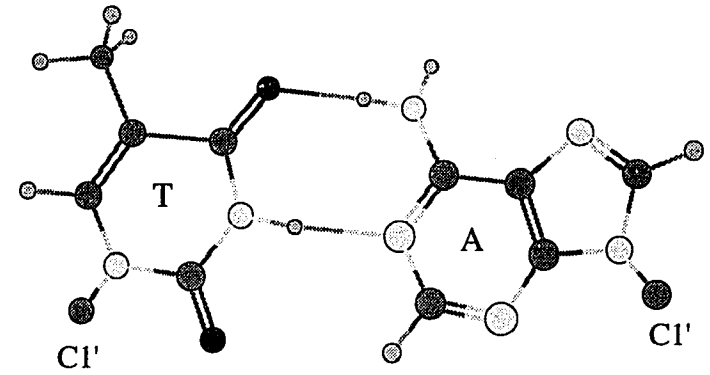
The third aim is the analysis of the geometry and energetic aspects of the complexes in detail. Molecular modeling of long sequence drug-DNA complexes will provide additional information about intercalation sites from which structural-functional relationships can be deduced.

1.2. DNA STRUCTURE

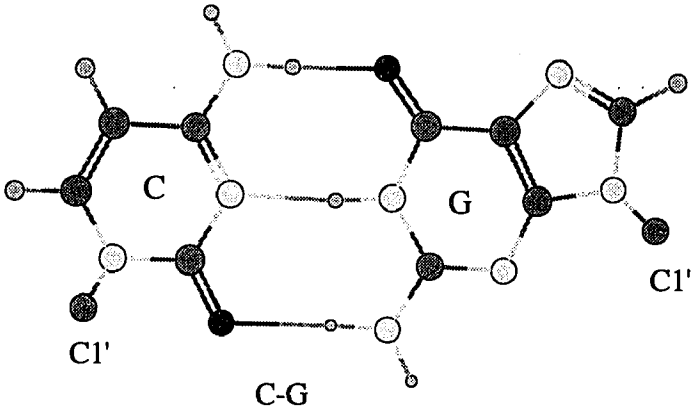
Nucleic acids are molecules that carry genetic information necessary for all living organisms. Their importance stems from the remarkable diversity of their functional roles.

The structure determination on the double helical DNA was first made by James Watson and Francis Crick in 1953 [1]. X-ray diffraction data obtained in Wilkins' laboratory by Rosalind Franklin in addition suggested that DNA was a helical molecule and is able to adopt several different conformations [2]. The major feature of the Watson-Crick model of DNA is that the DNA consisted of two antiparallel polynucleotide strands that wind around a common axis to form a double helix. The bases occupy the core of the helix while the sugar-phosphate backbone chains are coiled around its periphery. The planes of the bases are found to be nearly perpendicular to the helical axis and each base is hydrogen bonded to a base on the opposite strand to form a planar base pair. The most remarkable feature of the Watson-Crick model was that it accommodated only two

types of base pairs (**Figure 1-1**): adenine pairing with a thymine, guanine pairing with a cytosine and each in two orientations. This work ranks as one of science's major intellectual achievements; not only providing the structure of DNA, the central molecule of life, but also suggesting how hereditary information is encoded in the sequence of its bases and how that information is replicated and expressed.



T-A



C-G

Figure 1-1. Watson-Crick base-pairs

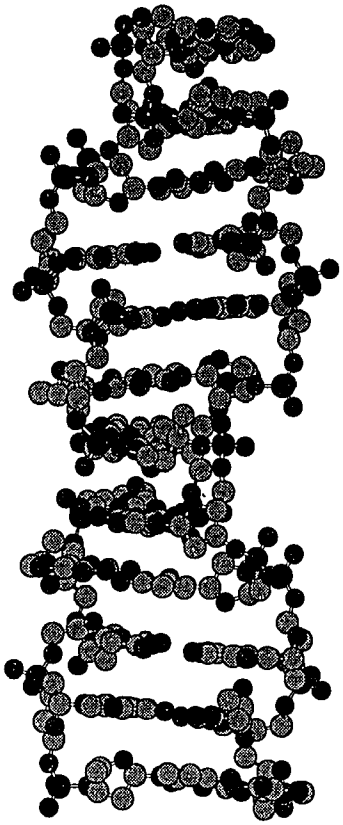


Figure 1-2. Overall structure of Z-DNA

(from Chen, 1993)

Figure 1-2. shows an overall view of a DNA structure. Deoxyribonucleic acid (DNA) is a linear polymer built up of monomeric units called nucleotides. A nucleotide is a nucleoside phosphorylated at one of the free sugar hydroxyls' groups. A nucleoside is a purine or pyrimidine base joined to a furanose sugar ring. The three-dimensional structure of a molecule is characterized by bond lengths, bond angles, and torsion angles. The torsion angles involve four atoms in sequence, A-B-C-D and it is defined as the angle between projected bonds A-B and C-D when looking

along the central bond either in direction B to C or in the opposite sense C to B and labeled as shown in **Figure 1-3.** and **Table 1-1.** [3]. Definitions of the helical parameters rise, roll and twist are illustrated in **Figure 1-4.** The rise is the distance along the local helix axis between successive base pairs, the roll is the angle between the planes defined by two successive base pairs, and the twist is the angular twine of the double helix between successive base pairs. Sugar ring can be puckered in an envelope (E) form with four atoms in a plane and the fifth atom is out by 0.5Å; or in a twist (T) form with two adjacent atoms displaced on opposite sides of a plane through the other three atoms. Atoms displaced from these three- or four-atom planes and on the same side as C5' are called endo; those on the opposite side are called exo (**Figure 1-5.**).

The DNA family consists of single strand, double strand, triple strand and four strand forms. The most common form is double helix DNA. The major helix types are right-handed helix A-DNA and B-DNA, and left-handed helix Z-DNA. Other than simple double helices, the structures also have a variety forms that include bulged loops, hairpin loops, cruciforms, catenanes, knots, and branches.

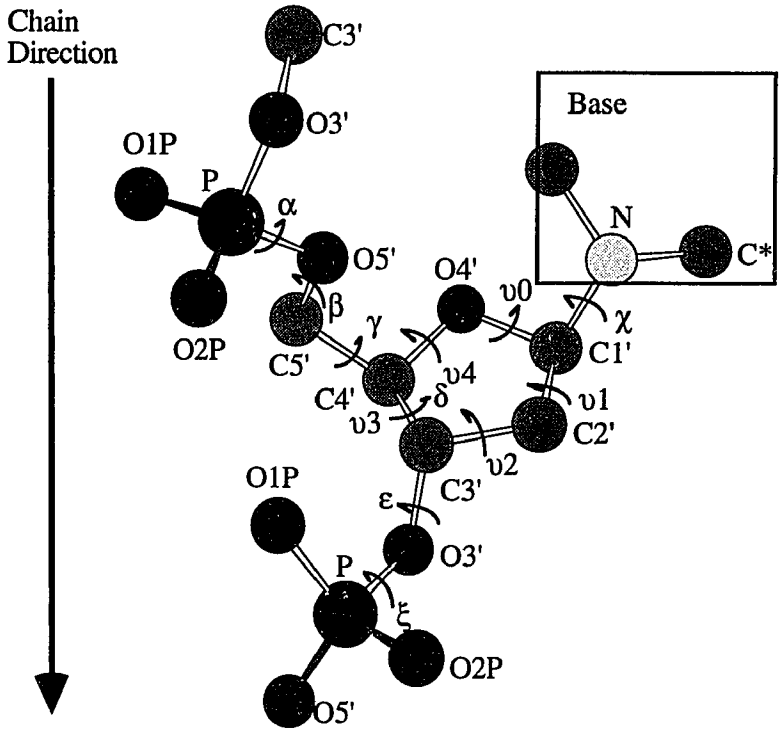


Figure 1-3. Definition of torsion angles in nucleotide

Table 1-1. Definition of Torsion Angles in Nucleotides

Torsion Angles	Atoms involved
α	$_{(n-1)}O_{3'}$-P-$O_{5'}$-$C_{5'}$
β	P-$O_{5'}$-$C_{5'}$-$C_{4'}$
γ	$O_{5'}$-$C_{5'}$-$C_{4'}$-$C_{3'}$
δ	$C_{5'}$-$C_{4'}$-$C_{3'}$-$O_{3'}$
ε	$C_{4'}$-$C_{3'}$-$O_{3'}$-P
ξ	$C_{3'}$-$O_{3'}$-P-$O_{5'}$
χ	$O_{4'}$-$C_{1'}$-N_1-C_2 (pyrimidines)
	$O_{4'}$-$C_{1'}$-N_9-C_4 (purines)

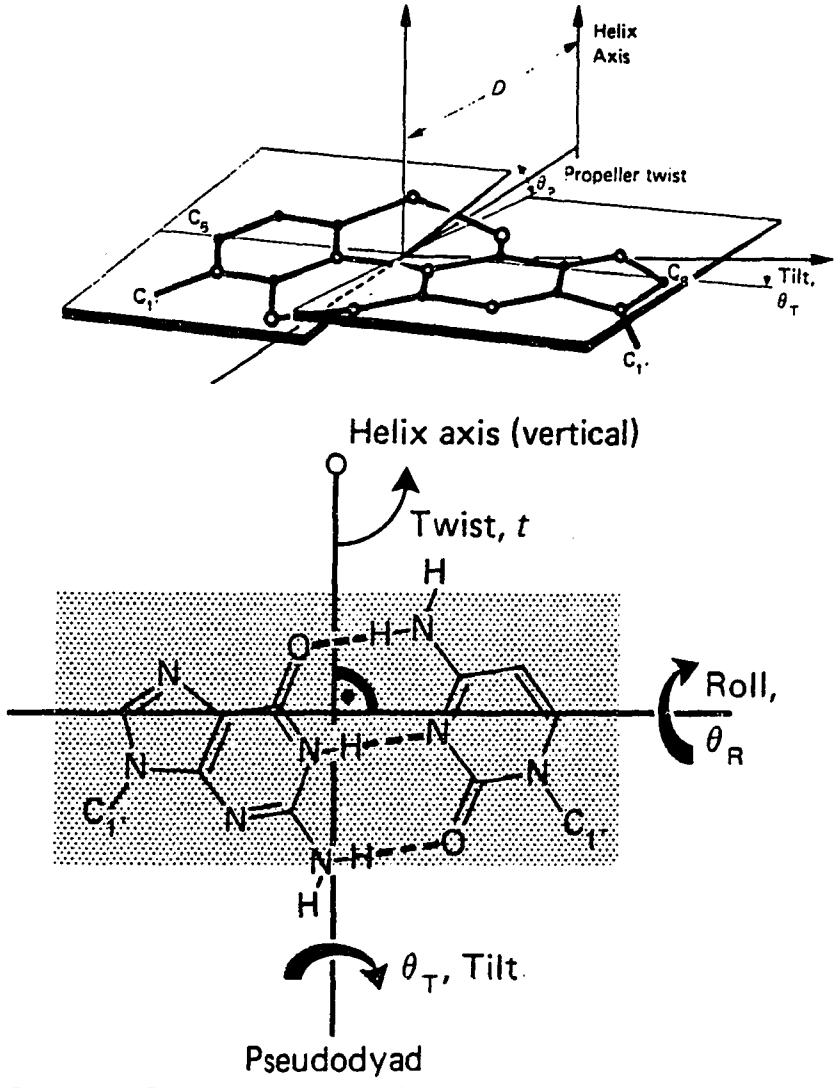


Figure 2-4. (Top) The two bases in a base-pair are not exactly coplanar but display propeller twist θ_p, defined as dihedral angle between individual base planes. Looking down the long axis of a base-pair, θ_p is defined positive if the near base is rotated clockwise with respect to the far one. (Bottom) The orientation of a base-pair with respect to the helix axis is given by rotational twist t, and by tilt θ_T and roll θ_R of a mean base-pair plane (stippled). Reference line for t and θ_T is the pseudodyad (see Figure 2-13), and for θ_R, a vertical to the pseudodyad passing approximately through C_6 (pyrimidine) and C_8 (purine) is taken. Signs of angles θ_T and θ_R is positive for clockwise rotation as indicated by curved, thick arrows. Reference line for 0° is paper plane (or vertical to helix axis). Positive rotation θ_T dips base to the right (cytosine) below paper plane, and for θ_R, atoms in major grove side (top of stippled area) are dipped below paper plane.

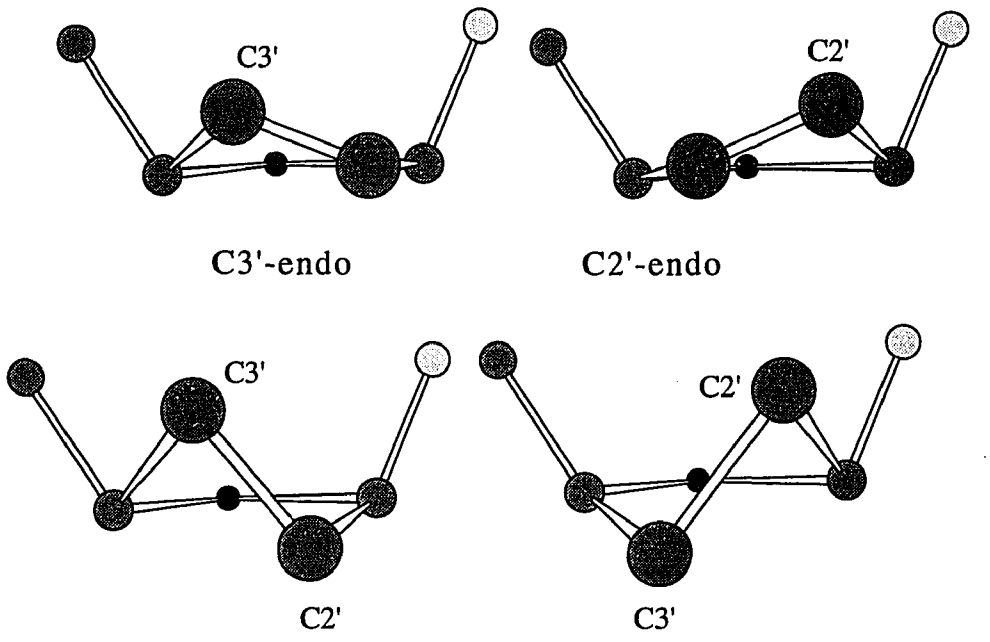


Figure 1-5. Conformations of sugar pucker
(from Chen, 1993)

1.3. DNA-LIGAND ASSOCIATION

A fundamental goal of the study of biological molecules is to understand the relationship between their structure and physicochemical properties and their mechanism of action. DNA-ligand association has a large impact on DNA structure. Small ligands, like metal ions, can greatly stabilize particular DNA helix conformations so that ion can induce conformational changes such as the B to Z transition [4,5]. Large ligand such as intercalating drugs on the other hand, can distort the DNA structure, and cause significant structural changes [6,7].

The important factors stabilizing the three dimensional structure of nucleic acids are:

1) Watson-Crick base pairing. Hydrogen bonding is both geometrically and electronically specific.

2) Base stacking is caused by electrical forces between the adjacent bases, van der Waals contacts and hydrophobic effects.

3) Electrostatic effects are interactions between the charged groups. Electronic complementarity plays an essential role in interactions between DNA or RNA and other molecules.

4) Hydration. In DNA water is not only a medium to keep the solutes dissolved but it also interacts with DNA and is responsible for the stabilization of secondary and tertiary structure.

1.3.1. METAL ION BINDING TO NUCLEIC ACIDS

The importance of metal ion-DNA interaction was first recognized when the antitumor activities of cis-dichlorodiamineplatinum(II) were discovered in 1969 [8]. People were aware of the necessity of charge balancing metal ions from the early model building prior to Watson and Crick. Metal ions are required by all living cells and their concentration range from 0.001 to a few hundreds mM in living cells. Some DNA replication process require divalent metal ions for catalysis [9,10]. Metal cation-nucleic acid complexes are extremely important for the biological action of nucleic acids, nucleotides, coenzymes, and nucleoside di- and triphosphates [4,10,11]. There is a great deal of interesting work on nucleic acid-metal ion interaction including: the possible modes of action of the

platinum anticancer agents [12], the errors in DNA replication induced by metal ions [12] and the high-resolution crystal structure of metal ions interacting with transfer RNA [13]. A number of review articles describing the stereochemistry of metal ion-DNA binding have been published [14,15,16].

The metal ions surrounding the DNA have a large impact on its structure. Nucleic acids are highly charged polyanions, the phosphate-phosphate electrostatic repulsion is diminished by the high dielectric constant of water and by the positively charged cations (or hydrated counterions). The conformation of nucleic acids in solution and in the crystal state is sensitive to the concentration and type of cation(s) present. Polyvalent cations tend to bridge nearby negatively charged groups, especially phosphates in DNA. Monovalent cations generally do not bind for long periods of time to specific sites but manifest their presence through a mobile cloud of charges [3]. Detailed structural information of the nucleic acid in solid state, and its dependence upon the cation and solvation can be shown, for example, by reducing humidity or increasing ionic strength both of which lead to a transition from the B to Z-form of DNA [17]. Divalent ions are much more effective than the monovalent alkaline metal ions in stabilizing DNA [3].

Nucleic acids can be thought of as offering four different binding sites for metal ions: phosphate oxygens, ribose hydroxyls for RNA and terminal residues, base ring nitrogens, and keto groups on the base. The dominate binding sites for transition metals are the phosphate oxygens and base nitrogens. Alkaline earth cations show a preference for phosphate

oxygens and sugar hydroxyls over base nitrogens. Alkaline metal ions also complex equally well to all kinds of other ligands [3]. But the affinity of metal ions for nucleotide-binding sites is rather specific [3].

From published DNA crystal structures, it can be found that the number of positively charged cations visualized is never enough to match the negative charge of the phosphates in DNA [18,19,20]. The charges must be neutralized so why are they not seen? The possibilities are: low-resolution models and disordered regions. Sodium ion is often believed to be present but it cannot easily be distinguished from water due to the fact that at low resolution much of solvent region is disordered.

Because mono- and divalent cations play an important role in stabilizing the DNA structure, our study focuses on the roles that K^+ and Mg^{2+} cations play in DNA structures.

1.3.2. ANTITUMOR DRUG INTERCALATION INTO DNA

Chemotherapy research in cancer focuses on two kinds of studies: inhibition of nucleic acid synthesis and interference with mitosis as a means of discriminating selectively against fast growing tumor cells. It is well established that the main cellular target of antitumor drugs is DNA [21,22].

It is also recognized that the modes of interaction of these drugs, with DNA as their receptor, are highly diverse. There are many ways that drugs can interact with DNA. They can react to form new bonds (covalent bond formation), they can bind non-covalently either on the major or minor groove, and they can intercalate. The specific mechanisms of drug binding are the basis of their antitumor activity.

Intercalation is considered as being the most significant mechanism for the interaction of the anthracylines daunomycin and adriamycin with DNA [23,24,25,26].

The concept of intercalation was first introduced by Lerman in 1961 [27]. It explained the reversible, non-covalent binding of compounds with extended planar aromatic chromophores to double-helical DNA. These compounds usually have a number of substituent groups or chains associated with the chromophore. A particularly interesting problem is the respective role of these different parts of the active compounds in the mechanism of their binding to DNA and the specificity of the interaction [28,29]. The most active species are electrophilic stacking agents [30]. The key point of intercalation is the insertion of such a chromophore between adjacent base pairs of DNA. The drug induces perturbations in the DNA structure, which disrupt the recognition properties of the nucleic acid [31]. An example of distortion of the DNA helical structure is illustrated in **Figure 1-6**. Some physical-chemical characteristics of DNA intercalation are: the increase in the length of duplex, the change of the unwinding angle and the increase in T_m of DNA in the complex [32].

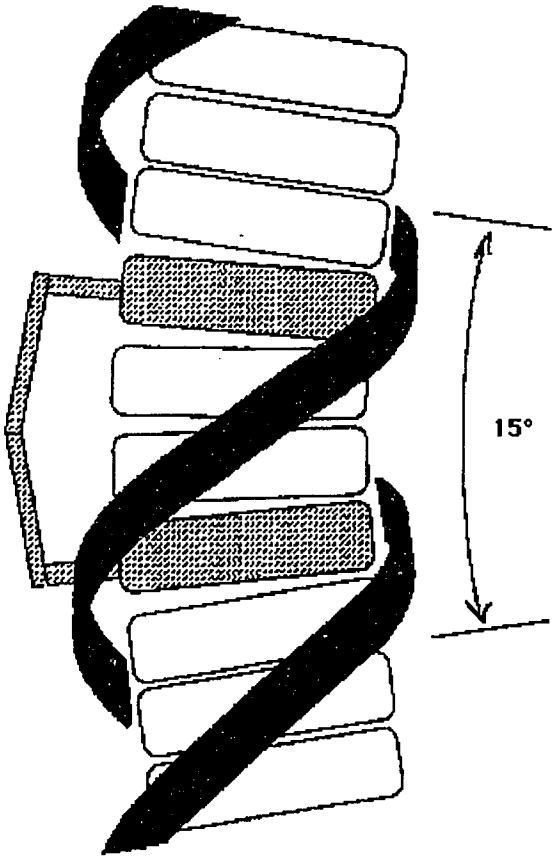


Figure 1-6. Schematic diagram of DNA-ditercalinium complex. The kink of DNA induced by the bis-intercalation is shown.

Many intercalating compounds are clinically used as anticancer drugs. Actinomycin D is a powerful antibiotic, and daunomycin is a potent chemotherapeutic agent. Ditercalinium elicits antitumor properties through a new mechanism of action. Ditercalinium appears to be unique in its ability to promote the enzyme-mediated incision of noncovalently damaged DNA [33,34,35]. The wide range of biological effects of intercalating drugs has been extensively documented [36,37]. Intercalation is currently considered to be a component for the anticancer action of these drugs, and

is probably the primary recognition event prior to the formation of a ternary drug-DNA-enzyme complex [31]. The processes of DNA transcription, translation, and replication can all be affected by intercalation. Three possible effects can be traced to the binding of intercalating drugs to DNA: (1) inhibition of DNA dependent enzymes, (2) frameshift mutagenesis, and (3) damage to the DNA.

Structural changes induced by the DNA-drug complex will determine the biological efficiency. Experiments show that there is DNA sequence specificity preference by a drug, and the minor changes in the structure of the drug substantially alter the antitumor activity [38,39]. The stereo-selectivity and biological activity can be effected by not only the presence or reactivity of particular atoms or groups, but also by the shape of the molecules.

1.4. MOLECULAR MODELING FOR DRUG DESIGN

In the past, drug design research involved screening thousands of compounds experimentally, an arduous and time-consuming process. Recently a new strategy has been developed. The basis of this strategy is the structure determination of the target proteins or DNA by x-ray crystallography or NMR. The three-dimensional structural information obtained from these studies provides insight to guide the design of drugs using the key and lock relationship with the target molecules [6,40].

X-ray crystal structures exist only for short sequences and are subject to end effects with 5'- and 3'-end nucleosides having conformations that are unconstrained because there are no succeeding residues. To

overcome the length constraints, molecular mechanics or computer modeling is used, in conjunction with x-ray crystallography.

A number of models have been derived for drug-DNA and drug-protein studies. They includes molecular mechanics, molecular dynamics, Monte Carlo simulations and free energy calculations [41,42,43]. One of the earliest computer modeling studies of nucleic acid structure using an atomic model was that by M. Levitt [44], in which energy minimization was used to optimize models for a 20 base pair segment of DNA.

Molecular mechanics is a method that uses a so-called force field to calculate approximate internal energies of a given molecule as a function of the coordinates of its atoms. The force field "knows" about molecular geometry and their statistical relevance, so that it can be considered as a physicochemical knowledge base [45].

The energy minimization searches used in these force fields have the advantage that bond lengths and angles can be perturbed in the location of minimum-energy structure, while the disadvantage is that the energy-minimum structure is frequently similar to the initial structure. The latter can be overcome by either systematically changing the input structure in order to locate the lowest energy configuration or by employing techniques that allow the structure to escape local minimum, such as molecular dynamics or Boltzmann jump techniques.

Several programs have been developed to conduct structure refinements. The AMBER [45] and CHARMM [46,47] programs are fundamentally similar and are used interchangeably. The force field used in this study was CHARMM, employed within the XPLOR macromolecular

refinement package [48] which is designed especially for three-dimensional structure determination and refinement using x-ray crystallographic diffraction data or nuclear magnetic resonance.

In considering the results of force field calculation, a major concern is the reliability of the results. The reliability depends critically on the physical property of the molecule inspected and the force field chosen. Local geometry, such as bond distances and bond angles are the most reliable. The energy-minimum structure is frequently similar to the initial structure. Large-scale geometric properties such as helix parameters, however, are much less reliable. The global conformation depends on complex physical properties of the molecule including hydration and counterions distribution. It is very difficult to incorporate these effects into the force fields. Knowing the limitation of molecular modeling will help us make sensible use of this method.

1.5. X-RAY CRYSTALLOGRAPHY

X-ray crystallography studies over the last thirty-five years have revolutionized our knowledge about molecular structures, especially in molecular biology. It was the tool used to discover the basic double-helical form of DNA [1]. Only after the three-dimensional structure of a macromolecule has been determined, does it became possible to have a full understanding of most biological processes at the atomic level. Although there are many methods which are available for 3D structure studies, x-ray crystallography is the only method that permits the accurate determination of the molecular architecture of compounds possessing a molecular weight

of up to at least 10^6 Daltons. A careful crystal structure analysis can provide information about bond lengths and angles, inter- and intra-molecular interactions, details of conformation, molecular packing, and hydrogen bonding.

The data from crystal structure analysis are unit cell constants and sets of atomic coordinates. The method consists of three major parts: crystallization, data collection, and structure solution. First, it is necessary to form crystals of the right size (0.2-0.4 mm) and of good quality. Second, data must be collected to obtain the scattering angles and the intensities of the diffracted beams. Third, using the Fourier transform methods or least-square methods the structure is refined.

1.5.1. CRYSTALLIZATION

The basis of crystallization by the vapor diffusion method, the method used in most of this works, is that the sample droplet contains the macromolecule with buffer, crystallizing agent, and additives equilibrated against a reservoir containing a solution of crystallizing agent at a higher (or lower) concentration than the droplet. Through the vapor phase, the concentration of salt or organic solvent in the reservoir equilibrates with the sample. The direction of the exchange depends on the vapor of the species, for those with a vapor pressure lower than water, the exchange occurs from the sample droplet to the reservoir [49].

Method: 10 to 50ml of mother liquor droplets are placed in multiple depression spot plates (Corning #7220). The samples are then sealed in transparent containers, such as plastic sandwich boxes, which contain

reservoirs of 30 to 50 ml of precipitating solution. The sample plates are held on the bottom of the reservoir by inverted halves of disposable Petri dishes.

The factors affecting crystallization of nucleic acids are temperature, pH, concentration and additives.

The widely used crystallizing agents are the MPD (2-methyl-2,4-pentanediol) and PEG (Polyethylene glycol) alcohols which are not volatile and therefore easy to handle. Nucleic acids are stable at a large range of temperatures (4°C to 35°C), pH variations play a much smaller role in nucleic acid crystallization than in protein crystallization. The most common buffer is Na cacodylate (pH range 5.0-7.0). Nucleic acid concentration is generally tried in the order of 0.2-4.0mM. Counter-ions are very important additives because nucleic acids are poly-electrolytes. The major families are the divalent (Mg^{2+}, Ca^{2+}) and the monovalent (Na^{+}, K^{+}, Li^{+}) and some polyvalent cations.

1.5.2. DATA COLLECTION AND STRUCTURE DETERMINATION

The goal of data collection is to determine the indices and to record the intensities of reflections.

The result of x-ray data collection is a list of intensities, each assigned indices hkl corresponding to its position in the reciprocal lattice [50]. The larger magnitude of the indices, hkl's, corresponds to the small magnitude d_{hkl} (interplanar spacing) in the spacing of real space planes. Meaning that the reflections far away from the origin (large hkl) come from sets of narrow spaced planes (small d), and carry information about

the fine details of the structure. On the other hand, the reflections close to the origin come from widely spaced lattice planes and carry information about larger features of the molecules in the unit cell.

After data collection, the raw intensities must be corrected to improve their consistency and to maximize the number of measurements with sufficient accuracy. If the data is measured from more than one crystal or on more than one film, scaling will be used to rescale the intensities of more than one set of data so that equivalent reflections have similar intensities.

Unit cell dimensions are determined by diffractometer software which search for reflections, measure their precise positions, and subsequently compute unit-cell parameters.

Space group is determined by examining unit cell parameters, intensity symmetry and for systematic absence of the diffraction data, then checking the *International Table for X-ray Crystallography* for systematic absences of all space groups.

If the structure studied relates to some known structure, molecular replacement is used for structure determination of orientation and position. The known model as a phasing model is used to calculate the initial phase for the working model. Patterson-map comparisons are used to find the best orientation of the model while structure-factor comparisons are used to find the best location of the model.

1.5.3. STRUCTURE REFINEMENT

The computer simulates the action of a lens, computing the electron density within the unit cell from the list of indexed intensities obtained from data collection.

The beauty of crystallography lies in the mathematical and geometric relationships between diffraction data and molecular images. The Fourier transform allows us to convert a Fourier-series description of the reflection intensities to a Fourier-series description of the electron density [51]. In other words, the structure factor F_{hkl} which represent the reflections of the diffraction pattern can be written as the integral of electron density over the unit-cell volume.

$$\mathbf{F}_{hkl} = \int_V \rho(x,y,z)\, e^{2\pi i\,(hx+ky+lz)}\, dV$$

Where the $\mathbf{F}_{hkl}$ is the structure factor that describes reflection hkl, $\rho(x,y,z)$ is the electron density at fractional coordinate (x,y,z), $e^{2\pi i\,(hx+ky+lz)}$ is the interference function for diffraction, and V is the unite cell volume.

$|F_{obs}|$ and $|F_{calc}|$ are the observed and calculated structure factors. The structure factor of $|F_{calc}|$ also can be expanded to show all the parameters included in the refinement as follows:

$$\mathbf{F}_{calc} = G \sum n_j f_j\, e^{2\pi i\,(hxj + kyj + lzj)}\, e^{-Bj[(\sin\theta/\lambda)]^2}.$$

Where G is an overall scale factor in putting all the F_{calc}'s on a numerical scale; n_j is the occupancy of atom j; f_j is its atomic scattering factor, x_j, y_j, and z_j are its fractional coordinates and B_j is its temperature factor, which

measures either how much the atom oscillates around the position specified in the model or how much disorder is present. It can be written as:

$$B_j = 8\pi^2 \langle u_j^2 \rangle$$

Where u_j is the rms displacement of atom j; sin θ/λ is the angle of the reflection (θ is the angle of the x-ray beam impinges on a reflection plane; λ is the wavelength of the x-radiation).

For macromolecules, isotropic (thermal motion of equal magnitude in all directions) temperature refinement is used and anisotropic (there are preferred directions of vibration for each atom) temperature refinement is used for small molecules. To obtain a reliable structure the number of measured data should well exceed the number of parameters, yet a description of anisotropic vibration needs six parameters but only one for isotropic, so most macromolecules are refined with isotropic temperature factors.

The |Fobs|s are proportional to the square roof of the intensity of reflection $(I_{obs})^{1/2}$. From the measured reflection intensities one can get the amplitudes $|F_{obs}|$s. Structure refinement is performed to compare the measured structure factor $|F_{obs}|$ with $|F_{calc}|$ from the current model, and seek for the best combination of parameters (atomic coordinates, occupancy and temperature factor) for all atoms in the model to minimize the difference between all the observed and calculated structure factors.

The overall agreement between the amplitudes of two sets of structure factors can described by the R-factor as follows:

$$\mathbf{R} = \sum ||F_{obs}| - |F_{calc}|| \,/\, \sum |F_{obs}|$$

This number indicates how well the calculated model fits the observed data. The smaller the R, the better agreement between observed and calculated intensities. The value of R ranges from 0 for perfect agreement approaching to 1. An R-factor greater than 0.5 implies very poor agreement and many models with R=0.5 will not respond to attempts at improvement unless more data is available. Macromolecules may refine from R=0.3 to 0.1, while small organic molecules commonly refine to better than R=0.05. This is because, in the structure of macromolecules, many solvent molecules may be highly disordered or in motion on the time scale of crystallography and there is frequently static disorder of the macromolecules themselves. Small molecules, on the other hand, tend to be well-ordered so the model can match the data well.

The quality of a crystallographic model is judged by whether the model is chemically, stereochemically and conformationally reasonable. In a "chemically" reasonable model, the criteria used are the rms deviations of all of the model's bond lengths and angles from an accepted standard set of values. Usually, rms deviations less than 0.02Å for bond lengths and 4° for bond angles are considered to be well refined models. A stereochemically reasonable model has no inverted centers of chirality. A conformationally reasonable model should be described by conformation angles, the torsion angles should be stable within a few degrees.

1.6. REFERENCES

1. Watson, J.D. and Crick, F.H. Nature, **1953**, 171, 737.

2. Wilkins, M.H.F. **1963**, 140, 941.

3. Saenger, W. Principles of Nucleic Acid Structure 1-556 (Springer-Verlag, Berlin, 1984).

4. Pezzano, H. and Podo, F. Chem. rev., **1980**, 80, 365.

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

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

7. Waring, M.J. Annu. Rev. Biochem, **1981**, 50, 159.

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

9. Kornberg, A. DNA synthesis (W. H. Freeman, San Franciso, 1974).

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

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

12. Barton, Lawrence and Teeter. 1980. In Metal ion-nucleic acid interactions, eds.,

13. Teeter, M.M., Quigley, G.J. and Rich, A. 1980. In Nucleic Acid-Metal Ion Interactions, Spiro, T.G. eds., John Wiley and Sons. New York.

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

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

16. 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.

17. Behe, M. Proc. Natl. Acad. Sci. U.S.A., **1981**, 78, 1619.

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

19. Dickerson, R.E., Drew, H.R. and Conner, B. 1981. In Meeting, Albany, N.Y., USA, April 26-29, 1981., Sarma, R.H. eds., Adenine Press. Guilderland, N.Y., USA.

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

21. 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).

22. Baguley, B.C. Anti-Cancer Drug Design, **1991**, 6, 1.

23. Pigram, W., Fuller, W. and Hamilton, L. Nature New Biol., **1972**, 235, 17.

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

25. Arcamone, F. In "x-ray crystallography and drug action" (Oxford Univ. press, 1984).

26. Chen, K.X., Gresh, N. and Pullman, B. J. Biomol. Struct. Dyn., **1985**, 3, 445.

27. Lerman, L.S. J. mol. Biol., **1961**, 3, 18.

28. Chen, K.X., Gresh, N. and Pullman, B. Mol. Pharmacol., **1986**, 30, 279.

29. Hui, X., Gresh, N. and Pullman, B. NAR, **1990**, 18, 1109.

30. Pullman, B. Adv. in Drug research, **1979**, 18, 113.

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

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

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

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

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

36. Berman, H.M., Stallings, W., Carrell, H.L., *et al.* Biopolymers, **1979**, 18, 2405.

37. Schwartz, H.S. Adv. Cancer Chemother, **1979**, 1, 1.

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

39. Dickerson, R.E. Nucleic Acids Res., **1989**, 17, 1797.

40. Hol, W.G.J. Nature, **1978**, 273, 443.

41. Neidle, S. Methods in enzymology, **1991**, 203, 433.

42. Karplus, M. and Pestko, G.A. Nature, **1990**, 347, 631.

43. Martin, Y.C. Methods in enzymology, **1991**, 203, 587.

44. Levitt, M. Proc. Natl. Acad. Sci. USA, **1978**, 75, 640.

45. Weiner, S.J., Kollman, P.A., Case, D.A., *et al.* J.A.C.S., **1984**, 106, 765.

46. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., *et al.* J. Comp. Chem., **1983**, 4, 187.

47. Karplus, M., III, C.B.B. and Brunger, A. 1986. In Current Communications in Molecular Biology: Computer Graphics and Molecular Modeling; Meeting, Cold Spring Harbor, N.Y., USA, De c. 10-13, 1985. Ix+145p, Fletterick, R.A.M.Z. eds., Illus.

48. Brunger, A.T. X-PLOR Manual, Version 2.1 (Yale University, 1990).

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

50. Glusker, J.P. and Trueblood, K.N. Crystal Structure Analysis (Oxford University Press, New York, Oxford, 1985).

51. Rhodes, G. Crystallography made crystal clear (Academic Press, Inc., 1993).

Chapter 2

CRYSTALLIZATION OF FOUR-STRANDED DNA

2.1. INTRODUCTION

Four-stranded DNA (G4) has recently been discovered as a novel DNA structure that forms under near-physiological condition, in which the four strands are held together by guanine Hoogsteen base-pairing [1,2]. The structure has the characteristic of a stack of G4 (four guanines) planes, or G quartets. The adjacent strands can be in either parallel or antiparallel orientation [3,4]. This family of structures are interconvertible, and the formation of a given structure is dependent on the primary sequence, polynucleotide concentration, temperature, and the concentration and identity of the surrounding ions [5]. Precursors for this structure are repetitive DNA sequences with G-rich motifs consisting of at least three contiguous guanines. These sequences can be found in evolutionarily or functionally conserved regions, such as the telomere and the switch-regions of immunoglobulin heavy chain genes [3,6].

Four-stranded polymers of nucleic acids held together by guanine quartet were first observed in gels and fibers many years ago [7]. Recently, there is a revitalized interest in these structure, largely due to the demonstration, by means of chemical probing, that they can form in solution under physiological conditions [1]. More direct proof of these structures should come from x-ray crystal diffraction and NMR studies of DNA oligomers [8,9]. The main feature of the structure, i.e., guanine quartet, remains the same, although the orientation of adjacent sugar backbones varies considerably. So far, very few studies of these structures by other physical and chemical methods, such as IR and fluorescence, have been undertaken.

In vitro studies with oligonucleotide resembling these G-rich sequences show that the four stranded structures manifest distinct mobilities in non-denaturing gel electrophoresis that differ from their unfolded counterparts. G4 DNA generated from intramolecular folding, in which the adjacent strands are in antiparallel orientation, has a faster migration mobility than the unfolded precursor [1,2]. Those resulting from intermolecular folding, in which the adjacent strands can be either antiparallel or parallel, migrate much more slowly than the unfolded oligonucleotides [1,3]. The overall structure of G4 four-stranded DNA generated from either inter- or intra- molecular folding are quite similar, although they may differ in sugar puckering and other details. The formation of the four-stranded structure requires Na^+ ion, and the structure can be further stabilized by K^+ ion [5,10]. The existence of the G4 structure in the cell has not yet been observed, but the strong evolutionary conservation of G-richness in certain biologically important regions compels one to believe that this structure does play a role in vivo [11].

In the past two years, several laboratories have begun to look for enzymes or proteins that interact with G4 DNA, in the hope that a knock out of these proteins in the cell will lead to some genetic lesion and thus provide clues on the function of these structure in vivo. W. Gilbert's lab at Harvard is working on several yeast proteins [2]. T. Cech's lab at the University of Colorado at Boulder, has found that a telomere binding protein from Oxytrticha can facilitate tetrameric G-rich oligonucleotide

folding into G4 DNAs [3]. Our lab was working on crystallization of the four-stranded DNA to obtain the detailed structure of G4.

2.2. METHOD

2.2.1. PURIFICATION OF OLIGONUCLEOTIDES

Two G-rich telomeric sequences were synthesized, d(TTGGGGTT) and d($T_2G_4T_4G_4T_2$) in the RCMI Sequence and Synthesis facility by Dr. Stephen Bobin using the phosphotriester chemistry method [12]. The oligonucleotide was purified by reversed-phase HPLC on a C18 column (4.6mm × 25cm), Flow: 1ml/min, Mobile phase: 30 min 10% to 35% of 30% acetonitrile/water in 0.1 TEAB. The major peak was collected and concentrated by lyophilization. According to the analytical HPLC, the purity was 99% for both samples.

2.2.2. FORMATION OF FOUR-STRANDED DNA

An important feature of the formation of four-stranded DNA is that G4-DNA complexation is promoted by sodium and inhibited by potassium [1]. The procedure for formation of G4-DNA is as follows:

1). Dissolve Oligomers d(TTGGGGTTTTGGGGTT) and d(TTGGGGTT) in 10μl of the TE buffer (10μl of 10mM Tris, pH 8.0, 1mM EDTA) at monomer DNA concentration 1mM.

2). Heat at 95^0 for 120s in a sealable eppendorf tube and chill on ice.

3). Add 10µl of TE + 1.9M NaCl + 0.1M KCl.

4). Draw up the solution into the central portion of a thin-walled 50µl capillary or micropipette.

5). Incubate the sample at 60^0 for 40 hours.

6). Cool to room temperature, release DNA solution into clean eppendorf tube, vortex briefly to collect liquid at the bottom of tube.

The G4-DNA is stable at room temperature and is not in rapid equilibrium with its single stranded form.

2.2.3. PURIFICATION OF G4

The G4-DNA $d(TTGGGGTT)_4$ was purified by reversed-phase HPLC on a C18 column (4.6mm × 25cm), Flow: 1ml/min, Mobile phase: 30 min 10% to 35% of 30% acetonitrile/water in 0.1 TEAB. The major peak was collected and concentrated by lyophilization. According to the analytical HPLC, the purity was 98%. The DNA stock solution was 3.26 mg/ml and was stored at -4°C.

2.2.4. CRYSTAL GROWTH

Since the conditions for growth of G4-DNA crystals was unknown, the statistically designed "factorial" experimental method for screening crystal growth conditions was used [13].

The factors we consided were pH, temperature, various mono- and divalent ions, concentration of the G4, and preciptating agent. Each factor divided to a few levels (pH=7.0, 8.0; room temperature and 4^0C). 20 difference combinations of the factors were set for crystallization.

2.3. RESULTS AND DISCUSSION

None of the crystallization settings gave good quality single crystal for x-ray diffraction. According to the crystal structure of four-stranded Oxytricha telomeric DNA (sequence d(GGGGTTTTGGGG)) [8], the crystals were grown from a solution that contained 1mM oligonucleotide, 10mM $MgCl_2$, 6mM spermine, 40 mM KCl, 20mM potassium cadodylate buffer (pH 7.0), 5% MPD equilibrated with a reservoir of 40% MPD. This condition is almost identical with one of our set up conditions although the sequence we used was different. Sometimes crystallization is an art more than science. We decided to discontinue the project as a result of the published x-ray crystal structure of four-stranded DNA from A. Rich' lab [8].

2.4. REFERENCES

1. Dipankar, S. and Gilbert, W. Nature, **1988**, 334, 364.

2. Dipankar, S. and Gilbert, W. Nature, **1990**, 344, 410.

3. Zahler, A.M. and Prescott, D.M. Nature, **1991**, 350, 718.

4. Rajagopal, P. and Feigon, J. Nature, **1989**, 339, 637.

5. Guo, Q., Lu, M. and Kallenbach, N.R. J. A. C. S., **1992**, 2451.

6. Kim, J., Cheong, C. and Moore, P.B. Nature, **1991**, 351, 331.

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

8. Kang, C., Zhang, X. and Rich, A. Nature, **1992**, 356, 126.

9. Smith, F.W. and Feigon, J. Nature, **1992**, 356, 164.

10. Oka, Y. Nucleic Acids Res., **1987**, 15, 8877.

11. Dipankar, S. and Gilbert, W. Methods in Enxymology, **1992**, 211, 191.

12. Voet, J.G. Biochemistry (John Wiley & Sons, Inc., 1990).

13. Murchie, A.I.H., Carter, W.A., Portugal, J. and Lilley, D.M. Nucleic Acids Res, **1990**, 18, 2599.

Chapter 3

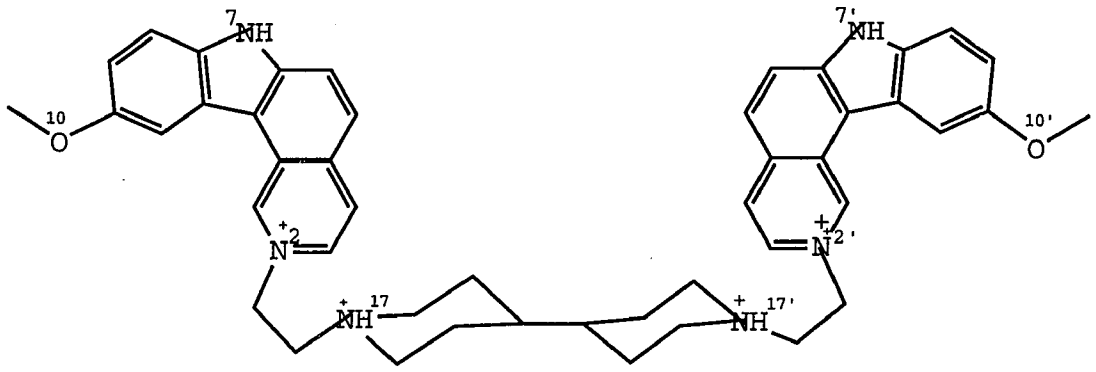
MODELING THE DITERCALINIUM-DNA INTERCALATION BY MOLECULAR MECHANICS

3.1. ABSTRACT

A molecular model has been derived for the bis-intercalation of ditercalinium into the central CpG sites of the dodecamer duplex of d(CGCGCGCGCGCG). The initial geometry of the model is from a previous crystallographic study of the ditercalinium-d(CGCG) complex [1]. The model's three dimensional structure has been investigated by means of a theoretical approach for modeling the conformational distortion of nucleic acids caused by the drug intercalation. The XPLOR program with Powell minimization was used to evaluate the internal energy of the drug-DNA complex and the interaction energies among the residues of the complex. In the best energy-minimized and stereo-chemically acceptable structure, the piperidinium chains of ditercalinium are located in the major groove of the right-handed oligonucleotide. Calculations show a distortion of helical axis, a local unstable interaction around the intercalation sites, and partially unstacked base pairs. The local and global geometric parameters and the interactive energies of the dodecamer complex were compared with the standard B-DNA structure and with crystal structure of the ditercalinium d(CGCG) complex. The proposed structure of the dodecamer complex is in good agreement with reported x-ray crystal structure data, although some structure features are different in the two structures. Following will be an account of the similarities and differences found by the interpretation of these studies.

3.2. INTRODUCTION

Ditercalinium, 2,2'-[4,4'-bipiperidine-1,1'-bis-(ethane-1,2-diy)]bis(10-methoxy-7H-pyrido[4,3-c]carbazolium) tetramethane sultanate is a synthetic bis-intercalating anti tumor agent which was discovered by Roques and Le Pecq in a systematic quest for improved chemotherapeutic agents [2,3]. The drug is made up of two identical 7H-pyrido[4,3-c] carbazole rings linked by symmetrical, rigid bipiperidine chains. The dimers bisintercalate into DNA with affinities as great as $10^{14}M^{-1}$ [4]. The drug contains four positive charges, two from the piperidinic nitrogens and two from the pyridinic nitrogens.



Ditercalinium

Figure 3-1. Ditercalinium, identical dimers linked by symmetrical rigid linker chains.

Ditercalinium is unique among known chemotherapeutic agents in that it causes cell death by an unprecedented mechanism. Unlike most other DNA binding chemotherapeutic agents, the mechanism of action of ditercalinium does not involve inhibition of the DNA replication or transcription processes. Instead the DNA-ditercalinium complexes provoke a malfunction of the DNA repair systems [5,6,7]. The drug induces SOS functions in certain

Escherichia coli strains [5]. The drug-DNA complex is recognized in vitro as a bulky covalent DNA lesions by the uvr ABC endonuclease repair system of Escherichia coli [6]. This observation led to the proposal [5] that such properties could stem from an undefined distortion of DNA structure by the bisintercalative binding of ditercalnium. The study of the three-dimensional structure of the DNA-ditercalinium complex should help us understand the unusual activity of this drug and its mode of intercalation.

The structure of ditercalinium-d$(CGCG)_2$ has been determined using x-ray crystal diffraction analysis at 1.7Å resolution. The space group is $P4_12_12$ with unit cell dimensions a=b=26.88Å and c=82.60Å [1]. The ditercalinium molecule bis intercalates at the two CpG steps of the DNA and the linker is located in the major groove. The geometric analysis revealed several interesting results, such as the complete unstacking of terminal cytosines and a 15^o kink at the site of intercalation. It has been suggested that the kink serves as the signal for induction of Escherichia coli uvr ABC repair pathway [8].

The structural information provided by x-ray crystallography may not be truly representative of the solution structure because of the short oligonucleotide sequence used in the crystallography. In the crystal structures of intercalating drug-DNA complexes the drug has always been found to intercalate near the end of the DNA sequence. The present study investigates the influence of the structural distortions observed in the crystal on an extended sequence of DNA using molecular modeling. In this extended DNA complex the drug is located away from the ends of the DNA sequence. The study will address two questions:

- What are the structural features of the ditercalinium-DNA complex that distinguish it from other chemotherapeutic agents
- Does the drug-induced distortion in the crystal structure persist in a longer DNA sequence?

For this purpose we used molecular mechanics to elucidate the DNA conformation and interaction energies.

3.3. METHODS

The molecular modeling begins by choosing or constructing an initial model. The model is then refined to its lowest energy conformation. Finally, the refined model is studied by geometric and energetic analysis.

3.3.1. MODEL BUILDING

The initial dodecamer complex model, ditercalinium-d(CGCGCGCGCGCG)$_2$ was constructed using the XPLOR [9] and HyperChem programs (HyperChem, 1992). Standard B-form DNA pieces were generated using HyperChem (HyperChem, 1992) and these segments were linked to the x-ray crystal structure of the drug complex [1]. The crystal structure of the ditercalinium-d(CGCG)$_2$ intercalation complex was used as a starting point. Two computer synthesized d(GCGCGC)$_2$ segments with the standard B-DNA conformation were linked on either end of the complex by superposition of the end base pairs of the complex and the segment. A least-squares fit was performed to make a better fit of the junction bases and then

the extra pair of the bases that overlapped at the ends of the segments was deleted. A double helical dodecamer DNA-ditercalinium complex, was obtained. **Figure 3-2** is a schematic diagram showing the model building and the residue numbers.

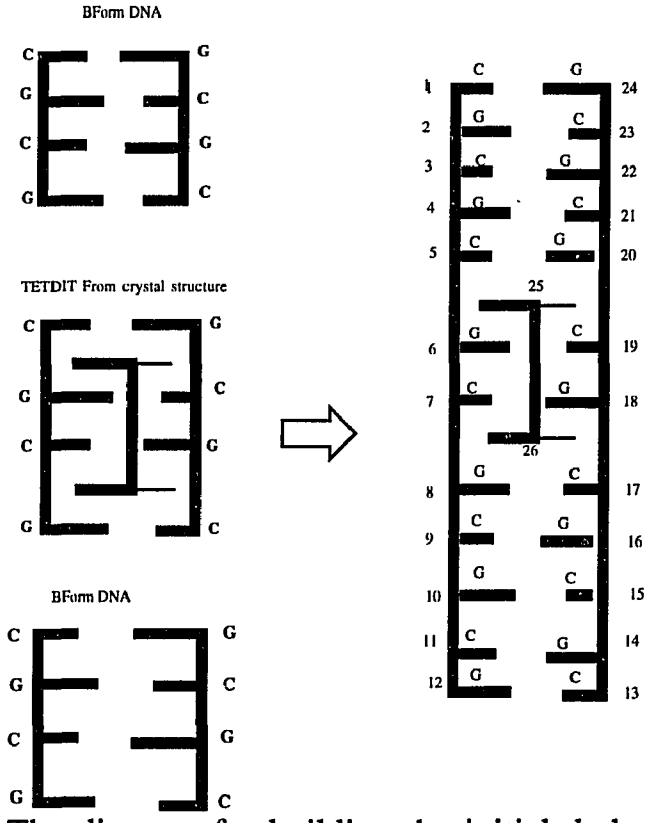


Figure 3-2. The diagram for building the initial dodecamer-ditercalinium complex. The residue numbers from 1 to 24 are DNA bases, and residues 25 and 26 are the two identical halves of ditercalinium.

3.3.2. ENERGY MINIMIZATION

The conformational energy expression employed in the program XPLOR for energy minimization consists of a series of component energies as shown below [9]:

$$E = \sum K_r(r - r_0)^2 + \sum K_\theta(\theta - \theta_0)^2 + \sum V_n/2 [1 + \cos(n\phi - \gamma)]$$
$$+ \sum [A_{ij} / R_{ij}^{12} - B_{ij} / R_{ij}^{6}] + C\sum q_i q_j / \varepsilon R_{ij}$$

where the summations represent the potentials for bond length stretch, bond angle distortion, torsional rotations, van der Waals and electrostatic interactions. Partial charges for ditercalinium were obtained from AM1 calculations. In the XPLOR minimization, the dielectric constant was four and distant independent, the cut-off of the van der Waals was 7.5Å, the cut-off of the electrostatic potential was 8.5Å, the factor for special 1-4 electrostatic interactions was 1.0 (here C= e_{14} = 1.0 see equation-20 from Brunger, 1990). Solvent effects were not explicitly considered in the calculation. Thus, the model complex is in vacuum environment. The model was energy minimized beginning with weak constraints on atoms very near the break and the kinking phosphates and with strong constraints on the rest of the complex. During the energy minimization residues 4 to 9, 16 to 21 and 25 to 26 were highly constrained by a harmonic force of 75 kcal/Mol/Å^2. A partially relaxed structure that constrains the crystal structure and the linking parts but does not constrain the other residues is obtained. After energy minimization of the partially relaxed structure, repeated refinements of the structure is achieved by slowly reducing the constraints to zero. **Figure 3-3** shows the energy minimized ditercalinium-d$(CG)_{12}$ complex structure.

The energy of interactions between drug and DNA bases, between parallel base pairs, and between adjacent base pairs were analyzed using XPLOR [9]. The geometrical parameters were calculated using Curves, a

program to calculate the helical parameters and axis of curvature for irregular nucleic acids [10]. The results are discussed in below.

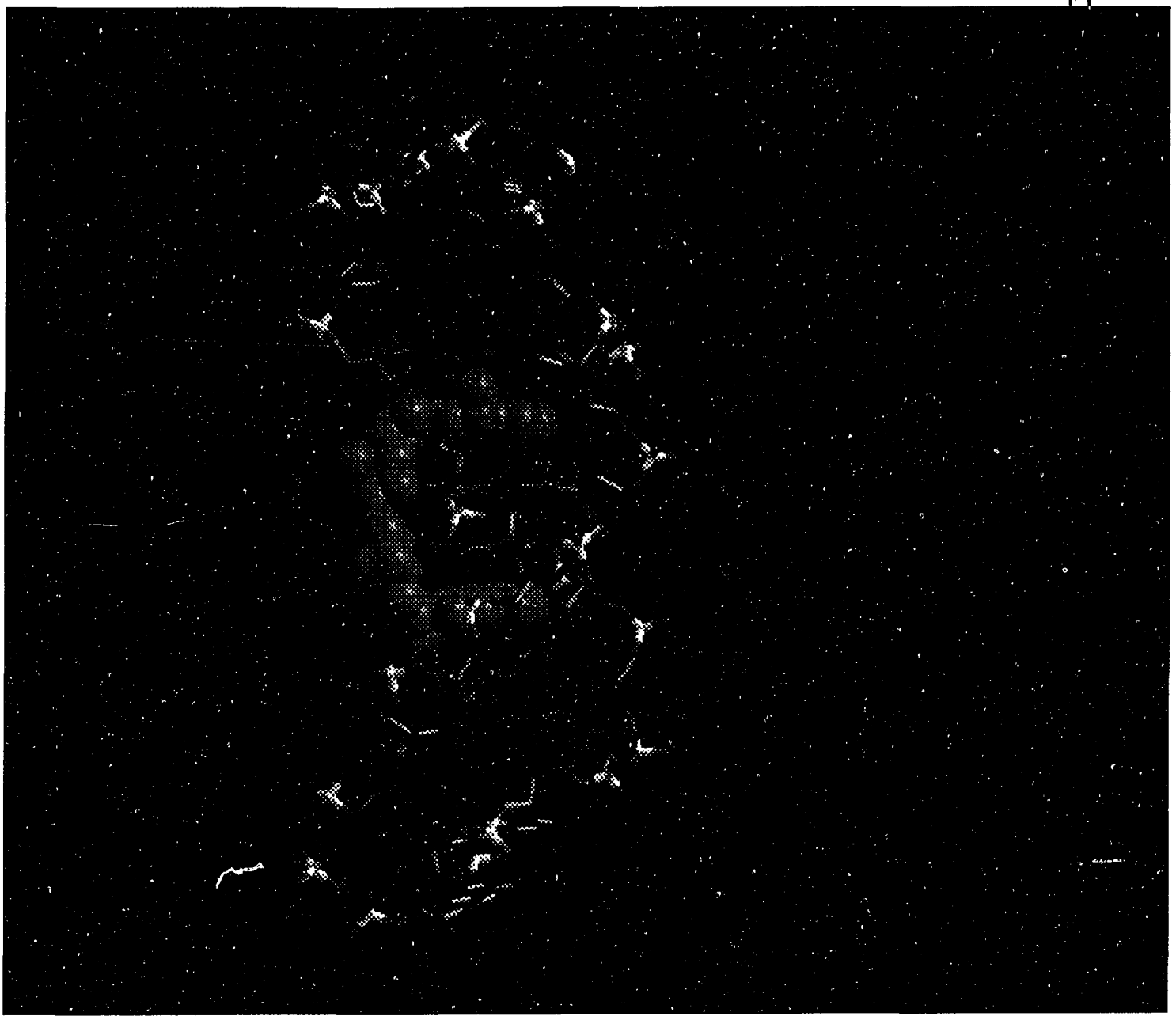


Figure 3-3. View of the ditercalinium-d$(CG)_{12}$ complex.

3.4. RESULTS AND DISCUSSION

In the final dodecamer complex model, the ditercalinium molecule remains bisintercalated at the two CpG steps of the DNA fragment and the linker lies in the major groove. The geometric analysis shows that the drug intercalation causes distortions on the DNA conformation such as the bend of the helix axis, the partial unstacking of the bases, and an asymmetry of the complex. The distortions of the DNA conformation result in an increase in the interaction energies among the nucleic acid residues. This energetic increase is offset by van der Waals and electrostatic interaction between the nucleic acid bases and the chromophores of the drug.

3.4.1. THE STABILIZATION FACTORS

The energy analysis shows that ditercalinium-DNA interaction is the major source of stabilization of the complex conformation. The intercalation increases the energy of the DNA and the drug by 80.4 and 24.4 kcal/Mol, respectively while the energy of the drug-DNA interaction decreases by 144.2 kcal/Mol. Thus the net change in energy upon intercalation is 39.6 kcal/Mol. **Figure 3-4.** illustrates the energy analysis of the chromophore-DNA interactions. Although only the chromophore 25-DNA interaction is showed on the figure, the chromophore 26-DNA interaction exhibits similar energetics. As the figure shows, the drug chromophores interact only with nearest three DNA base pairs and the interactions are asymmetric in energy for the two strands of the DNA.

Figure 3-4. also shows the energy decomposition of the chromophore 25-DNA interaction. The van der Waals (VDW) interactions and the

electrostatic interactions are the two primary components responsible for the stabilization energy.

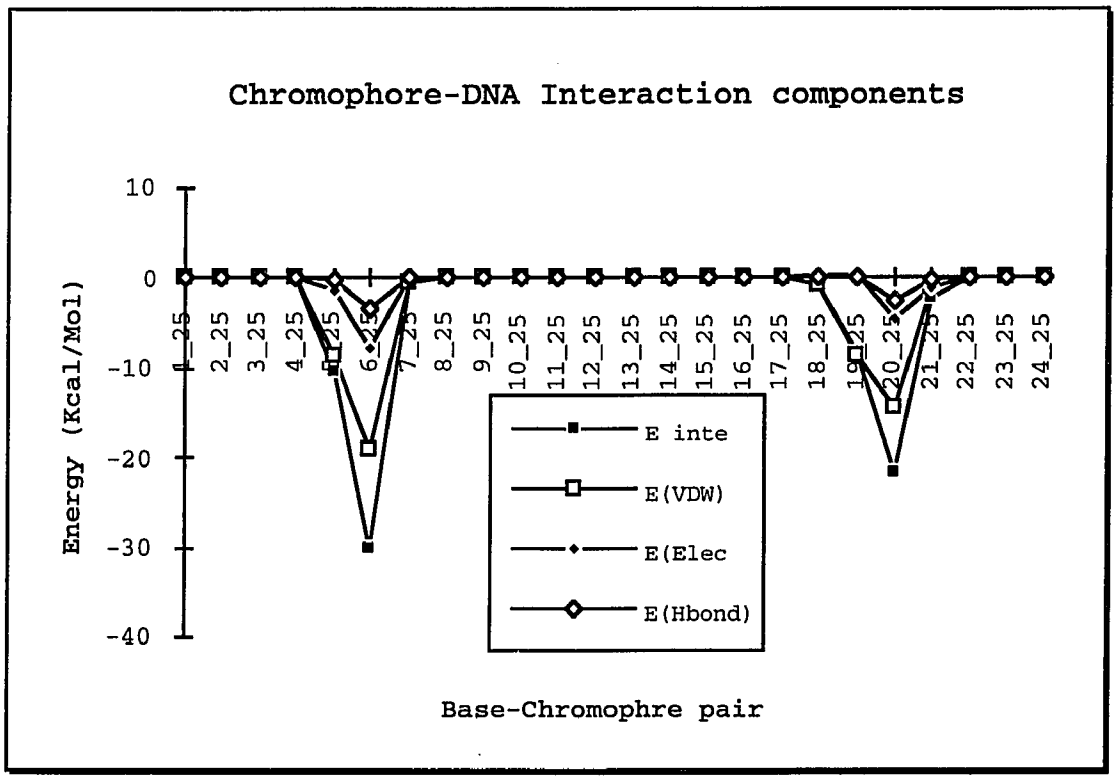


Figure 3-4 . The energy components of chromophore-DNA interactions.

3.4.1.a. van der Waals Interaction

The ditercalinium molecule forms 84 VDW contacts (**Table 3-1**) with the DNA fragment (the criterion for a contact is an inter-atomic distance of less than 3.5Å). The VDW contacts between the rigid linker chains and DNA fragment, and between the chromophores and the bases aid in stabilizing the structure. The ditercalinium molecule does not form hydrogen bonds to cytosines.

A striking feature is the selective interaction between the drug and guanine residues. Very few interactions are observed with the cytosine residues. This result is confirmed by the x-ray structure. Of the 84 VDW contacts, about 3/4 of them are with guanines and 1/4 are with cytosines.

An asymmetry is observed between the interaction of the two identical drug chromophores with the DNA. The chromophore interact with only the nearby eight bases (C5, G6, C7, G8, C17, G18, C19, and G20). N7 from both chromophores interact with G8 and G20 to form eight VDW contacts. N^+ and $N^{+'}$ also make contacts with DNA bases but the N2 (Dit25) is not involved. Chromophores 25 interacts with a phosphate group but the other one (Dit26) does not. Chromophore 25 is responsible for a greater reduction in energy.

Although the linker is hydrophobic, the interactions of the linker with the DNA do not appear to contain a large hydrophobic contribution. Most contacts are to the hydrophilic positions of the DNA, either O6 or N7. However, the calculation did not consider the solvent effects which is necessary while consider the hydrophobic effects. The hydrophobic interactions may make a significant contribution both to the base-chromophore and base-linker interactions in solvent environment.

Table 3-1 van der Waals contacts between DNA and ditercalnium

Residue	Atom		Dit atom		Distance	Residue	Atom		Dit atom		Distance
6	GUA O2P	25	DIT	HB1	2.161	18	GUA O1P	26	DIT	C3	3.214
6	GUA O2P	25	DIT	CB	3.162	18	GUA O5'	26	DIT	C3	3.400
6	GUA O2P	25	DIT	HC1'	2.906	18	GUA O5'	26	DIT	C4	3.061
6	GUA O4'	25	DIT	C4	3.236	18	GUA O4'	26	DIT	C4	3.129
6	GUA C4	25	DIT	C4	3.381	18	GUA O4'	26	DIT	C5	3.028
6	GUA C4	25	DIT	C12	3.463	18	GUA O4'	26	DIT	C12	3.401
6	GUA N3	25	DIT	C5	3.227	18	GUA N9	26	DIT	C4	3.417
6	GUA N3	25	DIT	C12	3.272	18	GUA C4	26	DIT	C12	3.443
6	GUA C2	25	DIT	C5	3.419	18	GUA C2	26	DIT	C13	3.387
6	GUA C2	25	DIT	C6	3.461	18	GUA N1	26	DIT	C16	3.330
6	GUA N2	25	DIT	C6	3.229	18	GUA C6	26	DIT	C17	3.453
6	GUA N2	25	DIT	C13	3.376	18	GUA O6	26	DIT	CD	3.077
6	GUA N1	25	DIT	C16	3.419	18	GUA C5	26	DIT	C17	3.458
6	GUA N1	25	DIT	C17	3.489	18	GUA N7	26	DIT	N2	3.460
6	GUA O6	25	DIT	HC2	3.267	18	GUA N7	26	DIT	C3	3.495
6	GUA O6	25	DIT	CD	3.145	18	GUA N7	26	DIT	N	2.991
6	GUA N7	25	DIT	N	2.880	18	GUA C8	26	DIT	C3	3.441
6	GUA N7	25	DIT	CD	3.383	18	GUA C8	26	DIT	C4	3.428
6	GUA N7	25	DIT	CD'	3.325	20	GUA O4'	25	DIT	N7	2.956
6	GUA N7	25	DIT	HC1'	3.489	20	GUA N9	25	DIT	N7	3.339
6	GUA N7	25	DIT	CC'	3.421	20	GUA C4	25	DIT	N7	3.272
8	GUA O4'	26	DIT	N7	2.849	20	GUA N3	25	DIT	C13	3.446
8	GUA N9	26	DIT	N7	3.377	20	GUA C2	25	DIT	C6	3.494
8	GUA C4	26	DIT	N7	3.168	20	GUA C2	25	DIT	C13	3.447
8	GUA N3	26	DIT	N7	3.459	20	GUA N2	25	DIT	C5	3.362
8	GUA C2	26	DIT	C6	3.320	20	GUA N2	25	DIT	C6	3.471
8	GUA C2	26	DIT	C13	3.461	20	GUA C5	25	DIT	C14	3.470
8	GUA N1	26	DIT	C13	3.385	20	GUA N7	25	DIT	C8	3.404
8	GUA C5	26	DIT	N7	3.367	20	GUA C8	25	DIT	C8	3.360
8	GUA C5	26	DIT	C14	3.376						
8	GUA N7	26	DIT	C8	3.279						
8	GUA C8	26	DIT	C8	3.421						

Table 3-1(cont.) van der Waals contacts between DNA and ditercalinium

Residue		Atom	Dit		atom	Distance
5	CYT	C2	25	DIT	C3	3.468
5	CYT	O2	25	DIT	C3	3.417
5	CYT	O2	25	DIT	C4	3.212
5	CYT	O2	25	DIT	C12	3.500
7	CYT	N1	26	DIT	C8	3.465
7	CYT	C6	26	DIT	C9	3.320
7	CYT	C2	26	DIT	C8	3.354
7	CYT	C4	26	DIT	C9	3.495
7	CYT	C4	26	DIT	C10	3.281
7	CYT	C5	26	DIT	C9	3.285
7	CYT	C5	26	DIT	O	3.387
7	CYT	C5	26	DIT	C10	3.357
17	CYT	O1P	26	DIT	CA	3.286
17	CYT	N3	26	DIT	C12	3.397
17	CYT	N3	26	DIT	C17	3.366
17	CYT	C4	26	DIT	C1	3.317
17	CYT	C4	26	DIT	N2	3.405
17	CYT	N4	26	DIT	C1	3.384
17	CYT	C5	26	DIT	N2	3.471
17	CYT	C5	26	DIT	CA	3.413
19	CYT	N3	25	DIT	C14	3.494
19	CYT	C5	25	DIT	C9	3.466

3.4.1.b. Electrostatic Energies

The positively charged nitrogens of ditercalinium are located near the floors of the major grooves of the DNA. Favorable charge-charge interaction within the grooves between the chromophores of ditercalinium and the bases of the DNA appear to be an important factor in the stability of the complex as shown by the interaction energy between the drug and the DNA (**Figure 3-4.**). It is interesting to note that the positive charges of the drug are not located near the negatively charged sugar-phosphate backbone (**Table 3-1**).

3.4.2. LOCAL INSTABILITY-BASE STACKING

A comparison of the dodecamer complex with the B-form DNA reveals information about base stacking perturbations upon complexation (**Figure 3-5.**) and the consequent increase in energy among the interactions of the DNA residues. The intercalation sites exhibit the highest energy and the increase of the energy gradually decreasing as we move away from the intercalation sites. In other words, the perturbation is most substantial around the intercalation sites energetic clays persist along the helix axis.

Although the intercalation sites exhibit the highest energy, these sites do not show the biggest difference in energy when compared to the standard B-form. The interaction energy difference between the dodecamer and the B-form is plotted in **Figure 3-6.** The greatest energy differences are at G6/C7 and G18/C19, which are the base pairs in between the two bound chromophore planes and within the intercalation sites. It is interesting that the largest base-base stacking distortion does not occur at the interactions sites but at the adjacent base pairs. The rigidity of the linker of ditercalinium may

be responsible for this effect. If the chromophores of ditercalinium overlap as much as possible with G6 and G18, the rigid linker may have to more or less distort the stacking with C7 and C19. The rigid linker limits the flexibility of the base pairs G6/C7 and G18/C19, and the adjacent bases G6-C7 as well as G18-C19 are pushed. The distortion at these sites cause local instability of DNA which might induce a different DNA deformation for the uvr ABC repair system of Escherichia coli to recognize.

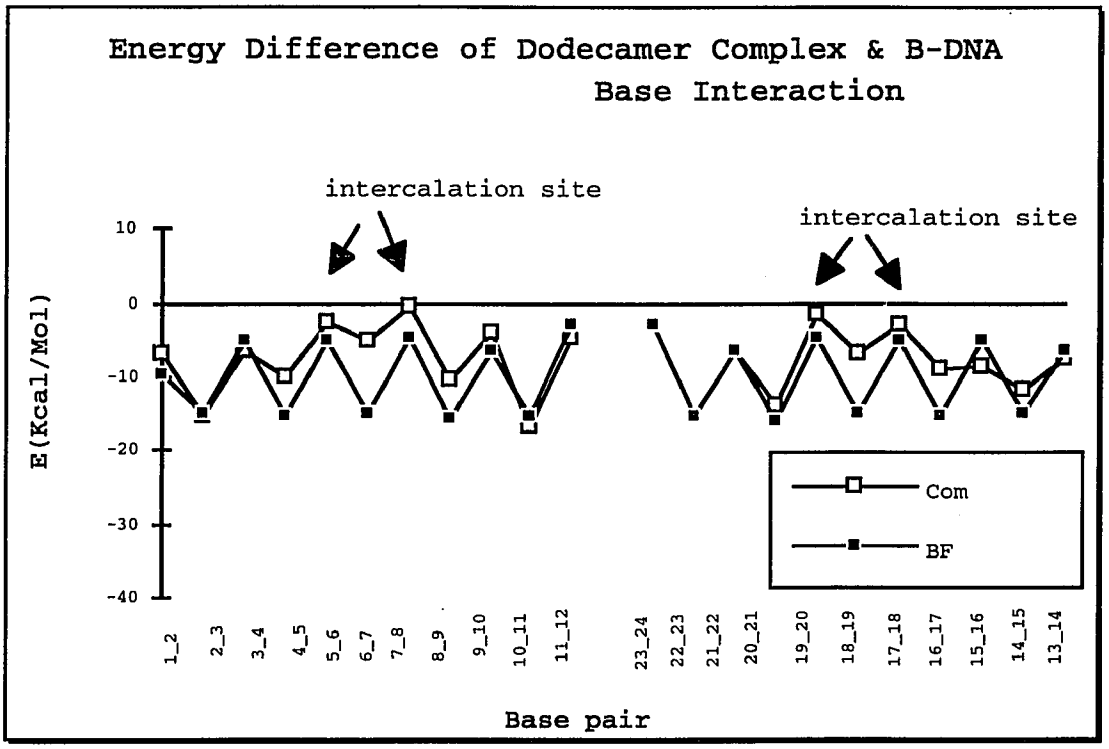


Figure 3-5. The comparison of base stacking interaction between the dodecamer complex and B-form DNA. The open square represents the dodecamer model and the solid square represents the standard B-DNA.

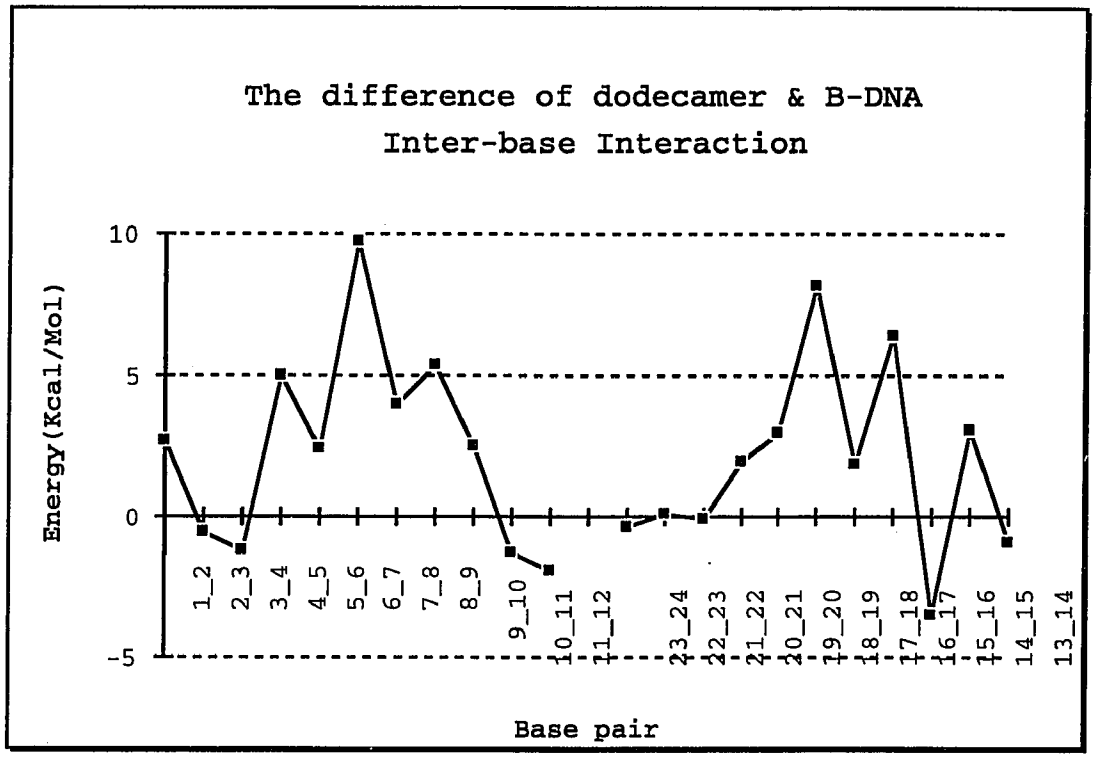


Figure 3-6. The base-base interaction difference of dodecamer complex and B-form DNA.

3.4.3. ASYMMETRY OF THE COMPLEX

The base-base interaction energy of the two strands of the DNA in the complex, plotted in **Figure 3-7**, shows that the two anti-parallel strands in the complex are not equivalent. Both geometric and energetic analysis show that the two chromophores of ditercalinium are in different environments and the two anti-parallel strands in the complex are conformationally non-equivalent. This result is in agreement with the x-ray crystallographic study but different from many other drug-DNA intercalation studies which shows that the two anti-parallel strands are conformationally equivalent, whether or not they are of self-complementary sequence [11,12]. Chinomycin, triostene-A, qamino acridine are all symmetrical drugs which form asymmetrical complexes. The

ditercalinium is a dimer molecule made up of two identical rings linked by symmetrical linking chains. The symmetrical ligand introduces asymmetry into its complex structures. It raises the question, unresolved by this study, what is the role of the symmetry in the structure/function relation, and can this asymmetric character be recognized by the Escherichia coli SOS repair system?

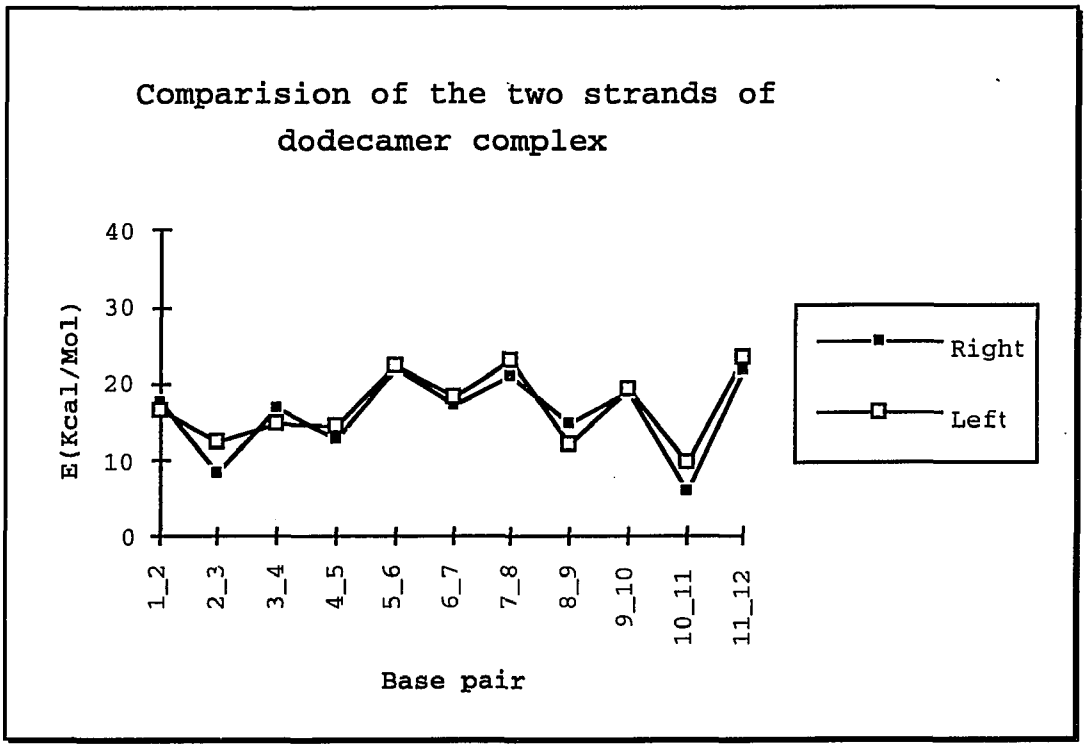


Figure 3-7. Energy vs. base pair for both DNA strands of the dodecamer complex. The two strands are not equivalent.

3.4.4. BACKBONE AND SUGAR PUCKER

The ditercalinium molecule slides between the nucleic acid base-pairs without disturbing the overall stacking pattern. However, because the base-pairs must separate vertically at the sites of intercalation, the sugar-phosphate

backbone is altered. The torsion angles and sugar puckers of the dodecamer complex are listed in **Table 3-2**. The intercalation causes the torsion angles and sugar puckers to change substantially with the most dramatic changes occurring at the drug intercalation sites. The α torsion angles are increased at the C3 and G4 sites; the β angles change significantly at G6; the γ of C5 increased more than three times compared to the others; the ε of the G18 is decreased to half. The regular helical structure is destroyed. The sugar puckering does not follow the C2'-endo for B-form but demonstrates various modes (C2'-exo as shown in Table 3-2).

One notable feature is the different conformational states of the C1'-N backbone (χ) torsion of the two strands. These angles would be expected to fall in the same conformational range if the symmetric dimer interacted with the DNA equivalently. The difference in the behavior of the backbone conformations of the two strands further demonstrates that the symmetric drug molecule intercalates in an asymmetric manner to form the drug-DNA complex, similar to the x-ray structure.

Table 3-2 The backbone torsion angles of dodecamer complex

Residue		χ	γ	δ	ε	ζ	α	β	Pucker
CYT	1	-130.51	.....	148.43	-171.55	-79.70	-76.90	171.86	C2'-endo
GUA	2	-119.54	59.10	136.39	177.25	-121.78	-65.87	-168.83	C2'-endo
CYT	3	-106.96	61.68	151.47	-174.47	-79.91	-95.51	170.72	C2'-endo
GUA	4	-96.36	57.54	117.16	-168.32	-67.36	-95.19	54.46	C1'-exo
CYT	5	-174.57	176.44	89.80	-153.79	-66.12	-63.55	-125.40	C2'-exo
GUA	6	-75.01	68.78	156.12	178.27	-177.50	35.81	66.98	C3'-exo
CYT	7	-115.20	58.67	107.79	153.93	-79.61	-72.41	-111.31	C1'-exo
GUA	8	-78.73	58.52	143.69	160.43	-91.53	-64.33	175.80	C2'-endo
CYT	9	-106.90	62.44	132.54	-160.25	-174.56	-51.49	132.65	C1'-exo
GUA	10	-128.82	61.65	147.37	172.63	-99.27	-78.56	-154.80	C2'-endo
CYT	11	-105.76	59.37	151.63	163.50	-95.58	-75.29	-161.02	C3'-exo
GUA	12	-103.38	57.30	152.53	.....	.....	.....	.....	C2'-endo
GUA	24	-95.97	60.38	158.25	.....	.....	.....	.....	C2'-endo
CYT	23	-107.37	62.35	150.52	129.97	-78.75	-73.22	-126.30	C3'-exo
GUA	22	-108.51	60.29	120.35	151.58	-97.18	-65.38	-162.90	C1'-exo
CYT	21	-95.34	60.71	146.26	179.41	-97.30	-70.78	167.98	C3'-exo
GUA	20	-94.06	60.07	122.12	104.15	-61.58	-75.39	-115.71	C1'-exo
CYT	19	-119.46	60.84	96.42	177.36	-81.38	-63.29	-155.22	O1'-endo
GUA	18	-76.42	67.63	156.53	-80.02	171.72	-76.39	108.57	C2'-endo
CYT	17	-170.15	59.01	85.21	-151.02	-70.74	-58.85	-135.70	C2'-exo
GUA	16	-109.41	57.27	131.79	-167.65	-159.05	-65.45	135.57	C1'-exo
CYT	15	-127.34	58.24	103.08	-176.97	-79.00	-72.97	169.44	O1'-endo
GUA	14	-128.81	58.48	112.72	175.50	-83.71	-73.09	-179.77	C4'-exo
CYT	13	-128.90	.....	148.89	-165.84	-77.88	-78.81	164.71	C3'-exo

3.4.5. THE END EFFECT

One of the most interesting findings from the modeling study is the interactions of the terminal cytosines with adjacent bases. In contrast the tetramer crystal structure terminal cytosines are "unstacked" with adjacent bases. The unstacking in the tetramer crystal structure is probably an end effect.

Figure 3-8. illustrates the comparison of the base-chromophore interaction energy of the dodecamer complex model and the tetramer crystal structure.

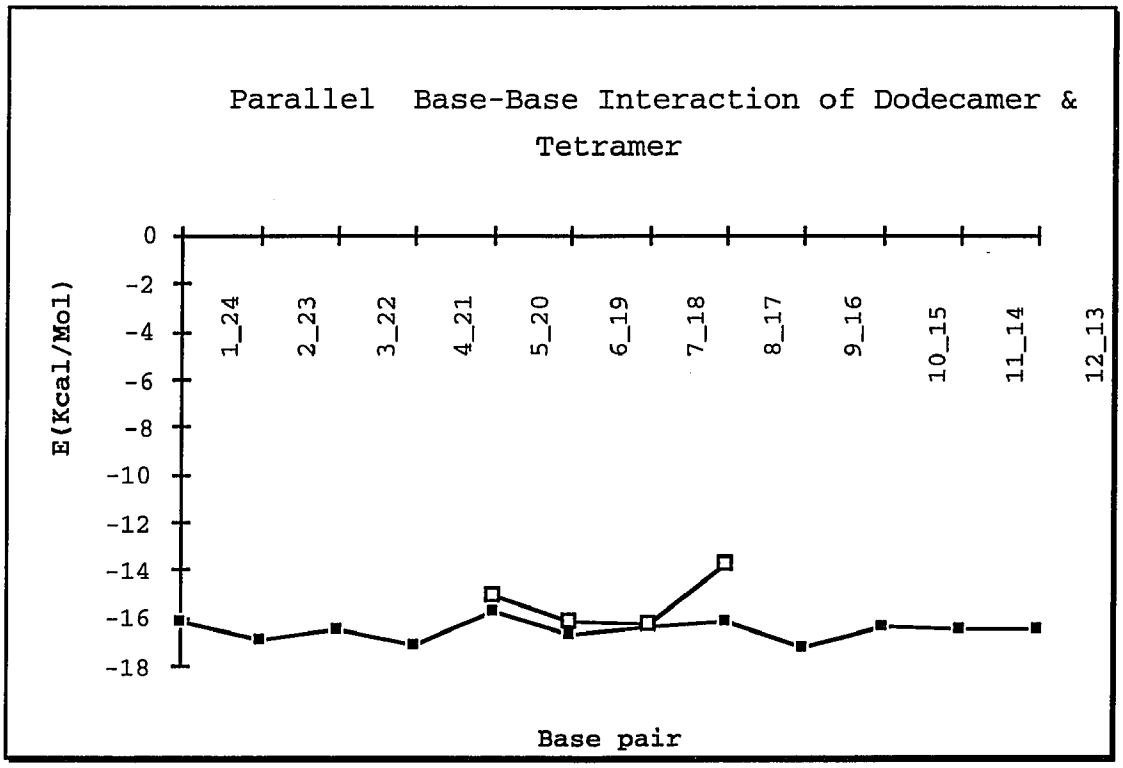


Figure 3-8. The comparison of parallel base-base interaction of the dodecamer complex and the crystal structure. The empty square represents the crystal structure and solid square represents the dodecamer model.

The "terminal cytosine" bases in the crystal structure are completely unstacked from the chromophores of ditercalinium (the terminal cytosines do not form a single VDW contact with this ditercalinium molecule). However, in the dodecamer complex model the feature is lost and there are van der Waals contacts between the drug and the cytosines (C5 and C17) as illustrated in **Table 3-1**. Consequently the base-base interaction energy at these sites is lower in the dodecamer complex model than those in the crystal structure as shown in **Figure 3-8** and **Figure 3-9**.

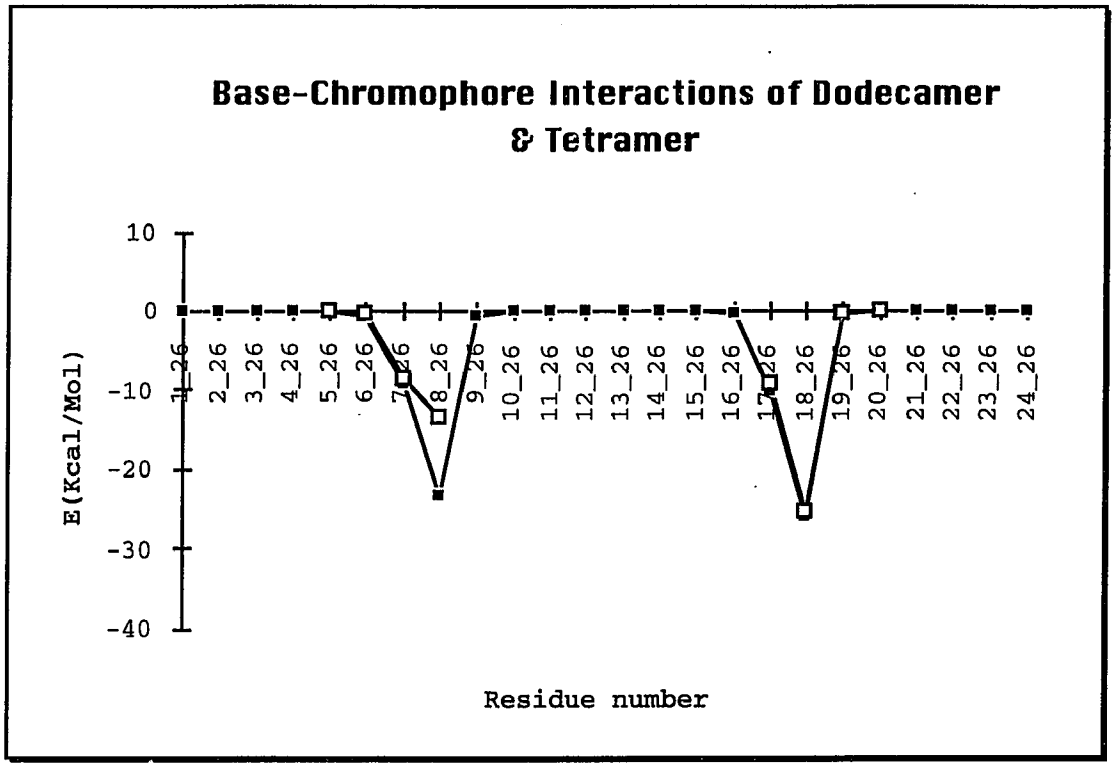


Figure 3-9. The comparison of ditercalinium-DNA interactions of dodecamer complex and the crystal structure. The chromophore of ditercalinium (Dit 26) has extensive interaction with guanines (G8 and G18) but much less interaction with cytosines (C7 and C17). The open square represents the crystal structure and the solid square represents the dodecamer model.

Other evidences to support the conclusion that the crystal structure has end effect are from the energetic and geometric analysis of the drug-DNA base interaction of the dodecamer complex (**Figure 3-8, Figure 3-9**), and the VDW interactions of the drug-DNA (**Figure 3-4, and Table 3-1**). These data all indicated that the ends of the tetramer complex should be similar. In fact, in the dodecamer complex the unstacking character is lost. This suggests that the "unstacking" observed from the crystal structure is not a significant

structure feature of the ditercalinium complex but is an end effect observed only in the crystal structure.

3.5. SUMMARY

This study suggests that both global and local conformational changes are involved upon ditercalinium intercalation in extended sequences. The structural features of the drug-DNA complex are the following: (1) The ditercalinium intensively interacts with the guanine residues but not the cytosine residues. (2) The intercalation of ditercalinium to the DNA causes long-range distortions of DNA conformation although the most significant changes occur around intercalation sites and the effects smooth out away from the sites. (3) The drug causes the DNA base-base interaction to have a local instability around the intercalation sites. (4) There is an apparent end effect in the crystal structure as indicated by the degree of the unstacking of the external terminal cytosines which is very different between the tetramer crystal structure and the dodecamer model. (5) An asymmetry is imposed by the intercalation of the drug. (6) Both the crystal structure and the model have about 15^0 bend at the same location.

Future Outlook

Further insight could be gained by studying a few ditercalinium analogs that are biologically inactive. These could be studied using molecular modeling and crystallography. Molecular modeling could be used to study the structural characteristics of active and inactive complexes, and these could be compared with the x-ray crystallographic results of the

structural distortions caused by active and inactive drug-DNA intercalation. The three-dimensional structural features of the two kinds of complexes (anti-tumor active and inactive) must be different. Systematic analysis and comparison of several of these complexes could help us understand more about ditercalinium. It is clear that longer DNA fragments for x-ray crystallography are desirable although it might be difficult to avoid the end effect problem.

3.6. REFERENCES

1. Gao, Q.I., Williams, L.D., Egli, M., *et al.* Proc. Natl. Acad. Sci. U. S. A., **1991**, 88, 2422.
2. Delaprat, D., Delbarre, A., Oberlin, R., *et al.* J. Med. Chem., **1980**, 23, 1336.
3. Pelaprat, D. and Pecq, J.B.L. J. Med. Chem., **1980**, 23, 1336.
4. Laugaa, P. **1985**, 24, 5567.
5. Lambert, B., Roques, B.P. and Le Pecq, J.B. Nucleic Acid Res., **1988**, 3, 1063.
6. Lambert, B., Jones, B.K., Rooques, B.P., Le Pecq, J.B. and Yeung, A.T. Proc. Natl. Acad. Sci. USA, **1989**, 86, 6557.
7. Lambert, B., Segal-Bendirdjian, E., Esnault, C., *et al.* Anticancer Drug Design, **1990**, 5, 43.
8. Sancar, A. and Sancar, G.B. Annu. Rev. Biochem., **1988**, 57, 29.
9. Brunger, A.T. X-PLOR Manual, Version 2.1 (Yale University, 1990).
10. Lavery, R. and Sklenar, H. J. Biomol. Struc. & Dynamics, **1988**, 6, 63.
11. Berman, H.M., Neidle, S. and Stodola, R.K. P.N.A.S., **1978**, 75, 828.
12. Neidle, S., Webster, G.D., Baguley, B.C. and Denny, W.A. Biochem Pharmacol, **1986**, 35, 3915.

Chpater 4

STRUCTURE DETERMINATION OF ORGANIC MOLECULES

4.1. INTRODUCTION

The structures of Spiro 3-methyl [5,5] undec-2-ene-1,7-dione, 2'-deoxy-2'-thiomethyl-5'methoxymethylthymidine and (2S,3R,4S,5S)-3-(Allyloxy)-4-[N-(benzyloxycarbonyl)-amino]-5,6-dihydroxy-2-[1(allyloxy)naphth-3-yl]-2-hexanol have been determined using the direct method process. The crystallographic data describing the solution process of these structures is given in the format prescribed for short structural reports in *Acta Crystallographica Section C*,; where these papers have been submitted for publication. Tables of atomic coordinates, thermal parameters, bond lengths, and angles, and the observed and calculated structure factor amplitudes are included in their entirety in their respective papers. A summary of the equipment and the software used to determine the structures of these three organic molecules is as follows:

Diffraction Analysis Equipment	
Diffractometer	Enraf Nonius CAD4 (Fine focus GX21, Regulator RF586) X-ray source CuK
Low temperature device	Enraf Nonius RF558
Software	
CAD4 (Enraf Nonius)	Diffractometer control and data collection
SDP	Data reduction, structure refinement
MolEN	Structure solution program

THE STRUCTURE OF SPIRO 3-METHYL [5,5] UNDEC-2-ENE-1,7-DIONE ($C_{12}H_{16}O_2$)

4.2. THE STRUCTURE OF SPIRO 3-METHYL [5,5] UNDEC-2-ENE-1,7-DIONE

4.2.1. ABSTRACT

Spiro 3-methyl [5,5] undec-2-ene-1,7-dione, $C_{12}H_{16}O_2$, M_r=192.26, triclinic, P_1, a=6.127(1), b=9.064(1), c=9.831(1)Å, V=526.2(1)Å^3, Z=2, Dx=1.21gcm^{-3}, λ(CuKα)=1.5406Å, μ=6.1cm^{-1}, F(000)=208, T=296K, R=5.1% for 1546 unique observed reflections.

4.2.2. EXPERIMENTAL

The compound was synthesized and crystallized as described in Huang, W. (1991). 0.4×0.5×0.5mm^3. space group P_1, Enraf-Nonius CAD4 diffractometer, Ni-filtered radiation, ω-scan method, Lorentz-polarization corrections, 1546 reflections with Fo^2>3.0 σ(Fo^2). Structure solved by the direct method (MULTAN78, Main et al., 1978), First E map revealed the positions of all non-H atoms; least-squares refinement using MolEN (ENRAF-NONIUS, 1990). Isotropic and then anisotropic temperature factors gave R=0.051 with unit weights. 192 parameters refined, H atoms positions were calculated by difference synthesis. Final cycle the largest shift $(\Delta/\sigma)_{max}$=0.050. Final difference synthesis $(\Delta\rho)_{max}$= 0.49e/Å^3 and $(\Delta\rho)_{min}$=-0.33e/Å^3, Scattering factors from *International Tables for X-ray Crystallography* (1974).

4.2.3. RESULTS

Table 4-1. gives an abbreviated form of the atomic parameters, Table 4-2. the bond distances, Table 4-3. the bond angles, and Table 4-4.

the general displacement parameters. Figure 4-1. shows the molecular structure.

Table 4-1. Atomic coordinates and equivalent isotropic thermal parameters. Numbers in parentheses indicate the uncertainty in the final digit(s).

Atom	x	y	z	B_{eq}(Å^2)
C4	0.0791(4)	0.2286(3)	0.1982(2)	3.56(5)
C3	0.0328(4)	0.3353(3)	0.3071(2)	3.85(5)
C5	-0.1380(4)	0.1313(3)	0.1351(3)	4.27(6)
C10	0.1466(4)	0.3186(3)	0.0834(3)	5.04(6)
C1	0.3249(4)	0.5124(3)	0.2690(3)	4.51(6)
O3	-0.1132(3)	0.3063(2)	0.3683(2)	5.80(4)
C2	0.1769(4)	0.4702(3)	0.3387(3)	4.37(6)
C7	0.0094(4)	-0.0881(3)	0.2831(3)	4.90(6)
C9	0.2724(4)	0.1290(3)	0.2721(3)	4.36(6)
C11	0.3481(5)	0.4272(3)	0.1425(3)	5.50(7)
C8	0.2046(4)	0.0157(3)	0.3663(3)	5.02(6)
O5	-0.2609(3)	0.1628(3)	0.0208(2)	7.11(6)
C6	-0.1938(4)	0.0025(3)	0.2146(3)	4.58(6)
C12	0.4825(5)	0.6461(3)	0.3096(4)	6.42(8)

$$B_{eq}= 4/3[a^2B(1,1)+b^2B(2,2)+c^2B(3,3)+ab(\cos\{\gamma\})B(1,2)+ac(\cos\{ß\})B(1,3) + bc(\cos\{\alpha\})B(2,3)\ ].$$

Table 4-2. Table of Bond Distances (Å)

Atom 1	Atom 2	Distance	Atom 1	Atom 2	Distance
C4	C3	1.530(3)	C1	C2	1.315(4)
C4	C5	1.543(3)	C1	C11	1.503(4)
C4	C10	1.526(4)	C1	C12	1.497(4)
C4	C9	1.550(3)	C2	H2	0.94(3)
C3	O3	1.221(3)	C7	C8	1.535(3)
C3	C2	1.456(3)	C7	C6	1.530(4)
C5	O5	1.221(3)	C7	H7"	1.03(3)
C5	C6	1.482(4)	C7	H7'	1.10(3)
C10	C11	1.527(4)	C9	C8	1.501(4)
C10	H10'	0.98(3)	C9	H9	1.09(3)
C10	H10"	0.97(3)	C9	H9"	1.05(2)
C11	H11'	1.07(3)	C6	H6"	1.05(2)
C11	H11"	1.00(3)	C6	H6'	1.02(3)
C8	H8'	1.15(3)	C12	H12A	0.87(3)
C8	H8"	0.94(3)	C12	H12C	0.97(4)
			C12	H12B	1.07(3)

Numbers in parentheses are estimated standard deviations in the least significant digits.

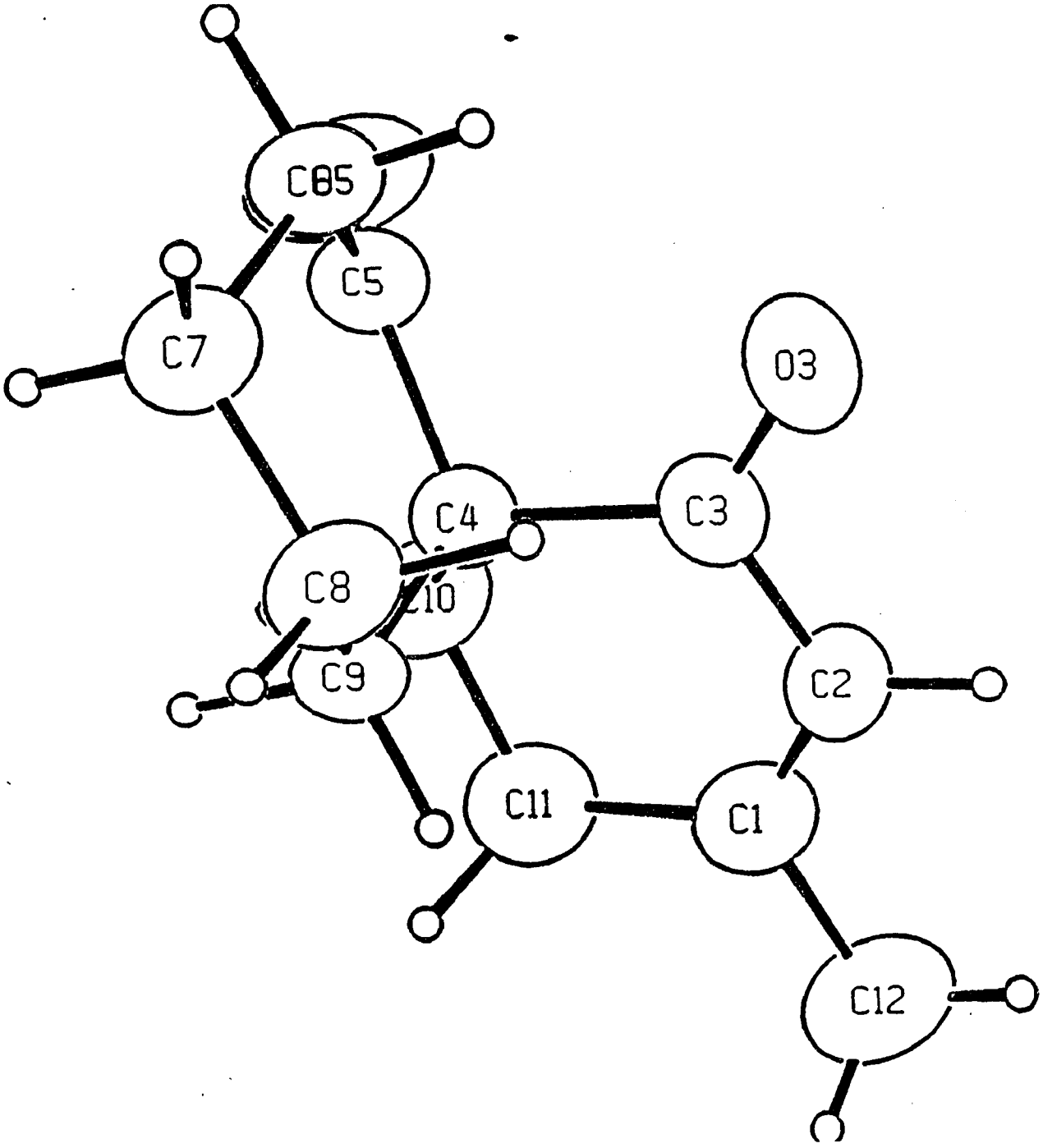
Table 4-3. Table of Bond Angles (°)

Atom 1	Atom 2	Atom 3	Angle	Atom 1	Atom 2	Atom 3	Angle
C3	C4	C5	108.7(2)	C11	C1	C12	115.3(3)
C3	C4	C10	108.6(2)	C3	C2	C1	124.0(2)
C3	C4	C9	109.1(1)	C3	C2	H2	114.(2)
C5	C4	C10	110.6(2)	C1	C2	H2	122.(2)
C5	C4	C9	109.3(1)	C8	C7	C6	109.4(2)
C10	C4	C9	110.5(2)	C8	C7	H7"	108.(1)
C4	C3	O3	121.5(2)	C8	C7	H7'	112.(1)
C4	C3	C2	117.0(3)	C6	C7	H7"	112.(1)
O3	C3	C2	121.5(2)	C6	C7	H7'	107.(2)
C4	C5	O5	119.1(2)	H7"	C7	H7'	109.(2)
C4	C5	C6	119.5(1)	C4	C9	C8	114.0(2)
O5	C5	C6	121.5(3)	C4	C9	H9	105.(1)
C4	C10	C11	112.5(1)	C4	C9	H9"	106.(1)
C4	C10	H10'	106.(2)	C8	C9	H9	114.(1)
C4	C10	H10"	107.(2)	C8	C9	H9"	115.(1)
C11	C10	H10'	110.(1)	H9	C9	H9"	102.(2)
C11	C10	H10"	110.(2)	C10	C11	C1	112.6(2)
H10'	C10	H10"	112.(2)	C10	C11	H11'	105.(2)
C2	C1	C11	121.7(2)	C10	C11	H11"	112.(1)
C2	C1	C12	123.1(2)	C1	C11	H11'	116.(2)
C1	C11	H11"	110.(2)	C7	C6	H6"	110.(1)
H11'	C11	H11"	100.(2)	C7	C6	H6'	111.(1)
C7	C8	C9	111.0(3)	H6"	C6	H6'	105.(2)
C7	C8	H8'	106.(1)	C1	C12	H12A	113.(2)
C7	C8	H8"	104.(2)	C1	C12	H12C	107.(2)
C9	C8	H8'	109.(1)	C1	C12	H12B	113.(2)
C9	C8	H8"	112.(2)	H12A	C12	H12C	108.(3)
H8'	C8	H8"	114.(2)	H12A	C12	H12B	109.(3)
C5	C6	C7	114.1(3)	H12C	C12	H12B	106.(3)
C5	C6	H6"	113.(1)				
C5	C6	H6'	104.(2)				

Table 4-4. General Displacement Parameter Expressions (B)

Name	B(1,1)	B(2,2)	B(3,3)	B(1,2)	B(1,3)	B(2,3)	B_{eq}(Å^2)
C4	2.89(8)	4.6(1)	3.03(8)	0.17(8)	0.43(7)	-0.67(8)	3.56(5)
C3	3.28(9)	4.9(1)	3.31(9)	0.75(8)	0.67(7)	-0.40(8)	3.85(5)
C5	2.96(9)	5.7(1)	3.7(1)	0.34(9)	0.08(8)	-0.86(9)	4.27(6)
C10	5.0(1)	6.4(1)	3.7(1)	-0.9(1)	1.24(8)	-0.7(1)	5.04(6)
C1	3.8(1)	4.6(1)	4.5(1)	0.29(9)	-0.01(9)	0.0(1)	4.51(6)
O3	5.43(7)	6.7(1)	6.08(8)	-0.22(7)	3.09(6)	-1.73(7)	5.80(4)
C2	4.5(1)	4.4(1)	4.1(1)	0.49(9)	0.80(8)	-0.72(9)	4.37(6)
C7	4.1(1)	4.8(1)	5.1(1)	0.3(1)	-0.1(1)	-0.6(1)	4.90(6)
C9	2.60(9)	5.1(1)	4.8(1)	0.43(8)	-0.09(8)	-1.29(9)	4.36(6)
C11	4.8(1)	6.7(1)	5.2(1)	-0.9(1)	1.96(9)	-0.6(1)	5.50(7)
C8	4.1(1)	5.0(1)	4.8(1)	0.7(1)	-1.0(1)	-0.6(1)	5.02(6)
O5	5.20(9)	9.3(1)	4.82(9)	-1.18(9)	-2.06(8)	1.12(9)	7.11(6)
C6	3.28(9)	5.2(1)	4.8(1)	-0.19(9)	0.34(9)	-0.5(1)	4.58(6)
C12	5.5(1)	5.4(1)	7.3(2)	-0.8(1)	0.0(1)	0.1(1)	6.42(8)

Figure 4-1. Structure of Spiro 3-methyl [5,5] undec-2-ene-1,7-dione



4.2.3. REFERENCES

The compound was synthesized by Huang, Wolin "A Dissertation for the Degree of Doctor of Philosophy" (CUNY), (1990)

International Tables for X-ray Crystallography (1974). Vol. IV, pp. 72-75

Main, P., Hull, S. E., Lessinger, L., Germain, G., Declercq, J.-P & Woolfson, M. M. (1978). MULTAN 78. *A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data.* Unives. of York, England, and Louvain, Belgium.

THE STRUCTURE OF 2'-DEOXY-2'THIOMETHYL-5'-METHOXYMETHYLTHYMIDINE ($C_{13}H_{20}N_2O_6S$)

4.3. STRUCTURE OF 2'-DEOXY-2'-THIOMETHYL-5'METHOXYMETHYLTHYMIDINE

4.3.1. ABSTRACT

2'-deoxy-2'-thiomethyl-5'-methoxymethylthymidine, $C_{13}H_{20}N_2O_6S$, M_r=521.62, orthorhombic, $P2_12_12_1$, a=5.972(1), b=15.510(1), c=17.107(1)Å, V=1584.5(1)Å^3, Z=4, Dx=1.189gcm^{-3}, l(CuKa)=1.5418Å, μ=6.1cm^{-1}, F(000)=704, T=296K, R=0.062 for 1726 unique observed reflections. The β-configuration was assigned to the compound because the thymine substituent at C-1' was elucidated to be cis to the substituent of C-4'. The S-methyl substituent at C-2' was determined to be trans to the thymine substituent at C-1' and cis to the 3'-hydroxyl which places this compound in the ribose series.

4.3.2. INTRODUCTION

2'-deoxy-2'-thiomethyl-5'-methoxymethylthymidine is an analog of 2', 3'-dideoxynucleosides which are used as effective drugs in the treatment of AIDS (Scheiner, P., Geer, A. et al. and Schinazi R. F., 1989, Robins, M. J., and Dalley N.K., 1992). Prof. Frank, R. 's laboratory is interested in the synthesis of 2'-deoxy-2'-thiothymidine derivatives. This synthetic procedure resulted in the formation of four isomers which were separated by radial chromatography. Since the furanose ring is flexible, NMR data could not be used to assign the stereochemistry of C-1' and C-2' of each isomer. X-ray crystallography was therefore needed to determine the anomeric configuration of these isomers.

4.3.3. EXPERIMENTAL

The compound was synthesized and crystallized as described in Geer, A. (1991). Orthorhombic crystal, $0.30 \times 0.30 \times 0.5 mm^3$. Enraf-Nonius CAD4 diffractometer, Ni-filtered radiation, ω-scan method, Lorentz-polarization corrections, linear decay, empirical absorption, 1700 reflections with $Fo^2 > 3.0\sigma(Fo^2)$. Structure solved by the direct method (MULTAN, 1978), first E map revealed the positions of all non-H atoms; least-squares refinement using MolEN (ENRAF NONIUS, 1990), isotropic and then anisotropic temperature factors gave R=0.062 with unit weights. 270 parameters refined, H atoms positions were calculated by difference synthesis. Final cycle the largest shift $(\Delta/\sigma)max=0.090$. Final difference synthesis $(\Delta r)_{max}=0.49$ e/A^3 and $(\Delta r)_{min}=-0.02$ e/A^3, Scattering factors from *International Tables for X-ray Crystallography* (1974).

4.3.4. RESULTS

Table 4-5. gives the non-hydrogen atomic parameters. Table 4-6. the bond distances, Table 4-7. the bond angles, and Table 4-8. the general displacement parameters. Figure 4-2. shows the molecular structure.

Table 4-5. Position Parameters and Their Estimated Standard Deviations

Atom	x	y	z	B_{eq}(Å^2)
S	1.0163(4)	0.1621(1)	0.5575(1)	3.34(4)
N1	1.038(1)	-0.0416(3)	0.4809(3)	2.1(1)
N3	1.004(1)	-0.0821(3)	0.3503(3)	2.7(1)
O'6	0.764(1)	-0.2134(3)	0.7929(3)	3.3(1)
O'4	1.2429(9)	-0.0712(3)	0.5942(3)	2.8(1)
O'5	0.8659(9)	-0.1136(3)	0.6954(3)	2.5(1)
C'1	1.165(1)	-0.0027(4)	0.5451(4)	2.3(1)
O2	1.294(1)	0.0058(4)	0.3887(3)	3.7(1)
C5	0.740(1)	-0.1372(4)	0.4434(4)	2.3(1)
O'3	1.307(1)	0.1013(3)	0.6921(3)	3.0(1)
C'2	1.020(1)	0.0541(4)	0.5976(3)	2.1(1)
C'4	1.224(1)	-0.0479(4)	0.6755(4)	2.2(1)
C6	0.856(1)	-0.0926(5)	0.4967(4)	2.5(1)
C2	1.126(1)	-0.0355(4)	0.4049(4)	2.5(1)
O4	0.725(1)	-0.1691(4)	0.3080(3)	4.3(1)
C7	0.546(1)	-0.1936(5)	0.4622(4)	3.4(2)
C'6	0.753(2)	-0.1919(4)	0.7129(4)	3.0(2)
C4	0.817(1)	-0.1310(5)	0.3630(4)	2.8(2)
C'3	1.124(1)	0.0435(4)	0.6789(4)	2.1(1)
C'7	0.664(2)	-0.1486(6)	0.8426(5)	5.3(2)
C'5	1.096(1)	-0.1164(5)	0.7187(4)	2.6(1)
C'8	0.280(2)	0.2907(6)	0.3930(5)	5.6(2)

$$B_{eq}= 4/3[a^2B(1,1)+b^2B(2,2)+c^2B(3,3)+ab(\cos\{\gamma\})B(1,2)+ac(\cos\{ß\})B(1,3) + bc(\cos\{\alpha\})B(2,3)\ ].$$

Table 4-6 Table of Bond Distances in Angstroms

S	C'2	1.810(6)	O'5	C'5	1.432(9)
N1	C'1	1.464(9)	C'1	C'2	1.53(1)
N1	C6	1.37(1)	O2	C2	1.22(1)
N1	C2	1.405(8)	C5	C6	1.34(1)
N3	C2	1.388(9)	C5	C7	1.49(1)
N3	C4	1.37(1)	C5	C4	1.453(9)
O'6	C'6	1.411(8)	O'3	C'3	1.434(9)
O'6	C'7	1.44(1)	C'2	C'3	1.532(9)
O'4	C'1	1.432(8)	C'4	C'3	1.541(9)
O'4	C'4	1.441(7)	C'4	C'5	1.50(1)
O'5	C'6	1.421(9)	O4	C4	1.241(9)

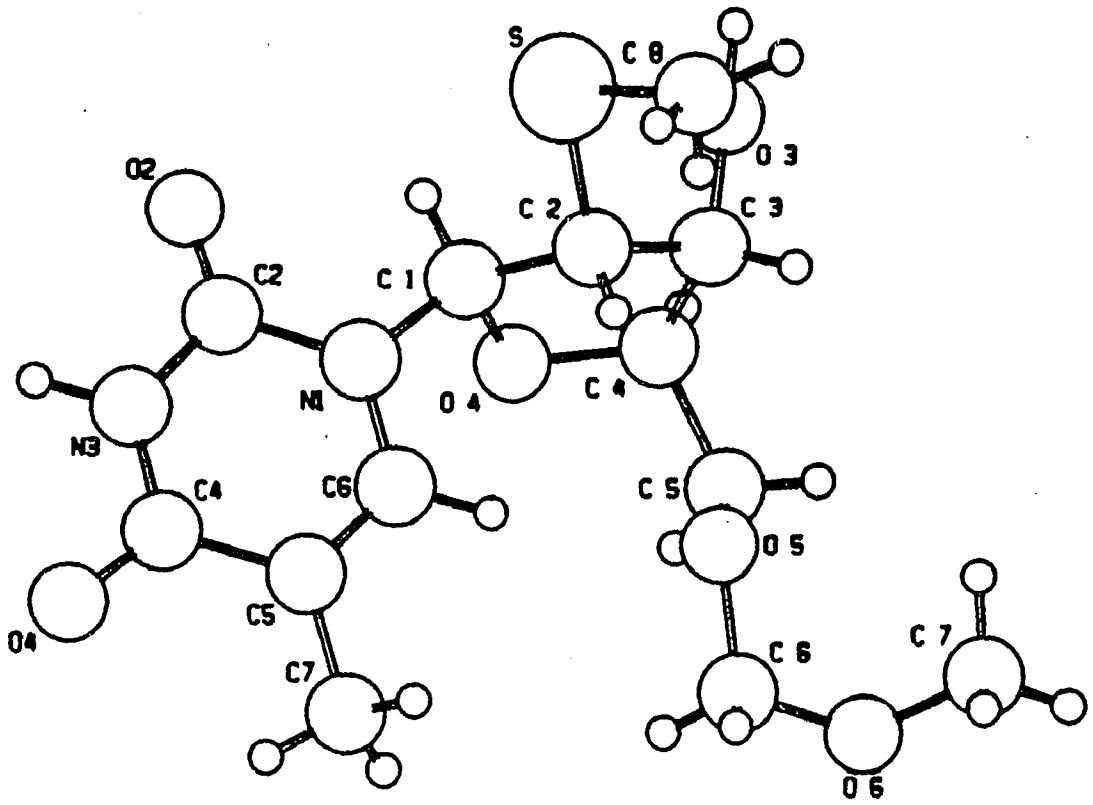
Table 4-7 Bond Angles (^{0})

Atom 1	Atom 2	Atom 3	Angle	Atom 1	Atom 2	Atom 3	Angle
C'1	N1	C6	119.9(5)	O'4	C'4	C'3	107.3(5)
C'1	N1	C2	118.2(6)	O'4	C'4	C'5	109.7(5)
C6	N1	C2	121.2(6)	C'3	C'4	C'5	115.7(6)
C2	N3	C4	127.7(6)	N1	C6	C5	125.0(6)
C'6	O'6	C'7	112.9(6)	N1	C2	N3	113.0(6)
C'1	O'4	C'4	110.8(5)	N1	C2	O2	123.5(6)
C'6	O'5	C'5	111.7(5)	N3	C2	O2	123.5(6)
N1	C'1	O'4	107.6(5)	O'6	C'6	O'5	112.7(6)
N1	C'1	C'2	112.7(6)	N3	C4	C5	116.4(6)
O'4	C'1	C'2	105.5(5)	N3	C4	O4	120.6(6)
C6	C5	C7	124.0(6)	C5	C4	O4	123.0(7)
C6	C5	C4	116.6(7)	O'3	C'3	C'2	112.7(5)
C7	C5	C4	119.3(6)	O'3	C'3	C'4	106.4(6)
S	C'2	C'1	108.5(4)	C'2	C'3	C'4	102.8(5)
S	C'2	C'3	116.6(4)	O'5	C'5	C'4	109.3(6)
C'1	C'2	C'3	104.0(6)				

Table 4-8 Table of General Displacement Parameter Expressions - B's

Name	B(1,1)	B(2,2)	B(3,3)	B(1,2)	B(1,3)	B(2,3)	B_{eq}(Å^2)
S	5.7(1)	2.24(6)	2.04(6)	0.29(9)	-0.38(9)	0.43(6)	3.34(4)
N1	3.3(3)	2.2(2)	0.9(2)	-0.8(2)	0.2(2)	-0.2(2)	2.1(1)
N3	4.5(3)	2.5(2)	1.1(2)	-0.8(3)	0.3(3)	-0.1(2)	2.7(1)
O'6	4.5(3)	2.8(2)	2.7(2)	0.3(2)	1.1(2)	0.8(2)	3.3(1)
O'4	4.3(3)	2.5(2)	1.6(2)	1.0(2)	0.2(2)	-0.2(2)	2.8(1)
O'5	3.0(2)	2.3(2)	2.2(2)	-0.3(2)	-0.2(2)	0.6(2)	2.5(1)
C'1	3.5(3)	2.2(2)	1.3(2)	-0.4(3)	-0.1(2)	-0.1(2)	2.3(1)
O2	5.2(3)	3.9(2)	1.9(2)	-1.4(3)	0.9(2)	-0.0(2)	3.7(1)
C5	3.3(3)	2.0(2)	1.7(2)	-0.1(3)	0.1(3)	-0.0(2)	2.3(1)
O'3	4.1(3)	2.7(2)	2.3(2)	-1.1(2)	-1.0(2)	-0.1(2)	3.0(1)
C'2	3.4(3)	1.6(2)	1.4(2)	0.4(3)	-0.1(3)	-0.2(2)	2.1(1)
C'4	3.5(3)	1.9(2)	1.3(2)	0.2(3)	-0.4(3)	0.1(2)	2.2(1)
C6	3.2(3)	2.9(3)	1.3(2)	-0.2(3)	0.2(3)	-0.1(2)	2.5(1)
C2	3.9(3)	2.2(3)	1.5(2)	-0.2(3)	0.3(3)	0.1(2)	2.5(1)
O4	7.0(4)	3.8(2)	2.1(2)	-1.4(3)	-0.6(3)	-0.8(2)	4.3(1)
C7	3.8(4)	3.5(3)	3.0(3)	-1.0(3)	0.1(3)	-0.4(3)	3.4(2)
C'6	4.2(4)	2.2(3)	2.6(3)	-0.4(3)	0.3(3)	0.1(2)	3.0(2)
C4	4.2(4)	2.5(3)	1.7(3)	-0.2(3)	-0.3(3)	-0.3(2)	2.8(2)
C'3	3.5(3)	1.3(2)	1.6(2)	-0.4(3)	-0.1(3)	-0.1(2)	2.1(1)
C'7	7.3(6)	5.0(5)	3.5(4)	1.0(5)	2.2(4)	-0.1(4)	5.3(2)
C'5	3.2(3)	2.5(3)	2.1(3)	-0.3(3)	-0.4(3)	0.5(2)	2.6(1)
C'8	7.8(6)	4.8(4)	4.0(4)	-3.3(4)	0.1(5)	0.0(4)	

Figure 4-2. Structure of 2'-deoxy-2'-thiomethyl-5'methoxymethylthymidine



4.3.5. REFERENCES

Geer, A. and Franck, R. W. (1991)

International Tables for X-ray Crystallography (1974)

Main, P., Hull, S. E., Lessinger, L., Germain, G., Declercq, J.-P & Woolfson, M. M. (1978). MULTAN 78. *A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data.* Unives. of York, England, and Louvain, Belgium.

Scheiner, P., Geer, A. etc. and Schinazi R.F., *J. Med. Chem.,* (1989), Vol. 32, No. 1, 73-76

Robins, M.J., and Dalley N.K., *J. Org. Chem.,* Vol. 57, No. 8, (1992),2357-2364

Shibuya, S., and Ueda, T., *J. Carbohyd. Nuc. Nuc.*, (1980) 7(1), 49-56.

Garner, P., and Park, J.M., *J. Org.. Chem., Vol. 55*, No. 12, (1990)3772-3787

THE STRUCTURE OF

(2S, 3R, 4S, 5S)-3-(ALLYLOXY)-4- [N-(BENZYLOXYCARBONYL)-AMINO]-5,6-DIHYDROXY-2-[1-(ALLYOXY)NAPHTH-3-YL]-2-HEXANOL

($C_{30}H_{35}NO_7$)

4.4. THE STRUCTURE OF (2S,3R,4S,5S)-3-(ALLYLOXY)-4-[N-(BENZYLOXYCARBONYL)-AMINO]-5,6-DIHYDROXY-2-[1(ALLYLOXY)NAPHTH-3-YL]-2-HEXANOL

4.4.1. ABSTRACT

The Structure of (2S,3R,4S,5S)-3-(Allyloxy)-4-[N-(benzyloxycarbonyl)-amino]-5,6-dihydroxy-2-[1-(allyloxy)naphth-3-yl]-2-hexanol, $C_{30}H_{35}NO_7$, M_r=521.62, orthorhombic, $P2_12_12_1$, a=9.181(2), b=10.116(2), c=31.383(2)Å, V=2914.63(1)Å^3, Z=4, Dx=1.189gcm^{-3}, l(CuKa)=1.5418Å, μ=6.1cm^{-1}, F(000)=1112, T=296K, R=0.052 for 2496 unique observed reflections. The 5S configuration was determined instead of the 5R which was considered as the necessary conformation for activity.

4.4.2. INTRODUCTION

Nogalamycin is a notable member of the anthracycline family and it is active against Gram-positive microorganisms, L1210 leukemia and KB cell carcinoma in vitro. Despite its promise, nogalamycin's unacceptable toxicity precluded its clinical use. A number of synthetic model studies have been reported (Sammes, P. G. et al, 1983; Kawasaki, M. et al, 1994). (2S,3R,4S,5S)-3-(Allyloxy)-4-[N-(benzyloxycarbonyl)-amino]-5,6-dihydroxy-2-[1-(allyloxy)naphth-3-yl]-2-hexanol is one of the products obtained during the synthesis of the CDEF-benzoxocin model of nogalamycin in professor R. Franck's lab. (Yin, H. et al, 1992)

4.4.3. EXPERIMENTAL

The compound was synthesized and crystallized as described in Yin, H. (1991). Orthorhombic crystal, $0.3 \times 0.4 \times 0.6 mm^3$, Enraf-Nonius CAD4 diffractometer, Ni-filtered radiation, ω-scan method, Lorentz-polarization corrections, linear decay, empirical absorption, h=0 to 10, k=0 to 11, l=0 to 35, 2496 reflections with $Fo^2 > 3.0\sigma(Fo^2)$. Structure solved by the direct method (MULTAN78, Main et al., 1978), First E map revealed the positions of all non-H atoms; least-squares refinement using MolEN (ENRAF-NONIUS, 1990). isotropic and then anisotropic temperature factors gave R=0.052 with unit weights, 343 parameters refined, H atoms positions were calculated by difference synthesis. Final cycle the largest shift $(\Delta/\sigma)max=0.200$. Final difference synthesis $(\Delta r)_{max}=0.38$ e/A^3 and $(\Delta r)_{min}=-0.01$ e/A^3, Scattering factors from *International Tables for X-ray Crystallography* (1974).

4.4.4. RESULTS

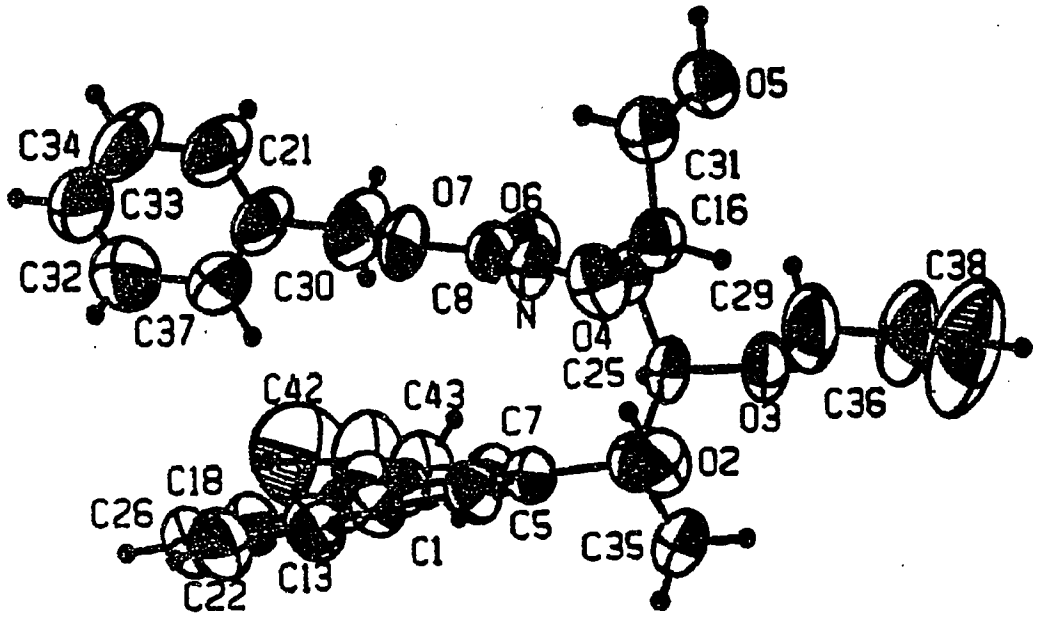
Table 4-6 contains the non-hydrogen atomic positions and thermal factors; Fig.4-3 the molecular structure.

Table 4-9. Atomic coordinates and equivalent isotropic thermal parameters

B_{eq}= 4/3[a^2B(1,1)+b^2B(2,2)+c^2B(3,3)+ab(cos{γ})B(1,2)+ac(cos{ß})B(1,3) + bc(cos{α})B(2,3)].

Atom	x	y	z	B_{eq}(Å2)
O1	-0.2774	-0.0422	0.2076	6
O2	0.2487	-0.2156	0.1036	5
O3	0.0714	-0.4043	0.0649	5
O4	0.1779	-0.0490	0.0377	5
O5	-0.0498	-0.1955	-0.0444	5
O6	-0.3256	-0.1668	0.0769	7
O7	-0.2827	0.0522	0.0867	6
N	-0.1060	-0.0758	0.0646	4
C1	0.0780	0.0880	0.1759	5
C5	0.1304	-0.0191	0.1511	5
C6	-0.0629	0.0774	0.1951	5
C7	0.0530	-0.1303	0.1456	4
C8	-0.2461	-0.0723	0.0758	5
C9	-0.1416	-0.0411	0.1884	5
C13	0.1575	0.2041	0.1822	6
C14	0.1092	-0.2491	0.1209	5
C16	0.0896	-0.1512	0.0196	4
C17	-0.0885	-0.1418	0.1648	5
C18	-0.1141	0.1832	0.2207	6
C20	-0.0351	-0.1936	0.0490	4
C21	-0.5696	0.2797	0.0939	8
C22	0.1054	0.3052	0.2067	7
C23	-0.3627	-0.1582	0.2030	8
C25	0.0055	-0.2928	0.0847	4
C26	-0.0343	0.2930	0.2265	7
C28	-0.4492	0.2148	0.1109	8
C29	-0.0264	-0.5016	0.0504	9
C30	-0.4334	0.0684	0.1001	8
C31	0.0289	-0.0967	-0.0223	5
C32	-0.3726	0.4112	0.1457	7
C33	-0.4946	0.4738	0.1304	6
C34	-0.5887	0.4108	0.1038	5
C35	0.1304	-0.3640	0.1520	6
C36	0.0499	-0.6128	0.0330	13
C37	-0.3510	0.2800	0.1360	13
C42	-0.5690	-0.0462	0.2345	10
C43	-0.5103	-0.1331	0.2223	12
C38	0.0524	-0.7156	0.0321	13

Figure 4-3. Structure of (2s, 3r, 4s, 5s)-3-(allyloxy)-4- [n-(benzyloxycarbonyl)-amino]-5,6-dihydroxy-2-[1-(allyoxy)naphth-3-yl]-2-hexanol



4.4.4. REFERENCES

Kawasaki, M. et al (1988) Tetrahedron, 44, 5727-5743

Main, P., Hull, S. E., Lessinger, L., Germain, G., Declercq, J.-P & Woolfson, M. M. (1978). MULTAN 78. *A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data.* Unives. of York, England, and Louvain, Belgium.

Sammes, P. G.; Bates, M. A.(1983) J. Chem. Soc., Chem. Commun., 896-898

Yin, H. and Franck, R. W. et al (1992) J. Org. Chem. Vol. 57, No. 2, 644-651

International Tables for X-ray Crystallography (1974), Vol.4,

Chapter 5

THE MAGNESIUM CATION'S INFLUENCE ON THE CRYSTAL STRUCTURE OF NUCLEIC ACIDS

5.1. ABSTRACT

The structure of $d(CGCGCG)_2$, crystallized in the presence of magnesium chloride has been determined at 1.3Å resolution. Orthorhombic crystals with a space group $P2_12_12_1$, a=17.99Å, b=30.97Å, and c=44.75Å, and with one DNA duplex per asymmetric unit were obtained. The structure was solved by molecular replacement and refined by the Powell minimization using the X-PLOR program [1]. The molecule adopts a Z-DNA conformation. The current model consists of six base pairs, 95 water molecules, and four magnesium ions. The R-factor is 17.1% for 5,411 reflections ($|F| > 1\sigma |F|$); 89% of possible observations between 5 and 1.3Å resolution. The root mean square (r.m.s.) deviations from ideal geometry is 0.01Å for bond lengths and 3.21° for bond angles, and the individual B-factors are highly restrained (r.m.s. deviations in B-factor between bonded atoms=2.0Å^2). The electron density of the bases is continuous everywhere at 1.2σ in a ($2|F_{obs}|-|F_{calc}|$) map (see section 1.5.3. for description of maps).

At a resolution of 1.3Å, the positions of magnesium ions were revealed with a high level of specificity. The general organization of the water molecules in the lattice and around the magnesium ions was also visualized. The structure was analyzed and compared to the $d(CGCGCG)_2$ structure that was crystallized in the presence of a mixed magnesium-sodium ion solvent environment [2,3]. The results showed an overall similarity between the two structures. However, the present structure does not have alternate phasphate conformations as do many other Z-DNA structures [3,4,5,6]. In the environment where magnesium is the only cation, the magnesium complexes show an extensive contact surface with the DNA, including more than three

times as many clearly defined indirect solvent-mediated hydrogen bonds to the DNA surface as in the mixed salt solvent structure. The higher charge and more specific fitting of the magnesium complexes may be the most important feature that stabilizes the conformation of this DNA hexamer.

5.2. INTRODUCTION

Among the divalent cations, Mg(II) is special in that it exists in living cells with a concentration of around 20 mM, has a profound influence on the stability of DNA and acts as a cofactor with enzymes whose activities involve DNA and RNA. It is not yet fully understood how magnesium interacts with DNA. Earlier studies have shown that the DNA base pair/Mg(II) binding rate is about 1:1 [7]. Crystal structure studies of Z-DNA have been able to provide some details of the magnesium ion binding to DNA [2,3,5,8] and NMR [9] studies have shown the specific binding sites of magnesium ions in solution. Previously, all reported crystallographic studies were carried out on crystals that were grown in a mixed salt environment (monovalent cations, divalent cations and other cations). This may account for the fact that insufficient counter ions were observed to neutralize the DNA charge.

Detailed information about the role cations play in stabilizing crystal structure and about solvent organization depends mainly on the resolution of the structure studied. A resolution of 1.3Å is sufficient to resolve adjacent atoms. Interpretable electron-density maps can be obtained with data out to 2.5 or 3Å resolution. High resolution allows us to visualize the structure in great detail and to gather a significant amount of information. The information includes: the organization of the surrounding solvent and ions in the crystal

lattice, the role of small ions around specific interaction sites, and the effects of different cations on the DNA conformation [4]. Low resolution reflects the fact that much of the structure is disordered. In most low resolution structures the solvent and ion environment is so disordered that these groups are not visualized in a meaningful way.

Since the work of Pohl and Jovin [10] on the effects of high ionic strength on alternating GC DNA polymers and the demonstration by x-ray crystallography [4] of the existence of a left-handed double helix for the hexamer $d(CGCGCG)_2$, the conditions under which transitions occur have been extensively studied [2,11,12,13]. These conditions are high ionic strength and have the presence of magnesium ions. Crystallization conditions are important when analyzing the fine structure details which we are interested in, such as cation-DNA interaction and solvation properties in DNA crystal structure. The work was undertaken to help in the understanding of the role cations play in stabilizing the DNA crystal structure and the way DNA interacts with cations.

The following questions are addressed in this chapter:

1) How does the presence of magnesium influence the structure of DNA?

2) What are the structural features of Z-DNA in a magnesium only environment?

3) How does magnesium interact with DNA?

4) Where are the preferred binding sites of magnesium on the DNA?

5) Will the solvation of DNA be affected by magnesium?

6) How are charges neutralized in the DNA crystal structure?

5.3. MATERIALS AND METHODS

5.3.1. SYNTHESIS AND PURIFICATION

The d(CGCGCG) oligomer was synthesized in the RCMI Sequence and Synthesis facility by Mr. Stephen Bobin using the phosphotriester chemistry method [14].

The oligonucleotide was purified by reversed-phase HPLC on a C18 column (4.6mm × 25cm), Flow: 1ml/min, Mobile phase: 30 min 10% to 35% of 30% acetonitrile/water in 0.1 TEAB. The major peak was collected and concentrated by lyophilization. According to the analytical HPLC, the purity was 99%. The DNA stock solution was 5.74 mg/ml and was stored at -4°C.

5.3.2. CRYSTALLIZATION

Crystals were grown by the vapor diffusion method using setting drops [15]. Many different ion concentrations and different metal ions (Li, Na, K, Rb, Cs and Mg) were tried. Only Rb and Mg give good enough quality crystals for x-ray single crystal diffraction. The best crystals were obtained from drops containing 1.71mM DNA, 26.3mM Mg-cacodylate buffer (pH7.0), and 5.2% 2-Methyl-2,4-pentanediol (MPD) placed over a reservoirs containing 6% MPD. Several small cubic shaped crystals formed after 3 days, and large crystals with 0.4 × 0.4 × 0.5mm dimensions formed in 7 days at room temperature.

5.3.3. DATA COLLECTION AND PROCESSING

Two data sets were collected at Yale University with the help of Dr. S. L. Chen, utilizing the multiwire area detector system from Hamlin Systems. The diffraction data were reduced to structure factor amplitudes with the use of software provided by the Resource for Crystallography. 27,405 observation reflections were observed out of which 5,619 unique structure factor amplitudes were obtained extending to a resolution of 1.3 Å. **Table 5-1** presents the data collection information.

Table 5-1 X-ray Data Collection Statistics

Merging R-factor based on I	0.05	
Resolution limits	1.3Å	2.0Å
Number of observations	27405	3582
Number of unique x-ray reflections collected	5619	1594
Average number of observations for each reflection	4.88	2.25
% complete	89	43

The lattice type is orthorhombic with unit cell dimensions of a=17.99Å, b=30.97Å, and c=44.75Å. To determine the space group x-ray diffraction data was examined for systematic absences. Along the a*, b*, and c* axis, only reflections with h, k or l=2n were observed. This coupled to similarity to unit cell dimensions of the related structures indicated a space group $P2_12_12_1$ with a unit cell volume of $2.49\times10^4Å^3$. This is consistent with that observed for various isomorphous Z-DNA hexamer crystals. The standard volume of each

Z-DNA base pair is approximately 1000Å^3 and since there were six base pairs in this oligonucleotide, the volume of each DNA molecule is therefore approximately 6000Å^3. Considering that in this space group there are four asymmetric cells in each unit cell, the volume of an asymmetric unit cell is therefore 6225Å^3. It appears that one DNA duplex exists per asymmetric unit. The structure factor phasing was determined by molecular replacement and the mixed salt d(CGCGCG)$_2$ structure was used as a phasing model for rigid body refinement.

5.3.4. STRUCTURE REFINEMENT

Position and temperature factor refinements of the DNA structure at 1.3Å resolution were carried out against the structure factors obtained from a d(CGCGCG)$_2$ model [2]. The refinement of our structure was based on an energy function approach [1]: an arbitrary combination of empirical and effective energies that used experimental information to restrain energy terms was implemented using the X-PLOR program. The structure was initially solved with 2.0Å data and subsequently refined with 1.3Å data. See **Table 5-2** for the summary of the crystallographic data. The model was subjected to energy minimization with the electrostatic energy term "off" to relieve the repulsion of close contacts. Water molecules were identified as well-defined peaks in the difference electron density map surrounding the refined DNA and with reasonable hydrogen bonding distances.

Table 5-2 The Summary of Crystallography Analysis

	Crystal 2	Crystal 1
Crystal	$d(CGCGCG)_2$	$d(CGCGCG)_2$
Contents	$MgCl_2$	$MgCl_2$
Crystal System	Orthorhombic	Orthorhombic
Space Group	$P2_12_12_1$	$P2_12_12_1$
Unit Cell (Å)	a=17.99	a=17.95
	b=30.97	b=30.91
	c=44.75	c=44.63
Unique Reflections	5,619	1,594
Reflections ($\|F\| > \sigma\|F\|$)	5,411	1,059
Data Coverage (%)		
Resolution (Å)	1.3	2.0
R-factor (%)	17.50	18.10
Total number of atoms	603	594
Number of water molecules	95	90
Number of magnesium ions	4	4
r.m.s. bond length (Å)	0.011	0.009
r.m.s. bond angle (degree)	3.208	3.182

The atomic coordinates of the mixed salt $d(CGCGCG)_2$ [2] were used as a starting model for the refinement. The model was initially refined without including any water or magnesium. The R-factor of this data against this initial model was 30.6% and quickly dropped to 27.7% after a 100 cycle rigid body refinement of the positions and orientations.

Since x-rays are scattered by electrons, and Mg^{2+} and O have almost the same number of electrons, all the difference map peaks were initially assigned to oxygen. The magnesium ions were identified by central "waters"

which were surrounded by six other waters with octahedral geometry and with distances of about 2.0Å. The repulsion from the non-bonding interaction of the water surrounding the central magnesium exhibited a specific and highly symmetric geometry.

During position refinement, the coordinates were checked and adjusted against a (2|Fobs| - |Fcalc|) and (|Fobs| - |Fcalc|) electron density map. In the computations of the electron density map, the phases are calculated from the model for Fcalc and the amplitudes from the intensity data for Fobs. The difference map of (|Fobs| - |Fcalc|) can be used for adding or deleting atoms. Depending on which of |Fobs| or |Fcalc| is larger, the Fourier terms can be either positive or negative. A positive density in the map implies that the contribution of the observed intensities (|Fobs|) are larger than the contribution of the model's |Fcalc|, and that the unit cell contains more electron density in this region than implied by the model. Therefor, atoms should be added to or moved toward the region to increase the electron density in this area. On the other hand, a region of negative density indicates that the model implies more electron density than the unit cell actually contains, and the atoms should be removed from or moved away from this region.

An overall anisotropic temperature factor followed by isotropic temperature factors for individual atoms were used for the temperature factor refinement. All temperature factors were assigned a starting value (B=10 for all non-hydrogen atoms and B=15 for hydrogens) with tightly restrained individual isotropic temperature factors. The restraint used was r.m.s. deviations=2.0Å^2 in the B-factor between bonded atoms and r.m.s. deviations=2.5Å^2 for second nearest neighbor atoms. From the temperature

factors computed during refinement, it was ascertained which atoms in the molecule had the most movement or disorder. Large temperature factors indicate that an atom has more freedom of movement, as calculated from $B_j = 8\pi^2\langle u_j^2\rangle$ [16] where the mean square of the displacement, u_j, is proportional to the temperature factor. After position and temperature factors were refined, water molecules were deleted from the coordinate list if their temperature factors exceeded 50Å^2. All residues were inspected on the Indigo graphics system using the CHAIN program (Baylor College of Medicine, 1991). This was to check to see that the geometry fit into the density and if there are any unusual or close contacts.

Late in the refinement, a region in the model which was located between two symmetry related molecules was noted as not showing significant electron-density in the maps. It produced an empty area in the model. The R factor (see section 1.5.3. for definition of R values) at this stage was 17.8% and could not be reduced any further. At that point four magnesium and 73 water molecules were included in the structure. Neither a 2Fo-Fc map nor a Fo-Fc map showed interpretable features. When the contour level was lowered from 2.5σ to 1.5σ, an interpretable 2Fo-Fc map appeared. The density regions in this problem area showed up as several sausage like images which were not well defined.

Is this a disordered solvent region? In macromolecular crystals, it is common to have a certain degree of disorder where the maps are not unambiguously interpretable. These features could be modeled as multiple strings of water molecules. There are two important reasons for uncertainty in atom positions: thermal motion and disorder, showing their effects on the final

B values and R factor. If the region is highly disordered or in motion, the result is atoms which become not well defined sausage like images or even invisible on the time scale of crystallography.

The empty region which was at the interface of two DNA duplexes related by crystallographic symmetry suggests that low resolution reflections may be responsible for this area. Low resolution data offer information about broad features of the whole model, while high resolution data give detailed information about the structure. Diffraction data can be randomly omitted without seriously affecting the result. However, reflections that are systematically excluded will affect the structure refinement. This refinement did not include the reflections which had resolutions lower than 5Å.

To investigate the "hole", data with resolution of 5Å to 1.3Å were retained for the refinement, but lower resolution data were included in the map calculation (resolution extended from 8Å to 1.3Å). Electron dense features corresponding to individual water molecules in the "hole" region were visualized once the resolution of the map was extended to 8Å. Thus, the final refined structure includes atoms throughout the volume of the asymmetric unit.

Model building, addition of water molecules, and a repeated position and temperature factor refinement reduced the R-factor to 17.1% for 5,411 reflections; 89% of a possible observation between 5Å and 1.3Å resolution, with root-mean-square deviations from target values of 0.011Å and 3.208° for bond lengths and bond angles, respectively. Since there was magnesium in the molecule and the Mg-water distance is approximately 2.0Å, water peaks were retained if they were between 1.8Å and 3.5Å of nitrogen or oxygen atoms and

not closer than 2.5Å to carbon atoms. Water molecules were assigned with the help of a peak search programs (Quigley, personal communication). At the earlier stages of refinement, broken density at some of the DNA aromatic ring surfaces was found. As the refinement neared completion, the broken area no longer existed and the 2|Fobs|-|Fcalc| map became continuous everywhere on the DNA.

The final atomic model consisted of 111 residues: residue 1 to 12 are the DNA residues, 13 to 16 are magnesium, and 20 to 115 are water molecules. Details of the crystallographic refinement are given in **Table 5-3.** The stereo view of the complex shown in **Figure 5-1.**

Table 5-3 Progress of Refinement

Stage	R-Factor	Solvent Included	r.m.s. Bond* Length(Å)	r.m.s. Bond* Angle(°)
1	0.306	0		
2	0.260	0	0.026	3.606
3	0.197	$64H_2O$, 4Mg	0.013	3.288
4	0.186	$62H_2O$, 4Mg	0.012	3.326
5	0.180	$68H_2O$, 5Mg	0.012	3.318
6	0.179	$73H_2O$, 4Mg	0.012	3.274
7	0.171	$95H_2O$, 4Mg	0.011	3.208

*deviation from model target values.

The quality of the refined structure can be assessed using various criteria. The R-factor and temperature factors are examined here.

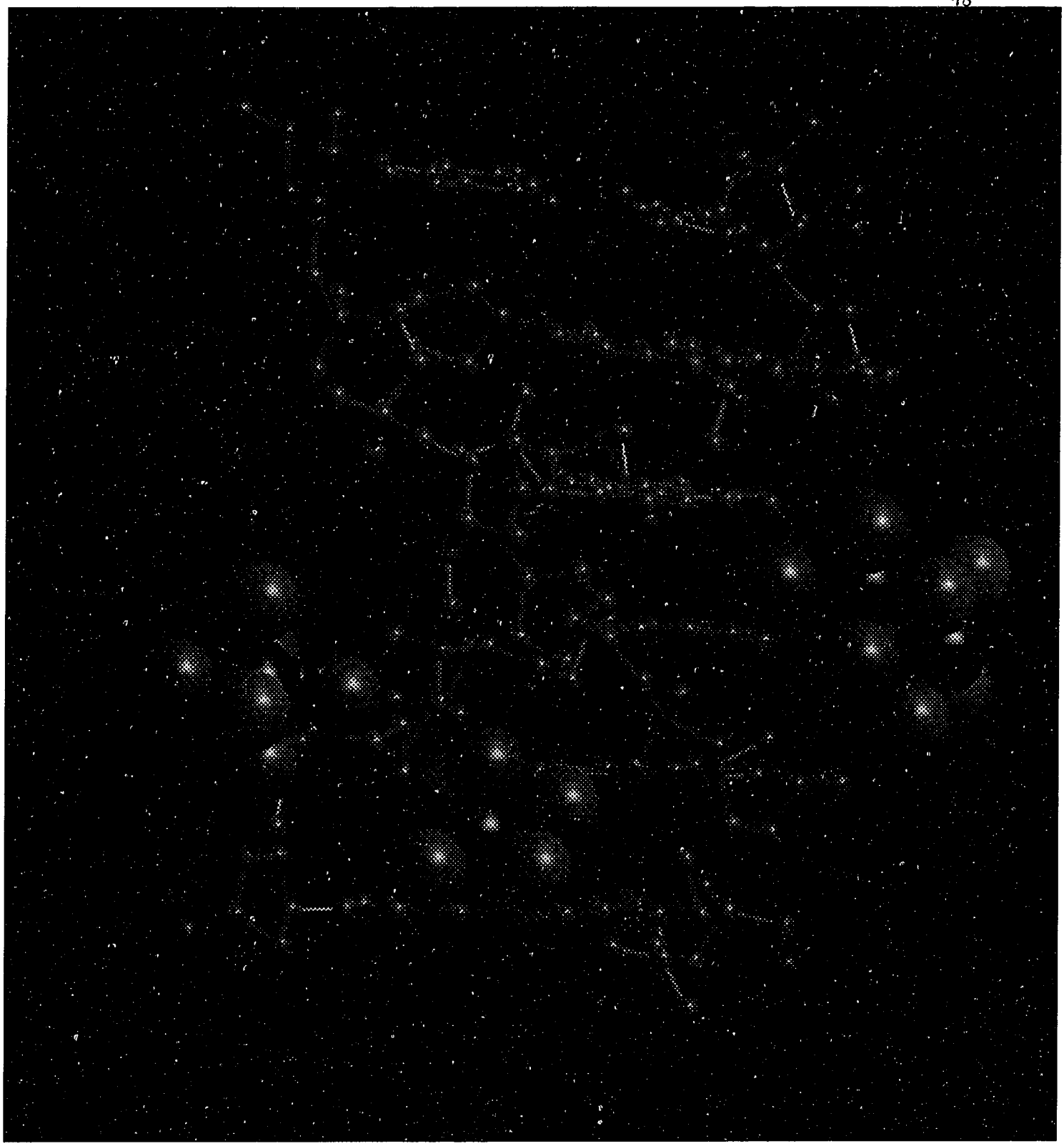


Figure 5-1. View of the magnesium ion binding to Z-DNA.

Figure 5-2 illustrates the R-factor as a function of resolution in a so called Luzzati plot [17]. Luzzati's method estimates the precision of an atom's location in a refined crystallographic model. The numbers to the right of each smooth curve on the Luzzati plot are theoretical estimates of the average uncertainty in the positions of atoms in the refined model or the r.m.s. errors in the atom's position. The Luzzati plot implies that the upper limit of error in atomic position is less than 0.22Å which means that the uncertainty of the atomic positions in the model is on the average less than 0.22Å, about one-eighth of a carbon-carbon covalent bond length.

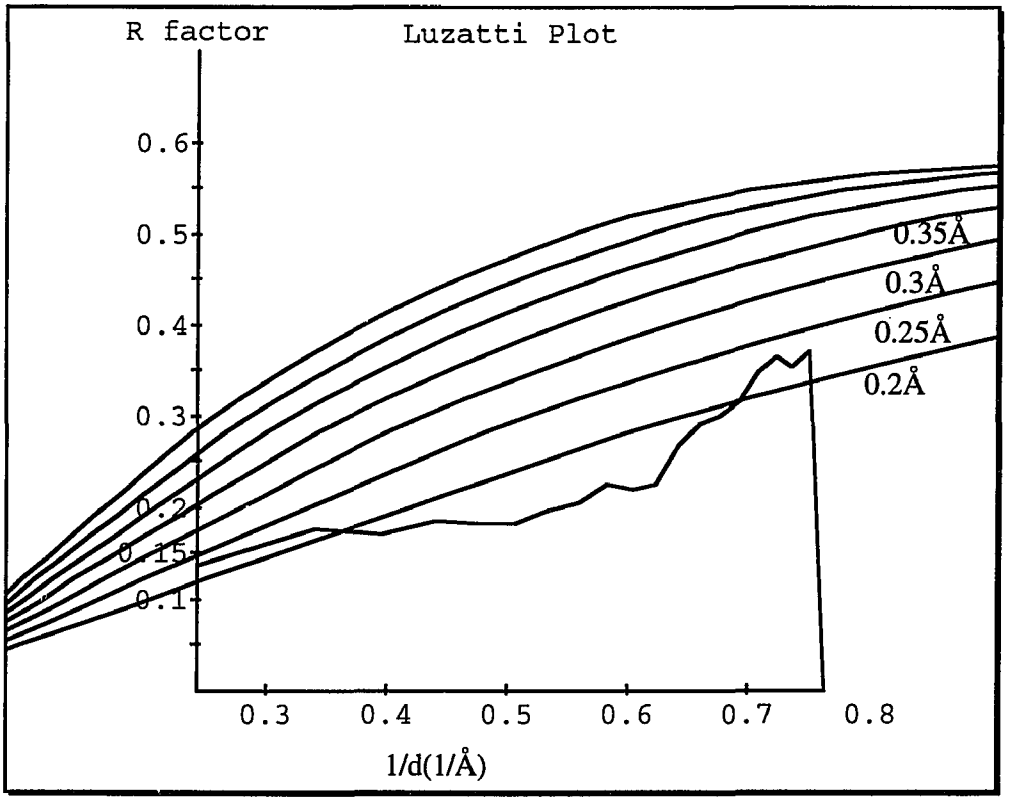


Figure 5-2. R-factor versus resolution, theoretical estimates of the r.m.s. positional errors in atomic coordinates according to Luzzati(1952) are plotted. The coordinate error estimated from this plot is less than 0.2Å with an upper limit of 0.23Å.

The isotropic temperature factor can be plotted as a function of the residue number (see **Figure 5-3** and **Figure 5-4**). **Figure 5-3** shows that the temperature factors of the backbone phosphate groups (between 8 to 11 $Å^2$, μ=0.32 to 0.37Å) are larger than the average B values of DNA bases (about 6 $Å^2$, μ=0.28Å). **Figure 5-4.** show that most of the water molecules have B values between 10 to 45 $Å^2$ (μ=0.36-0.75Å). The numbers here imply that the DNA bases are relatively rigid and that the solvent has much more freedom to move or is slightly disordered. Waters which are in direct contact with DNA are found to have smaller B-factors due to the restriction

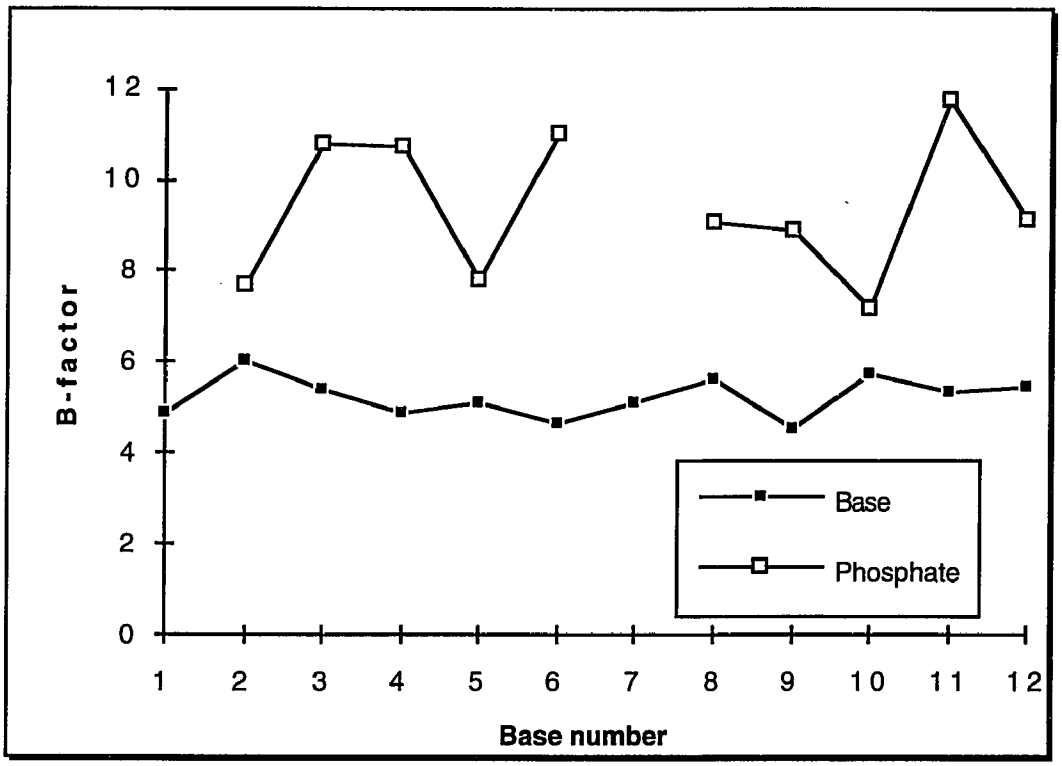


Figure 5-3 The average isotropic temperature factor in the magnesium only structure. The temperature factors are from our final refined structure. The mean values of the DNA bases (residue 1 to 12) and backbone phosphate groups temperature factors are plotted against the residue number.

that applies to the water molecule. The B value of the magnesium ions are between 17 and 36 $Å^2$ ($17 < B(Mg^{2+}) < 36Å^2$, $0.46< \mu(Mg^{2+})<0.68Å$). As shown in **Figure 5-5.** the average backbone B-value in the mixed salt structure are higher than those in the magnesium only structure, which means that the backbone of the magnesium only DNA structure is less flexible and has fewer micro arrangements than the mixed salt's. The B value does not necessarily imply that the atoms are actually "vibrating" over an r.m.s. Δr of 0.5Å but rather that there is a mixture of large scale libration type motions and more importantly small variations in the ordering of the crystal. The higher B's of the phosphates clearly indicates that they exhibit more micro arrangements than the bases. The high B's of water environment are also an indication of its greater micro variability.

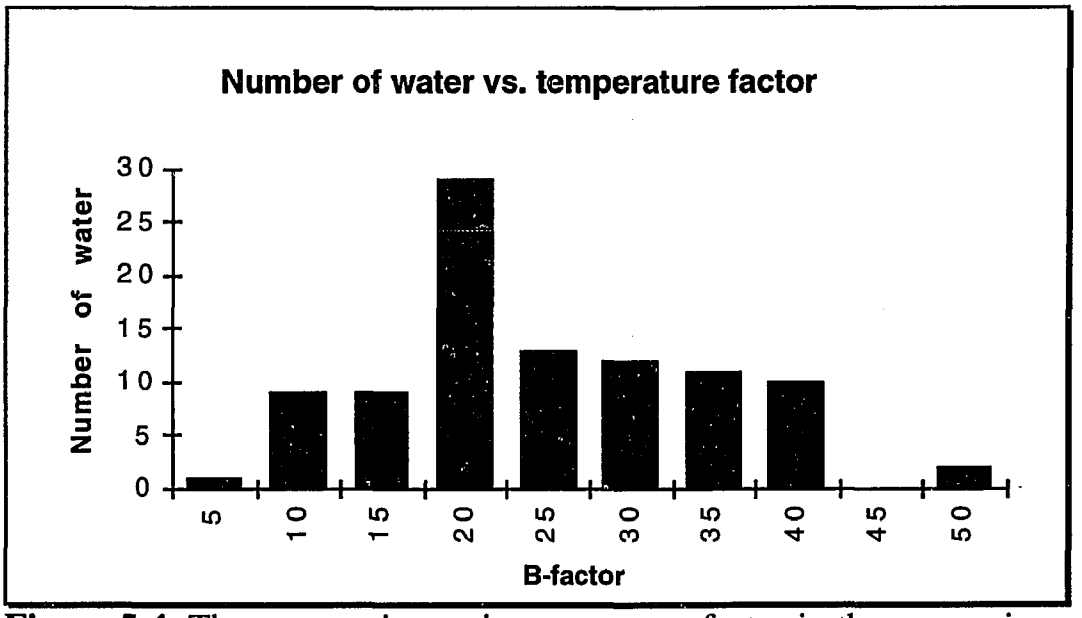


Figure 5-4 The average isotropic temperature factor in the magnesium only structure. The temperature factors are from our final refined structure. The number of water molecules vs. temperature factor. More than 90% of the water molecules fall into a reasonable range from 10 to 40 $Å^2$ and about 30% of the water have B value 20 $Å^2$.

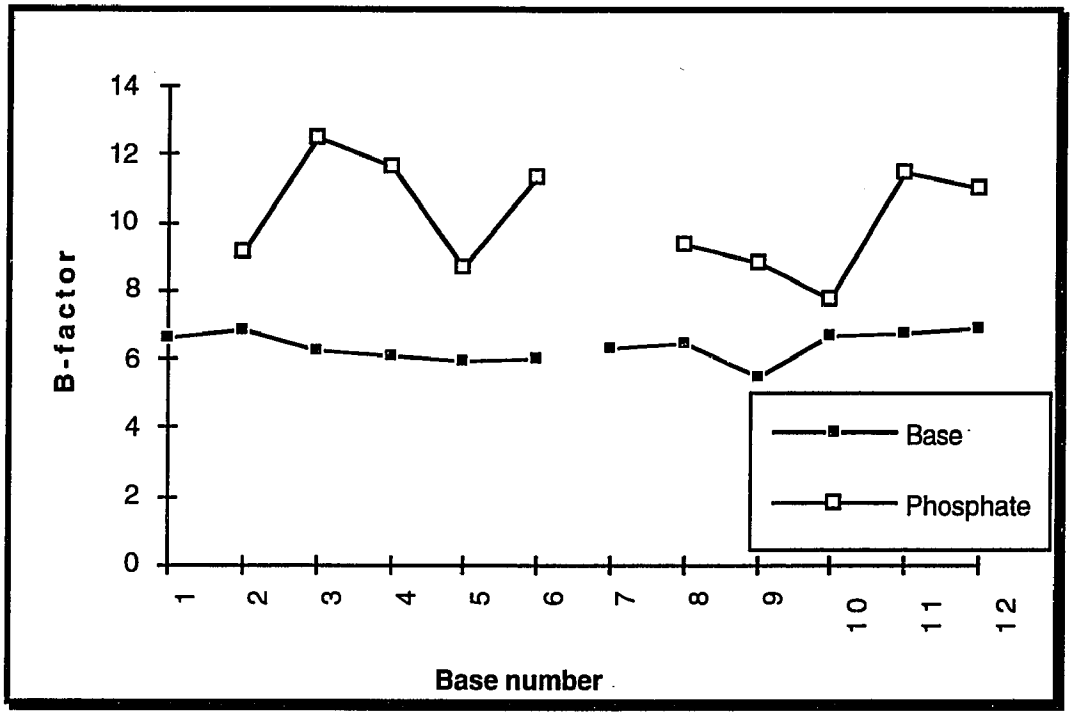


Figure 5-5. The average isotropic temperature factor vs. DNA base number in the mixed salt structure. The temperature factors of the DNA bases are about 6.5 $Å^2$ and the backbone phosphate groups are between 9 to 13 $Å^2$.

5.4. RESULTS AND DISCUSSION

5.4.1. NUCLEIC ACID STRUCTURE

5.4.1.a Overall View

There is a great deal of similarity between our structure and that of the mixed salt model [2]. The evidence is that the root mean square (r.m.s.) deviation between the entire molecule and that of the mixed salt model is less then 0.3Å (**Figure 5-6**).

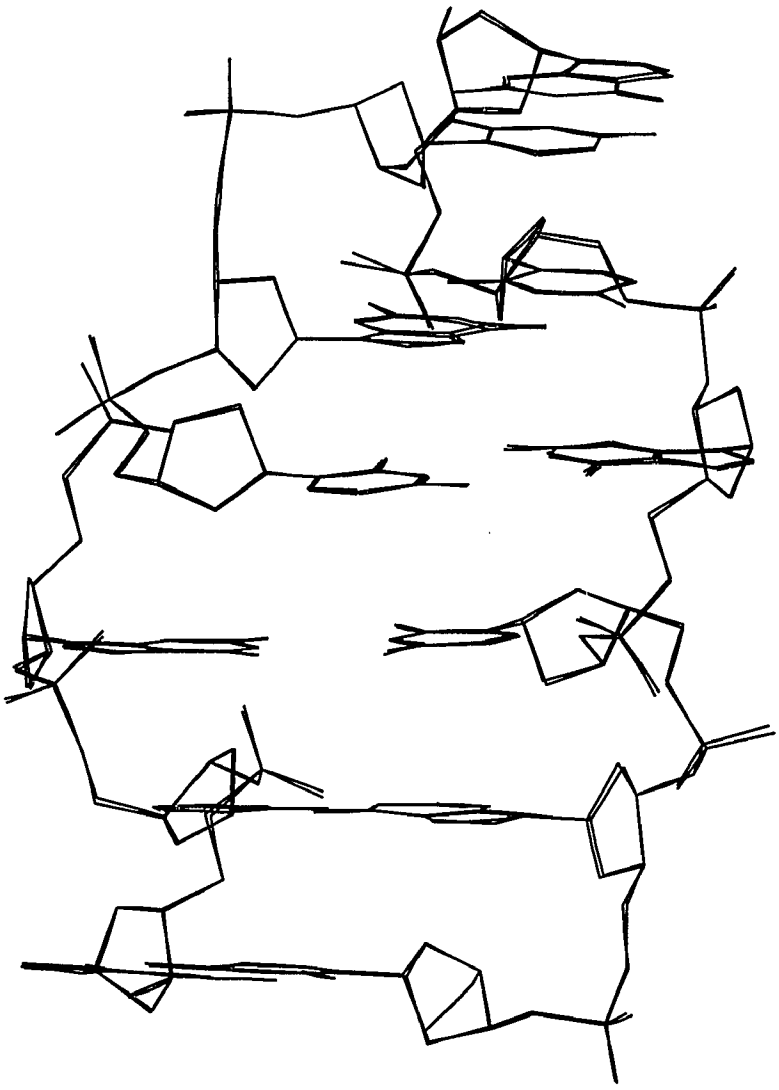


Figure 5-6. Stick figure of overlay of Mg-DNA and mixed salt-DNA structure.

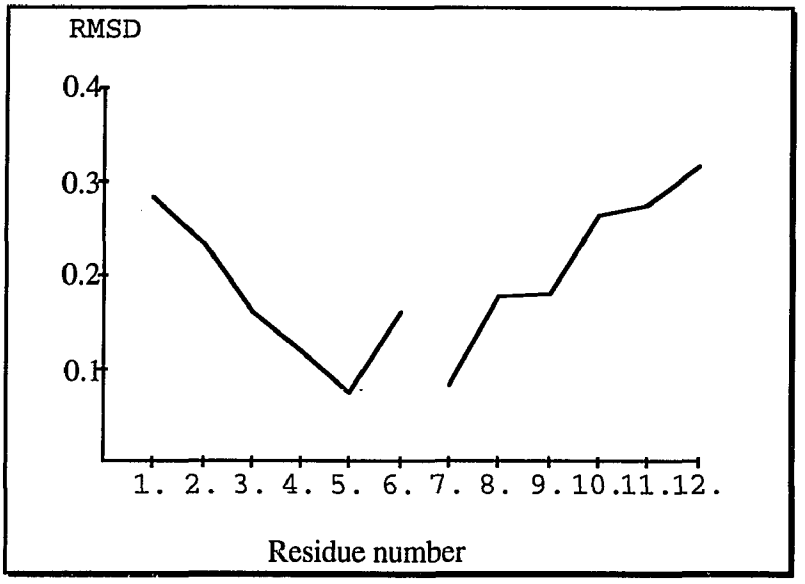


Figure 5-6 The comparison of the R.M.S. differences between magnesium-d(CGCGCG)$_2$ and the mixed salt-d(CGCGCG)$_2$. The average of r.m.s. deviation for the entire molecule is about 0.2Å.

To analyze the differences in geometry between the two structures the program "Curves" [18] was used. The nucleotide parameter (see section 1.2. for explanation of parameter) averages of our structure were found to be: rise=3.59Å, roll=2.81°, twist=25.58° compared to rise=3.58Å, roll=2.73°, twist=25.54° for the mixed salt model. The backbone conformations are similar as indicated by the fact that their torsion angles are about the same. **Figure 5-7** shows the stick figure overlay of the two structures for comparison. The sugar puckering conformations are slightly different at 2GUA as shown in **Table 5-4**. **Table 5-5** shows the r.m.s. of the magnesium structure from the standard bases (root-mean-square deviations from an accepted set of values based on the geometry of small organic molecules). The

magnesium structure has smaller r.m.s. than of the mixed salt structure. This may reflect the fact that the magnesium only structure has a slightly lower resolution so minor distortions are less evident or the fact that the mixed salt crystal varies more, possibly due to apparent distortions associated with the ZI and ZII disordering.

Table 5-4. Comparing the sugar puckering of magnesium $d(CGCGCG)_2$ to the mixed salt $d(CGCGCG)_2$ model.

Base	magnesium $d(CGCGCG)_2$ Pucker	mixed salt $d(CGCGCG)_2$ Pucker
1 CYT	C2' -endo	C2' -endo
2 GUA	C3' -endo	C4' -exo
3 CYT	C2' -endo	C2' -endo
4 GUA	C3' -endo	C3' -endo
5 CYT	C2' -endo	C2' -endo
6 GUA	C2' -endo	C2' -endo
12 GUA	C2' -endo	C2' -endo
11 CYT	C2' -endo	C2' -endo
10 GUA	C3' -endo	C3' -endo
9 CYT	C2' -endo	C2' -endo
8 GUA	C4' -exo	C4' -exo
7 CYT	C2' -endo	C2' -endo

Table 5-5. Comparison of the least squares fitting of standard bases (standard/Mg only structure and standard/mixed salt model).

		magnesium $d(CGCGCG)_2$	mixed salt $d(CGCGCG)_2$
Base	Residue	r.m.s.(Å)	r.m.s.(Å)
CYT	1	0.029	0.036
GUA	2	0.051	0.062
CYT	3	0.017	0.028
GUA	4	0.045	0.065
CYT	5	0.020	0.023
GUA	6	0.050	0.067
GUA	12	0.053	0.067
CYT	11	0.014	0.036
GUA	10	0.037	0.074
CYT	9	0.029	0.022
GUA	8	0.055	0.056
CYT	7	0.029	0.033

Table 5-6. The summary of magnesium-$d(CGCGCG)_2$ conformation parameters

Base	Sugar	Backbone Orientation	
	Pucking	χ	γ
CYT	C2'-endo	anti	+sc
GUA	C3'-endo	syn	-sc

The influence of magnesium on the conformation of $d(CGCGCG)_2$ is summarized in **Table 5-6.** It shows that in general, sugar puckering modes obey the normal C2'-endo for cytosine and C3'-endo for guanine with the

exception of C4'-exo at G8. The orientations about the glycol bond (connecting base-sugar, torsion angle χ) are anti for cytosine, syn for guanine. The conformation about the C4'-C5' bond (connecting phosphate sugar, torsion angle γ) is +sc and -sc for cytosine and guanine, respectively (see chapter 1. Introduction and Background part for definitions).

5.4.1.b The Backbone Conformation

In previous Z-DNA structures [3,4,5,6,19] two different conformations (ZI and ZII) of the GpC phosphate group were identified. The ZI and ZII conformations are shown in **Figure 5-8**. In the ZII conformation the phosphate group of ZI is rotated such that it is bridged to a guanine by forming hydrogen bonds with hydrated ions. The ZII conformation is only observed when it is stabilized by specific ion interactions (here a magnesium hydrated complex). The magnesium structure has only one conformation (ZI). The explanation for this is obtained from a detailed analysis of the DNA backbone. At regions of P3 and P9, where the phosphates are rotated to the ZII conformation, there appears to be more water around each of the phosphate groups in the magnesium structure compared to the mixed salt structure. These water molecules form specific hydrogen bonds with O1P and O2P of C3 and C9, which bridge from the backbone to the solvent network. These hydrations stabilize the ZI conformation so that it is the only conformation that exists in the magnesium only structure (see **Table 5-7** for more information of water-DNA interaction analysis).

5.4.2. SOLVATION-STRUCTURE OF CATIONS AND SOLVENT

5.4.2.a Magnesium-DNA

Our structure contains four magnesium complexes per DNA molecule. The magnesium ions bond directly and indirectly to the surface of Z-DNA.

At physiological pH, there are ten negative charges per asymmetric unit in $d(CGCGCG)_2$ but there are only four magnesium ions in this structure providing a total positive charge of eight. How are the charges neutralized in the Z-DNA crystal structure? The negative charges must be neutralized, either by specific, strong cation coordination or by nonspecific, weak cation binding.

So far, from the reported crystal structures, the positive charges of the visualized cations are never sufficient to neutralize the negative charges of the phosphate groups. This experiment was designed such that magnesium is the only metal cation introduced in the crystallization set up. In mixed salt structure, the missing cations may be due to the presence of monovalent cation (sodium). The presence of sodium ions appears to shift those sites where magnesium is stable. This may be due to the fact that two magnesium ions in a region would result in too much positive charge so they are disordered.

The explanation for the unsaturated positive charge may be because the electrostatic neutralization is delocalized in both cases. We may account for the missing charge from some combination of protons and disordered magnesium ions. There are limitations on the crystallographic models. Even at a resolution of 1.3Å, there is still uncertainty in the final models even with an R-factor of 17.1%. Presumably the large solvent channel region in between

the two symmetry related DNA duplex was highly disordered or in motion and thus invisible on the time scale of crystallography. This less well defined region may be the location of an additional metal ion and it is possible that metal ions and water are both partially occupying the same region.

Even though magnesium is the only source of metal ion introduced in our structure, there is no clear evidence that all the positive charges that are neutralizing the phosphates come from these metal ions. To investigate how charges are neutralized in the DNA crystal structure, one may have to use low temperature to slow down the molecular motion or use the neutron diffraction experiments to identify the hydrogens.

Shown in **Figure 5-1** are the positions of the magnesium-water complexes located on the surface of the Z-DNA.

Our structure and the mixed salt structure show a difference of magnesium cation locations and the number of hydrogen bond contacts between the magnesium complexes and the DNA.

There are two types of magnesium-water complexes in this structure. In type one, two magnesium ions form a dimer coordinated with ten oxygen or nitrogen atoms. The two magnesium ions are linked by two water bridges and are about 2.8Å apart. Gessner [2] has reported this kind Mg-H_2O dimer complex in the mixed salt structure but it does not have a good octahedral geometry. Chen [3] reported that in the dimer complex one of the magnesium is half occupancy at low temperature. The second type is a central magnesium forming an octahedral complex coordinated to six water molecules (**Figure 5-1**). None of the magnesium in these complexes is in contact with the Z-DNA surface directly but through the hydration shell.

Table 5-7 shows the favorable interactions of the magnesium-water complexes with DNA and water. Hydrogen bonds were assumed when the distance between the donor and acceptor atoms was less then 3.2Å and the geometry was acceptable.

The **Mg(13)-Mg(14)** hydrated complex is located in the groove between two symmetry related molecules and in contact with nitrogens and keto oxygens of the bases. The hydrated cations bridge base atoms of opposing strands across the surface on Z-DNA. The magnesium ion is coordinated directly with a base nitrogen N_7 of G6 and indirectly through water mediated hydrogen bonds to the keto oxygen atom O_6 of G6, G10 and G12. Besides intermolecular hydrogen bonding, there is intermolecular bonding connecting the symmetry related (parallel strand) DNA molecules' bases #G6, #C5 and some symmetry related water molecules.

The **Mg(15)** hydrated group was located in the groove between base pairs G4-C9 and C5-G8. The hydrogen bonding connects the opposite chains of the phosphate backbone, bonding to the two free oxygen atoms (O_1P and O_2P) of G6 and G10. Additionally, the complex hydrogen bonds to the keto oxygen atoms (O_6) of G4 and G8, and forms an intermolecular contact to a symmetry related (parallel strand) DNA base N_7 of #G8.

Table 5-7 Hydrogen bond contacts with Mg-H_2O complexes

a. Interaction with (Mg(13)-Mg(14)) dimer complex

Residue	Atom	Residue	Atom	Dist (Å)	Residue	Atom	Residue	Atom	Dist (Å)
10	O6(G)	23	O	2.642	#5	O1P(C)	29	O	2.573
12	O6(G)	21	O	2.647	#6	O6(G)	29	O	2.778
45	O	25	O	2.720	#6	N7(G)	29	O	3.155
45	O	27	O	3.158	#6	N7(G)	27	O	2.976
47	O	22	O	2.662	#62	O	27	O	2.859
47	O	28	O	2.821	#6	O6(G)	26	O	2.931
25	O	106	O	3.159	#6	N7(G)	26	O	2.991
23	O	107	O	2.789	#43	O	28	O	2.851
29	O	106	O	2.677					

b. Interaction with Mg(15) complex c. Interaction with Mg(16) complex

Residue	Atom	Residue	Atom	Dist (Å)	Residue	Atom	Residue	Atom	Dist (Å)
10	O1P(G)	33	O	2.620	5	O2P(C)	39	O	2.859
4	O6(G)	31	O	2.913	8	O1P(G)	41	O	2.842
6	O2P(G)	32	O	3.008	9	O1P(C)	36		2.416
6	O2P(G)	30	O	2.842	10	O2P(G)	37	O	2.644
8	O6(G)	32	O	2.844	10	O2P(G)	38	O	2.569
32	O	91	O	2.927	39	O	163	O	2.876
62	O	34	O	2.919	60	O	40	O	2.775
33	O	91	O	2.960	60	O	41	O	3.182
31	O	112	O	2.983	#6	O1P(G)	40	O	2.495
30	O	81	O	2.973	#6	O5' (G)	40	O	3.169
35	O	90	O	2.801	#6	O4' (G)	37	O	2.910
34	O	91	O	3.141	#121	O	41	O	2.764
32	O	82	O	3.040					
42	O	32	O	3.086					
#8	N7(G)	34	O	2.756					
#72	O	35	O	3.102					
#30	O	112	O	2.878					

The **Mg(16)** forms an interstrand phosphate to phosphate contact in the structure. The bonding involves oxygen atoms (O_1P and O_2P) of the phosphate group and ribose ($O_{4'}$, and $O_{5'}$) of residues C5, G8, C9, G10, and #G6.

In our magnesium only structure, the locations of magnesium are highly specific so that the hydrated cations' spatial fitting in the groove of the DNA double helix results in the complexes' having the possibility to form more hydrogen bonds than in the mixed salt structure. In the equilibrium between B-DNA and Z-DNA in poly(dC-dG), the cobalt hexamine ion formed five hydrogen bonds to DNA, resulting in an effect of 4 orders of magnitude in stabilizing the Z-form structure [20] as compared to the Mg^{2+}-Na^{+} mixed structure. As Table 5-6 suggested, the hydrogen bonding possibilities for hydrated cation complex Mg(13)-Mg(14), Mg(15) and Mg(16) are 17, 17, and 12, respectively. One can predict that the magnesium structure should be much more effective in stabilizing the Z-DNA structure due to more hydrogen bonds interacting with the surface of DNA.

In summary, in the present structure the hydrated cation complexes show an extensive contact surface, with hydrogen bonds to the phosphate groups and bases of DNA. The most strinking finding is that the crystal structure shows almost no direct ligation (only one out of 46) between the cations and the DNA. The principles are still poorly understood, but it appears that these contacts effect the geometry of the phosphate backbone and bases, which in turn permits the formation of the stable conformation.

5.4.2.b DNA Solvation

DNA is heavily hydrated. In the primary hydration shell the preferred hydration sites usually are around phosphate groups, sugar oxygens, and base nitrogens and oxygens [21,22]. The solvent sites of the magnesium $d(CGCGCG)_2$ and mixed salt $d(CGCGCG)_2$ are compared here. Generally, the preferred solvent sites are about the same, but the degree of solvent bonding is different. The former has more attached water molecules than the latter. According to the hydrogen bonding frequency, the water-DNA interactions are classified to five groups as in the scheme described in **Figure 5-8**. site one (O_1P and O_2P) , site two (O_6, O_2, and N_7), site three ($O_{3'}$, and $O_{5'}$), site four (N_2 and N_4), and site five ($O_{4'}$), in the order of decreasing contacts. **Table 5-8** compares the number of water-DNA interactions for the magnesium structure and the mixed salt structure. Phosphate $O_{3'}$ and $O_{5'}$ are even less hydrated. Most water molecules are associated with backbone anionic phosphate oxygens and DNA surface sites.

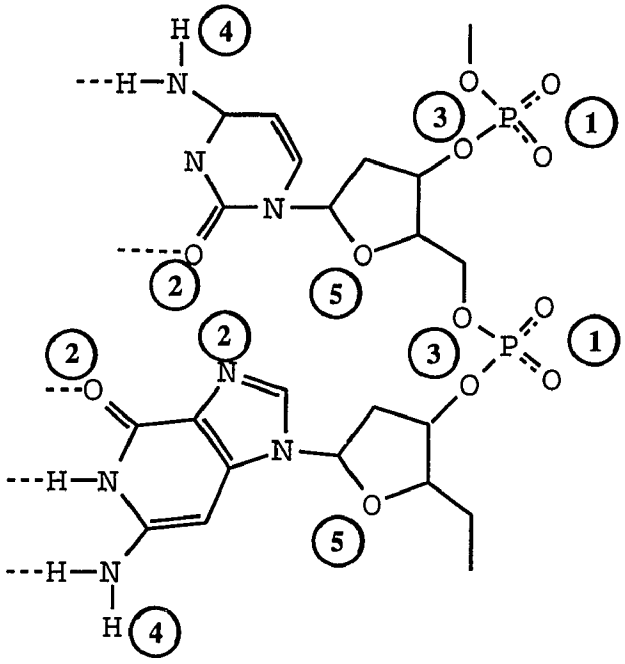


Figure 5-8. Preferred Hydration Sites in the Mg-DNA Structure.

The data showed that, overall, the magnesium structure has more water interacting with DNA (97 water molecules involved in 105 possible hydrogen bond contacts) than in the mixed salt model (78 water molecules involved in 83 possible hydrogen bond contacts). The most highly hydrated sites are phosphate oxygen O_1P and O_2P. There are 45 water molecules that interact with these atoms, averaging about four water molecules per phosphate compared to 35 water molecules in the mixed salt structure. The less hydrated sites are nitrogen and keto oxygen atoms, and the number of bound solvents is 30 versus 25 that are found in the mixed salt structure.

Figure 5-9 illustrates the comparison of associated water at each binding site. The data shows more extensive water-DNA interactions in the magnesium structure at every binding site. The result of more water molecules

associated with DNA structure means that the water molecules show higher order in the magnesium structure. It was expected that the water molecules around the DNA should be more organized as a consequence of magnesium complexes being a better fit onto the DNA surface.

Another interesting observation arose when special attention was paid to the regions of P3 and P9, the locations that are rotated in the ZII conformation. There were more attached water molecules around the phosphate groups P3 and P9 in the magnesium structure compared to mixed salt structure. These water molecules form hydrogen bonds to link the O_1P and O_2P and the solvent network. This plays an important role in stabilizing the conformation. In the mixed salt crystal those are more disordered so the solvent is harder to visualized. The effects of the solvent disorder may be one of the reasons causing unidentified charges in crystal structures.

Table 5-8. The comparison of water-DNA interactions of magnesium-d(CGCGCG)$_2$ and mixed salt-d(CGCGCG)$_2$.

	Magnesium-d(CGCGCG)$_2$					Mixed salt-d(CGCGCG)$_2$				
Base	Site 1	Site 2	Site 3	Site 4	Site 5	Site 1	Site 2	Site 3	Site 4	Site 5
C1	0	1	2	1	0	0	2	2	2	0
G2	3	4	2	1	0	2	4	0	0	0
C3	5	2	1	1	0	3	1	0	1	0
G4	5	3	2	1	1	2	1	1	1	1
C5	4	1	1	1	0	5	1	1	0	0
G6	4	5	1	1	1	3	3	1	0	1
C7	0	2	1	1	0	0	1	4	0	0
G8	6	3	0	1	0	4	3	0	0	0
C9	3	2	1	1	0	2	2	0	1	0
G10	6	3	1	1	0	6	3	1	1	0
C11	4	2	2	1	0	5	1	2	2	0
G12	5	2	2	1	0	3	3	1	0	0
#HB	45	30	16	12	2	35	25	13	11	2
Total	105					83				
#H2O	45	25	13	12	2	32	23	10	11	2
Total	97					78				

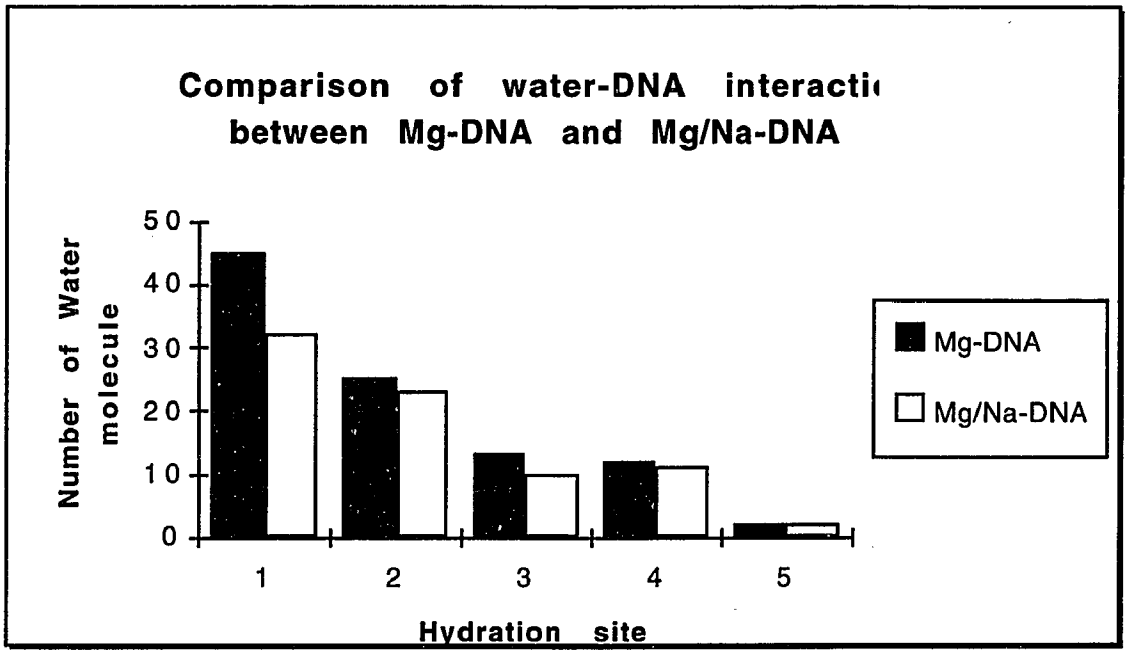


Figure 5-9. The comparison of water-DNA interactions. At each hydration site, there are more attached water molecules in the magnesium-d(CGCGCG)$_2$ structure than in the mixed salt-d(CGCGCG)$_2$.

5.5. CONCLUSIONS

Features of the non-covalent cation-DNA interactions that stabilize DNA crystal structure have been inferred from the x-ray crystallographic structure of the hexanucleotide $d(CGCGCG)_2$ at high resolution. Cation binding sites and DNA solvation have been visualized in great detail. The crystal structure displays substantial solvent and solute order; the hydrated magnesium complexes exhibit specific contacts with DNA; and there is only one GpC phosphate conformation (ZI). This contrasts with the identical hexamer crystallized using higher ionic strength mixed salt buffer. This structure exhibits greater solvent and solute disorder, less well defined cation-DNA contacts and two conformations (ZI and ZII) of several of the phosphate groups of the nucleic acid. It is the first time the dehydrated magnesium-water dimer complex is observed with certainty.

This work contributes to two aspects of our understanding of nucleic acid structure. First, it shows that the left-handed Z-form can be crystallized under relatively low ionic strength condition. The crystal structure clearly shows a better-defined solvent and solute structure than the high ionic strength structure. Secondly, specific magnesium-DNA interactions stabilize the ZI conformation. The cation-DNA interactions are more intensive in the magnesium only structure than in the mixed salt structure, so the structure should be more stable than the mixed salt structure. In the mixed salt (high cation strength) structure, electrostatic neutralization appears non-specific and no dehydration of the cations is observed. On the other hand, the pure magnesium structure (low cation strength) needs more specific contacts to stabilize the crystal by increasing the direct coordination of Mg-DNA binding

and by organizing the water-DNA contacts in greater order as reflected in the lower B-factors of the backbone phosphate groups. The magnesium ions effect on the fine geometry of the DNA and the water-mediated contacts to the phosphate groups appear to contribute part of the specificity.

Future outlook

From this study, it is clear that the role of the magnesium ions play in the formation of a stable interface with the DNA and in the contribution to the specificity of the DNA geometry or the conformational adjustments are critical. The specificity of the Mg-DNA interactions are not only direct between the magnesium ions and the DNA, but also a combination of indirect structural effects and solvent-mediated interactions. To understand the contribution of DNA structure and hydrations in this context, it would be interesting to extend the crystallographic studies to the structural comparison of the free DNA its target protein or drug.

5.6. REFERENCES

1. Brunger, A.T. X-PLOR Manual, Version 2.1 (Yale University, 1990).

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

3. Chen, S.L. "A Dissertation for the Degree of Doctor of Philosophy" (CUNY), (1993).

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

5. Wang, A.H.J., Quigley, G.J., Kolpak, F.J., *et al.* Science, **1981**, 211, 171.

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

7. Felsenfeld, G. Ann. Rev. Biochem., **1967**, 36, 407.

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

9. Rose, D.M. Proc. Natl. Acad. Sci. U.S.A., **1980**, 77, 6289.

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

11. Holbrook, S.R., Wang, A.H.-J., Rich, A. and Kim, S.H. J. Mol. biol., **1986**, 187, 429.

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

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

14. Voet, J.G. Biochemistry (John Wiley & Sons, Inc., 1990).

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

16. Glusker, J.P. and Trueblood, K.N. Crystal Structure Analysis (Oxford University Press, New York, Oxford, 1985).

17. Luzzati, V. Acta Crystallog., **1952**, 5, 801.

18. Lavery, R. and Sklenar, H. J. Biomol. Struc. & Dynamics, **1988**, 6, 63.

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

20. Behe, M. Proc. Natl. Acad. Sci. U.S.A., **1981**, 78, 1619.

21. Quigley, G.J. *The Role of Hydrogen Bonding in Nucleic Acid Conformation* 1-121 (1986).

22. Quigley, G.J., Teeter, M.M. and Rich, A. Proc. Nat. Acad. Sci. USA, **1978**, 75, 64.

Bibliography

Chapter 1

1. Watson, J.D. and Crick, F.H. Nature, **1953**, 171, 737.

2. Wilkins, M.H.F. **1963**, 140, 941.

3. Saenger, W. Principles of Nucleic Acid Structure 1-556 (Springer-Verlag, Berlin, 1984).

4. Pezzano, H. and Podo, F. Chem. rev., **1980**, 80, 365.

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

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

7. Waring, M.J. Annu. Rev. Biochem, **1981**, 50, 159.

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

9. Kornberg, A. DNA synthesis (W. H. Freeman, San Franciso, 1974).

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

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

12. Barton, Lawrence and Teeter. 1980. In Metal ion-nucleic acid interactions, eds.,

13. Teeter, M.M., Quigley, G.J. and Rich, A. 1980. In Nucleic Acid-Metal Ion Interactions, Spiro, T.G. eds., John Wiley and Sons. New York.

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

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

16. 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.

17. Behe, M. Proc. Natl. Acad. Sci. U.S.A., **1981**, 78, 1619.

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

19. Dickerson, R.E., Drew, H.R. and Conner, B. 1981. In Meeting, Albany, N.Y., USA, April 26-29, 1981., Sarma, R.H. eds., Adenine Press. Guilderland, N.Y., USA.

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

21. 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).

22. Baguley, B.C. Anti-Cancer Drug Design, **1991**, 6, 1.

23. Pigram, W., Fuller, W. and Hamilton, L. Nature New Biol., **1972**, 235, 17.

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

25. Arcamone, F. In "x-ray crystallography and drug action" (Oxford Univ. press, 1984).

26. Chen, K.X., Gresh, N. and Pullman, B. J. Biomol. Struct. Dyn., **1985**, 3, 445.

27. Lerman, L.S. J. mol. Biol., **1961**, 3, 18.

28. Chen, K.X., Gresh, N. and Pullman, B. Mol. Pharmacol., **1986**, 30, 279.

29. Hui, X., Gresh, N. and Pullman, B. NAR, **1990**, 18, 1109.

30. Pullman, B. Adv. in Drug research, **1979**, 18, 113.

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

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

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

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

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

36. Berman, H.M., Stallings, W., Carrell, H.L., *et al.* Biopolymers, **1979**, 18, 2405.

37. Schwartz, H.S. Adv. Cancer Chemother, **1979**, 1, 1.

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

39. Dickerson, R.E. Nucleic Acids Res., **1989**, 17, 1797.

40. Hol, W.G.J. Nature, **1978**, 273, 443.

41. Neidle, S. Methods in enzymology, **1991**, 203, 433.

42. Karplus, M. and Pestko, G.A. Nature, **1990**, 347, 631.

43. Martin, Y.C. Methods in enzymology, **1991**, 203, 587.

44. Levitt, M. Proc. Natl. Acad. Sci. USA, **1978**, 75, 640.

45. Weiner, S.J., Kollman, P.A., Case, D.A., *et al.* J.A.C.S., **1984**, 106, 765.

46. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., *et al.* J. Comp. Chem., **1983**, 4, 187.

47. Karplus, M., III, C.B.B. and Brunger, A. 1986. In Current Communications in Molecular Biology: Computer Graphics and Molecular Modeling; Meeting, Cold Spring Harbor, N.Y., USA, De c. 10-13, 1985. Ix+145p, Fletterick, R.A.M.Z. eds., Illus.

48. Brunger, A.T. X-PLOR Manual, Version 2.1 (Yale University, 1990).

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

50. Glusker, J.P. and Trueblood, K.N. Crystal Structure Analysis (Oxford University Press, New York, Oxford, 1985).

51. Rhodes, G. Crystallography made crystal clear (Academic Press, Inc., 1993).

Chapter 2

1. Dipankar, S. and Gilbert, W. Nature, **1988**, 334, 364.
2. Dipankar, S. and Gilbert, W. Nature, **1990**, 344, 410.
3. Zahler, A.M. and Prescott, D.M. Nature, **1991**, 350, 718.
4. Rajagopal, P. and Feigon, J. Nature, **1989**, 339, 637.
5. Guo, Q., Lu, M. and Kallenbach, N.R. J. A. C. S., **1992**, 2451.
6. Kim, J., Cheong, C. and Moore, P.B. Nature, **1991**, 351, 331.
7. Arnott, S. and Selsing, E. J. Mol. Biol., **1974**, 88, 509.
8. Kang, C., Zhang, X. and Rich, A. Nature, **1992**, 356, 126.
9. Smith, F.W. and Feigon, J. Nature, **1992**, 356, 164.
10. Oka, Y. Nucleic Acids Res., **1987**, 15, 8877.
11. Dipankar, S. and Gilbert, W. Methods in Enxymology, **1992**, 211, 191.
12. Voet, J.G. Biochemistry (John Wiley & Sons, Inc., 1990).
13. Murchie, A.I.H., Carter, W.A., Portugal, J. and Lilley, D.M. Nucleic Acids Res, **1990**, 18, 2599.

Chapter 3

1. Gao, Q.I., Williams, L.D., Egli, M., *et al.* Proc. Natl. Acad. Sci. U. S. A., **1991**, 88, 2422.

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

3. Pelaprat, D. and Pecq, J.B.L. J. Med. Chem., **1980**, 23, 1336.

4. Laugaa, P. **1985**, 24, 5567.

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

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

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

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

9. Brunger, A.T. X-PLOR Manual, Version 2.1 (Yale University, 1990).

10. Lavery, R. and Sklenar, H. J. Biomol. Struc. & Dynamics, **1988**, 6, 63.

11. Berman, H.M., Neidle, S. and Stodola, R.K. P.N.A.S., **1978**, 75, 828.

12. Neidle, S., Webster, G.D., Baguley, B.C. and Denny, W.A. Biochem Pharmacol, **1986**, 35, 3915.

Chapter 4

1. The compound was synthesized by Huang, Wolin A Dissertation for the Degree of Doctor of Philosophy (CUNY), **1990**.

2. International Tables for X-ray Crystallography, 1974, Vol. IV, pp. 72-75

3. Main, P., Hull, S. E., Lessinger, L., Germain, G., Declercq, J.-P & Woolfson, M. M. ,**1978**, MULTAN 78. *A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data.* Unives. of York, England, and Louvain, Belgium.

4. Geer, A. and Franck, R. W., **1991**.

5. Scheiner, P., Geer, A. etc. and Schinazi R.F., J. Med. Chem., **1989**, Vol. 32, No. 1, 73-76

6. Robins, M.J., and Dalley N.K., J. Org. Chem., Vol. 57, No. 8, **1992**, 2357-2364

7. Shibuya, S., and Ueda, T., J. Carbohyd. Nuc. Nuc., **1980**, 7(1), 49-56.

Garner, P., and Park, J.M., J. Org.. Chem., Vol. 55, No. 12, **1990**, 3772-3787

8. Kawasaki, M. et al., Tetrahedron, **1988**, 44, 5727-5743

9. Sammes, P. G.; Bates, M. A., J. Chem. Soc., Chem. Commun., **1983**, 896-898

10. Yin, H. and Franck, R. W., J. Org. Chem., **1992**, Vol. 57, No. 2, 644-651

Chapter 5

1. Brunger, A.T. X-PLOR Manual, Version 2.1 (Yale University, 1990).

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

3. Chen, S.L. (1993).

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

5. Wang, A.H.J., Quigley, G.J., Kolpak, F.J., *et al.* Science, **1981**, 211, 171.

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

7. Felsenfeld, G. Ann. Rev. Biochem., **1967**, 36, 407.

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

9. Rose, D.M. Proc. Natl. Acad. Sci. U.S.A., **1980**, 77, 6289.

10. Pohl, F.M. and Jovin, T.M. J. mol. biol., **1972**, 67, 375.

11. Holbrook, S.R., Wang, A.H.-J., Rich, A. and Kim, S.H. J. Mol. biol., **1986**, 187, 429.

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

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

14. Voet, J.G. Biochemistry (John Wiley & Sons, Inc., 1990).

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

16. Glusker, J.P. and Trueblood, K.N. Crystal Structure Analysis (Oxford University Press, New York, Oxford, 1985).

17. Luzzati, V. Acta Crystallog., **1952**, 5, 801.

18. Lavery, R. and Sklenar, H. J. Biomol. Struc. & Dynamics, **1988**, 6, 63.

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

20. Behe, M. Proc. Natl. Acad. Sci. U.S.A., **1981**, 78, 1619.

21. Quigley, G.J. *The Role of Hydrogen Bonding in Nucleic Acid Conformation* 1-121 (1986).

22. Quigley, G.J., Teeter, M.M. and Rich, A. Proc. Nat. Acad. Sci. USA, **1978**, 75, 64.