Order Number 9304642

NMR studies of single-site phosphate ethylated oligodeoxynucleotides

Cahill, Sean Michael, Ph.D.

City University of New York, 1992

U·M·I

300 N. Zeeb Rd.
Ann Arbor, MI 48106

NMR STUDIES OF SINGLE-SITE PHOSPHATE ETHYLATED OLIGODEOXYNUCLEOTIDES

by

SEAN MICHAEL CAHILL

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

1992

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

6/15/92
Date

M S Brodo
Chair of Examining Committee

September 3, 1992
Date

Norma Schulz
Executive Officer

Ruth Stark
Supervisory Committee

The City University of New York

Abstract

NMR STUDIES OF SINGLE-SITE PHOSPHATE ETHYLATED OLIGODEOXYNUCLEOTIDES

by

Sean Michael Cahill

Advisor: Professor Michelle S. Broido

Phosphate alkylation of DNA is often the primary event associated with the introduction of certain mutagenic and carcinogenic alkylating agents into the cellular environment. The presence of phosphotriester lesions in DNA has been shown to interfere with template-directed DNA polymerization and the observed reduction in template activity depends on the stereochemistry of the phosphotriester moiety. In order to determine the structural and stability changes in DNA induced by phosphate alkylation and to correlate the observed changes with the stereochemistry of alkylation, a series of ^{1}H and ^{31}P nuclear magnetic resonance (NMR) studies have been performed on an eight-base self-complementary oligonucleotide in its duplex state, $\{d(GGAATTCC)\}_2$, and two of its single-site phosphate ethylated analogues, $(\underline{S}_p,\underline{S}_p)$- and $(\underline{R}_p,\underline{R}_p)$-$\{d(GGA(ethyl)ATTCC)\}_2$. In addition, a series of computer simulations were performed to evaluate the effectiveness of different structure refinement methods for determining nucleic acid conformations in solution using interproton distance data derived from two dimensional nuclear Overhauser effect (2D NOESY) experiments.

The nonexchangeable proton resonances (except for the C5' proton resonances), the imino proton resonances and the phosphorus resonances of each octamer duplex and the stereochemistry of each phosphotriester moiety were assigned using 1D and 2D NMR experiments. The ^{1}H chemical shift differences between the native octamer duplex and its ethylated analogues are localized to nucleotides that flank the phosphotriester lesion and are more pronounced for the $\underline{R}_p$-analogue which contains a phosphotriester ethyl group oriented into the major groove of the helix. Upfield shifts of ^{31}P resonances corresponding to the GpA

and ApT phosphodiesters and the ApA phosphotriester within the $\underline{R}_p$- and $\underline{S}_p$-analogues were observed. The $\underline{R}_p$-duplex has a lowered duplex stability relative to the native octamer duplex and $\underline{S}_p$-duplex as monitored by the temperature dependence of the imino proton linewidths. However, only small differences were observed in the base-pair lifetimes and the ^{31}P spin-lattice relaxation rates between the native octamer duplex and its ethylated analogues.

The computer simulations demonstrate that the energy minimization and molecular dynamics structure refinement methods, when used with a set of interproton distance data which mimics that derived from typical 2D NOESY experiments on nucleic acids, can accurately reproduce many of the conformational features of a target structure such as deoxyribose ring conformations and base-related helicoidal parameters. The interproton distance data obtained on the octamer duplexes by 2D NOESY experiments was incomplete compared to the interproton distance data used in the computer simulations, primarily because of crosspeak overlap. Although the lack of a complete set of interproton distances prevented the use of refinement methods to generate accurate structures for the octamer duplexes, the available interproton distance data suggests that the $\underline{R}_p$-duplex adopts a different conformation in the vicinity of the phosphotriester lesion relative to the native octamer duplex and the $\underline{S}_p$-duplex. Overall, the lowered duplex stability, the chemical shift differences, and the interproton distance differences of the ($\underline{R}_p$,$\underline{R}_p$)-{d(GGA(ethyl)ATTCC)}$_2$ duplex relative to the native octamer duplex and the ($\underline{S}_p$,$\underline{S}_p$)-{d(GGA(ethyl)ATTCC)}$_2$ duplex suggest that a phosphotriester moiety oriented into the major groove induces a conformational change which destabilizes DNA duplexes in solution.

This thesis is dedicated to those people who have inspired me during my years of education and research: my family, my friends, my teachers, my advisor, Michelle Broido and my greatest inspirations, Linda and Hannah.

Acknowledgements

There are many people who have helped me during my years of graduate study. Professor Michelle Broido, my advisor, has helped me from the beginning. Michelle encouraged my interest in NMR and has given me the benefit of her knowledge in this field. I am very much indebted to her abilities as both a teacher and a researcher and to her patience.

I would like to thank the people who have been involved with the NMR group over the years: Sam, Don, Chen, Paula, John and Mike. I have received much help and inspiration from each. I would like to acknowledge every member of the Hunter College Chemistry Department, Cliff, Mike and the rest, who has made my experience there invaluable. I am very grateful to Jerry Zon for synthesizing and purifying the DNA oligomers that were used in this study and to Andy Byrd for helping me implement various 2D NMR pulse sequences on the JEOL NMR spectrometer. I would also like to thank the people at the Graduate School, Sophie especially, for helping me along the way.

I am indepted to my committee members: Professors Ruth Stark, Gary Quigley, Fred Naider, Maria Tomasz, Peter Lipke and Robert Bittman. All of my committee members have given me a great deal of useful advice. I wish to thank Gary Quigley for his expertise in using the XPLOR refinement programs and for his many worthwhile suggestions. I also wish to sincerely acknowledge Ruth Stark who went out of her way to encourage and help me in the final stages of my thesis and its presentation.

Finally, I would like to thank my wife, Linda; she has given me a tremendous amount of encouragement over the years, especially during the times I needed it most.

Thank you everyone.

Table of Contents

List of Abbreviations

Å	Angstrom
A	adenine
a_{ij}	crosspeak volume
ac	anticlinical
ap	antiperiplanar
$a(\tau_m)$	matrix of mixing coefficients
B, B_i	external magnetic field, component in i direction
C	cytosine
CD	circular dichroism
CORMA	COmplete Relaxation Matrix Analysis
D	delay interval
DNA	deoxyribonucleic acid
D_2O	deuterium oxide
e	electron charge
E, E_i	energy, energy of state i
EDTA	ethylenediaminetetraacetic acid, disodium salt
Et	ethyl
f	frequency
F_i	force exerted on atom i
FID	free induction decay
fs	femptosecond
FT	Fourier transform
G	guanine
h	hour
ħ	Planck's constant divided by 2π

H, H_i	Hamiltonian operator, operator for interaction i
HOD	water, residual ^1H in D_2O
HPLC	high pressure liquid chromatography
Hz	Hertz
I	nuclear quantum number
I_i	signal intensity at time i
iPr	isopropyl
ISPA	isolated spin pair approximation
J	scalar coupling constant
$J(\omega)$	spectral density
k	harmonic force constant, Boltzmann's constant
k_i	rate constant for process i
K	degrees Kelvin
kcal	kilocalorie
l_{ij}, u_{ij}	lower, upper bounds
m	mass
M	molar
M_o, M_i	equilibrium magnetization, magnetization vector in i direction
Me	methyl
MHz	megahertz
ml	milliliter
mM	millimolar
mm	millimeter
ms	millisecond
N	north furanose ring conformer, number of particles
N_i	population of state i
NMR	nuclear magnetic resonance

NOE	nuclear Overhauser effect
nm	nanometer
ns	nanosecond
OD	optical density
P	pseudorotation phase angle
PDB	protein data bank
ppm	parts per million
ps	picosecond
q	atomic charge
r, r_{ij}	distance, internuclear vector for spins i and j
R	relaxation matrix
$\underline{R}_p$	R phosphotriester diastereomer
RF	radiofrequency
s	second
S	solvent, scale factor, spectrum, signal, south furanose ring conformer
sc	synclinical
$\underline{S}_p$	S phosphotriester diastereomer
sp	synperiplanar
t, t_1, t_2	time, time dimensions in NMR experiments
T	thymine, absolute temperature
T_m	transition temperature
TSP	3-trimethylsilylpropionate, sodium salt
T_1	longitudinal or spin-lattice relaxation time
UV	ultraviolet
W_i	relaxation probability i
1D, 2D	one dimensional, two dimensional
2D COSY	two dimensional correlated spectroscopy

2D NOESY	two dimensional nuclear Overhauser effect spectroscopy
$\alpha,\beta,\gamma,\delta,\epsilon,\zeta,\chi$	torsion angle variables
$\nu_0,\nu_1,\nu_2,\nu_3,\nu_4$	deoxyribose ring torsion angle variables
α,β	quantum mechanical spin states
β	pulse interval
γ	gyromagnetic ratio
$\Delta\nu$	linewidth
Δ^+,Δ^-	positive, negative interproton distance error
ϵ	dielectric constant
η	nuclear Overhauser enhancement factor, CURVES inclination
θ	notation for angles, CURVES tip
λ	diagonal matrix of eigenvalues of the rate matrix
μ	magnetic dipole moment
μl	microliter
μM	micromolar
ν	frequency
ρ	external relaxation rate, CURVES roll
σ	cross-relaxation rate, CURVES opening
τ	CURVES tilt
τ_c	rotational correlation time
τ_{ex}	exchange time
τ_m	mixing time
τ_{op}	base pair opening lifetime
ϕ	dihedral angle
χ	CURVES buckle
χ	matrix containing eigenvectors of the rate matrix
Ω	CURVES twist
ω	CURVES propeller twist
ω_i	angular frequency i
°C	degrees Centigrade

List of Tables

List of Figures

CHAPTER 1 Introduction

1.1 Phosphate Alkylation of DNA

The importance of deoxyribonucleic acid (DNA) to living systems is undisputed. DNA is the primary constituent of the chromosome which contains the molecular design essential for most living organisms. Survival of the cell and the organism as a whole depends on the ability of the cell to successfully direct the flow of information carried by DNA to macromolecular end products in a precise temporal and spatial fashion. This process is known as gene expression and one of the primary goals of molecular biology is to develop an understanding of the factors that influence gene expression and how these influences are related to the functional operation of the cell and organism.

Gene expression often requires the specific interaction of proteins and/or cofactors with segments of DNA. Any factor that impairs the ability of DNA to serve as a template for such interactions can potentially lead to a diseased state. Modification of DNA function by alkylating agents is a well-established consequence of the introduction of specific mutagens and carcinogens into the cellular environment. DNA alkylation leads to both mutagenic and inactivating DNA alterations and numerous biochemical studies have focussed on correlating the biological effects of DNA alkylation with the site of alkylation.[1,2] The sites on DNA which have been found to be susceptible to alkylation include the nitrogen atoms of the bases, the oxygen atoms of the bases and the phosphate oxygen atoms of the sugar-phosphate backbone.[3,4]

Although the precise role of phosphate alkylation in regard to *in vivo* mutagenesis and carcinogenesis remains to be established, some studies suggest that the presence of a phosphotriester lesion may alter cellular activity. Nitrosourea compounds, which constitute a class of highly potent carcinogens, have been shown to be highly specific toward the phosphate backbone of DNA. For example, 65-75% of the *in vitro* alkylation events of poly-d[AT] by N-methyl-N-nitrosourea and N-ethyl-N-nitrosourea are found as phosphotriesters.[5,6] Although the formation of a phosphotriester can result in backbone

breakage,[7] it appears that, in general, phosphotriester derivatives are very stable under physiological conditions.[8] In particular, DNA phosphotriesters produced *in vivo* in eukaryotic systems are persistent lesions and are probably not repaired enzymatically.[9] These factors suggest that phosphotriesters might be biologically significant lesions and that the role of phosphate alkylation in inducing DNA malfunction should be of considerable interest to studies of mutagenesis and carcinogenesis at the molecular level.

1.2 Biological Consequences of DNA Phosphate Alkylation - *In Vitro* Studies

Numerous studies have been conducted on phosphotriester adducts in order to understand the functional and structural consequences of phosphate alkylation. Because an alkyl phosphotriester group is asymmetric and can have either an $\underline{R}_p$ or $\underline{S}_p$ configuration (Figure 1.1), the properties of a phosphotriester adduct may be a function of the site, the moiety and the stereochemistry of alkylation. One diastereomer results in an alkyl group which points into the major groove of the duplex and the other diastereomer results in an alkyl group which is oriented away from the helix into the bulk solvent (Figure 1.2).

A number of studies have been performed to determine if phosphate alkylation alters the ability of DNA to serve as a functional template. To avoid confusion in interpretation which might result from the presence of other lesions in native DNA, synthetic oligomers have been employed to identify the effects that the site, the moiety, and the stereochemistry of phosphate alkylation have on the structural and functional properties of DNA. Miller *et al.*[10] studied the catalytic efficiency of DNA polymerase I on the decamer d[CCAAGATTGG] and its two diastereomeric ethyl phosphotriester analogues, $\underline{S}_p$-d[CCAAG(ethyl)ATTGG] (isomer I) and $\underline{R}_p$-d[CCAAG(ethyl)ATTGG] (isomer II). Terminal deoxynucleotidyl transferase was used to add oligodeoxyadenylate tails to the 3' end of each decamer and the resulting oligomers were tested as templates for enzymatic extension by *E. coli* DNA polymerase I with d[TTTTTTTTCC] as a primer. It was found that the rate and extents of polymerization directed by the modified templates were 25% (isomer I) and 50%

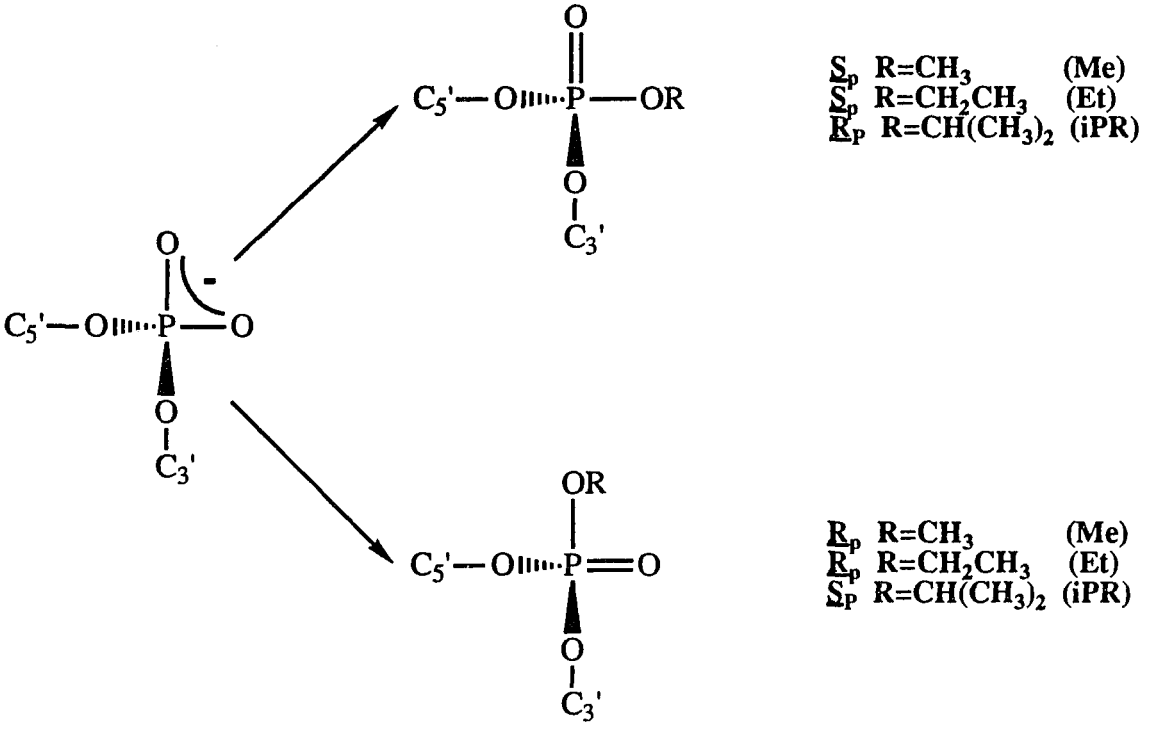
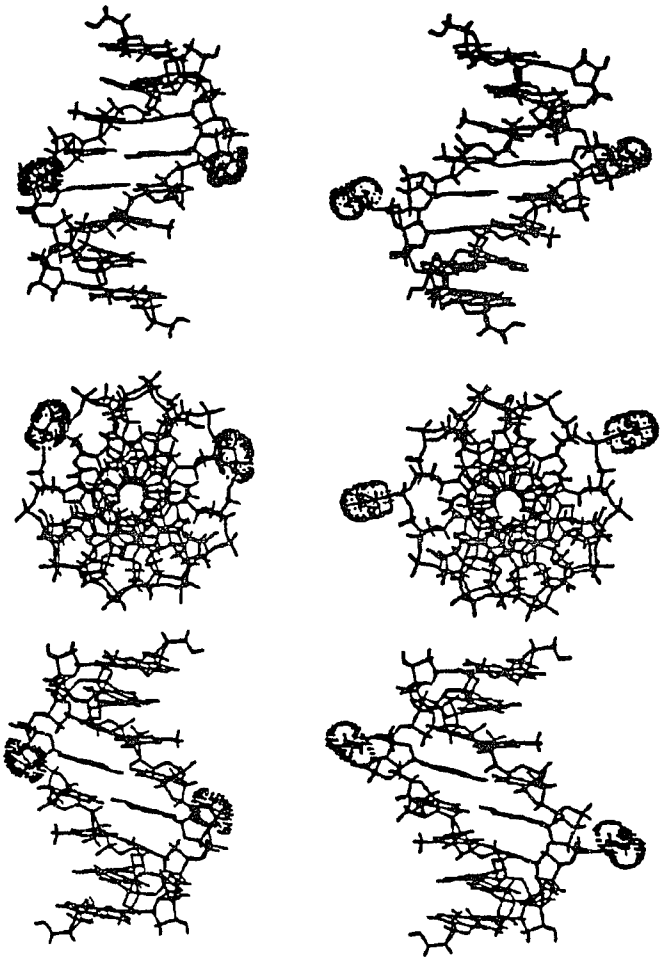


Figure 1.1 Diastereomeric alkyl phosphotriesters. Phosphate alkylation introduces chirality into the phosphorus atom resulting in a phosphotriester that has an absolute configuration of $\underline{R}_p$ or $\underline{S}_p$.



R_P	$R=CH_3$	(Me)	S_P	$R=CH_3$	(Me)
R_P	$R=CH_2CH_3$	(Et)	S_P	$R=CH_2CH_3$	(Et)
S_P	$R=CH(CH_3)_2$	(iPr)	R_P	$R=CH(CH_3)_2$	(iPr)
	I			II	

Figure 1.2 B-DNA helical form of $\{d[GGAATTCC]\}_2$ containing a single ethyl phosphotriester on each strand of the duplex (the ethyl group of the phosphotriester is shaded). One stereochemistry (I) allows the alkyl group to point directly into the major groove whereas in the opposite stereochemistry (II), the alkyl group points away from the major groove of the helix.

(isomer II) less than those of the unmodified template, suggesting that both the presence and the orientation of the ethyl group modifies the interaction of the DNA replication protein with its template.

Gallo *et al.*[11] synthesized the $\underline{R}_p$ and $\underline{S}_p$ diastereomers of an ethyl phosphotriester modified *Eco*RI recognition sequence, d[GGAA(ethyl)TTCC]. The diastereomers were separated by reversed-phase high pressure liquid chromatography (HPLC) and the absolute configuration of the phosphotriester for each diastereomer was determined by a chemo-enzymatic method. The two diastereomers were totally resistant toward cleavage by *Eco*RI endonuclease under conditions that led to cleavage of the unmodified duplex suggesting that a single phosphotriester lesion can interfere with enzymatic activity on the DNA template.

Majumdar and Adhya[12] examined the effect of phosphate alkylation on the binding of Gal repressor to the *gal* operon. The extent of repressor binding was assayed by measuring the mobility of single-site phosphate ethylated operon fragments on polyacrylamide gels in the presence of Gal repressor. Twenty-three different phosphate sites were found which interfere with repressor binding when individually ethylated. However, it was not determined if a configurational dependence was present on the decreased binding of repressor to the ethylated operon fragments.

Weinfeld and Livingston[13] prepared the single-site phosphate ethylated decamer d[TTT(ethyl)TCTATTT] and the resulting diastereomeric mixture was tested for the ability to prime DNA synthesis catalyzed by *E. coli* polymerase I when annealed to a single-stranded plasmid template. DNA polymerase I catalyzed the transfer of additional nucleotides to the 3'-terminus of the unmodified and the modified decamers with equal effectiveness. This result indicated that a phosphotriester lesion, at a position seven nucleotides removed from the site of template activity, does not interfere with enzymatic function on DNA.

Collectively, the *in vitro* studies discussed above suggest that formation of a single phosphotriester lesion can modify the ability of a DNA fragment to serve as a functional template if the phosphotriester is localized to the site of template activity and has a

stereochemistry that enables the phosphotriester alkyl group to point into the major groove of the duplex. It cannot be determined from these studies if the modification in template activity is due to one or more of the following factors: a structural change in the template, unfavorable steric interactions between the protein and the DNA template due to the presence of an alkyl group on the phosphate backbone, or the loss of important DNA-protein electrostatic interactions as a result of charge neutralization at the site of phosphate alkylation. However, it should be noted that if the modification in DNA template activity were solely due to charge neutralization, then it would be expected that the reduction in template activity would not depend on the stereochemistry of phosphotriester formation in contrast to the results of Miller *et al.*[10] The results of biophysical studies on single-site phosphate alkylated oligodeoxynucleotides has helped determine what factors are responsible for the modification in template activity of phosphate alkylated DNA (*vide infra*).

1.3 DNA Conformation: Nomenclature and Symbols

The basic repeating unit of DNA is the nucleotide. A nucleotide consists of three components: a furanoside-type sugar (β-D-2'-deoxyribose), a phosphate group which is attached to the 5' position of the sugar and a heterocyclic aromatic base attached to the C1' position of the sugar via a β-glycosyl C1'-N linkage (Figure 1.3).[14] The heterocycles for DNA are the purine bases, adenine (A) and guanine (G) and the pyrimidine bases, cytosine (C) and thymine (T). The atom numbering scheme is indicated in Figure 1.3.

The conformation of a nucleic acid is generally described in terms of the torsion angle variables $\alpha,\beta,\gamma,\delta,\epsilon,\zeta,\nu_0,\nu_1,\nu_2,\nu_3,\nu_4$ and χ (Figure 1.4).[14] The endocyclic torsion angles of the sugar are denoted ν_0,ν_1,ν_2,ν_3 and ν_4. Because of potential energy barriers, the endocyclic torsion angles adopt only certain values that restrict the number of sugar ring conformers found in DNA structures. The preferred conformations or puckering modes can be conveniently represented by a pseudorotation cycle (Figure 1.5).[15] In nucleotide structure analysis, two ranges of pseudorotation phase angles are preferred; the C3'-*endo* (N

conformation) at $0° \leq P \leq 36°$ and C2'-*endo* (S conformation) at $144° \leq P \leq 180°$.[14]

The conformation of the sugar-phosphate backbone is defined by the torsion angles $\alpha,\beta,\gamma,\delta,\epsilon$ and ζ. Although a torsion angle can assume any value between 0° and 360°, it is often convenient to describe a structure with torsion angle ranges as shown in (Figure 1.6).[14] Unfavorable steric contacts restrict the conformationally allowed regions for each torsion angle. For instance, conformational analysis of polynucleotides indicate that each of the torsion angles α,β,γ and ζ assumes three preferred values (+synclinical (+sc), -synclinical (-sc) and antiperiplanar (ap)), δ assumes two preferred values (+sc and ap) and ϵ assumes two preferred values (ap and -synperiplanar (-sp)).[14]

The base orientation about the glycosyl C1'-N bond is given by the torsion angle χ. The torsion angle χ can reside in two distinct conformational ranges, the *anti* orientation and the *syn* orientation illustrated in (Figure 1.7).[14]

The inherent flexibility of the torsion angles within DNA leads to considerable variation of the conformations adopted by DNA structures. Even though DNA exhibits considerable polymorphism, it has been shown that most DNA structures exist in either the A, B, or Z conformational forms.[14] The structural form taken by DNA has been shown to depend on its sequence and on its environment (i.e., relative humidity, ionic strength, and nature of solvent).[14] The mean double-helix parameters for the three conformational families of DNA are summarized in Table 1.1.[15]

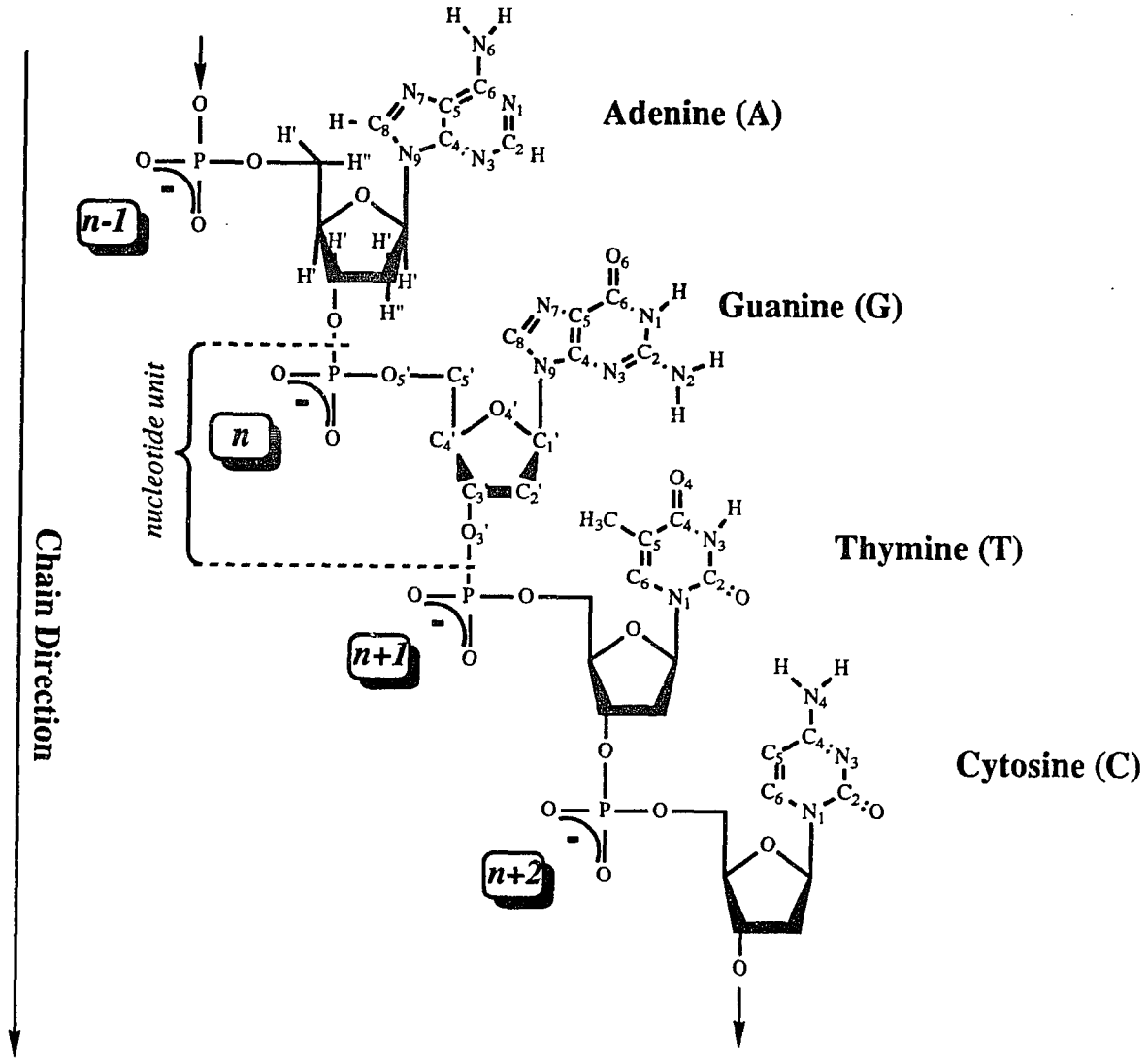
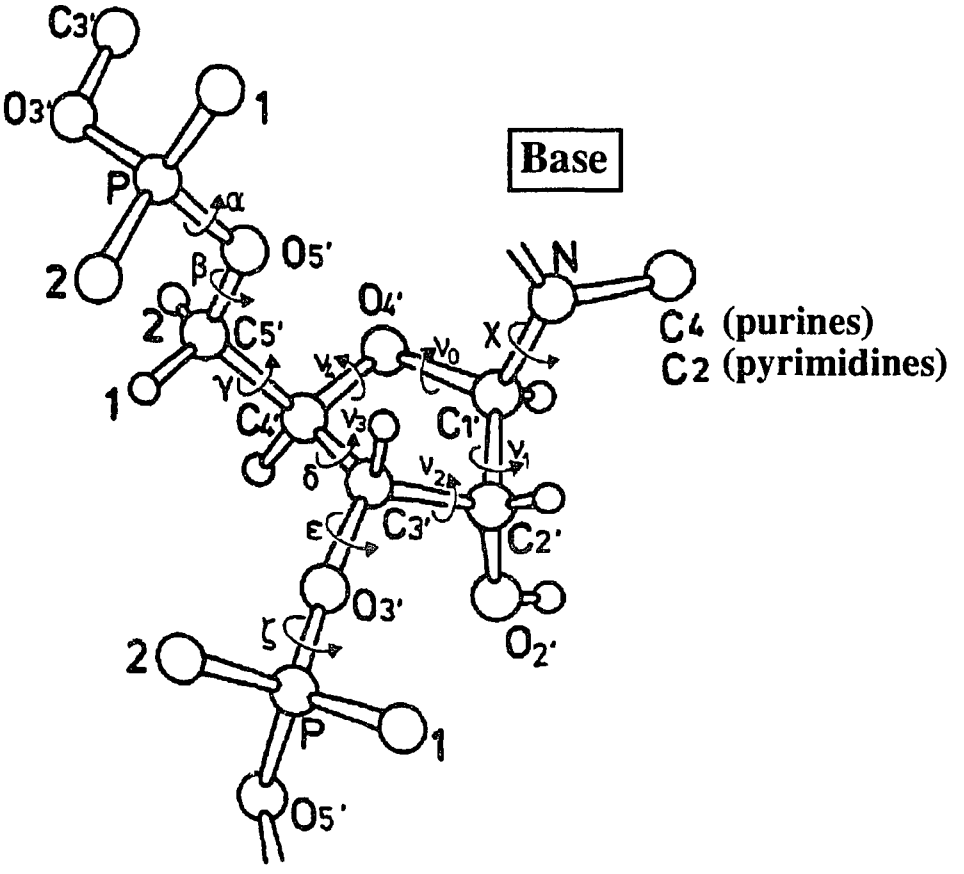


Figure 1.3 Fragment of deoxyribonucleic acid (DNA) with sequence adenine (A), guanine (G), thymine (T), cytosine (C), linked by 3',5' phosphodiester bonds. Chain direction is from 5'- to 3'- end as shown by arrow. Atomic numbering scheme for the deoxyribose ring is indicated in guanine nucleotide.



α	$_{(n-1)}O_{3'}$-P-$O_{5'}$-$C_{5'}$	ν_0	$C_{4'}$-$O_{4'}$-$C_{1'}$-$C_{2'}$
β	P-$O_{5'}$-$C_{5'}$-$C_{4'}$	ν_1	$O_{4'}$-$C_{1'}$-$C_{2'}$-$C_{3'}$
γ	$O_{5'}$-$C_{5'}$-$C_{4'}$-$C_{3'}$	ν_2	$C_{1'}$-$C_{2'}$-$C_{3'}$-$C_{4'}$
δ	$C_{5'}$-$C_{4'}$-$C_{3'}$-$O_{3'}$	ν_3	$C_{2'}$-$C_{3'}$-$C_{4'}$-$O_{4'}$
ε	$C_{4'}$-$C_{3'}$-$O_{3'}$-P	ν_4	$C_{3'}$-$C_{4'}$-$O_{4'}$-$C_{1'}$
ξ	$C_{3'}$-$O_{3'}$-P-$O_{5'(n+1)}$		

χ $O_{4'}$-$C_{1'}$-N_1-C_2 (pyrimidines)
$O_{4'}$-$C_{1'}$-N_1-C_4 (purines)

Figure 1.4 Definition of torsion angles for deoxyribonucleotide chain. Figure taken from Reference 14.

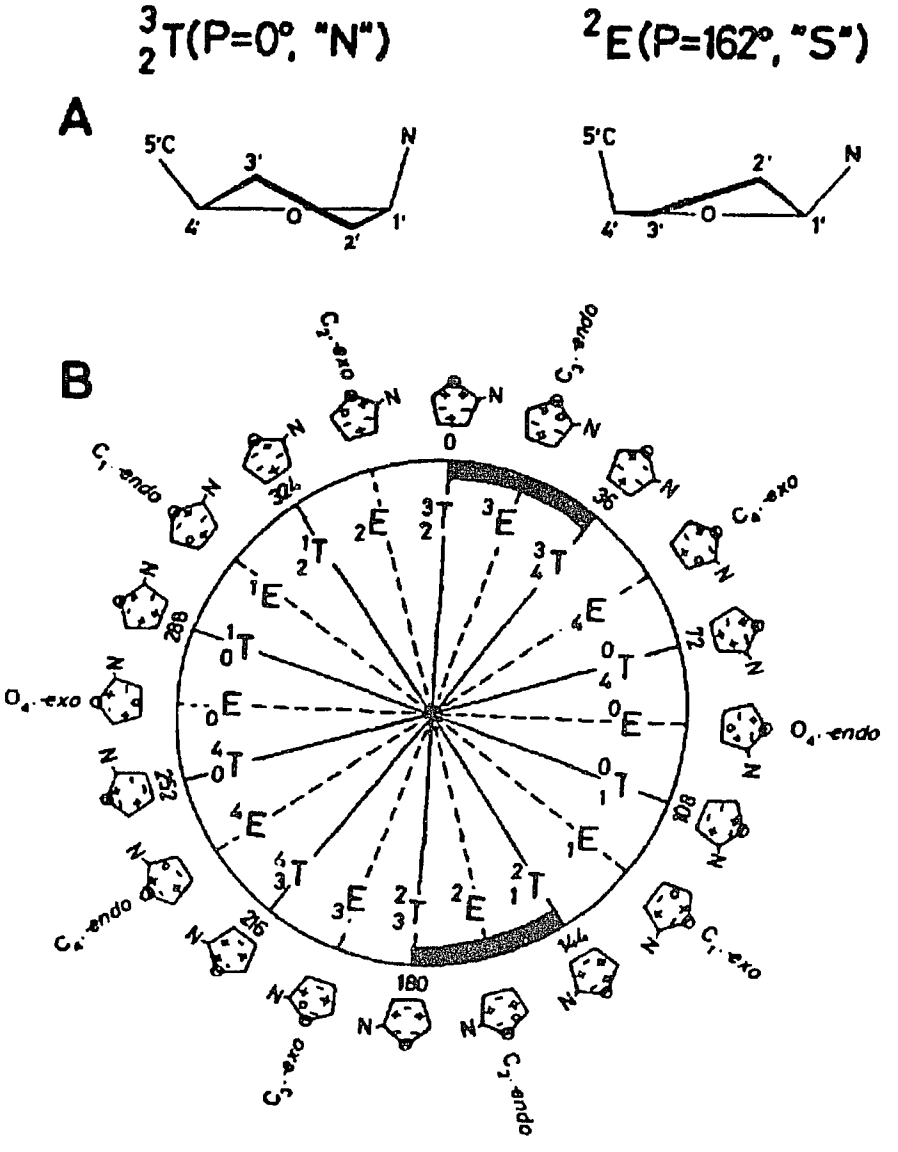


Figure 1.5 A. Illustration of two furanose ring puckers representing the preferred north (N) and south (S) conformers. B. Pseudorotation cycle of the furanose ring in nucleosides. Values of phase angles are given in multiples of 36°. Envelope (E) and twist (T) forms alternate every 180°. Thick lines inside circle indicate the preferred pseudorotational regions. Figure taken from Reference 15.

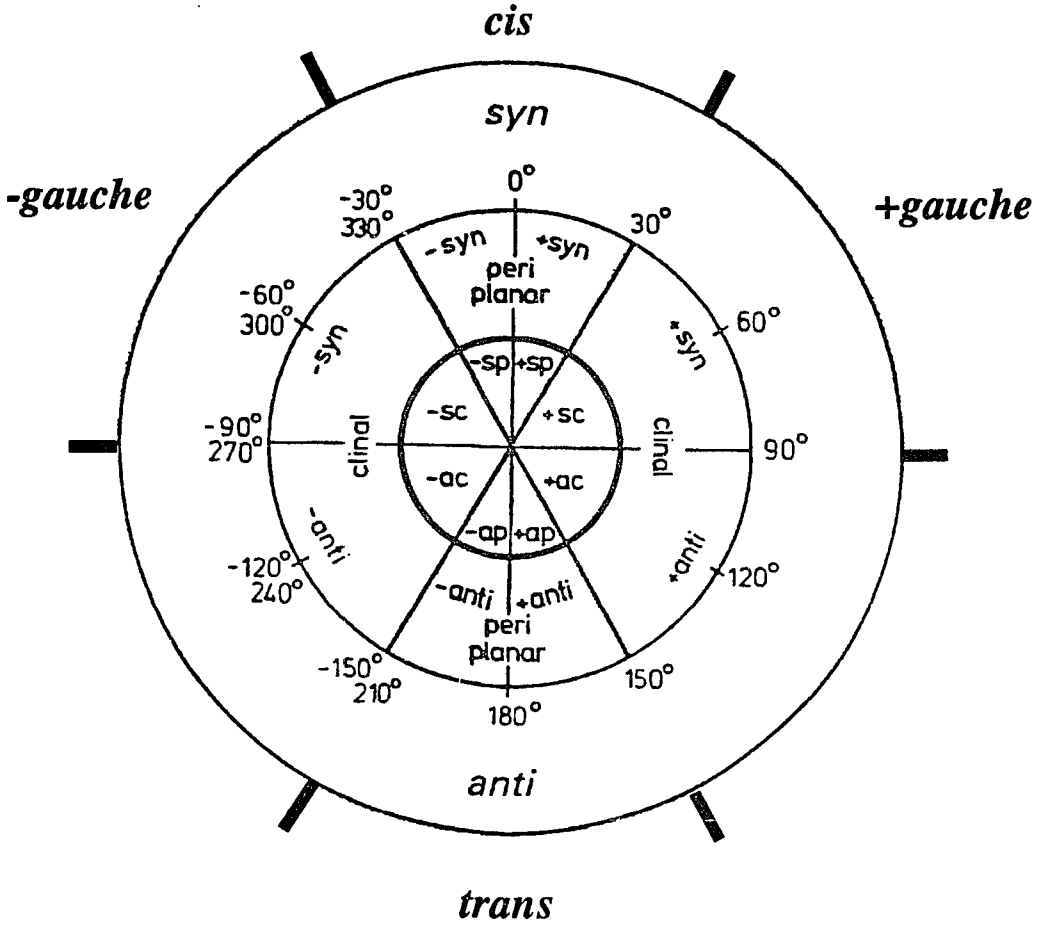


Figure 1.6 Figure of torsion angle ranges defined by Kyne and Prelog (syn or synperiplanar, anti or antiperiplanar, +synclinical, -synclinical, +anticlinical, -anticlinical). Torsion angle nomenclature (cis, trans, +gauche, -gauche) that is also in use is indicated in the figure. Figure taken from Reference 14.

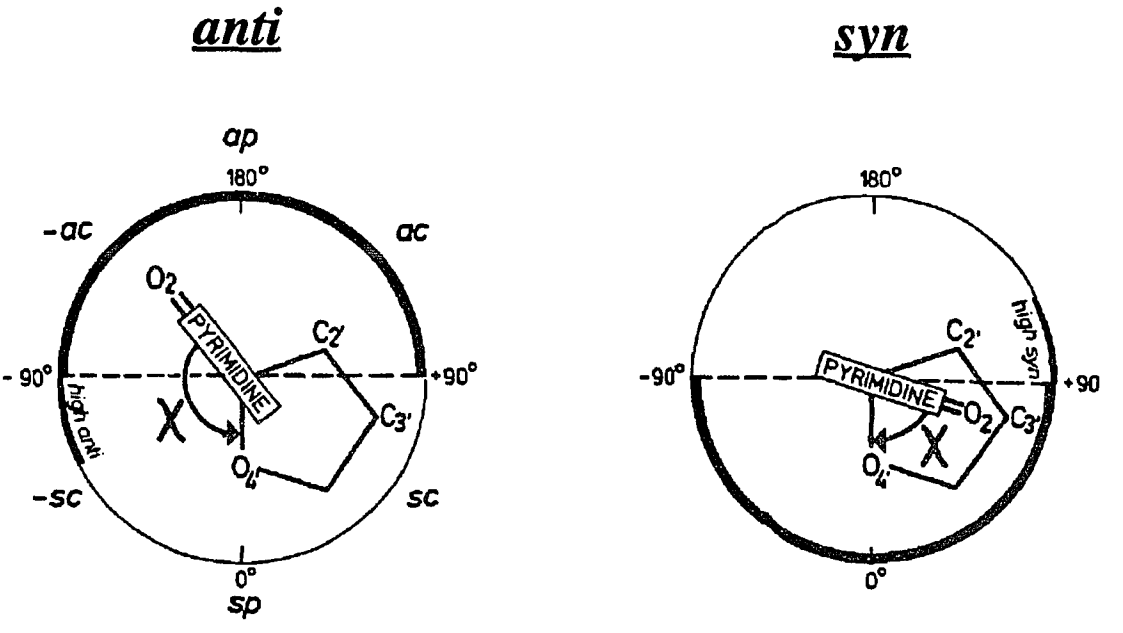


Figure 1.7 Definition of anti and syn conformational ranges shown for a pyrimidine nucleoside. χ is defined as the torsion angle O4'-C1'-N1-C2. Figure taken from Reference 14.

Table 1.1 Mean double-helix parameters for A-DNA, B-DNA, and Z-DNA. Table taken from Reference 15.

Helix parameter	A-DNA	B-DNA	Z-DNA
Handedness	right	right	left
Bases per turn	10.9	10.0	12.0
Height per base	2.9 Å	3.4 Å	GC 3.5 Å CG 4.1 Å
Pitch	31.6 Å	34.0 Å	45.6 Å
Glycosyl angle	anti	anti	C anti G syn
Sugar pucker	3'C endo	2'C endo	C 2'C endo G 3'C endo
Base pairing	Watson-Crick	Watson-Crick	Watson-Crick
Repeating helix unit	1 base pair	1 base pair	2 base pairs
Twist per base pair	33°	36°	GC -51° CG - 9°
Axis displacement[b]	4 Å	0	-3 Å
Major groove	very deep	deep, wide	shallow
Minor groove	shallow	deep, narrow	very deep
Base inclination	13°	-2°	9°
Base roll	6°	-1°	3°
Propeller twist	15°	12°	4°

1.4 Biophysical Studies of Phosphate Alkylated Oligodeoxynucleotides

Several biophysical studies have focussed on determining the ability of phosphotriester analogues to form stable duplexes and on determining the structural distortions and the stereochemical dependence of such distortions that result from phosphate alkylation. Brennan *et al.*[16] obtained an X-ray structure of the phosphotriesterified nucleotide, adenosine 5'-O-(diethyl phosphate) and found structural features that were different from those of nonalkylated nucleotides. Because the 5'-phosphate is triesterified, adenosine 5'-O-(diethyl phosphate) possesses three sets of phosphodiester linkages. The three sets of phosphodiester linkages (α,ζ) were found to display conformations ((-sc, ap), (+sc, -sc), (ap, ap)) which are different from the preferred conformation of right-handed polynucleotide helices (α=-sc, ζ=-sc). The ribofuranose was found in the C2'-*endo*-C1'-*exo* conformation (pseudorotation=145°) which represents a slight sugar pucker distortion from the standard C3'-*exo*-C2'-*endo* (pseudorotation=180°) conformation which is commonly found in right-handed B-form DNA helices. It was noted that the sugar pucker distortion increases the purine ring to ethyl group distance in order to avoid unfavorable steric interactions between the two groups. The authors suggested, based on the structural parameters observed for the phosphotriesterified nucleotide, that single-site phosphate alkylation would tend to decrease the degree of base-base stacking in the vicinity of the phosphoesterification site.

Several studies have been reported on phosphate-alkylated dinucleotides. Miller *et al.*[17] synthesized the methyl and ethyl phosphotriester derivatives of [TT] and d[AA] which, in each case, resulted in a mixture of diastereomers that was used without separation in a series of ultraviolet (UV), circular dichroism (CD) and ^{1}H nuclear magnetic resonance (NMR) measurements. The base proton chemical shift positions of each phosphotriester analogue relative to the corresponding nonalkylated dinucleotide indicated that the phosphate alkyl group decreased the degree of base-base stacking for each dinucleotide studied. This result was supported by a decrease in the hypochromicity in the UV spectra and a decrease in the rotational strength of the CD spectra of the phosphotriesters. However, the positions of the

absorption minima and maxima of the UV and CD spectra of the phosphotriester analogues were identical to their corresponding nonalkylated dinucleotides indicating that the phosphotriesters adopt right-handed stacked conformations in aqueous solution such that the nucleosidyl residues of these dimers are mainly in an *anti* conformation.

Weinfeld *et al.*[18] synthesized the eight diastereomers derived from phosphate ethylation of the dinucleoside monophosphates, d[AT], d[GT], d[CT] and d[TT], and separated each pair of diastereomers by HPLC. CD measurements indicated that base-base stacking was present in all of the phosphate-ethylated dinucleoside monophosphates and all dimers displayed temperature dependent CD spectra consistent with an equilibrium between a right-handed helical stacked state and an unstacked state. Although they were unable to assign the absolute stereochemistry of each phosphotriester, the authors inferred the absolute configuration of the phosphorus atom by its stacking ability as determined by the temperature dependence of the CD spectrum and by its HPLC retention time. The authors argued that the diastereomer within each set that has the ethyl group in proximity to the bases should have a decreased base stacking and this decrease in base stacking should increase the chromatographic retention time and decrease the temperature dependence of the CD absorbance. The relative HPLC retention times for the diastereomers suggested a convenient method for confirming the stereochemistry of single-site phosphate alkylated oligomers (*vide infra*).

Kan *et al.*[19] used ^{1}H NMR, UV and CD spectroscopy to study the mode of base-base stacking, the handedness and the backbone conformation of d[GA] and its diastereomeric ethyl phosphotriester analogues. The absolute configurations of the two diastereomers were determined by ^{1}H-{^{1}H) nuclear Overhauser effect (NOE) experiments. An NOE was observed between the G1-H3' proton and the methylene protons of the ethyl phosphotriester for one isomer indicating a spatial proximity of less than 5 Å between the two protons and verifying that this isomer possessed an ethyl group oriented toward the bases of the dinucleotide. The relative chromatographic retention times of each diastereomer were consistent with the results

of Weinfeld *et al.*[18] The spectroscopic measurements on d[GA] and its two phosphotriester analogues indicated that the base-base overlap for the isomer with ethyl group oriented toward the bases was reduced relative to both the nonalkylated dinucleotide and the isomer with ethyl group oriented away from the bases.

The experimental results obtained to date on mononucleotides and dinucleotides suggest that conformational changes (i.e. changes in base-base stacking and changes in the sugar-phosphate backbone) occur upon triesterification and that such changes appear to be dependent on the stereochemistry of the alkylation. Several studies have been reported on the effect of single-site phosphate alkylation on polynucleotides. Weinfeld and Livingston's[13] study of the diasteromeric mixture of d[TTT(ethyl)TCTATTT] indicated that the slope of the CD vs. temperature curve for the phosphate ethylated decamer was nearly identical to that of the unmodified decamer. This result implied that there was little or no disruption to base stacking in the decamer upon the introduction of a single phosphotriester lesion. However, that this result does not appear to support the dinucleotide studies that revealed decreased base-base stacking upon phosphate alkylation may be because CD is not a sensitive enough technique to distinguish distortion about one internal phosphate out of nine if there is destacking only at the site of ethylation.

Praminik and Kan[20] examined the temperature variation of the imino proton NMR resonances and ^{31}P NMR resonances of the self-associated duplex $\{d[CCAAGATTGG]\}_2$ and the isolated single-site phosphate ethylated analogues, $(\underline{R}_p,\underline{R}_p)$-$\{d[CCAAG(ethyl)ATTGG]\}_2$ and $(\underline{S}_p,\underline{S}_p)$-$\{d[CCAAG(ethyl)ATTGG]\}_2$. The measurements indicated that the presence of a single phosphotriester in the decamer sequence did not disrupt the normal Watson-Crick base pairs of the DNA duplex. However, the temperature variation of both the imino proton linewidths and chemical shifts indicated that phosphate ethylation destabilized the duplex and that destabilization was greater for the diastereomer with an ethyl group oriented into the major groove.

Summers *et al.*[21] examined the stability and the structural characteristics of the

phosphate ethylated oligonucleotides, ($\underline{R}_p$,$\underline{R}_p$)-{d[GGAA(ethyl)TTCC]}$_2$ and ($\underline{S}_p$,$\underline{S}_p$)-{d[GGAA(ethyl)TTCC]}$_2$ (referred to as R- and S-A(et)T), that were synthesized and separated by Gallo *et al.*[11] The ^{1}H NMR chemical shift assignments for these compounds were determined for all signals (except for a few H5' and H5" signals) using proton-proton two dimensional nuclear Overhauser effect spectroscopy (^{1}H-{^{1}H} 2D NOESY). The absolute configuration of the phosphotriester in each oligonucleotide was also determined by 2D NOESY measurements. An NOE was observed between the A4-H3' proton and both the methylene and methyl protons of the ethyl phosphotriester for one isomer indicating a spatial proximity of less than 5 Å between the protons and verifying that the isomer contained an ethyl group oriented toward the major groove of the octamer. This result confirmed the stereochemical assignments based on both chemo-enzymatic and HPLC methods. The ^{1}H chemical shifts were, in general, similar to those reported for the unmodified, parent duplex which had been shown to adopt a B-like conformation in solution by Broido *et al.*[22] Differences in chemical shift were primarily localized to the A(et)T region of the duplex, presumably reflecting minor structural distortions within the phosphotriester modified analogues relative to the unmodified octamer. Furthermore, the ^{1}H chemical shift differences were not the same for each diastereomer suggesting that the structural distortions could be stereochemically dependent. Variable temperature ^{31}P NMR, ^{1}H NMR and UV measurements indicated that the R-A(et)T duplex (ethyl group oriented toward the major groove) melts ≈10°C lower than either the S-A(et)T (ethyl group oriented away from the major groove into the bulk solvent) or parent duplexes. Furthermore, melting curves were obtained for the oligomers over a range of salt concentrations and the T_m values for the $\underline{S}_p$ and $\underline{R}_p$ diastereomers were significantly different (≈10°C) at all salt concentrations supporting the hypothesis that duplex destabilization by phosphate alkylation is primarily a consequence of steric effects rather than electrostatic effects.

Lawrence *et al.*[23] and Broido *et al.*[24] examined the alkylated analogues ($\underline{S}_p$,$\underline{S}_p$)- and ($\underline{R}_p$,$\underline{R}_p$)-{d[GGA(isopropyl)ATTCC]}$_2$ (referred to as S- and R-A(iPr)A), ($\underline{S}_p$,$\underline{S}_p$)- and

($\underline{R}_p$,$\underline{R}_p$)-{d[GGAA(isopropyl)TTCC]}$_2$ (referred to as S- and R-A(iPr)T) and ($\underline{S}_p$,$\underline{S}_p$) and ($\underline{R}_p$,$\underline{R}_p$)-{d[GG(ethyl)AATTCC]}$_2$ (referred to as S- and R-G(et)A) of the duplex {d[GGAATTCC]}$_2$. The stereochemistry of the phosphotriester for each analogue was determined by ^{1}H-{^{1}H} 2D NOESY experiments. The assignment based on the 2D NOESY data was consistent with the HPLC behavior and T_m measurements for each diastereomeric pair. In general, the phosphotriester analogues containing an inward pointing alkyl group displayed an NOE between the alkyl group and the nonexchangeable base proton on the adjacent (*n-1*) residue. This type of NOE was absent for the analogues containing an alkyl group oriented away from the major groove. The duplex stability for the isopropylated analogues was monitored with optical measurements, ^{31}P NMR and ^{1}H NMR. The apparent order of duplex stability as inferred by the melting temperature measurements was: S-A(iPr)T (20°C) < S-A(iPr)A (22°C) < R-A(iPr)A (24°C) < R-A(iPr)T (28°C) ~ Parent (28°C) and is consistent with the results of Summers *et al.*[21] on A(et)T, indicating that the duplex stability for those analogues in which the phosphotriester alkyl group is oriented into the major groove is measurably less than that of the parent or of the analogues containing an alkyl group oriented away from the major groove. The ^{1}H chemical shift differences between each analogue and the unmodified octamer indicated that the chemical shift positions of several protons were dependent upon both the site of esterification and the stereochemistry of the phosphotriester. This observation is consistent with the results of Summers *et al.*[21] which indicated that the structural distortions within a single-site phosphate alkylated oligomer were localized to the site of the phosphotriester lesion and were stereochemically dependent.

Broido and Mezei[25] performed a computational (AMBER[26,27,28]) analysis of the effects of site-specific phosphate alkylation in the DNA oligomer {d[GGAATTCC]}$_2$. Molecular mechanics calculations were performed on fourteen analogues of the octamer, each bearing a single ethyl phosphotriester of known stereochemistry and each initially in an "idealized" B-conformation. Results indicated that each oligonucleotide remained in a B-DNA conformation, and that although there was only minimal effect on the aromatic bases, the

presence of a phosphotriester disturbed the sugar-phosphate backbone in complex ways as determined by both geometric features (backbone torsion angles) and energetic features (interaction energies between groups of atoms). Although no strong correlation was found between trends in various energetic contributions and the alkylation site stereochemistry, the total minimum energies were lower for the $\underline{S}_p$-diastereomers than for the $\underline{R}_p$-diastereomers by ~1-4 kcal/mole. This difference in total minimum energy is consistent with experimental observations of the relative stability of phosphate alkylated oligomers and probably reflects the disruptive nature of a substitution which points into the major groove of a double helix.

The results of the experimental and computational studies on single-site phosphate alkylated oligonucleotides indicate that although esterification of the sugar-phosphate backbone of DNA does not inhibit the formation of a Watson-Crick double helical structure, it does alter physical properties of the duplex (e.g., T_m and base-base stacking) and the changes appear to be related to both the site and the stereochemistry of alkylation. Some of these studies have also correlated the HPLC and stability behavior of phosphate alkylated oligomers with the stereochemistry of the phosphotriester group. However, there is no conclusive evidence to date concerning the specific structural changes that take place upon phosphate alkylation and there still does not exist a strong correlation between the change in a specific molecular event (e.g., decreased DNA polymerase I activity upon single-site phosphate alkylation of a DNA template) and a specific perturbation (e.g., change in template structure) with respect to phosphate alkylation.

Since knowledge of the effects of phosphate alkylation on the physical properties of DNA is still limited, we have carried out a series of ^1H and ^{31}P NMR studies on the self-complementary octamer, {d[GGAATTCC]}$_2$ and two ethylated analogues, ($\underline{S}_p$,$\underline{S}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ and ($\underline{R}_p$,$\underline{R}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ (Figure 1.8). The central six residues of the octamer sequence, GAATTC, comprise the *Eco*RI restriction site.[29] The *Eco*RI restriction endonuclease hydrolyzes the phosphodiester bond between the guanylic and adenylic acid residues resulting in a 5'-phosphate. In addition to the specific interaction of

the *Eco*RI endonuclease with the canonical sequence, the enzyme also binds DNA in a nonspecific manner that does not result in hydrolysis of the DNA suggesting that the canonical sequence contains sufficient information for specific recognition by the restriction endonuclease protein.[30]

The structure of the *Eco*RI restriction site has been the subject of numerous investigations. Dickerson and coworkers solved the structure of a double-helical dodecamer $\{d[CGCGAATTCGCG]\}_2$ by single-crystal X-ray analysis and found that the innermost part of this sequence comprising the *Eco*RI recognition displays B-type DNA structural characteristics.[31,32,33,34] Detailed examination of the crystal structure revealed several structural features that appeared to be related to base sequence. The solution structure, conformational flexibility and stability of sequences containing the *Eco*RI restriction site have been extensively characterized by one and two dimensional NMR methods.[22,35,36,37,38,39,40,41,42] In general, the solution data indicates that sequences containing the restriction site form duplexes with 2-fold symmetry in a B-DNA conformation and that certain NMR parameters (e.g., ^1H and ^{31}P chemical shifts) demonstrate the existence of sequence-specific variations within the restriction site.

Because it has been shown that the activity of the *Eco*RI restriction endonuclease on the octamer template is affected by the presence of a phosphotriester lesion,[11] structural studies on phosphate alkylated analogues of the octamer should allow one to correlate the changes in template activity with perturbations that result from structural, steric and/or electrostatic effects. The collection of studies reported on single-site phosphate alkylated analogues of the octamer sequence have demonstrated that the presence of a phosphotriester lesion affects the octamer structure and stability as determined by changes in various NMR parameters (e.g., ^1H chemical shift differences between the analogues and the parent) and the observed changes depend on the site and stereochemistry of the phosphotriester group. Therefore, using the results of these studies on the octamer sequence as a foundation, we have used 1D and 2D ^1H NMR methods and 1D ^{31}P NMR methods to obtain a more detailed

understanding of the effects of single-site phosphate ethylation on the structure, conformational dynamics and stability of the octamer, $\{d[GGAATTCC]\}_2$. In particular, imino proton linewidths were measured as a function of temperature to determine the stability of each duplex and base-pair lifetime measurements and ^{31}P relaxation measurements were made to determine the dynamical properties of each duplex. Furthermore, ^1H and ^{31}P chemical shift measurements and interproton distances derived from ^1H-$\{^1H\}$ 2D NOESY experiments were used to elucidate the structural consequences of phosphate alkylation. In complementary studies, we have used a series of computer simulations to determine how accurately DNA structures can be generated using data obtained from the ^1H-$\{^1H\}$ 2D NOESY experiments presented here.

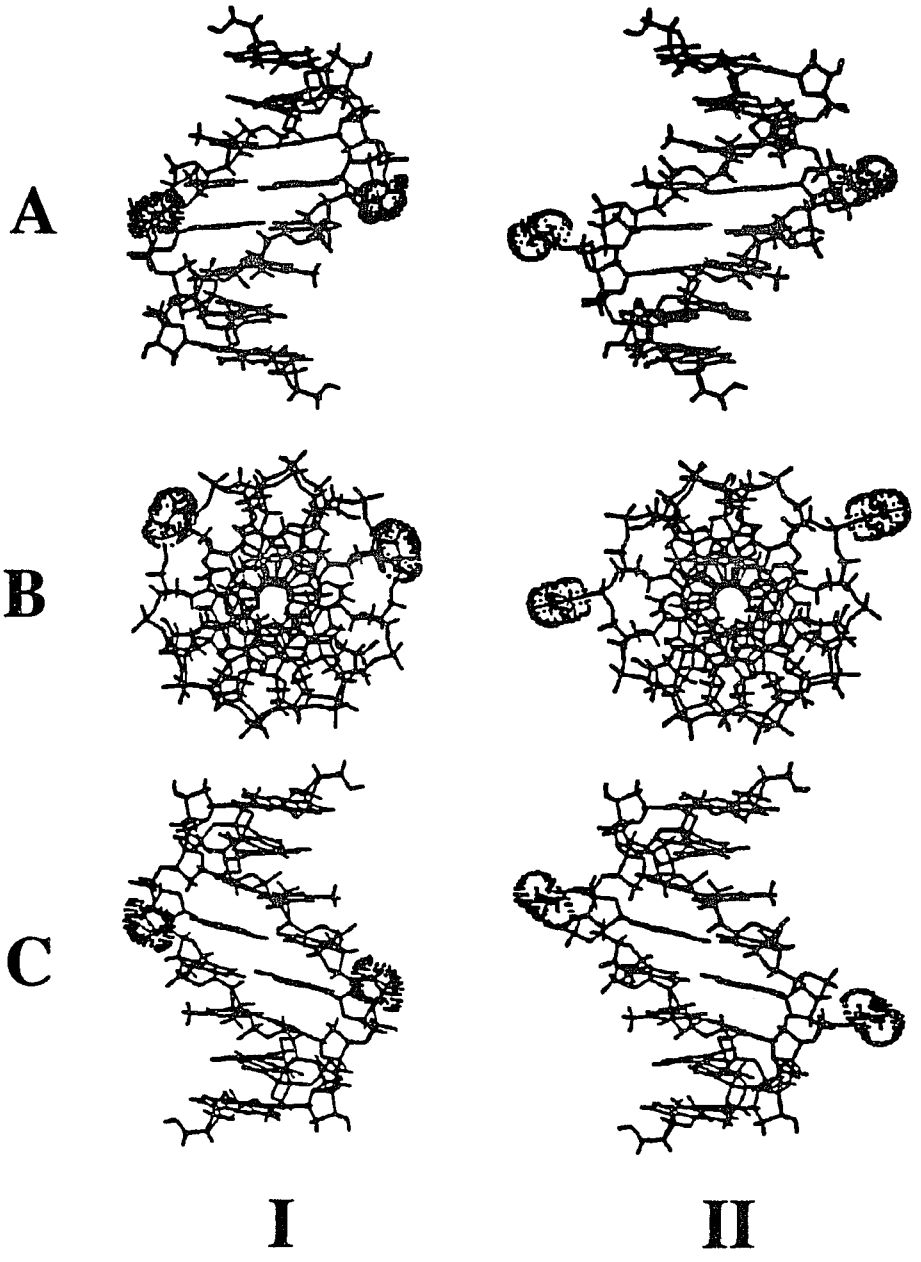


Figure 1.8 B-DNA helical forms of the two single-site ethylated analogues of {d-[GGAATTCC]}$_2$; (I) ($\underline{R}_p$,$\underline{R}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ and (II) ($\underline{S}_p$,$\underline{S}_p$)-{d[GGA(ethyl)ATTCC]}$_2$. (**A**) is view down minor groove, (**B**) is view down helical axis and (**C**) is view down major groove. The ethyl moiety of the phosphotriester group is shaded.

CHAPTER 2 Fundamental Principles

2.1 Nuclear Magnetic Resonance: Basic Principles

An atomic nucleus is composed of elementary particles (protons, neutrons, etc.). Each elementary particle possesses its own spin angular momentum and depending on the number of each type of elementary particle present, an atomic nucleus may have a net non-zero spin angular momentum represented by the quantity, $I\hbar$. I is the nuclear quantum number and is found in integer or half-integer units depending on the atomic nucleus. A nucleus that has a net spin angular momentum will have a magnetic dipole moment

$$\mu = \gamma \hbar I \qquad [2.1]$$

that is collinear with the nuclear spin angular momentum. The gyromagnetic ratio, γ, is a constant whose magnitude depends on the internal structure of the nucleus and is thus a characteristic of each nuclear species.

When placed in an external magnetic field, $\mathbf{B}_o$, the dipole moment will obtain an energy (the Zeeman energy) that depends on its orientation in the magnetic field,

$$E = -\mu \cdot B_o. \qquad [2.2]$$

Quantitization of the spin angular momentum dictates that only a discrete number of energy levels are available to a dipole moment in a magnetic field and a collection of dipoles will occupy the available energy levels according to a Boltzmann distribution

$$N(E_i) = e^{(-E_i/kT)} \qquad [2.3]$$

where E_i represents the energy of some state i, $N(E_i)$ is the fraction of spins occupying the energy state i, T represents the spin temperature and k is Boltzmann's constant. The energy levels available to a nuclear spin can be perturbed by its interaction with local magnetic fields. The interaction energy of a nuclear spin system with local fields may be expressed in terms

of the Hamiltonian operator

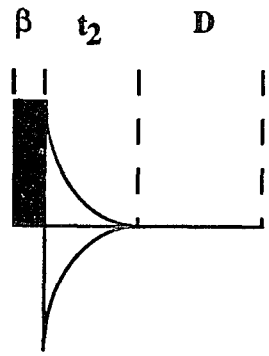
$$H = H_D + H_J + H_Q + H_{CR} \qquad [2.4]$$

where H_D represents the direct through-space interaction energy between two magnetic dipoles, H_J represents the through-bond (scalar) interaction energy between two magnetic dipoles, H_Q represents the interaction energy between the quadrupole moment of the nucleus and the electric field gradient at the nucleus, and H_{CR} represents the interaction energy between the nuclear magnetic moment and a magnetic field generated by the molecular magnetic moment associated with the charge distribution within the molecule.

Nuclear magnetic resonance spectroscopy involves inducing transitions between the different energy states available to a nuclear spin system by applying a radiofrequency field (RF) and detecting the resultant absorption of electromagnetic energy from this field. The groups of Bloch[43] and Purcell[44] were the first to report NMR absorption in bulk matter. After the absorption of energy from the RF field, the nuclear spin system "relaxes" (gives up energy to its surroundings, the spin-lattice) in order to reestablish an equilibrium distribution of spin populations within the available energy states in accordance with the Boltzmann distribution. Because spontaneous emission of RF energy is very inefficient (timescale$\approx 10^{-25}$ s^{-1})[45] the relaxation of a nucleus by stimulated emission depends on the coupling of the nuclear moment with local magnetic fields that fluctuate at the site of the nucleus as a result of molecular motion. The properties of a given NMR signal (i.e., chemical shift, spin-spin coupling, linewidth and spin-lattice relaxation time) are therefore a function of the magnetic and electronic environment of the nucleus as well as the motional characteristics of the nucleus in its environment. Although it is beyond the scope of this manuscript to discuss the various experimental techniques that are used within the field of nuclear magnetic resonance and the different formalisms by which nuclear magnetic resonance can be described, it should be noted that there are several excellent references that are devoted to these topics.[46,47,48,49,50]

2.2 One Dimensional NMR Spectroscopy

The one dimensional (1D) NMR experiments employed for the studies discussed in this manuscript are based on the pulsed Fourier transform (FT) method.[51] The basic single pulse NMR experiment is shown below and consists of three basic time intervals; a pulse interval (β), an acquisition interval (t_2), and a delay interval (D),



The sample is placed in a static magnetic field, $\mathbf{B}_o$, where the nuclear spins polarize resulting in a net macroscopic magnetization (longitudinal), **M**, in the direction of $\mathbf{B}_o$ (usually assumed to be oriented along the z-axis). A radiofrequency field, $\mathbf{B}_1$, in the form of a pulse is applied along an axis perpendicular to $\mathbf{B}_o$. The RF pulse induces transitions between different energy levels available to the spin system which results in the formation of transverse magnetization in the plane perpendicular to $\mathbf{B}_o$. After the RF pulse is applied, the transverse magnetization precesses at frequencies corresponding to the stimulated energy transitions inducing a voltage in a transverse coil. The time evolution, $S(t_2)$, of the coil voltage is collected over the acquisition interval, t_2. The transverse magnetization decays with time due to relaxation processes and so the time evolution of the detected voltage is often referred to as the free induction decay (FID). The delay interval allows the nuclear spin system to reestablish thermal equilibrium. The entire cycle (pulse, acquire, and delay) can be repeated and the FIDs collected are co-added in order to enhance the signal-to-noise ratio of the resulting spectrum.[52]

The frequency (ω) domain NMR spectrum is obtained by a Fourier transformation of the time domain signal

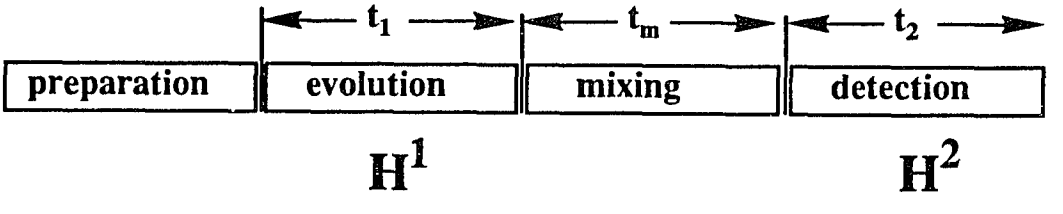
$$S(\omega) = \int_{-\infty}^{+\infty} S(t) e^{-i\omega t} dt. \qquad [2.5]$$

Quite often the time domain signal , S(t), is multiplied (apodized) by another time domain function before Fourier transformation in order to mathematically enhance the sensitivity or the resolution of the frequency domain spectrum. For example, it has been demonstrated that multiplication of the time-domain signal by a decaying exponential function will de-emphasize the noise components in the tail of the FID and lead to enhanced sensitivity of the frequency domain spectrum.[52]

2.3 Two Dimensional NMR Spectroscopy

The elucidation of spatial structures of biological macromolecules using NMR spectroscopy is often based on a variety of two dimensional (2D) experiments.[15] A 2D NMR experiment yields a data set which is a function of two independent frequency variables, ω_1 and ω_2. A peak in a 2D NMR spectrum indicates the presence of an interaction between a nucleus resonating at one frequency (ω_1) and a nucleus resonating at the other frequency (ω_2). The interaction (Hamiltonian) which "connects" the resonances of the two frequency dimensions depends on the experimental scheme employed.[53]

The general concept of 2D NMR spectroscopy was first suggested by Jeener in 1971.[54] All 2D NMR experiments are characterized by the four distinct time intervals shown below: a preparation period, an evolution period (t_1), a mixing period (τ_m), and a detection period (t_2).



The preparation period allows the spin system to be prepared in a suitable initial state. The preparation period usually consists of a delay period to allow the nuclear spin system to achieve thermal equilibrium followed by a pulse or sequence of pulses to create a nonequilibrium spin system. During the evolution period, the nonequilibrium spin system evolves under the influence of a Hamiltonian, H^1, that is designed to allow only certain interactions (e.g., scalar) to occur. At the end of this period, the spin system assumes a state which depends on the operative Hamiltonian, H^1, and on the elapsed time period, t_1. The evolution period is followed by a mixing period which allows the exchange of "information" between different spins by the application of a pulse or sequence of pulses. During the detection period, the spin system will evolve under the influence of the operative Hamiltonian, H^2. During the detection period, a signal is collected over the time period t_2.

The entire cycle is repeated with the systematic variation of t_1 so that the detected signal is measured as a function of two independent variables, $S(t_1,t_2)$. The time domain signal is converted by two dimensional Fourier transformation into a 2D frequency domain spectrum, $S(\omega_1,\omega_2)$, which is then subject to subsequent analysis. Because the spin system evolves under two different Hamiltonians (H^1 and H^2) during the two time periods, a 2D NMR experiment allows for the separation of spin interactions into orthogonal frequency domains.

2.3.1 Correlated Spectroscopy

Spin-spin coupling is the result of an interaction between two magnetic dipoles. The coupling between two nuclear spins in a molecule originates as a second order effect from the hyperfine coupling between the nuclear magnetic moments and the bonding electrons and consequently is the result of a through bond interaction. The spin-spin coupling energy between two spins, I_i and I_j is given by

$$H_J = I_i \cdot J \cdot I_j \qquad [2.6]$$

where J is known as the scalar coupling constant. The strength of the coupling is reflected in the magnitude of J and will depend on a number of factors including the size of the magnetic moments of the two coupled nuclei and the number and types of bonds separating the coupled nuclei. Spin-spin coupling will, in general, be observable only for nuclei separated by a four or fewer bonds. Because the coupling constant is also sensitive to the geometry of the coupling path, spin-spin coupling can be a useful parameter for molecular structure determination.

2.3.1.1 Two Dimensional Homonuclear Chemical-Shift Correlated Spectroscopy

Two dimensional homonuclear chemical-shift correlated spectroscopy (2D COSY) is an experiment with operative Hamiltonians, H^1 and H^2, designed such that the data obtained indicate spins that are connected by homonuclear (e.g., ^1H-(^1H)) spin-spin coupling networks.[55] The typical result of a 2D COSY experiment is shown in Figure 2.1A. The signal intensity that falls on the diagonal of the contour plot consists of resonances that would be observed in the corresponding 1D NMR spectrum. The signal intensity located off the diagonal (crosspeak) arises from spin-spin coupling between nuclei with resonances on the diagonal. Thus from a single 2D COSY experiment, those distinct resonances that are related by spin-spin coupling can be identified.

A number of variants of the homonuclear chemical shift correlation experiment exist that either improve or enhance the information content of the COSY experiment. For example, the TOCSY experiment allows through-bond connectivities to be made between spins that are separated by more than a few bonds and whose scalar coupling is so small as to prevent its detection by a normal COSY experiment.[56]

2.3.1.2 Two Dimensional Heteronuclear Chemical-Shift Correlated Spectroscopy

Two dimensional heteronuclear chemical-shift correlated spectroscopy enables one to map out all spins that are connected by heteronuclear (e.g., ^1H-(^{31}P)) scalar coupling

networks. The typical result of a 2D heteronuclear COSY experiment is shown in Figure 2.1B. Note that a 2D heteronuclear COSY spectrum does not contain signal intensity that falls on the diagonal of the contour plot. The crosspeaks in the 2D spectrum correlate nuclei that are scalar coupled.

2.3.2 Nuclear Overhauser Effect Spectroscopy

The nuclear Overhauser experiment is capable of providing detailed information on internuclear distances and thus can be a powerful tool for the structural elucidation of macromolecules.[57] The nuclear Overhauser effect between two nuclei is the result of a through-space dipolar interaction which enables the two nuclei to undergo cross-relaxation.[58] The dipolar coupling energy between two spins, $\mathbf{I}_i$ and $\mathbf{I}_j$, is given by

$$H_D = -\frac{\hbar^2 \gamma_i \gamma_j}{r_{ij}^3}[I_i \cdot I_j - \frac{3(I_i r_{ij})(I_j r_{ij})}{r_{ij}^2}]. \qquad [2.7]$$

The dipolar interaction essentially "connects" the energy levels between two different spin populations and the nuclear Overhauser enhancement (η) between two nuclei is the relative change in intensity of a given spin resonance when another spin resonance is perturbed, that is, when its spin population is redistributed. The cross-relaxation rate and hence, the nuclear Overhauser enhancement, is a function of the distance between the two nuclei and the reorientational frequency of the internuclear vector (see Section 2.4).

The operative Hamiltonians in the 2D nuclear Overhauser effect (2D NOESY) experiment results in a spectrum which maps out all spins connected by dipolar (e.g., ^{1}H-{^{1}H}) coupling networks.[59] A typical 2D NOESY spectrum is shown in Figure 2.1C. The signal intensity that falls on the diagonal of the contour plot consists of resonances that are identical in the two orthogonal frequency domains. During, the mixing period of a 2D NOESY experiment, two dipolar-coupled spins undergo cross relaxation and consequently, the

crosspeaks arising are those resonances in the two orthogonal frequency domains that represent interacting nuclei. The crosspeak intensity in a 2D NOESY spectrum will depend on the cross-relaxation rate and the mixing time employed during the experiment. The determination of internuclear distances from NOE crosspeak intensities is discussed in Section 2.5.

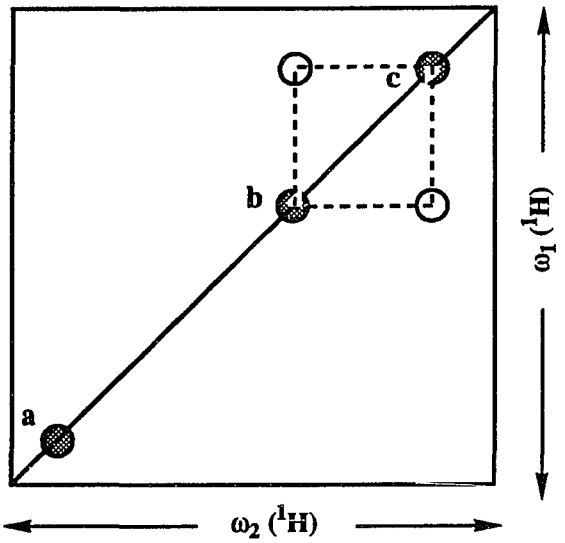
Figure 2.1 Cartoons of various two dimensional NMR experiments.

A. Representative contour plot of a ^{1}H-{^{1}H} COSY experiment: the solid circles (●) represent diagonal peaks, the open circles (○) represent COSY crosspeaks and the dashed lines (---) show the connectivity between the diagonal peaks b and c and the crosspeaks. The crosspeaks in this example would represent a scalar interaction between the proton spin belonging to peak b and the proton spin belonging to peak c.

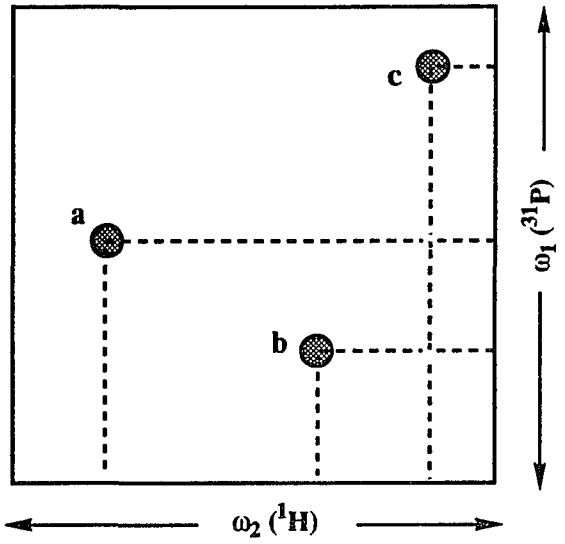
B. Representative contour plot of a ^{1}H-{^{31}P} heteronuclear chemical-shift correlation experiment: the solid circles (●) represent heteronuclear correlation crosspeaks and each dashed line (---) represents the heteronuclear scalar connectivity between a ^{1}H spin with a signal in the ω_2 domain and a ^{31}P spin with a signal in the ω_1 domain.

C. Representative contour plot of a ^{1}H-{^{1}H} NOESY experiment: the solid circles (●) represent diagonal peaks, the open circles (○) represent NOESY crosspeaks and the dashed lines (---) show the connectivity between the diagonal peaks a and c and the crosspeaks. The crosspeaks in this example would represent a dipolar interaction between the proton spin belonging to peak a and the proton spin belonging to peak c.

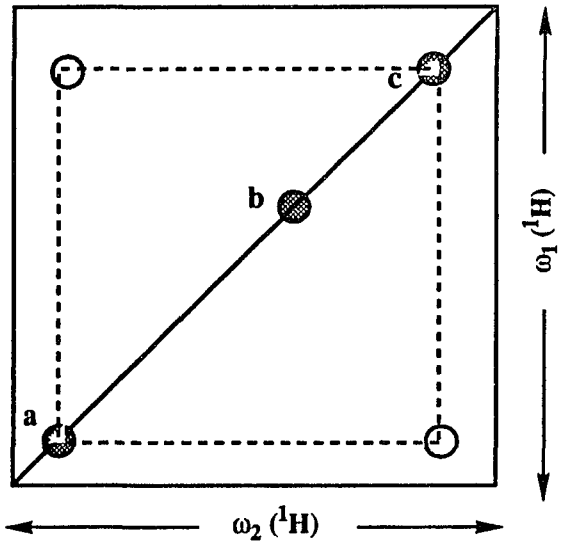
A



B



C



2.4 Complete Relaxation Matrix Analysis

Complete relaxation matrix analysis (CORMA) is an algorithm for calculating 2D NOE intensities from an assumed molecular structure and an assumed motional model for that molecular structure.[59,60,61,62] In a multispin system, the overall relaxation rate of nuclear z magnetization of the i^{th} spin, $\langle M_{zi} \rangle$ is a function of its external relaxation rate (ρ_i) and its cross-relaxation rate (σ_{ij}) with the surrounding nuclei, j:[59,63]

$$\frac{d\langle M_{zi}\rangle}{dt} = -\rho_i(\langle M_{zi}\rangle - M_{0i}) - \sum_{i \neq j} \sigma_{ij}(\langle M_{zj}\rangle - M_{0j}). \qquad [2.8]$$

The cross-relaxation rates, σ_{ij}, lead to a distribution of magnetization within the spin system while the external relaxation rates, ρ_i, lead to a loss of magnetization towards the environment. The relaxation rate can be conveniently written in matrix form as

$$\frac{dM}{dt} = -RM \qquad [2.9]$$

where **R** is the relaxation matrix with ρ_i as diagonal elements and σ_{ij} as off-diagonal elements and **M** is the magnetization vector describing the deviation from thermal equilibrium ($M=M_z-M_0$). Equation 2.9, describing the recovery of magnetization towards equilibrium, has the solution

$$M(\tau_m) = e^{-R\tau_m} M(0) = a(\tau_m) M(0) \qquad [2.10]$$

where $\mathbf{a}(\tau_m)$ is the matrix of "mixing coefficients", a_{ij}, which are proportional to the 2D NOESY crosspeak intensities for magnetic nuclei i and j at the mixing time τ_m. The CORMA approach to calculating 2D NOESY intensities involves expressing the exponential in Equation 2.10 as a series expansion

$$a(\tau_m) = e^{-R\tau_m} = 1 - R\tau_m + \frac{1}{2}R^2\tau_m^2 \cdots. \qquad [2.11]$$

The rate matrix **R** can be represented as a product of matrices

$$R = \chi\lambda\chi^T \qquad [2.12]$$

where χ is the matrix containing the eigenvectors of the rate matrix and λ is the diagonal matrix containing the eigenvalues of the rate matrix. Since λ is diagonal, the series expansion of Equation 2.11 becomes

$$a(\tau_m) = 1 - \chi\lambda\chi^T\tau_m + \frac{1}{2}\chi\lambda\chi^T\chi\lambda\chi^T\tau_m^2 \cdots = \chi e^{-\lambda\tau_m}\chi^T. \qquad [2.13]$$

Therefore, once the rate matrix **R** is diagonalized, the calculation of cross peak intensities at a given mixing time, $a_{ij}(\tau_m)$, requires only a single matrix-matrix multiplication.[62]

The calculation of 2D NOESY cross peak intensities using Equation 2.13 requires the construction of an appropriate rate matrix **R** for the multispin system. When the spin population of any level deviates from equilibrium, a flux of spin populations results in such a direction as to return the system to equilibrium. For an AX (two weakly coupled nuclear spins) system, the population flux is proportional to the appropriate population differences and the relaxation transition probabilities (refer to Figure 2.2) are as follows:

$$\frac{d}{dt}\begin{pmatrix} N_1 \\ N_2 \\ N_3 \\ N_4 \end{pmatrix} = \begin{pmatrix} -W_1 & W_1^X & W_1^A & W_2 \\ W_1^X & -W_S & W_0 & W_1^A \\ W_1^A & W_0 & -W_S & W_1^X \\ W_2 & W_1^A & W_1^X & -W_Q \end{pmatrix} \begin{pmatrix} N_1 - N_1^0 \\ N_2 - N_2^0 \\ N_3 - N_3^0 \\ N_4 - N_4^0 \end{pmatrix} \qquad [2.14]$$

where

W_1^A is the single quantum transition probability for the spin A

W_1^X is the single quantum transition probability for the spin X

W_2 is the double quantum transition probability

W_0 is the zero quantum transition probability

$W_Q = W_1^A + W_1^X + W_2$

$W_S = W_1^A + W_1^X + W_0$

N_i is the spin population of level i at time t

and N_i^0 is the equilibrium spin population of level i.

A normal mode analysis of the relaxation processes for an AX spin system reveals that the external relaxation rate (ρ_i) and the cross relaxation rate (σ_{ij}) are related to the interlevel transition probabilities, W_k, as follows:[58]

$$\rho_i = W_2 + 2W_{1i} + W_0 \qquad [2.15]$$

$$\sigma_{ij} = W_2 - W_0 \qquad [2.16]$$

The spin transition probabilities can be calculated if the perturbing Hamiltonian is known.[64] When the transition inducing Hamiltonian is represented by the dipole-dipole interaction (Equation 2.7), the transition probabilities are

$$W_{1i} = \frac{3}{20}\frac{\gamma_i^2\gamma_j^2\hbar^2}{r_{ij}^6}\frac{\tau_c}{1+\omega_i^2\tau_c^2} \qquad [2.17]$$

$$W_0 = \frac{1}{10}\frac{\gamma_i^2\gamma_j^2\hbar^2}{r_{ij}^6}\frac{\tau_c}{1+(\omega_i-\omega_j)^2\tau_c^2} \qquad [2.18]$$

$$W_2 = \frac{3}{5}\frac{\gamma_i^2\gamma_j^2\hbar^2}{r_{ij}^6}\frac{\tau_c}{1+(\omega_i+\omega_j)^2\tau_c^2} \qquad [2.19]$$

where τ_c is the correlation time characteristic of the of the dipole-dipole field fluctuations between spin i and spin j. Substituting Equations 2.17-2.19 into Equations 2.15 and 2.16, shows that for a homonuclear (e.g, ^{1}H-{^{1}H}) system of spins undergoing isotropic motion, the external relaxation rate (ρ_i) and the cross-relaxation rate (σ_{ij}) are given by

$$\rho_i = \frac{\hbar^2\gamma^4}{10}\sum_{j\neq i}\frac{1}{r_{ij}^6}\left[\tau_c + \frac{3\tau_c}{1+(\omega\tau_c)^2} + \frac{6\tau_c}{1+4(\omega\tau_c)^2}\right] \qquad [2.20]$$

and,

$$\sigma_{ij} = \frac{\hbar^2\gamma^4}{10}\frac{1}{r_{ij}^6}\left[\frac{6\tau_c}{1+4(\omega\tau_c)^2} - \tau_c\right] \qquad [2.21]$$

The relaxation matrix **R** for a multispin system can be completely constructed with the terms of Equations 2.20 and 2.21 if all of the internuclear vectors, r_{ij}, and the correlation time τ_c for each internuclear vector are known. Therefore, from an assumed structural model (i.e., all ^{1}H-{^{1}H} r_{ij} are known) and from an assumed motional model (τ_c for the dipole-dipole vectors is known), the CORMA algorithm can determine the complete set of 2D NOESY intensities expected for a given experimental mixing time. The CORMA generated NOE intensities can be used to evaluate the effectiveness of both the assumed structural and motional models in reproducing the experimentally derived NOE intensities. Furthermore, this process can be used in a refinement procedure in order to generate a structural model of a macromolecule that is consistent with the experimental NOE data.

level 1 spin A is α, spin X is α
level 2 spin A is α, spin X is β
level 3 spin A is β, spin X is α
level 4 spin A is β, spin X is β

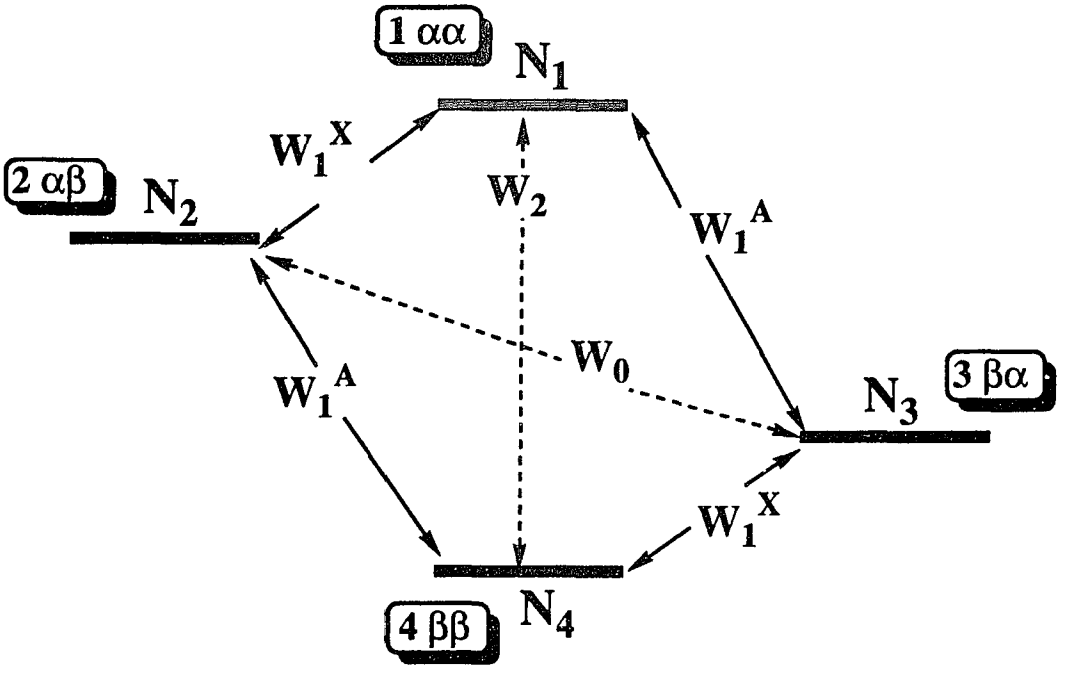


Figure 2.2 Energy level diagram for an AX spin system. The W's represent transition probabilities between different energy levels and the N's represent the populations of each energy level.

2.5 Determination of Interproton Distances From NOE Experiments

Interproton distances determined by NOE measurements can be used for three-dimensional structure determination.[57,58] The presence of a dipole-dipole interaction between two spins (Equation 2.7) enables the two nuclei to undergo cross relaxation. Cross relaxation between two spins results in a 2D NOESY crosspeak whose intensity (crosspeak volume), a_{ij}, will depend on the cross relaxation rate between the two spins and the mixing time employed during the experiment (see Equations 2.11 and 2.21) as follows

$$a_{ij} \propto f(\omega, \tau_c, \tau_m) r_{ij}^{-6} \qquad [2.22]$$

where $f(\omega, \tau_c, \tau_m)$ is a function of the correlation time (τ_c), the mixing time (τ_m) and the spectrometer frequency (ω). When the slow motion ($\omega\tau_c > 1$) condition is satisfied, the crosspeak volumes are given by

$$a_{ij} \propto \tau_c \tau_m r_{ij}^{-6} \qquad [2.23]$$

which is derived from Equations 2.11 and 2.21 under conditions of slow motion and short mixing time. An unknown distance (r_{ij}) can thus be estimated relative to a fixed distance (r_{lm}) from the ratio of the measured crosspeak volumes, a_{ij} and a_{lm}, at a given mixing time according to the equation

$$r_{ij} = \left(\frac{a_{lm}}{a_{ij}}\right)^{1/6} r_{lm} \qquad [2.24]$$

This equation is derived from Equation 2.23 assuming that the internuclear vector r_{ij} has the same correlation time as the vector r_{lm} and assuming that both internuclear vectors undergo isotropic reorientation.

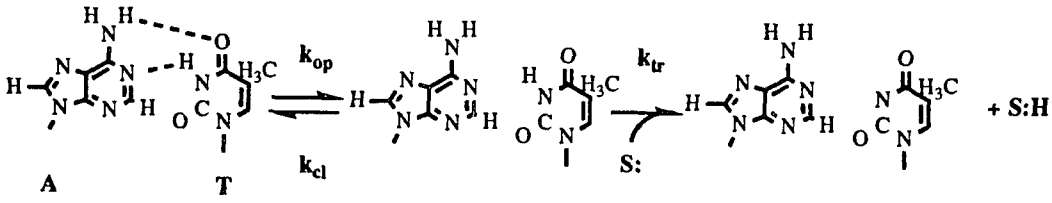
It should be noted that Equation 2.24 is only valid if τ_m is sufficiently short.

Otherwise, the presence of other spins in the vicinity of the (i,j) spin pair can modify the NOE intensity (a_{ij}) by second and higher order dipolar relaxation effects (i.e., "spin diffusion" as a result of higher order terms in Equation 2.11). As the mixing time increases, the higher order terms become more appreciable and the effect of spin diffusion on the NOE intensity becomes larger. Therefore, proton-proton distance determinations using Equation 2.24 require, as a precondition, the estimation of the spin diffusion limits which may be different for different protons.

2.6 Theory of Imino Proton Exchange in Nucleic Acids

The kinetics of imino proton exchange with solvent (S) from a hydrogen-bonded base pair (AH···B) is usually discussed in terms of a two state model with the base pair either in a closed state or an open state and where exchange can only occur via the open state [65,66,67,68].

$$AH\cdots B \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} AH+B+S \xrightarrow{k_{tr}} A+B+SH$$



This model for imino proton exchange is characterized by three rate constants: k_{op}, the rate of base-pair opening; k_{cl}, the rate of base-pair formation from the open state; and k_{tr}, the rate of transfer of the imino proton to the solvent from the open state.

Two possibilities occur depending on whether k_{cl} is smaller or larger than the rate of proton transfer from the open state, k_{tr}. In the first case, $k_{tr} \gg k_{cl}$, the transfer occurs every time the base-pair opens. The overall transfer rate then measures k_{op}, or its reciprocal, the base-pair lifetime, τ_o:

$$k_{ex} = k_{op} = \tau_o^{-1}. \quad [2.25]$$

The reciprocal of the overall transfer rate constant is thus equivalent to the base-pair lifetime

$$\tau_{ex} = \tau_o. \quad [2.26]$$

In the second case, $k_{cl} > k_{tr}$, the net transfer rate is equal to the transfer rate from the open state multiplied by the fraction of time that the base-pair is open

$$k_{ex} = k_{tr}\left(\frac{k_{op}}{k_{cl}}\right). \quad [2.27]$$

In this case, the net exchange time is

$$\tau_{ex} = \tau_{tr}\left(\frac{k_{cl}}{k_{op}}\right) = \tau_o\left(\frac{k_{cl}}{k_{tr}}\right). \quad [2.28]$$

The general formula for the exchange time is

$$\tau_{ex} = \tau_o\left(1 + \frac{k_{cl}}{k_{tr}}\right) \quad [2.29]$$

which simplifies to Equation 2.26 for the base-pair opening limited case ($k_{tr} > k_{cl}$) or Equation 2.28 for the exchange limited case ($k_{cl} > k_{tr}$). The transfer rate , k_{ex}, can be increased by the presence of an exchange catalyst. It has been shown that the transfer rate, k_{tr}, is proportional to the concentration of exchange catalyst in the solvent[69]

$$k_{tr} = R[B] \quad [2.30]$$

where [B] is the catalyst concentration and R is a constant characteristic of the catalyst and the solvent conditions. Substituting Equation 2.30 into 2.29 gives

$$\tau_{ex}=\tau_o(1+\frac{k_{cl}}{R[B]})=\tau_o+\frac{D}{[B]} \qquad [2.31]$$

where $D=\tau_o k_{cl}/R$. Hence, τ_{ex} vs. $[B]^{-1}$ is linear and the intercept at infinite concentration of catalyst (1/[B]=0) gives the base-pair lifetime, τ_o.

The exchange time, τ_{ex}, at each catalyst concentration can be obtained from measurements of the longitudinal-relaxation rate ($1/T_1$) of the imino proton, the linewidth ($\Delta\upsilon$) of the imino proton or the extent of magnetization transfer from water to the imino proton. The linewidth of an imino resonance, $\Delta\upsilon$, is related to the exchange time by

$$\tau_{ex}=\frac{1}{\pi\Delta\nu} \qquad [2.32]$$

with

$$\Delta\nu=\Delta\nu^b-\Delta\nu^0 \qquad [2.33]$$

where $\Delta\upsilon^b$ is the measured linewidth at a given catalyst concentration and $\Delta\upsilon^0$ is the linewidth in the absence of added catalyst.

2.7 Molecular Dynamics and Energy Minimization

As for any system, the structure and dynamics of a macromolecular system are governed by the thermodynamic properties of the system and its surroundings. Although many experimental methods exist to characterize the structural and dynamical features of a macromolecule, these features can be simulated by computational methods if the thermodynamical properties of the system and its surroundings are accurately known.[70] Molecular dynamics is one such computational method and involves solving the classical

equations of motion for a system of atoms by numerical integration over a period of time.[71,72]

The classical equation of motion for a system of particles is

$$\frac{\partial^2 r_i}{\partial t^2} = -\frac{\nabla E(r_i)}{m_i} \qquad [2.34]$$

where the index i runs through all the atoms in the system under simulation, m_i is the mass of atom i and r_i is the position of atom i. The gradient $\nabla E(r_i)$ is derived from an empirical energy function that has the form

$$E_{total} = E_{bond} + E_{angl} + E_{dihe} + E_{impr} + E_{VdW} + E_{hbond} + E_{elec}. \qquad [2.35]$$

The term

$$E_{bond} = \sum_{ij} k_{r_{ij}} (r_{ij} - r_{oij})^2 \qquad [2.36]$$

describes the covalent bond energy, where k_{rij} is the harmonic force constant for bond stretching between two atoms i and j, r_{ij} is the bond distance between the two atoms, and r_{oij} is the equilibrium bond distance between the two atoms. The term

$$E_{angl} = \sum_{l} k_{\theta_l} (\theta_l - \theta_{ol})^2 \qquad [2.37]$$

describes the deformation energy of angles between covalent bonds to a given atom, where $k_{\theta l}$ is the angular deformation force constant for a bond angle θ_l and θ_{ol} is the equilibrium bond angle value. The term

$$E_{dihe} = \sum_{m} k_{\Phi_m} [1 + \cos(n\Phi_m - \gamma)] \qquad [2.38]$$

describes the intrinsic deformation energy for twisting about an axis through covalently bonded atoms, where $k_{\phi m}$ is the dihedral force constant for dihedral angle ϕ_m, n is the periodicity and γ is the phase. The E_{impr} is an energy term in the form of E_{dihe} and is used to maintain either chirality or planarity within a group of atoms. The term

$$E_{VdW}=\sum_{ij}\left(\frac{A}{r_{ij}^{12}}\right)-\left(\frac{B}{r_{ij}^{6}}\right) \qquad [2.39]$$

is the Lennard-Jones potential function which represents the nonbonded Van der Waals interaction energy between pairs of atoms, i and j, where A is a repulsion term and B is an attraction term. The term

$$E_{elec}=\sum_{ij}\frac{q_i q_j}{\epsilon r_{ij}} \qquad [2.40]$$

is the Coulombic potential function used to represent the electrostatic interaction energy between pairs of charged atoms (q_i and q_j), where r_{ij} is the distance between the charged atoms and ϵ is the dielectric constant of the medium. The term

$$E_{hbond}=\sum_{ij}\left(\frac{C}{r_{ij}^{12}}\right)-\left(\frac{D}{r_{ij}^{10}}\right) \qquad [2.41]$$

represents an explicit hydrogen bond potential used to represent the interaction energy between a hydrogen, i, bonded to a donor heavy atom and an acceptor atom, j, where C is a repulsion term and D is an attraction term.

Once the potential energy functions are parameterized, the total force exerted on atom i, F_i, is given by the negative gradient of the total potential energy with respect to the position of atom i,

$$F_i = -\frac{\partial E_{total}}{\partial r_i}. \qquad [2.42]$$

If the position of a particle is known at time t, then the position after a time interval, Δt, is given by integrating the classical equation of motion given above. The numerical solution to the equation of motion can be approximated as[73]

$$x_i(t+\Delta t) = x_i(t) + v_i(\Delta t) - \nabla E_{total}\frac{\Delta t^2}{2}m_i \qquad [2.43]$$

Thus, the new position, $x_i(t+\Delta t)$, of particle *i* after the time interval, Δt, is given by Equation 2.43 using the position $x_i(t)$ and the velocity $v_i(t)$ of particle *i* and the force, F_i, on particle *i*. Calculating the new position of each particle at the end of the time step completes one time cycle of the molecular dynamics simulation and can be repeated iteratively to produce a trajectory, that is, a collection of coordinates and velocities which describe the movement of the system of atoms through phase space. Analysis of the movement of atoms as a function of time can then be used to determine the structural and dynamical features of the system under study.

Energy minimization is another method that is often employed in structural studies.[74] Although energy minimization is not a dynamical method (there is no time coordinate), it can be used to find the coordinates of a system that minimizes the total empirical energy function (Equation 2.35). Because macromolecules have a large number of degrees of freedom (~3N where N is the number of particles), energy minimization procedures applied to such systems rely on an iterative method to step along the potential surface defined by Equation 2.35 in a direction that decreases the energy. Consequently, this procedure produces a structure in a local minimum energy consistent with the potential function that is employed.

2.8 Conformational Analysis of DNA Structures

The conformational microheterogeneity of an irregular nucleic acid segment cannot be readily described by the full set of torsion angles described in Section 1.3. Lavery and Sklenar have developed an algorithm CURVES that is capable of providing a complete and unique helicoidal description of an irregular DNA structure.[75,76] CURVES calculates both the local helical parameters of individual bases or base pairs and the junction parameters between successive bases or base pair steps (Figure 2.3). The basis of the CURVES methodology is the definition of a function which describes the departure of a DNA segment from perfect helical symmetry in terms of the curvature of the DNA helical axis and in terms of changes in the position of successive bases with respect to the helical axis. Minimization of this geometric function yields the curved helical axis of the DNA segment and provides a unique helical description where both types of irregularities (i.e., in the helical axis and in the base positions) have been "smoothed" in a least-squares fashion. Since the function is constructed so as to take into account simultaneously the positions of all the bases, the final description of any one base thus depends on the position of its neighbors. Consequently, this approach results in a detailed summary of the heterogeneities present within a DNA segment and a much more coherent view of the overall conformation than that obtained with any purely local parameters such as the backbone torsion angles.

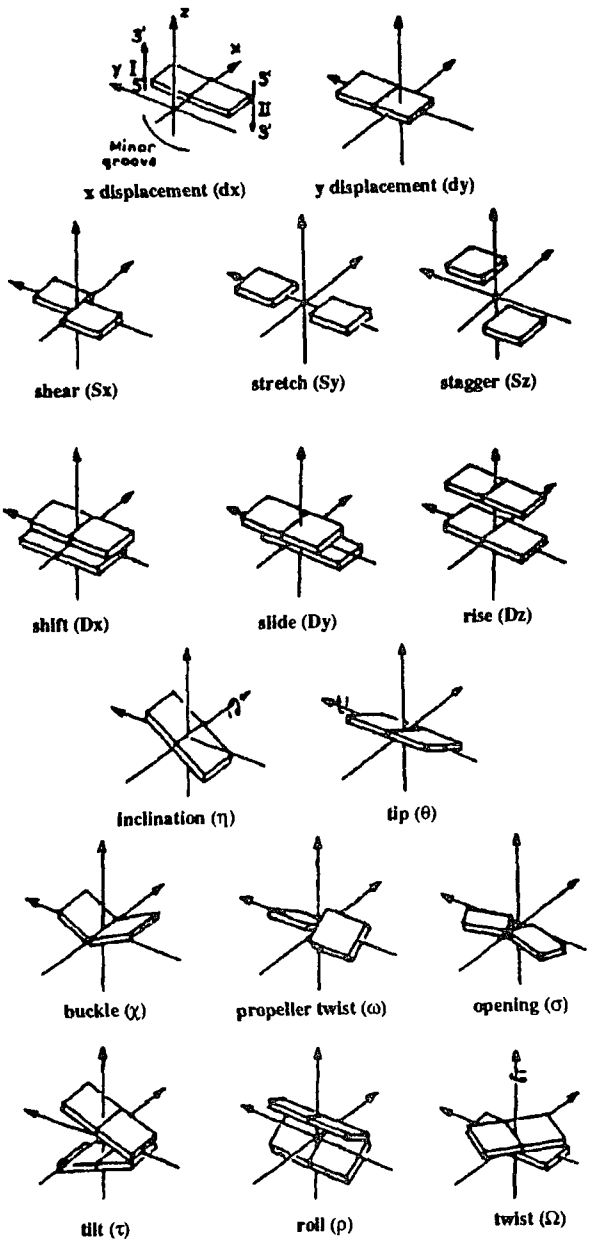


Figure 2.3 Schematic representation of the helicoidal parameters used in the CURVES program.

CHAPTER 3 Materials and Methods

3.1 Sample Section

3.1.1 Deoxyoligonucleotide Samples

Each of the deoxyribonucleotides listed below were provided by Dr. Gerald Zon (Applied Biosystems, Foster City, CA).

Sample	OD_{260}-units
d[GGAATTCC]	400
(R_p)-d[GGA(ethyl)ATTCC]	110
(S_p)-d[GGA(ethyl)ATTCC]	140

In general, each oligomer was obtained by automated (Applied Biosystems Model 380B) synthesis using phosphoramidite reagents. Each synthesis was followed by removal of the final 5'-DMT protecting group, ammoniolytic (conc. NH_4OH, 25°C, 1 h) release of material from the long-chain alkylamine controlled-pore glass support and then base/backbone-deprotection under mild conditions (conc. NH_4OH, 25°C, 48 h). The oligodeoxynucleotides were isolated by reversed-phase HPLC methods.[11] Each oligomer appeared >95% pure as judged by HPLC.

3.1.2 NMR Sample Preparation

Each oligomer was dissolved in 0.5 ml 10 mM phosphate buffer (pH=7.0) containing 0.02% NaN_3 and 100 μM EDTA. The pH of the resulting solution was adjusted to 7.0 with either 1 mM NaOH or 1 mM HCl. The pH was measured with a Corning Digital 109 pH meter equipped with a glass pH electrode (Wilmad, Buena, NJ). No corrections were made for deuterium isotope effects on pH.

Each sample was lyophilized repeatedly; after each lyophilization, the sample was dissolved in 99.96% D_2O (Aldrich Chemical Company, Inc., Milwaukee, WI) and the pH was

checked and adjusted if necessary to maintain a sample pH of 7.0. After 3-4 lyophilization cycles, the sample was dissolved in 99.996% D_2O (Cambridge Isotope Laboratories, Woburn, MA) and transferred to a 5 mm (o.d.) NMR tube (Wilmad 528-PP) under N_2 gas.

Samples in H_2O were prepared by lyophilizing the corresponding D_2O sample, dissolving the sample in a 0.5 ml H_2O/D_2O (80%/20%) solution and transferring the sample to a Wilmad 528-PP 5 mm NMR tube.

3.2 NMR Spectroscopy Section

All NMR experiments were performed on the JEOL GX-400 spectrometer of the CUNY NMR Facility at Hunter College. The spectrometer operates at a ^1H resonance frequency of 400 MHz (161 MHz for ^{31}P). All ^1H NMR experiments and ^1H-(^{31}P) indirect detection experiments were performed with a JEOL 5 mm ^1H-(X) probe. The 1D ^{31}P NMR experiments were performed with a JEOL 10 mm high frequency tunable broadband probe.

Chemical shifts of all ^1H resonances were reported relative to an external TSP (3-trimethylsilylpropionate, sodium salt) standard (TSP chemical shift at 0.0 ppm). Chemical shifts of all phosphorus resonances are reported relative to an external 50 mM phosphate (pH=7.0) standard (phosphate chemical shift at 0.0 ppm).

3.2.1 One Dimensional ^1H NMR Spectroscopy

One dimensional ^1H spectra in D_2O were acquired with a normal single pulse sequence

pulse-acquire-delay.

The residual solvent peak (HOD) was suppressed by saturating its frequency during the delay interval (homogated decoupling). Typically, 8K complex data points were acquired over a 4000 Hz spectral width. Sixty-four FIDs were signal averaged with a 3.5-4.0 s recycle delay (acquisition interval and delay interval) and were Fourier transformed after exponential multiplication with 0.5-1.0 Hz line broadening.

Because of rapid exchange between the solvent and the imino protons (thymine H3 and guanine H1), imino proton NMR resonances were observable only when the sample was dissolved in H_2O. One dimensional ^1H spectra in H_2O were obtained with a $1\bar{1}$ pulse sequence[77]

45° pulse-τ-45° pulse-acquire-delay.

The $1\bar{1}$ pulse sequence is characterized by an excitation profile with zero spectral density at the RF carrier position and maximal spectral density at a frequency

$$f_{max} = \frac{1}{2\tau} \qquad [3.1]$$

from the RF carrier position. For maximal excitation of the imino region (~3200 Hz downfield of the water signal) coupled with minimal excitation of the water signal, the RF carrier was placed at the H_2O resonance and τ was set to 0.155 ms. Typically, 8K complex data points were acquired over a 10,000 Hz spectral width. One hundred and twenty-eight FIDs were signal averaged with a 2.5-3.0 s recycle delay and were Fourier transformed after exponential multiplication with 1.5 Hz line broadening.

3.2.2 One Dimensional ^1H-{^1H} Nuclear Overhauser Effect Experiments

One dimensional NOE experiments can be used to study the dipolar interactions (Section 2.3.2) within a spin system when the signals in an NMR spectrum are sufficiently resolved. This was the case for the imino proton resonances in the unmodified octamer and its two single-site phosphate ethylated analogues. One dimensional truncated NOE experiments of the imino protons were acquired with the pulse sequence

τ_{NOE}-45° pulse-τ-45° pulse-acquire-delay

with selective saturation of a given imino resonance during the τ_{NOE} period. The RF carrier was placed at the H_2O resonance in order to avoid its excitation (Section 3.2.1) and typically 8K complex data points were acquired over a 10,000 Hz spectral width. Two hundred and

fifty-six FIDs were signal averaged with a 5 s recycle delay. Difference spectra were obtained by subtracting the FID from a reference FID accumulated with selective saturation of a frequency within the imino spectral region corresponding to noise so as to avoid Bloch-Siegert shift artifacts.[78] The difference FID was Fourier transformed after exponential multiplication with a 10 Hz line broadening. τ_{NOE} values were 10, 50, 100, 250, 500, 750, 1000, and 1250 ms.

3.2.3 Base-Pair Lifetime Measurements

Base-pair lifetime determinations were made by measuring the linewidths of the imino protons as a function of phosphate concentration. Twelve data points were obtained in which a 5.0 M phosphate solution (pH=7.0) was added in 5 μl aliquots for each data point and a 1D ^{1}H NMR spectrum was taken after each phosphate addition using a $1\bar{1}$ pulse sequence (512 scans/spectrum). The linewidth of each imino proton was determined using the NMR1 automated curve-fitting routine. The NMR1 package (New Methods Research, Inc., Syracuse, NY) employed is located at the College of Staten Island (Department of Chemistry). The exchange time (τ_{ex}) of an imino resonance at each phosphate concentration was calculated from the linewidth according to Equation 2.32. The linewidth in the absence of added phosphate catalyst was determined by extrapolating the measured linewidth (at phosphate concentrations of 10, 60 and 110 mM) to zero phosphate concentration. The τ_{ex} was fitted as a function of $1/[PO_4^{3-}]$ to Equation 2.31 using a linear least-squares program and the base-pair lifetime (τ_o) was taken as the y-intercept ($1/[PO_4^{3-}]=0$) of the linear fit. The linear least-squares analysis was performed using the PROPHET package (distributed by BBN Laboratories Inc., Cambridge, MA; sponsored by National Center for Research Resources) on the SUN 3-60 at Hunter College.

3.2.4 One Dimensional ^{31}P NMR Spectroscopy

One dimensional ^{31}P spectra were acquired with a normal single pulse sequence with

continuous broadband noise decoupling ($\gamma_H B_2 \approx 20$ kHz) of the protons. Typically, 2K complex data points were acquired over a 2000 Hz spectral width. Two hundred and fifty-six FIDs were signal averaged with a 4.0 second recycle delay and were Fourier transformed after exponential multiplication with 3 Hz line broadening.

3.2.5 Nonselective ^{31}P Spin-Lattice Relaxation Measurements

Nonselective ^{31}P spin-lattice (T_1) relaxation measurements were determined by the inversion recovery method using the pulse sequence

180° pulse-τ-90° pulse-acquire-delay

with continuous broadband decoupling ($\gamma_H B_2 \approx 20$ kHz) of the protons. Typically 2K complex data points were acquired over a 2000 Hz spectral width. One hundred and twenty-eight FIDs corresponding to a single τ value were signal averaged with a 10 s recycle delay and were Fourier transformed after exponential multiplication with 3 Hz line broadening. Eighteen τ values were chosen between 10 ms and 20 s and the signal intensity (I_τ) for each ^{31}P resonance was measured for each τ period. A single exponential was used to model the I_τ vs τ data:

$$I_\tau = I_o(1-2e^{-\tau/T_1}) \qquad [3.2]$$

where I_o is the signal intensity for a given resonance at time τ=0. A regression procedure was used to fit the I_τ vs τ data to Equation 3.2 in order to determine the unknown coefficients T_1 and I_o. The regression analysis was performed using the PROPHET package on the SUN 3-60 at Hunter College.

3.2.6 Two Dimensional ^{1}H-{^{1}H} NOESY Experiments

2D NOESY spectra were acquired with the pulse sequence

90° pulse-t_1 period-90° pulse-τ_{mix}-90° pulse-acquire-delay

with mixing times of 2, 50, 100, 200 and 500 ms. Homogated decoupling during the delay interval was employed to remove the residual solvent (HOD) peak. Typically, two sets of 1K

complex data points were acquired over a 4000 Hz spectral width in order to obtain phase-sensitive mode data using States-Haberkorn-Ruben phase cycling.[79] Thirty-two FIDs, preceded by 4 dummy scans, were signal averaged with a 3.0 s recycle delay for each t_1 value. The t_1 was initially set to 0.01 ms and incremented by 0.250 ms to achieve a 4000 Hz spectral width in the t_1 dimension. A total of 384 complex t_1 values were accumulated during the course of a single 2D experiment. In the t_2 dimension, a 5 Hz Gaussian line broadening was applied prior to Fourier transformation. The time-domain signal in the t_1 dimension was zero-filled to 1K data points and multiplied by a 90° shifted sine squared bell filter and a 5 Hz Guassian line broadening prior to Fourier transformation.

3.2.7 Absolute-Value Two Dimensional ^{1}H-{^{1}H} COSY Experiments

2D COSY spectra were acquired with the pulse sequence

90° pulse-t_1 period-90° pulse-acquire-delay

with homogated decoupling during the delay interval to remove the residual solvent (HOD) peak. Typically, one set of 1K complex data points were acquired over a 4000 Hz spectral width, 64 FIDs, preceded by 2 dummy scans, were signal averaged with a 0.75 s recycle delay for each t_1 value. The t_1 was initially set to 0.01 ms and incremented by 0.250 ms to achieve a 4000 Hz spectral width in the t_1 dimension. A total of 512 t_1 values were accumulated during the course of a single 2D experiment. In the t_2 dimension, a sine bell filter was applied over 1K data points prior to Fourier transformation. The time-domain signal in the t_1 dimension was zero-filled to 1K data points and multiplied by a sine bell filter over 512 data points prior to Fourier transformation.

Two-dimensional TOCSY spectra were acquired with the pulse sequence

90° pulse-t_1 period-(MLEV17 pulse)$_n$-acquire-delay

with n adjusted in order to produce a 70 ms spin lock period. Homogated decoupling during the delay interval was employed to remove the residual solvent (HOD) peak. Typically, two sets of 1K complex data points were acquired over a 4000 Hz spectral width in order to obtain

phase-sensitive mode data using States-Haberkorn-Ruben phase cycling.[79] Ninety-six FIDs, preceded by 4 dummy scans, were signal averaged with a 0.75 s recycle delay for each t_1 value. The t_1 value was initially set to 0.1ms and was incremented by 0.250 ms to achieve a 4000 Hz spectral width in the t_1 dimension. A total of 384 complex t_1 values were accumulated during the course of a single 2D experiment. In the t_2 dimension, a 3 Hz Lorentzian line broadening was applied prior to Fourier transformation and a fifth order polynomial baseline correction was applied to each t_2 spectrum to flatten the baseline. The time-domain signal in the t_1 dimension was zero-filled to 1K data points and multiplied by a 90° shifted sine squared bell filter prior to Fourier transformation.

3.2.8 Two Dimensional ^{1}H-{^{31}P} Heteronuclear Chemical-Shift Correlation Spectroscopy

^{1}H-{^{31}P} 2D COSY spectra were acquired with the pulse sequence[80]

^{1}H (180° pulse-τ)$_n$----------t_1 period-90° pulse-acquire-delay

^{31}P ----------90° pulse----------90° pulse----------

with ^{1}H homogated decoupling during the delay interval to remove the residual solvent (HOD) peak. During the evolution period, t_1, the ^{31}P magnetization was transferred to ^{1}H as a result of J_{HP} coupling. ^{31}P signals that do not originate from transfer were suppressed by phase cycling and by presaturation of the ^{1}H resonances using a series of n 180° ^{1}H pulses spaced by 100 ms. Typically, two sets of 1K complex data points were acquired over a 4000 Hz spectral width in order to obtain phase-sensitive mode data using States-Haberkorn-Ruben phase cycling.[79] Two hundred and fifty-six FIDs were signal averaged with a 0.75 s recycle delay for each t_1 value. The t_1 was initially set to 0.01 ms and incremented by 1 ms to achieve a 1000 Hz spectral width in the t_1 (^{31}P) dimension. A total of 100 complex t_1 values were accumulated during the course of a single 2D experiment. In the t_2 dimension, a 7 Hz Gaussian line broadening was applied prior to Fourier transformation. The time-domain

signal in the t_1 dimension was zero-filled to 256 data points and multiplied by a 90° shifted sine squared bell filter and a 10 Hz Gaussian line broadening prior to Fourier transformation.

3.3 Computation Section

3.3.1 NMR Data Processing

1D NMR data sets were processed using either the JEOL PLEXUS software supplied with the NMR spectrometer or the FTNMR software package provided by Dr. Dennis Hare (Hare Research, Woodinville, WA). All 2D data sets were processed with the FTNMR software package. FTNMR version 5.2 is installed on a DEC microVAX-II and FTNMR version 5.82 is installed on a SUN 3-60; both computers are located in the RCMI (Research Centers in Minority Institutions; sponsored by National Center for Research Resources) Graphics Laboratory at Hunter College.

3.3.2 Complete Relaxation Matrix Analysis

The CORMA package is installed on the DEC microVAX-II at Hunter College. CORMA version 1.3 was provided by Dr. Thomas L. James and Dr. Brandan A. Borgias (Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, CA).

3.3.3 Energy Minimization and Molecular Dynamics Computations

All energy minimization and molecular dynamics simulations were carried out with the program XPLOR on either the VAXstation 3100 in the Hunter College Chemistry Department or the CRAY-YMP at the Pittsburgh Supercomputing Center (sponsored by National Science Foundation). XPLOR version 2.1 was kindly provided by Dr. Axel Brunger (The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University). The force field and topology parameters (atomic masses and charges) for the DNA structures were taken from the PARAMH19E.DNA and

TOPNAH19E.DNA files provided with the XPLOR package. The topology and parameter sets treat all hydrogens explicitly and the force field uses an explicit hydrogen-bonded potential where hydrogen-bonds are modeled by a specific short-range directional interaction.

3.3.4 CURVES

The CURVES package is installed on the DEC microVAX-II at Hunter College. CURVES version 2.1 was provided by Richard Lavery (Institut de Biologie Physico-Chemique, Paris, France) and Heinz Sklenar (Central Institute of Molecular Biology, Academy of Sciences of the GDR, Berlin, Germany). The DNA input geometry was in protein data bank (PDB) file format.

3.3.5 DOCK

The DOCK package was used for the interactive display of DNA structures on the Evans and Sutherland PS-390 display subsystem located in the RCMI Graphics Laboratory at Hunter College. DOCK version 5F is installed on the DEC microVAX-II and was kindly provided by Robert K. Stodola, Frank J. Manion and Helen M. Berman (Fox Chase Cancer Center, Philadelphia, PA) and William P. Wood, Jr. (Smith Kline and French, Philadelphia, PA). The DNA input geometry was in PDB format.

CHAPTER 4 Computer Simulations

4.1 Restrained Energy Minimization and Restrained Molecular Dynamics: A Model Study Using the Dickerson Dodecamer

Nuclear magnetic resonance spectroscopy is often used to determine the structure of macromolecules in solution. A variety of NMR experiments can be used to determine structural parameters such as torsion angles (derived from scalar coupling constants) and internuclear distances (derived from NOE data). Refinement techniques such as restrained energy minimization, restrained molecular dynamics,[81,82] COMATOSE (complete matrix analysis torsion optimized structure)[62] and distance geometry[83,84] are frequently used to optimize macromolecular structures based on NMR-derived structural parameters. In order to effectively use a structural refinement method, it is important to understand the limitations of the refinement procedure for generating a macromolecular structure from the available NMR data. To this end, we have analyzed the effectiveness of the restrained energy minimization method and the restrained molecular dynamics method to refine a nucleic acid structure based on NOE-derived interproton distance data. These methods are based on the use of an energy restraint in order to generate a structure which satisfies the distance constraints imposed by experimentally determined NOE intensities.[82] A series of energy minimization and molecular dynamics calculations were performed using simulated NOE-derived interproton distances obtained from a model structure. The conditions for determining the set of interproton distance data and the errors inherent in the distance data were chosen to mimic conditions present in the experiments discussed in Chapter 5. The results of these computer simulations were used in Chapter 5 to evaluate the effectiveness of 2D NOESY experiments on the parent and its ethylated analogues for defining the solution structure of these molecules.

4.2 The Starting Structure and Target Structure

The goal of these calculations is to determine what conformational features of an

irregular nucleic acid (the target structure) can be reproduced by the structural refinement of an assumed starting structure using either restrained energy minimization or restrained molecular dynamics. The structure obtained from x-ray coordinates for the dodecamer, $\{d[CGCGAATTCGCG]\}_2$, was used as the target structure (<T>) for these simulations. Heavy atom coordinates for the dodecamer were obtained from the Protein Data Bank (2BNA, deposited by H. Drew and R. Dickerson, 11/81) and protons were added in standard geometries using the XPLOR HBUILD subroutine. A standard B-form geometry of the dodecamer was used as the starting structure (<S>) for the simulations. The atomic coordinates for the B-form dodecamer were obtained with a program provided by Dr. Gary Quigley (Hunter College).

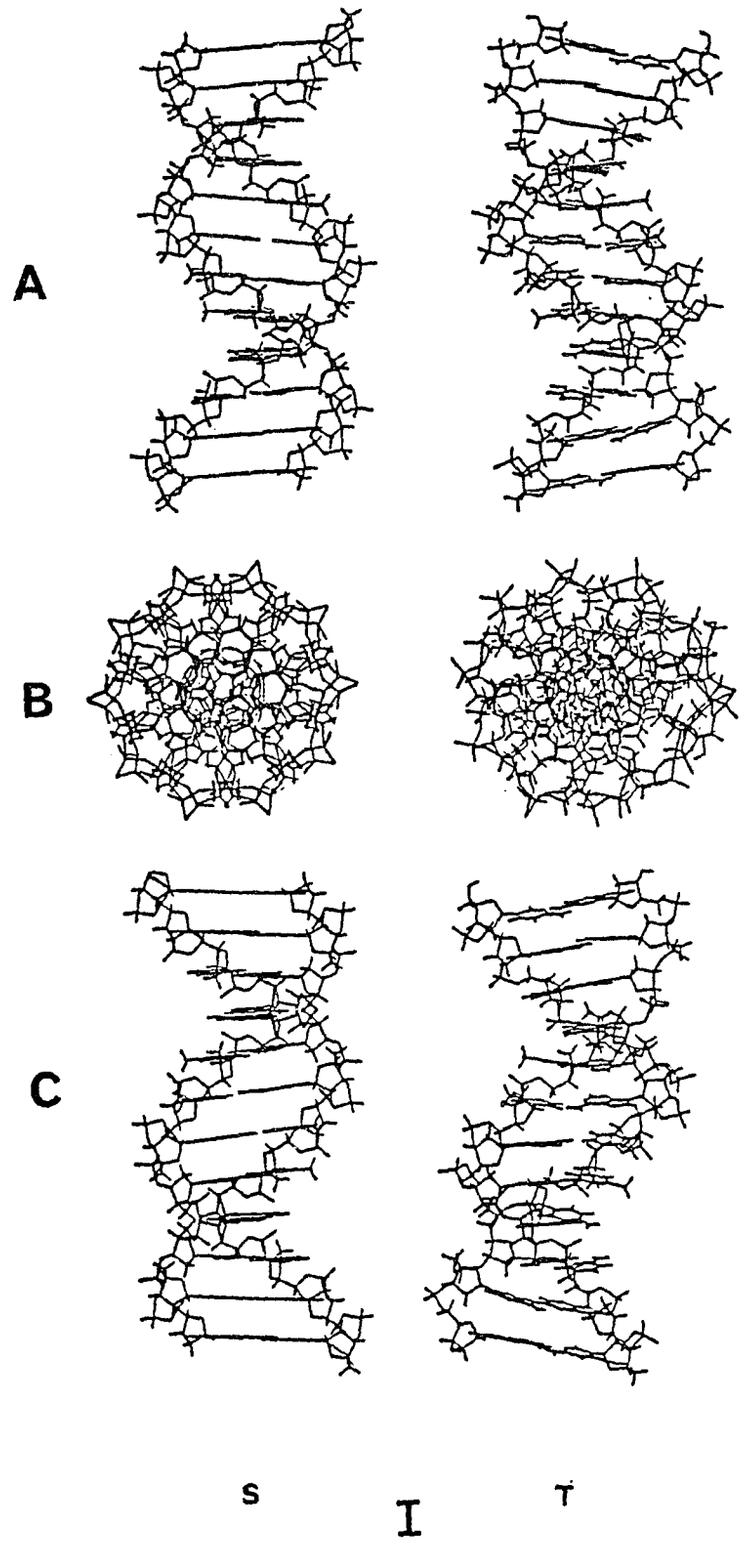
Different perspectives of the start and target structures are shown in Figure 4.1. Although the DOCK generated pictures display the spatial location of all atoms in each molecule, the structural irregularities of each nucleic acid are better defined by the CURVES generated pictures which emphasize the specific orientation of successive nucleotides and the non-linearity of the helical axis. The helicoidal parameters for the start and target structures were determined using the CURVES algorithm described in Section 2.8. The CURVES generated helicoidal parameters for the start and target structures are given in Appendix I. A comparison of the helicoidal parameters for the two structures reveals that although both structures adopt B-form conformations, the target structure displays conformational features characteristic of an irregular nucleic acid structure. For example,

- sugar puckering is not constant over the target structure (the pseudorotation phase angle varies between 25° and 195°);
- the overall helix of the target structure is not straight but rather is bent by ~19°;
- the base pair propeller angle appears to be greater for the A:T base pairs in the center than for the G:C base pairs at the ends of the target structure;
- large buckling angles are observed for the base pairs that flank the central (AATT) sequence of the target structure.

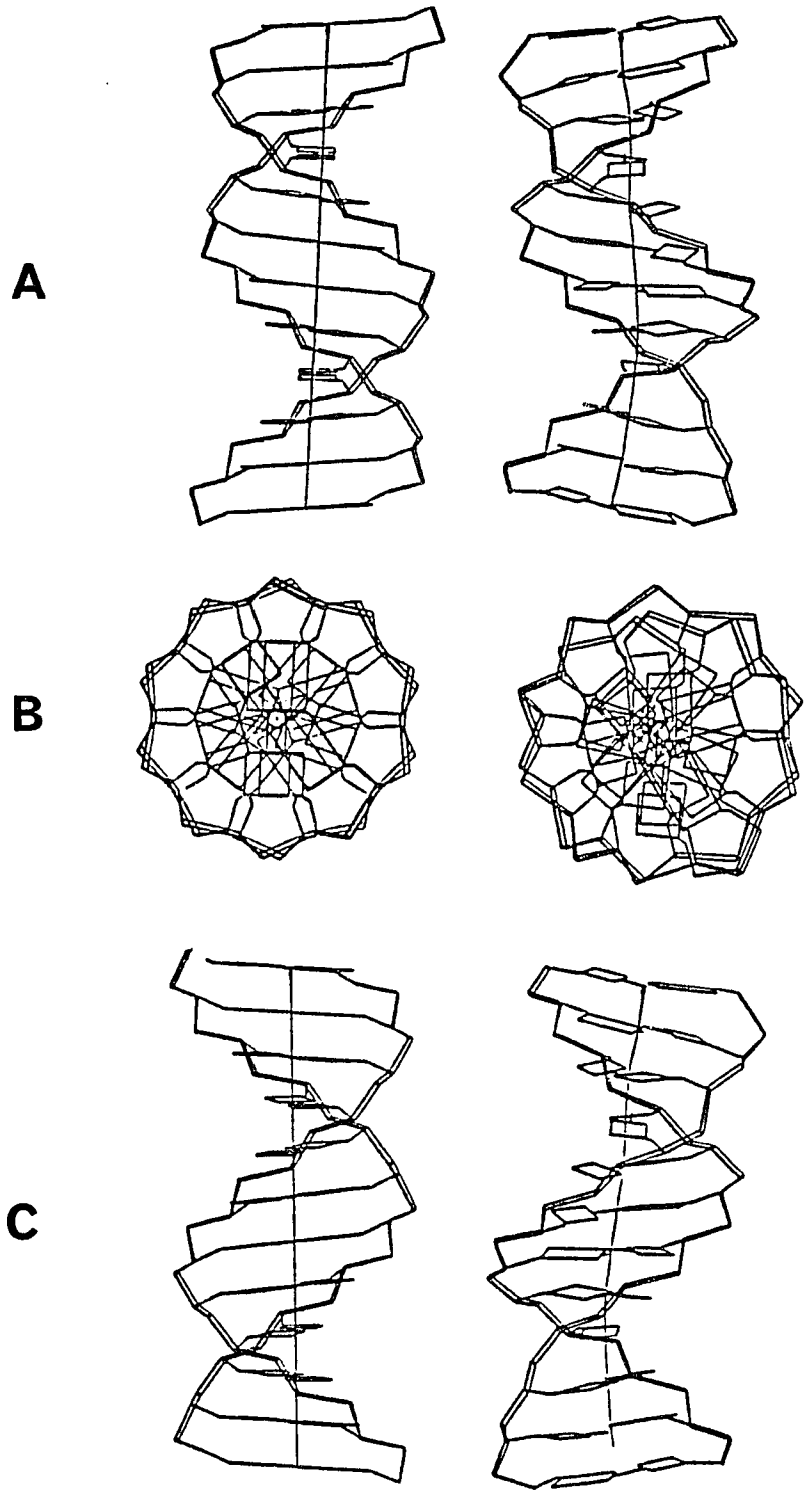
The structures obtained from restrained energy minimization and restrained molecular dynamics refinement of the starting structure were similarly analyzed using the CURVES program to determine if the conformational features found for the target structure are effectively incorporated into the refined structures.

Figure 4.1 DOCK (I) and CURVES (II) generated pictures of the starting structure <S> and target structure <T>. The DOCK generated picture of each molecule is shown as an all-atom skeleton model in which each vector corresponds to a bond between a pair of atoms (Dreiding model). The CURVES generated picture of each molecule is a representation in which the sugar-phosphate backbone is shown as a ribbon and each base is shown as a rectangular plane.

A. View of major groove
B. View down helical axis
C. view of minor groove



I



II

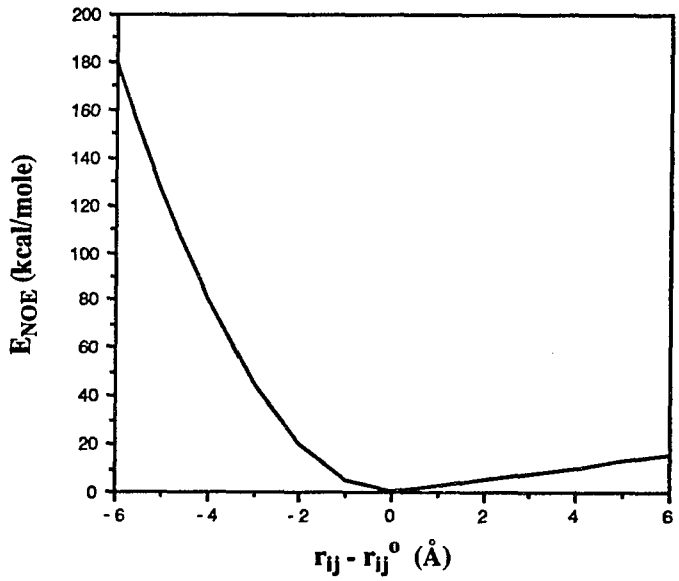
4.3 Generation of Restraints Employed in the Refinement of the Starting Structure

Distance constraints that simulate an NOE-derived interproton distance set were included as input for the restrained energy minimization and restrained molecular dynamics simulations. Interproton distances derived from the target structure were used as the distance constraints and were obtained using the XPLOR module DISTANCE. The following criteria were used to decide which interproton distances should be included in the simulations:

- only ^1H-^1H distances less than 5.0 Å were included; under typical experimental conditions NOE interactions can be observed only between protons separated by less than 5 Å;
- all 5',5" sugar protons and the exocyclic amino and imino protons were excluded; these protons were either unresolved, unassigned or undetected in the 2D NOESY experiments discussed in Chapter 5.

Using these criteria, a total of 717 distance constraints were found between 168 protons in the target structure. Of these, 131 ^1H-^1H pairs were at a separation of $r \leq 2.5$ Å, 168 pairs were at a separation of $2.5 Å < r \leq 3.5$ Å, 284 pairs were at a separation of $3.5 Å < r \leq 4.5$ Å, and 134 pairs were at a separation of $4.5 Å < r < 5.0$ Å.

Once each interproton distance has been determined, a corresponding energy restraint (E_{NOE}) is added to the empirical energy function (Equation 2.35) used for the energy minimization or molecular dynamics procedure. The NOE-restraint function should be designed so that the refinement procedure will bring the protons i and j within an upper bound, u_{ij}, but not closer than a lower bound, l_{ij}, in the refined structure where the upper and lower bounds represent the positive and negative error estimates, respectively, of the NOE-derived interproton distance, r_{ij}°. The skewed biharmonic energy function is capable of producing this behavior[82] and has the form



where,

$$E_{NOE}=C_1(r_{ij}-r_{ij}^{o})^2 \text{ for } r_{ij}>r_{ij}^{o} \qquad [4.1]$$

$$E_{NOE}=C_2(r_{ij}-r_{ij}^{o})^2 \text{ for } r_{ij}<r_{ij}^{o} \qquad [4.2]$$

and where r_{ij} is the interproton distance of the structure being refined, r_{ij}^{o} is the target (NOE derived) interproton distance, and C_1 and C_2 are force constants for the biharmonic energy term. The force constants are given by Equations 4.3 and 4.4

$$C_1=\frac{kTS}{2\Delta_{ij}^{+2}} \qquad [4.3]$$

$$C_2=\frac{kTS}{2\Delta_{ij}^{-2}} \qquad [4.4]$$

where k is the Boltzmann constant, T is the absolute temperature, S is a scale factor and Δ^{+}_{ij}

and Δ^{-}_{ij} are the positive (upper bound) and negative (lower bound) error estimates of the target interproton distance r^{o}_{ij}. The derivation of each force constant (Equations 4.3 and 4.4) requires a positive and negative error estimate (Δ^{+}_{ij} and Δ^{-}_{ij}) for each NOE-derived interproton distance. In general, distances derived from 2D NOESY data are estimated relative to a fixed distance according to Equation 2.24. At short mixing times, a 2D NOESY crosspeak represents cross-relaxation between a pair of protons and the crosspeak intensity is directly proportional to the inverse sixth power of the interproton distance. At longer mixing times, spin diffusion due to indirect cross-relaxation pathways can contribute to the intensity of a NOESY crosspeak intensity leading to errors in the calculated interproton distance. An analysis of the systematic error introduced by spin diffusion made by Clore and Gronenborn[85] indicated that the apparent value of an NOE-derived interproton distance will become closer to that of the reference distance as the extent of spin diffusion increases. Therefore, an NOE-derived interproton distance that is smaller than the reference distance will be systematically overestimated whereas an NOE-derived interproton distance that is larger than the reference distance will be systematically underestimated.

In B-form DNA there are an average of 3.4 neighboring protons within a 3.0 Å radius of each proton. The presence of neighboring protons provides indirect cross-relaxation pathways that can lead to spin diffusion. In order to quantitate the errors introduced by spin diffusion on interproton distance determinations in a multi-spin system like B-DNA, CORMA calculations (Section 2.4) on the dodecamer target structure were performed at a number of mixing times and correlation times. The CORMA calculation allows the determination of the complete set of 2D NOESY crosspeak intensities with the presence of spin diffusion explicitly included. Distances derived from the CORMA-simulated NOE intensities for the target structure (the C5-C6 distance of 2.5 Å was used as the reference distance), were compared with the actual distances obtained from the target structure. An analysis of the errors in the distance determination:

(r_{ij} derived from the target structure) -

(r_{ij} derived from CORMA calculation on the target structure)

indicates that the following errors would be expected for NOE-derived distances obtained at mixing times of 100-200 ms and an isotropic correlation time of 4 ns (conditions appropriate to the NOE analysis of the parent octamer and its ethylated analogues discussed in Chapter 5):

NOE-Derived Distance	Positive Error
r ≤ 2.5 Å	+0.10 Å
2.5 Å < r ≤ 3.5 Å	+0.25 Å
3.5 Å < r ≤ 4.5 Å	+0.50 Å
4.5 Å < r	+1.00 Å

For each distance range, spin diffusion results in the calculated distance being underestimated relative to the actual distance and, therefore, the above errors represent a positive error estimate (Δ^{+}_{ij}) to the NOE restraint. Note that as the NOE-derived interproton distance becomes larger than the reference distance, the positive error for the derived distance becomes larger because the effects of spin diffusion causes this distance to be underestimated (i.e., an actual interproton distance of 4.0 Å would, in general, be experimentally underestimated by ~0.5 Å so that its NOE-derived distance would be ~3.5 Å). Figure 4.2 demonstrates the effect of spin diffusion on the sequential (*n*)H8/H6↔(*n+1*)H1' interproton distances derived from the CORMA-simulated NOE intensities for the target structure at different mixing times. It is clear from this figure that at longer mixing times, the calculated distances converge toward the reference distance and the sequence-specific variation of the (*n*)H8/H6↔(*n+1*)H1' interproton distance disappears. The energy restraint used for structure

refinement also requires a lower bound (negative) error estimate, Δ^{-}_{ij}. The lower bound error represents the degree to which the NOE-derived interproton distances are overestimated. Because most NOE-derived interproton distances within DNA oligomers are larger than the reference distances available (1.8 Å for the 2'↔2" reference or 2.5 Å for the C-H5↔C-H6 reference), it is safe to assume that the derived distances are unlikely to be overestimated and therefore we have arbitrarily chosen -0.1 Å as the lower bound error for each category of distance ranges.

It is also important to note that Equation 2.24 is only valid if the reference internuclear vector (r_{lm}) has the same effective correlation time, τ_c, as the measured internuclear vector, r_{ij}. A difference in τ_c between the reference internuclear vector and the measured internuclear vector also could result in an error in the distance determination for the measured internuclear vector. The motional model assumed for this study is an isotropically reorienting rigid molecule. This motional model is therefore characterized by a single correlation time. It should be noted that this type of model may not be a valid assumption for DNA fragments[86] but errors in the distance determination associated with differences in τ_c among the internuclear vectors in a structure were not evaluated here.

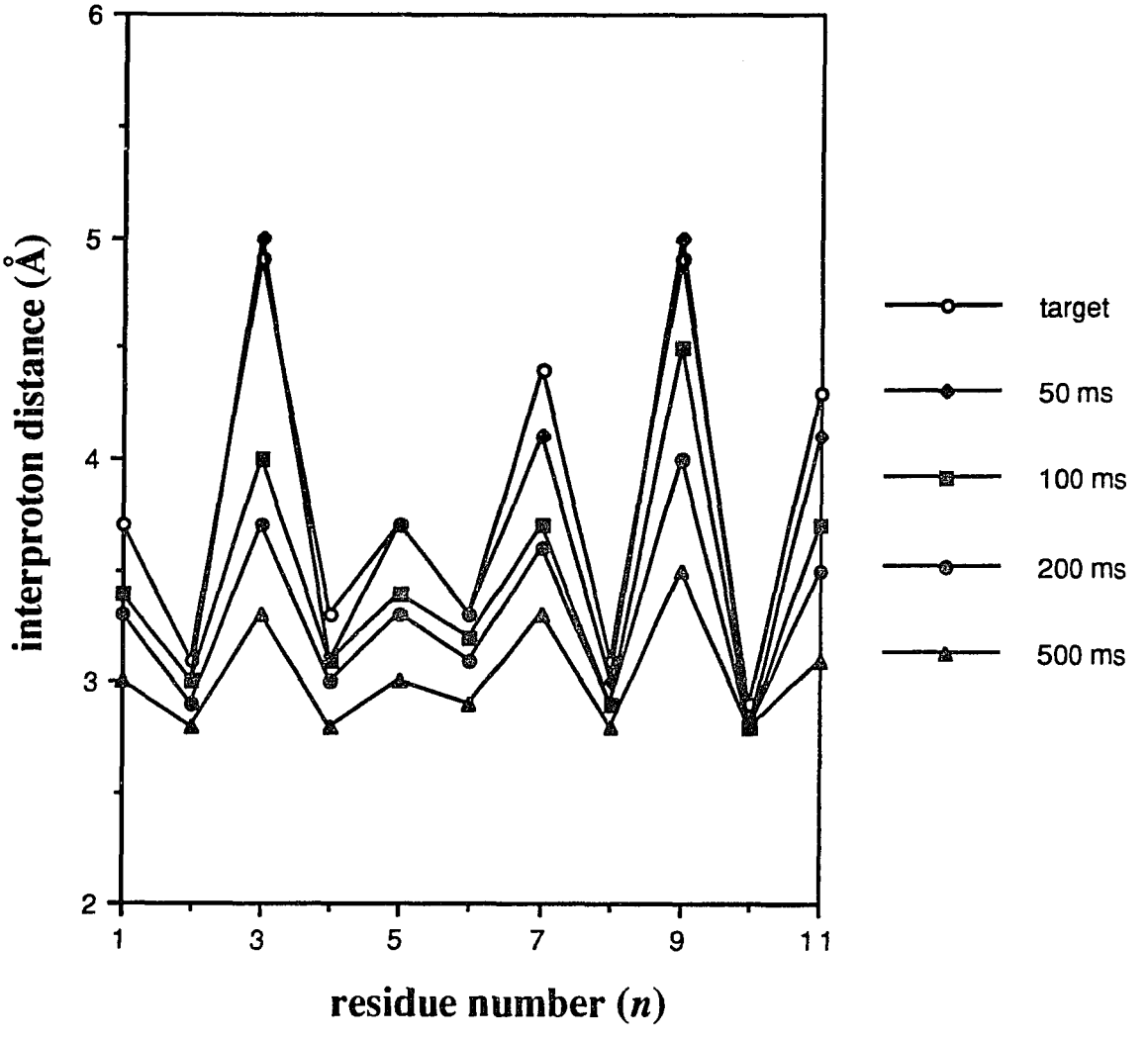


Figure 4.2 The sequential (*n*)H8/H6-(*n+1*)H1' interproton distance as a function of residue number. The open circle (○) represents the interproton distances for the target structure. The filled symbols (●,▲,■,♦) represent the interproton distances derived from the CORMA-simulated NOE intensities at different mixing times.

4.4 Restrained Molecular Dynamics and Restrained Energy Minimization Calculations

The computational time of a molecular dynamics or energy minimization calculation depends on the number of particles in the system. Because of limited computational resources, all calculations were performed *in vacuo* and thus water molecules and counterions were not explicitly included. To approximate the effects of solvent, a 1/r dielectric term was used in the electrostatic potential function (Equation 2.40) and to approximate the effects of counterion condensation, the net charge on each phosphate group was reduced to -0.32 e.[82] The computational time was further reduced by limiting the number of pairwise interactions used for the nonbonded potential energy functions (Equations 2.39 and 2.40); the nonbonded interactions were switched off smoothly to zero by using a cubic switching function between 9.5 Å and 10.5 Å with pairwise interactions up to 11.5 Å included in the nonbonded list.

Energy minimizations were performed with the MINIMIZE POWELL module within XPLOR. Typically, an energy minimization was carried out until the total energy gradient dropped below 0.001 kcal mole^{-1} Å^{-1}. Molecular dynamics calculations were performed with the DYNAMICS VERLET module within XPLOR. Before a given molecular dynamics simulation was performed, the nearest local minimum in the neighborhood of a structure was found by conjugate gradient minimization with a soft harmonic constraint

$$E_{harm} = \sum k(r_i - r_i^{ref})^2 \qquad [4.5]$$

where $\mathbf{r}_i^{ref}$ represents the atomic coordinates of the structure and k_h represents the harmonic force constant (~ 25 kcal mole^{-1} Å^{-2}). The system was then prepared for molecular dynamics at 300 K by a heating and equilibration period. Heating was accomplished with initial velocities assigned from a Maxwellian distribution at 300 K to all of the atoms in the structure. Equilibration of the system for molecular dynamics was performed for 10,000 steps with a 0.5 fs timestep (total time=5.0 ps) where the temperature was maintained in the vicinity of 300 K by rescaling the velocities of all atoms every 0.1 ps. Bond lengths involving hydrogen atoms were kept fixed using the SHAKE algorithm.[87] Twenty picoseconds of

"production" molecular dynamics (20,000 steps with a 1 fs timestep) was performed on the equilibrated structure with velocity rescaling and coordinate sampling every 0.1 ps. Averaged dynamics structures were obtained for analysis by averaging the coordinates over the last 15 ps of trajectories (i.e., from 5 to 20 ps) and subjecting the averaged structure to conjugate gradient minimization with a soft harmonic constraint (harmonic force constant ~ 25 kcal $mole^{-1}$ $Å^{-2}$).

Restrained dynamics was performed as described above except an NOE energy term representing each NOE-derived interproton distance (Equations 4.1 and 4.2) was added to the potential energy function during the equilibration period by increasing the scale factor, S, in Equations 4.3 and 4.4 from 0 to 3 in 0.5 increments every picosecond during the first 6 ps. The restraint energy (E_{NOE} term) force constant for each proton-proton pair was calculated for a given scale factor using Equations 4.3 and 4.4. The following table lists the force constant calculated for each category of distance errors using a scale factor of 3:

Distance Error (Å)	**Force Constant (kcal mol^{-1} $Å^{-2}$)**
0.10	89.4
0.25	14.3
0.50	3.6
1.00	1.8

It should be noted that a C-H covalent bond has a force constant of ~100 kcal mol^{-1} $Å^{-2}$. Therefore, proton pairs that have an NOE-derived distance error of ±0.1 Å will have a restraint energy that approximates that found in a C-H bond for a scale factor of 3. Although

the scale factor for the restraint energy function can be arbitrarily chosen, if the force constant is approximately 50 kcal mol^{-1} $Å^{-2}$, a violation of 0.1 Å (a distance violation, $r_{ij}-r_{ij}^{o}$, represents the difference between the target interproton distance and the corresponding interproton distance in the structure during the dynamics simulation) will add kT/2 to the total potential energy at 300 K. Because the average kinetic energy of an atom per degree of freedom is kT/2, this implies that during the simulation, violations of the order of 0.1 Å are tolerated. An advantage of a low value for the force constant is that the molecule will have more freedom to find low-energy conformations during the simulation. However, if the force constant is made too small by decreasing the scale factor, the molecular dynamics simulation may not be able to generate a structure that is consistent with the NOE-derived interproton distance data. Evidently, the size of the scale factor needs to be chosen so that it is large enough to overcome "inaccuracies" in the force field used in the simulation but small enough so that a sufficient amount of conformational flexibility is maintained.

4.5 Analysis of Structures Obtained From the Computer Simulations

A series of energy minimization and molecular dynamics calculations were carried out either in the presence (i.e., restrained) or absence (i.e., free) of NOE restraints according to Section 4.4. The following nomenclature is used to represent the structure derived in each case:

Structure	Nomenclature
start	<S>
target	<T>
start + free minimization	<SFM>
start + restrained minimization	<SRM>
start + free dynamics	<SFD>
start + restrained dynamics	<SRD>

DOCK plots of each structure are shown in Figure 4.3. In each case, except for the free dynamics structure <SFD>, the resulting structure remains close to a B-DNA conformation. In the free dynamics calculation, the duplex melts because the simulation conditions are evidently not accurate enough to keep the duplex structure together. The presence of NOE restraints sufficiently overcomes the deficiencies in the simulation conditions to keep the restrained structures, <SRM> and <SRD>, close to a B-DNA conformation.

The root mean square (r.m.s.) difference between any two structures, A and B was calculated, after a least squares superposition of the two structures, with the expression

$$r.m.s. = \sqrt{\sum_i (r_{i_A} - r_{i_B})^2} \qquad [4.6]$$

where r_{iA} is the coordinate of atom i of structure A and r_{iB} is the coordinate of atom i of structure B. The r.m.s. differences between various combinations of structures for all atoms are given in Table 4.1 and the calculated energy terms for each structure are given in Table 4.2. The NOE energy represents the sum of the restraint energies (E_{NOE}) for all proton pairs

in the restraint list. The restraint energies were calculated using Equations 4.1 and 4.2 and a scale factor of 3. The NOE energy for the starting structure is ~8.5 kcal/mole and this energy is substantially reduced after both restrained minimization (E_{NOE}~20 cal/mole) and restrained dynamics (E_{NOE}~40 cal/mole). The reduction in E_{NOE} indicates that both refinement procedures, restrained minimization and restrained dynamics, are capable of substantially reducing the NOE-restraint energy. That is, the internuclear distances for all proton pairs in the restraint list successfully converge to the simulated NOE-derived internuclear distances during the refinement procedure. It is interesting to note that although the NOE-restraint energy is approximately zero after either refinement procedure, the r.m.s. difference between the <T> and refined structures is only reduced by ~10% compared to the r.m.s difference between <T> and <S> (r.m.s. <T>-<S>=1.48 Å, r.m.s. <T>-<SRM>=1.33 Å, r.m.s. <T>-<SRD>=1.30 Å). This could arise because some regions of the target structure become precisely defined by the distance restraints (e.g., the sugar ring geometry) while other regions, not defined by the distance restraints (e.g., the sugar-phosphate backbone) deviate significantly from the target structure. Note that E_{elec} for the refined structures is ~100 kcal mol^{-1} less than that for the target structure. The E_{elec} term is dominated by charge-charge interactions involving the phosphate groups of the DNA backbone. The conformation of the backbone will be largely determined by electrostatic interactions and because it is not defined by the distance constraints, it is free to deviate from the target structure during the refinement procedure.

Table 4.3 summarizes some of the CURVES-derived conformational parameter deviations between the target structure, the starting structure, and the structures obtained from restrained minimization and restrained dynamics of the starting structure. Examination of Table 4.3 reveals that both restrained minimization and restrained dynamics reduce the deviations in most conformational parameters, torsion angles and sugar ring puckers between the starting structure and the target structure, the only exception being the α (P-O5') torsion angle. The lack of improvement in the α torsion angle after restrained minimization or dynamics is probably due to the lack of NOE restraints along the sugar phosphate backbone.

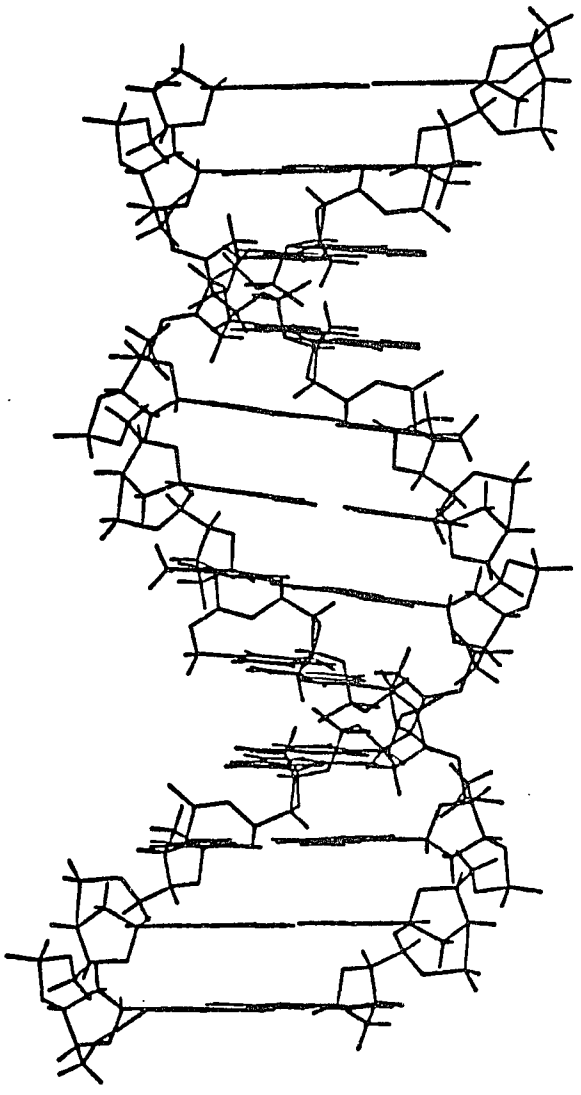
Figure 4.4 and Figure 4.5 displays the sequential (*n*)H8/H6↔(*n+1*)H1' interproton distance and some of the CURVES-derived conformational parameters as a function of the residue number within each structure. The starting structure exhibits no sequence-dependent variation in any of the helicoidal parameters because it is in an idealized B-DNA conformation whereas the target structure exhibits sequence-dependent variations in all of the helicoidal parameters. Examination of Figure 4.4 indicates that the <SRD> and <SRM> structures reproduce the sequence-dependent variations in the (*n*)H8/H6↔(*n+1*)H1' interproton distance whereas the <SFM> and <SFD> structures do not. Examination of Figure 4.5 indicates that although the exact value of each helicoidal parameter in the target structure is not adopted by the <SRD> and <SRM> structures, the refined structures do follow the same trend as the target structure for most of the sequence-dependent irregularities in each helicoidal parameter. The base-pair opening helicoidal parameter is an exception and is consistent with the fact that no NOE-derived interproton distances (e.g., the exocyclic amino and imino protons) that serve to define the hydrogen-bonding geometry between base-pairs were used in the simulation.

Table 4.3, Figure 4.4 and Figure 4.5 also indicate that, although some conformational parameters (e.g., propeller twist) are improved by performing restrained dynamics compared to restrained minimization, the majority of conformational parameters, torsion angles and sugar puckers become well defined (r.m.s deviations decreases) during restrained minimization and are not significantly improved after restrained dynamics. Although molecular dynamics would be required to convert, for example, an A-DNA structure to a B-DNA structure using distance constraints,[82] a dynamics method is apparently not necessary to reproduce many of the local conformational features of an irregular B-DNA structure (<T>) from an idealized B-DNA structure (<S>). This suggests that the energy barriers to most of the geometrical changes that must be crossed to cause the transition between <S> and <T> are relatively small and can be crossed using an energy minimization method with the distance restraints employed in this study.

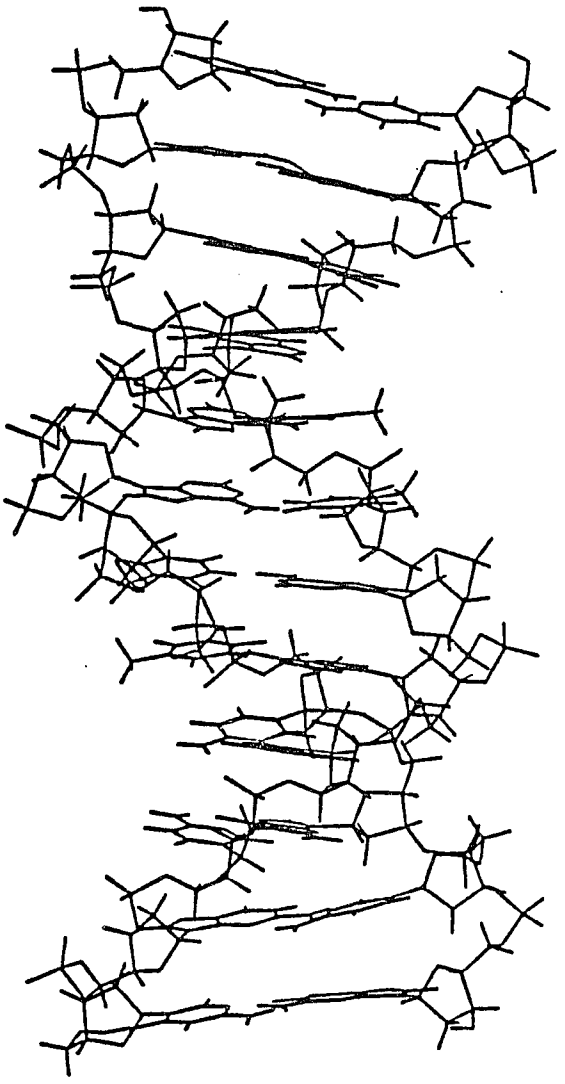
These results demonstrate the limitations of the restrained energy minimization and restrained molecular dynamics refinement procedures for generating the structural features of a DNA fragment from a set of simulated NOE-derived distances. The restraint function employed distance data and distance errors that would be obtained from a typical 2D NOESY experiment. It was found that both refinement methods, using this distance data, succeeded in defining the sugar ring conformations and many of the base-related structural irregularities present in the target structure. However, some of the sugar-phosphate backbone torsion angles were not accurately defined during the simulations. These results are used in Chapter 5 to evaluate the ability of NOE-derived interproton distances for the parent and its ethylated analogues to define the structural features of these molecules in solution.

Figure 4.3 Dock plots of:

A. <S> (I) and <T> (II)
B. <SFM> (I) and <T> (II)
C. <SRM> (I) and <T> (II)
D. <SFD> (I) and <T> (II)
E. <SRD> (I) and <T> (II)

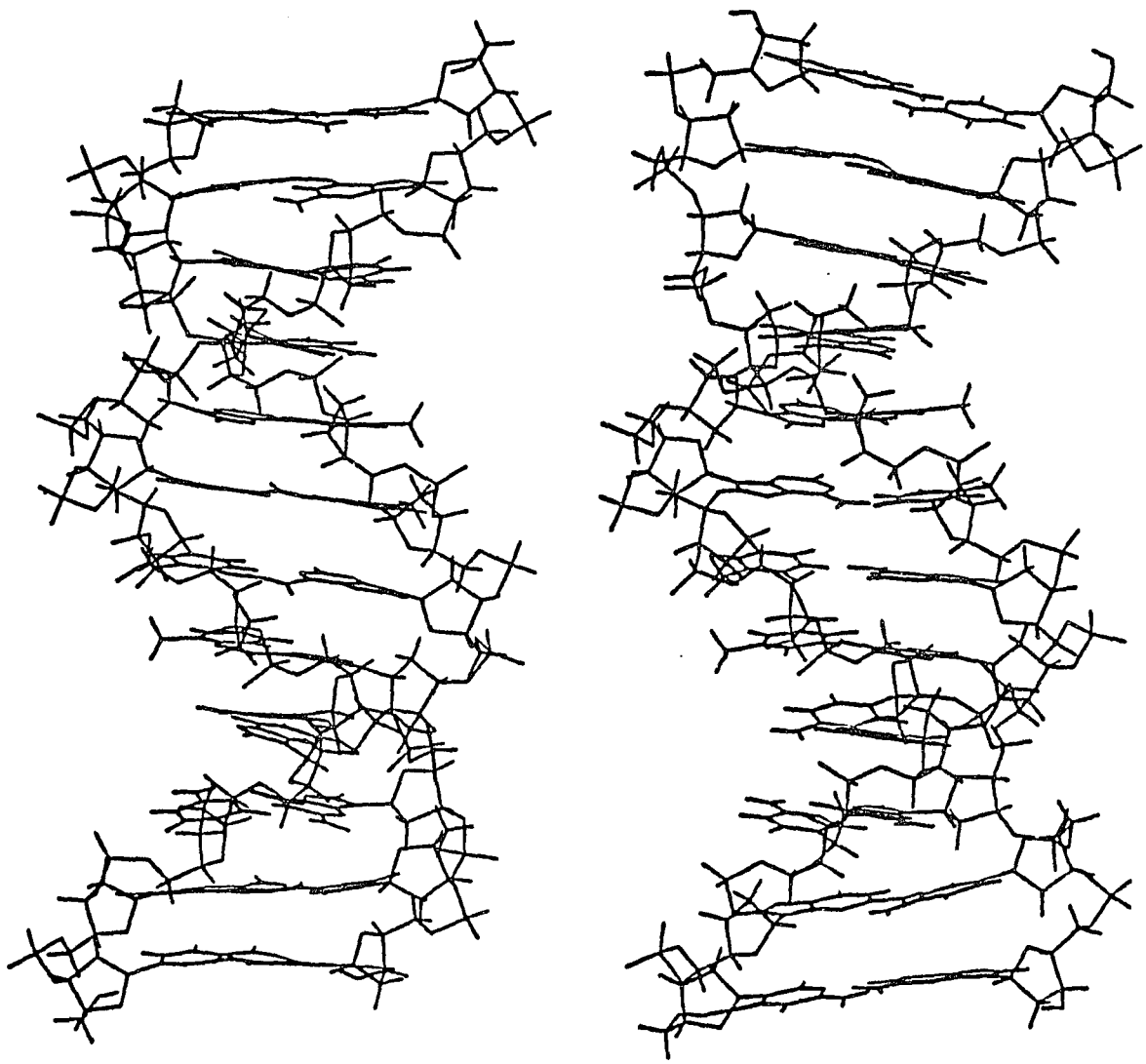


I



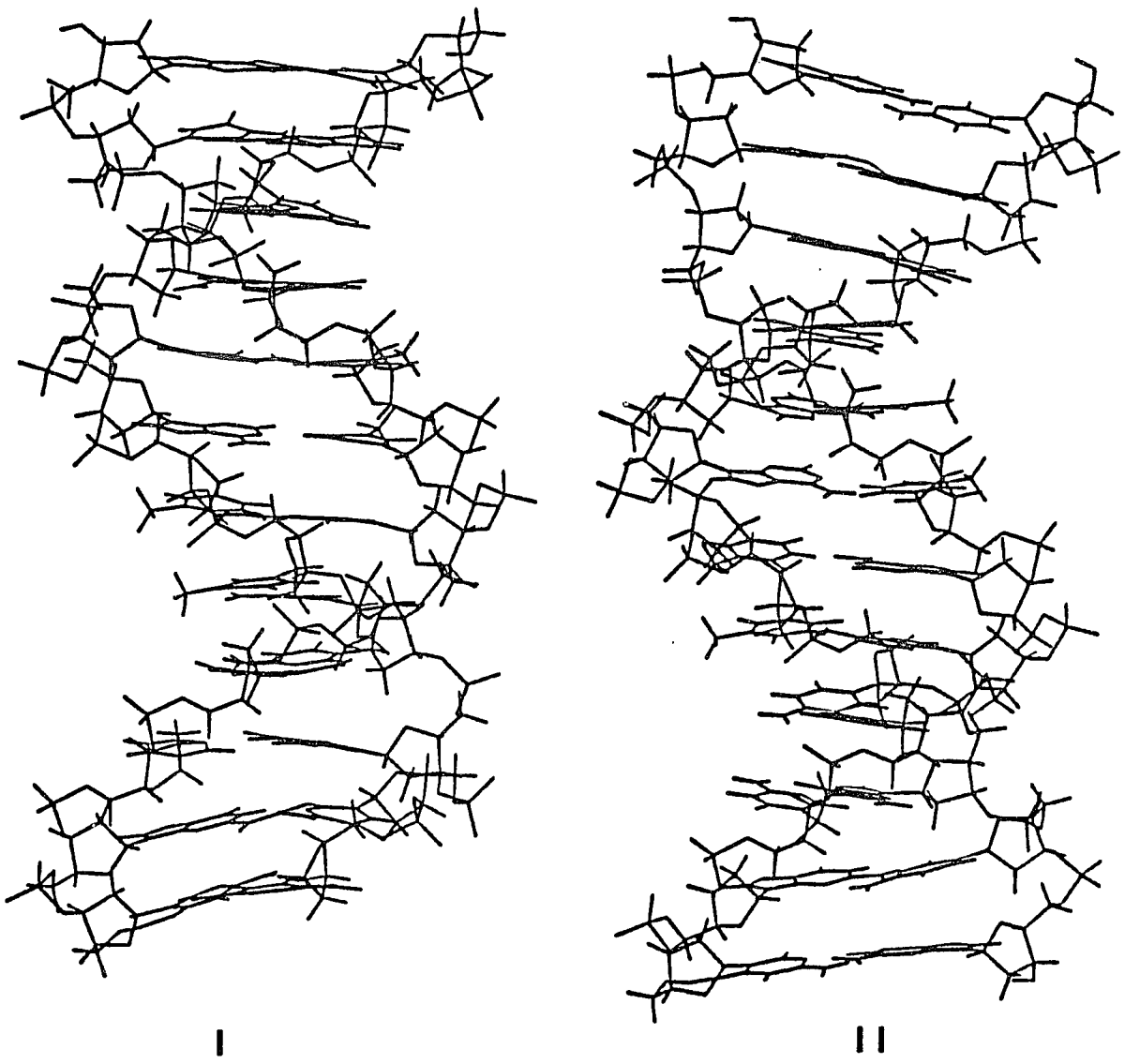
II

A



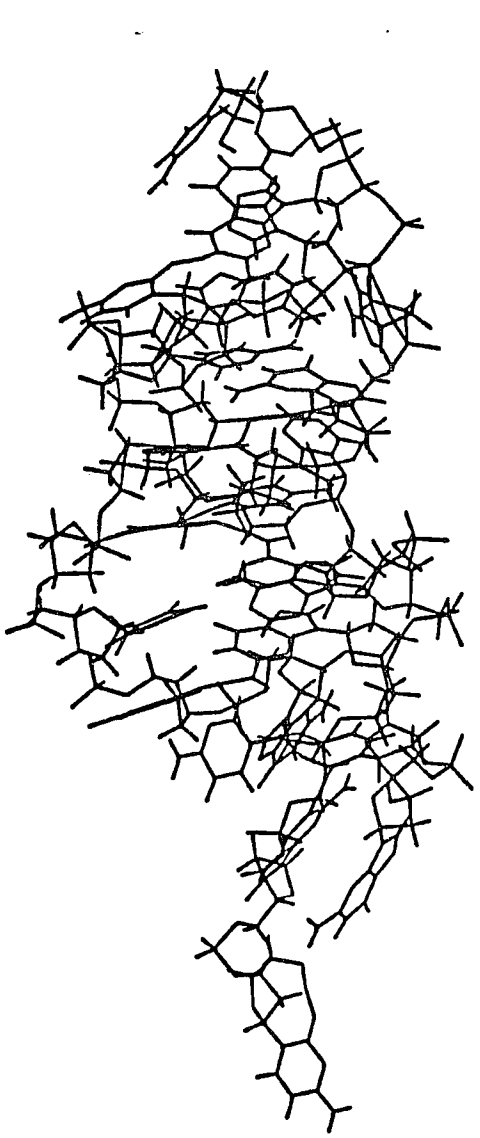
I II

B

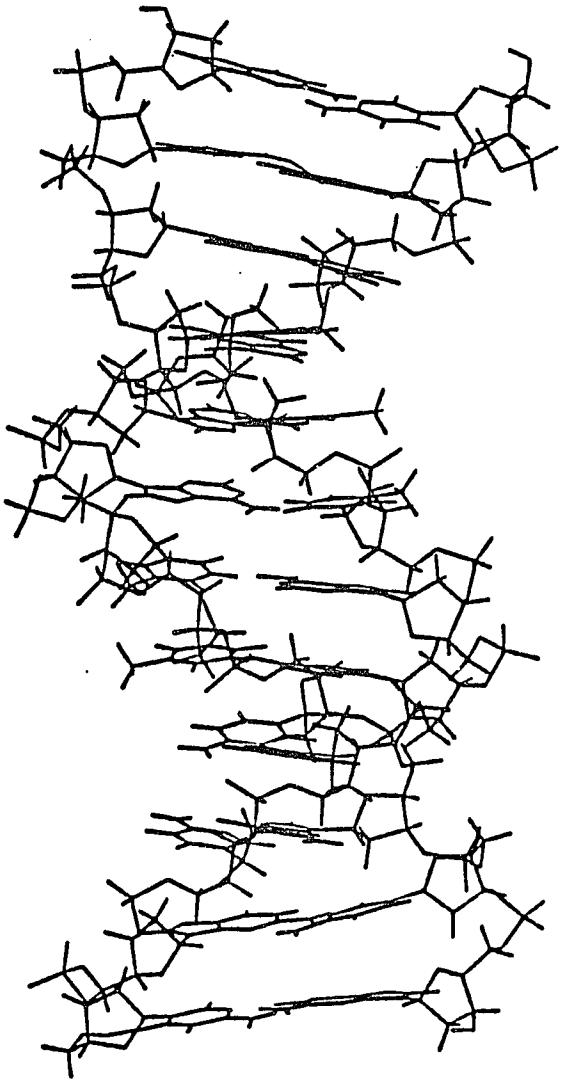


I II

C

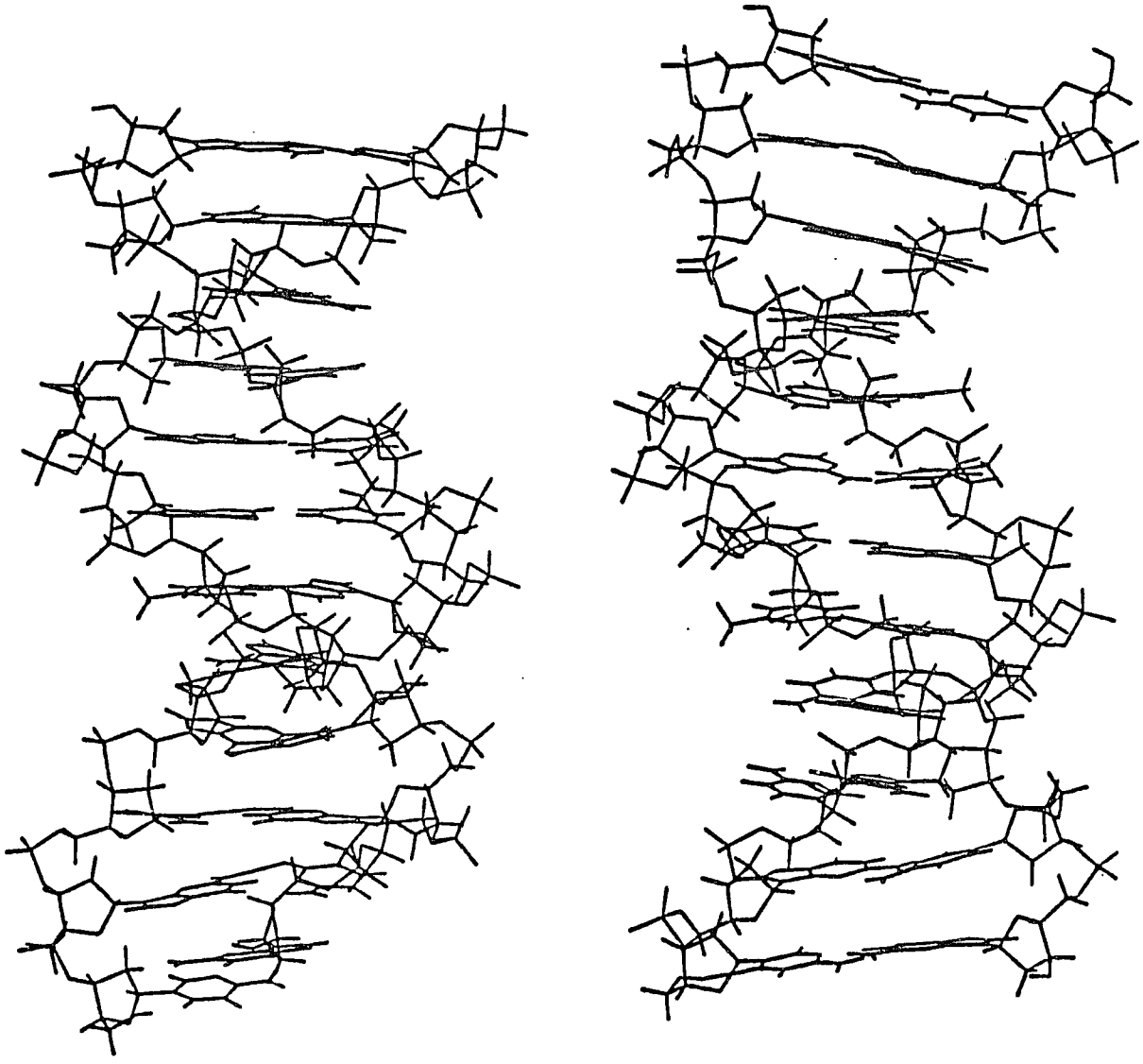


I



II

D



I II

E

Table 4.1 Root mean square differences (Å) between various combinations of structures for all atoms.

	<S>	<T>	<SFM>	<SRM>	<SFD>	<SRD>
<S>						
<T>	1.48					
<SFM>	0.98	1.64				
<SRM>	1.18	1.33	0.95			
<SFD>	4.31	4.62	3.95	4.09		
<SRD>	1.48	1.30	1.12	0.86	3.97	

Table 4.2 Calculated energy terms (Equation 2.35) for various structures. The NOE energy term corresponds to that calculated from Equations 4.3 and 4.4 with scale factor 3 for all interproton interactions employed in the refinement. The total energy and the energy gradient does not include the NOE energy term.

	Total Energy	Energy Grad.	Bond	Angle	Dihedral	Improper	VDW	Electrostatic	H-Bond	NOE
<S>	40.59	1.98	4.46	174.04	424.72	1.02	-292.56	-182.82	-88.28	8585.52
<T>	-22.71	1.44	7.91	188.87	383.82	3.03	-362.35	-172.24	-71.75	0.32
<SFM>	-313.64	0.32	7.29	196.54	294.31	9.74	-397.70	-317.58	-106.24	5761.93
<SRM>	-227.30	1.06	8.61	185.97	326.58	12.32	-396.55	-274.93	-89.30	20.75
<SFD>	-319.90	1.97	21.67	206.17	304.94	14.54	-424.05	-321.57	-121.58	20371.69
<SRD>	-202.01	1.34	10.44	181.97	325.70	7.73	-393.11	-250.09	-84.65	42.61

Table 4.3 CURVES-derived conformational parameter deviations (in degrees) for various structures.

	<S>-<T>	<SRM>-<T>	<SRD>-<T>
inclination	3.0	3.3	2.2
tip	3.3	2.8	3.5
opening	7.0	8.0	8.2
tilt	2.6	2.6	1.9
roll	6.4	4.3	3.8
buckle	8.1	9.5	4.7
propeller twist	16.6	9.5	4.5
twist	4.5	4.2	3.4
chi	26.4	8.6	5.8
gamma	15.9	13.7	14.4
delta	34.7	6.2	4.3
epsilon	44.6	16.5	11.8
zeta	38.8	14.6	13.9
alpha	16.3	22.1	22.1
beta	49.6	14.4	14.0
total torsion	34.5	14.4	13.4

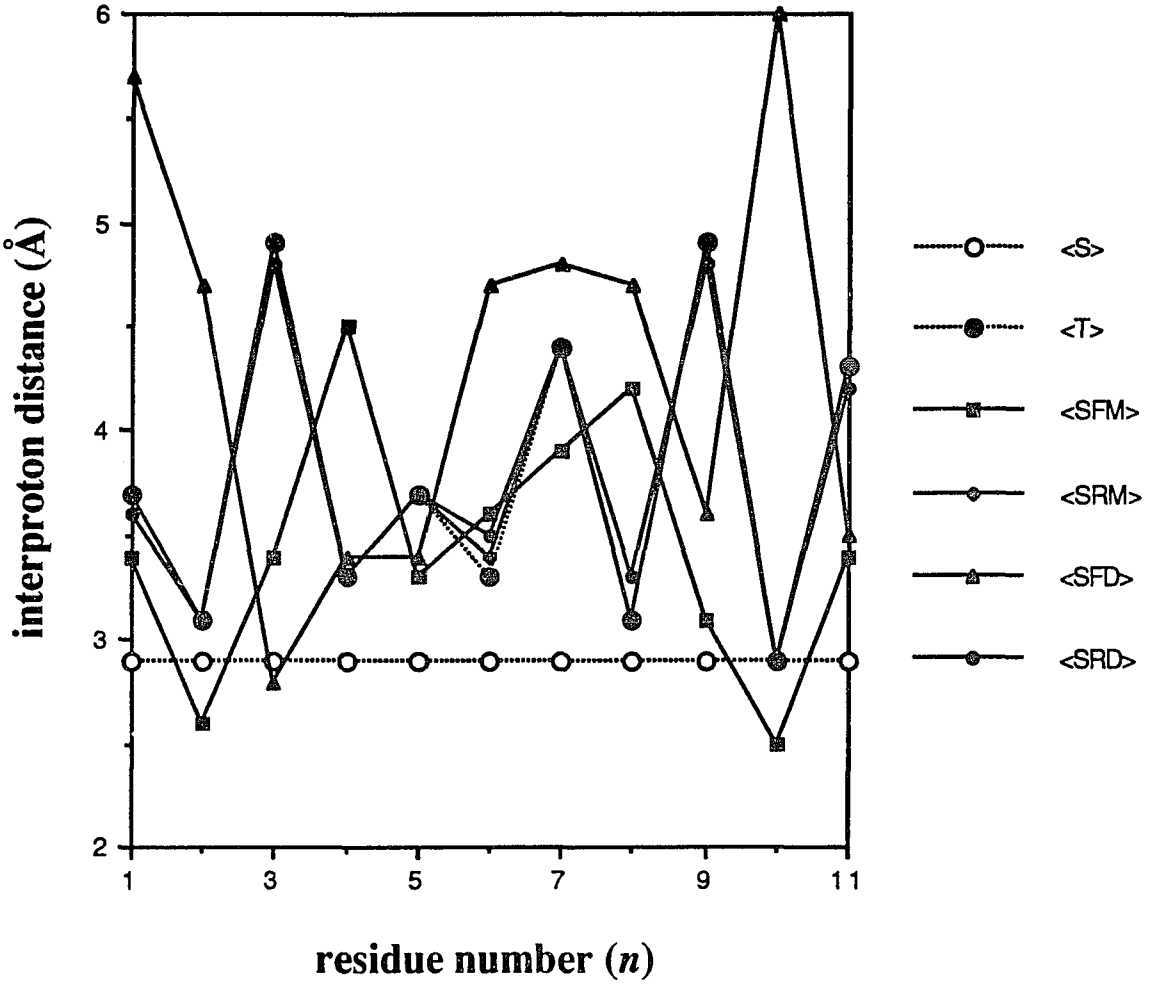
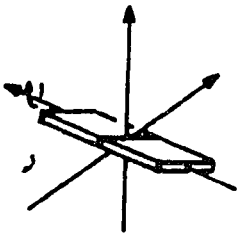
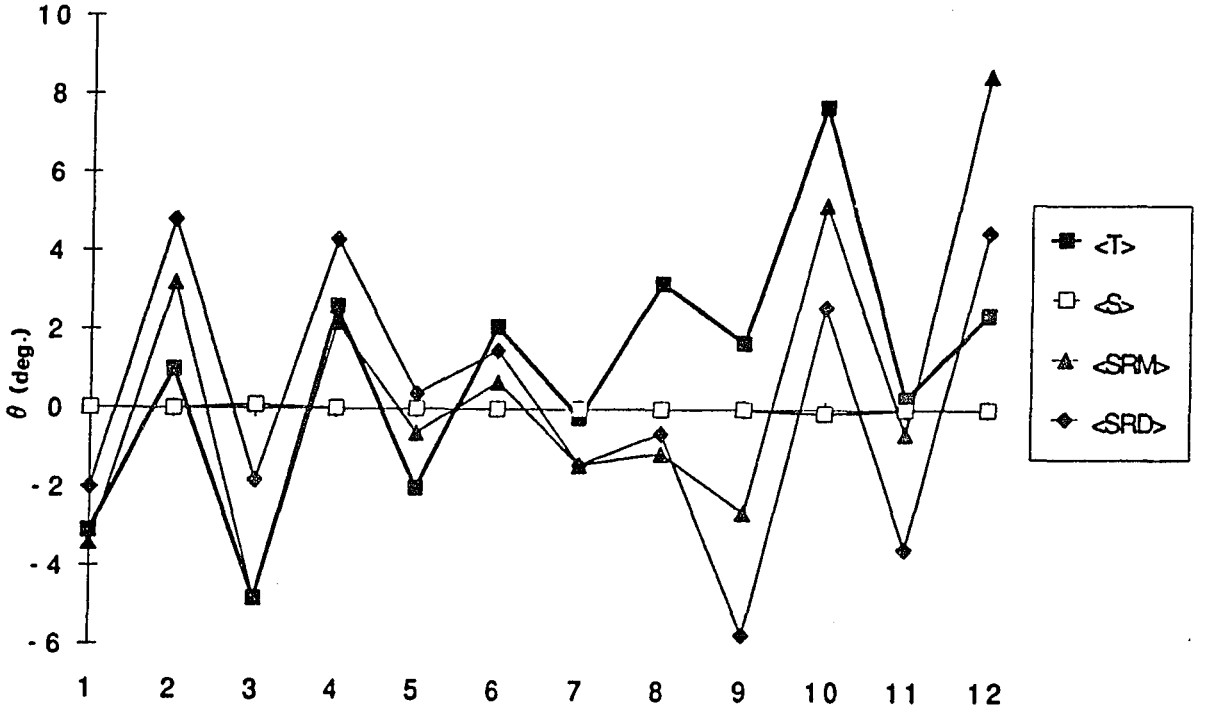


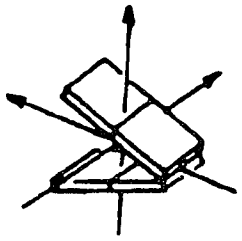
Figure 4.4 The sequential (*n*)H8/H6↔(*n+1*)H1' interproton distance as a function of residue number for various structures.

Figure 4.5 CURVES-derived conformational parameters for various structures as a function of either the base-pair (i.e., 1 is the C1:G24 base-pair) or the base-pair:base-pair interaction (i.e., 1 is the C1:G24-G2:C23 interaction).

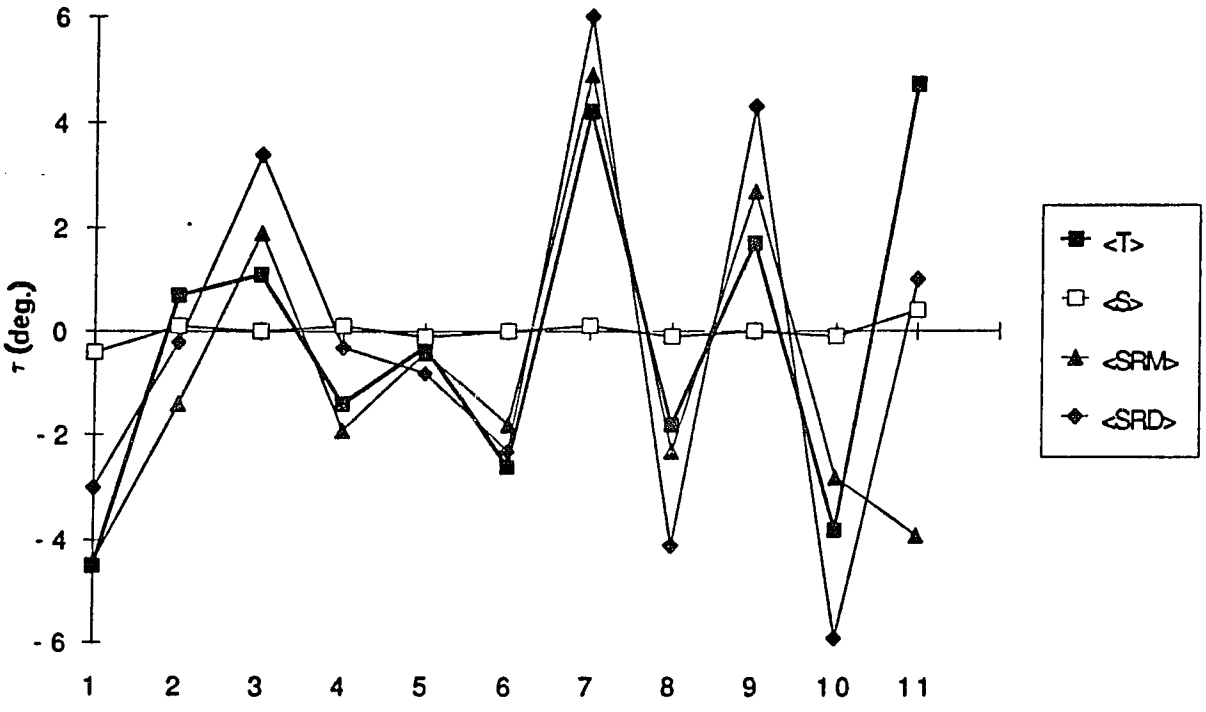


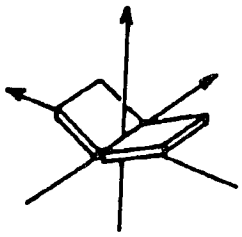
Tip (θ)



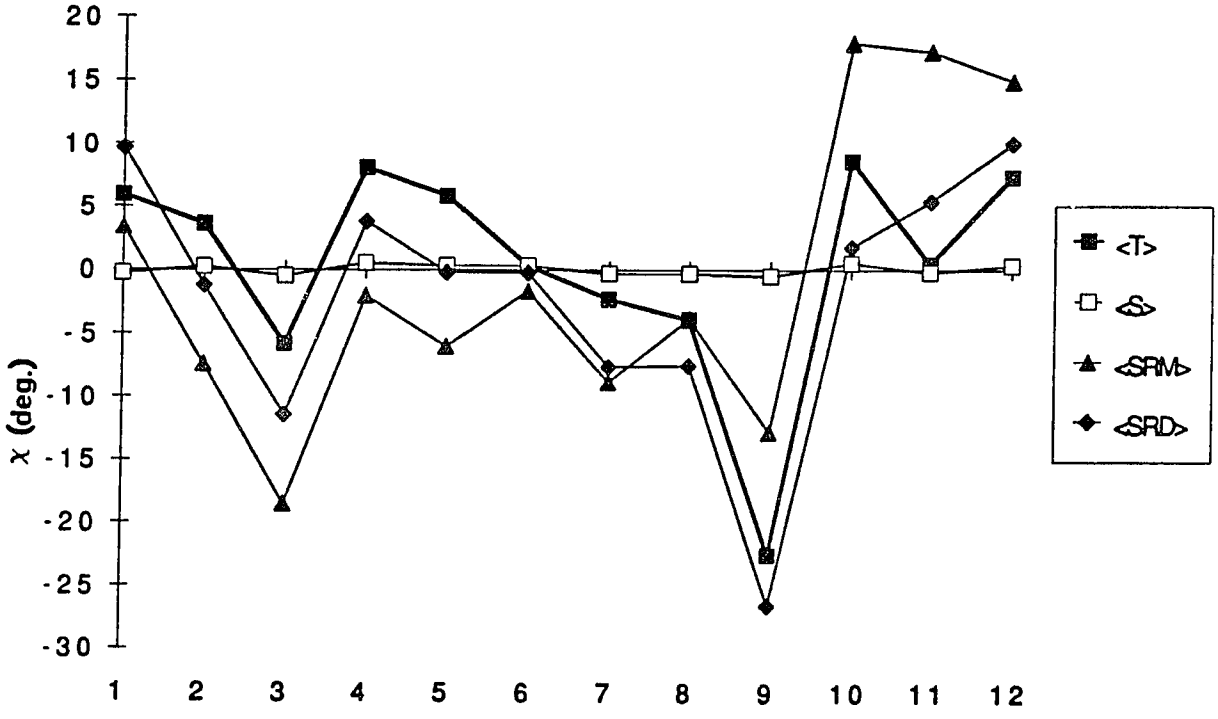


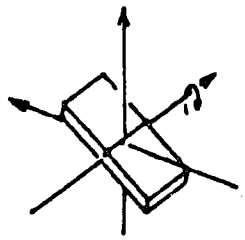
Till (τ)



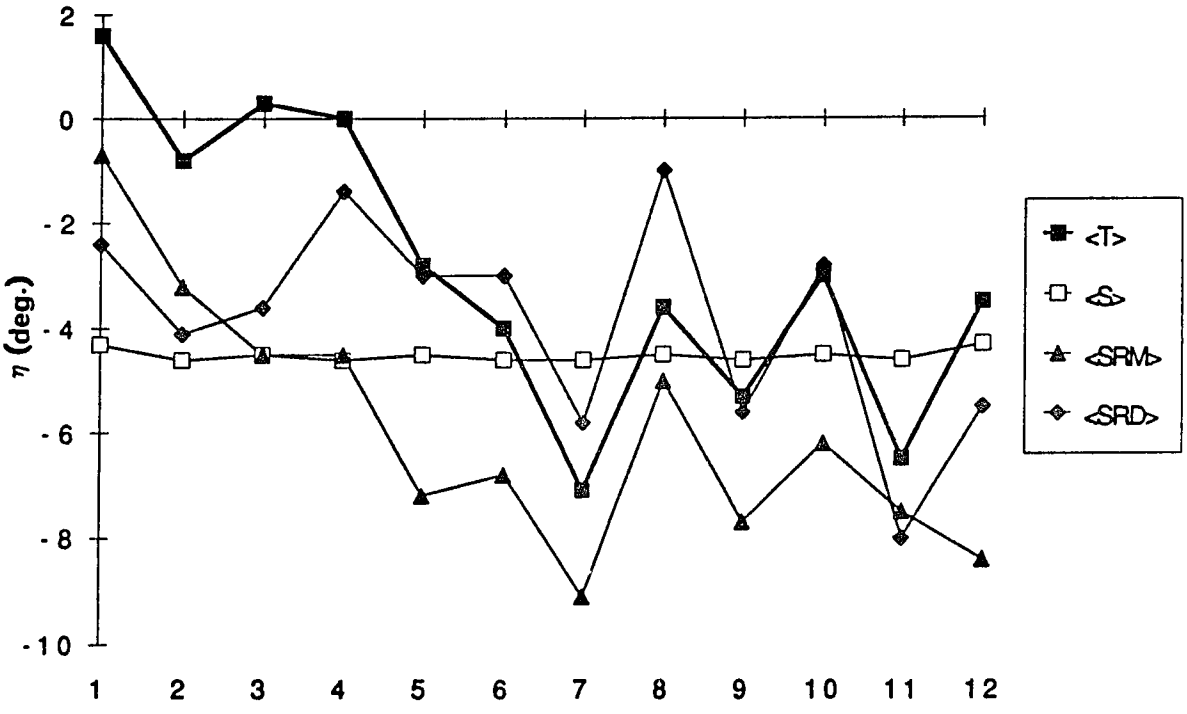


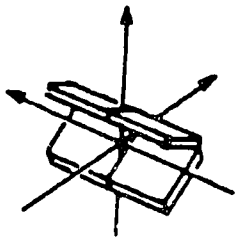
Buckle (x)



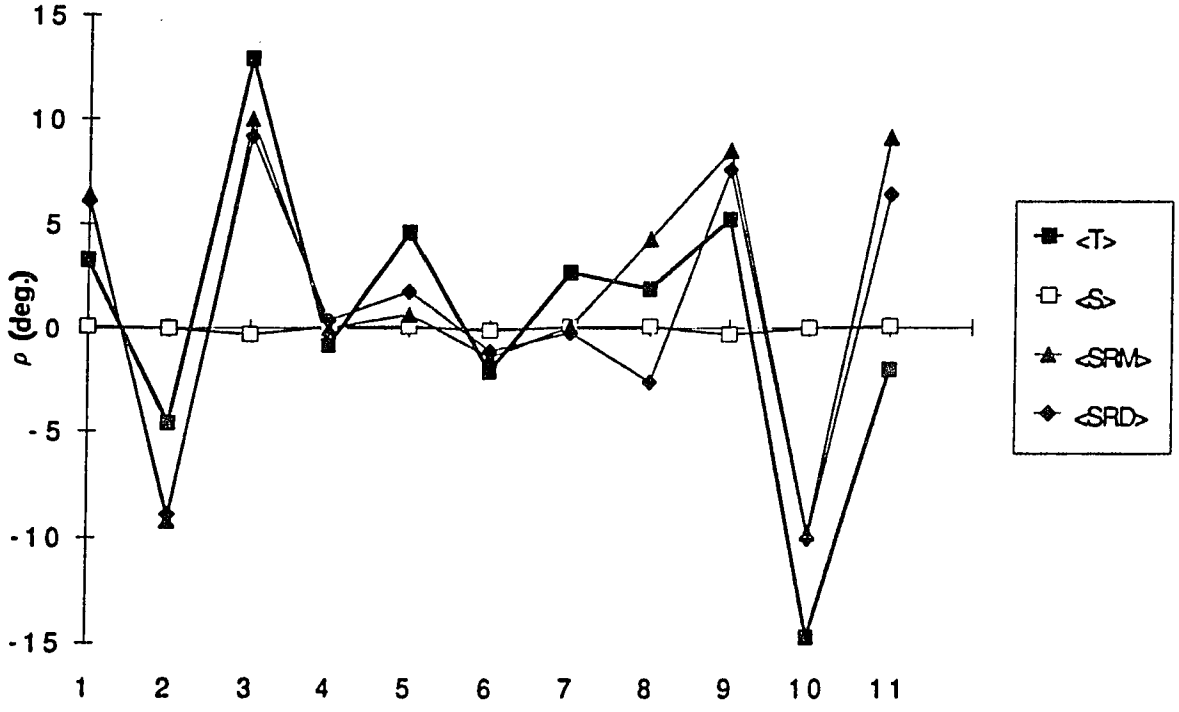


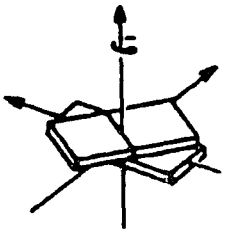
Inclination (η)



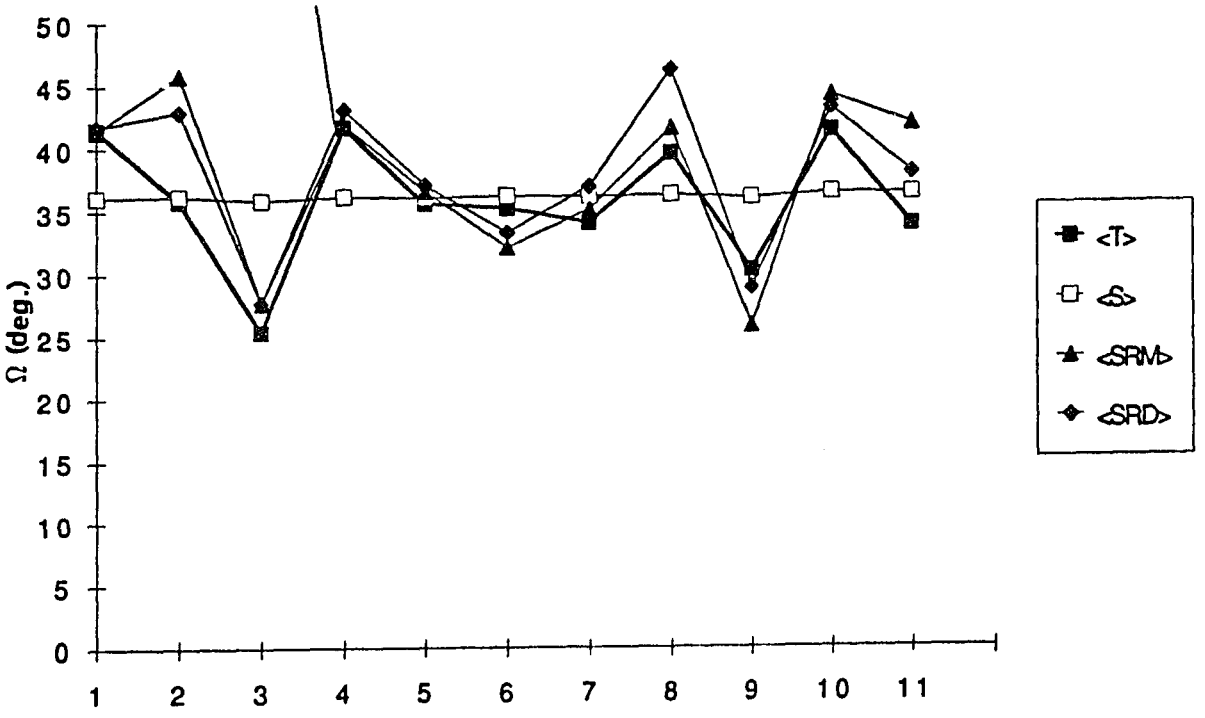


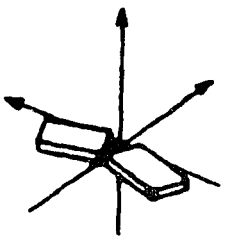
Roll (ρ)



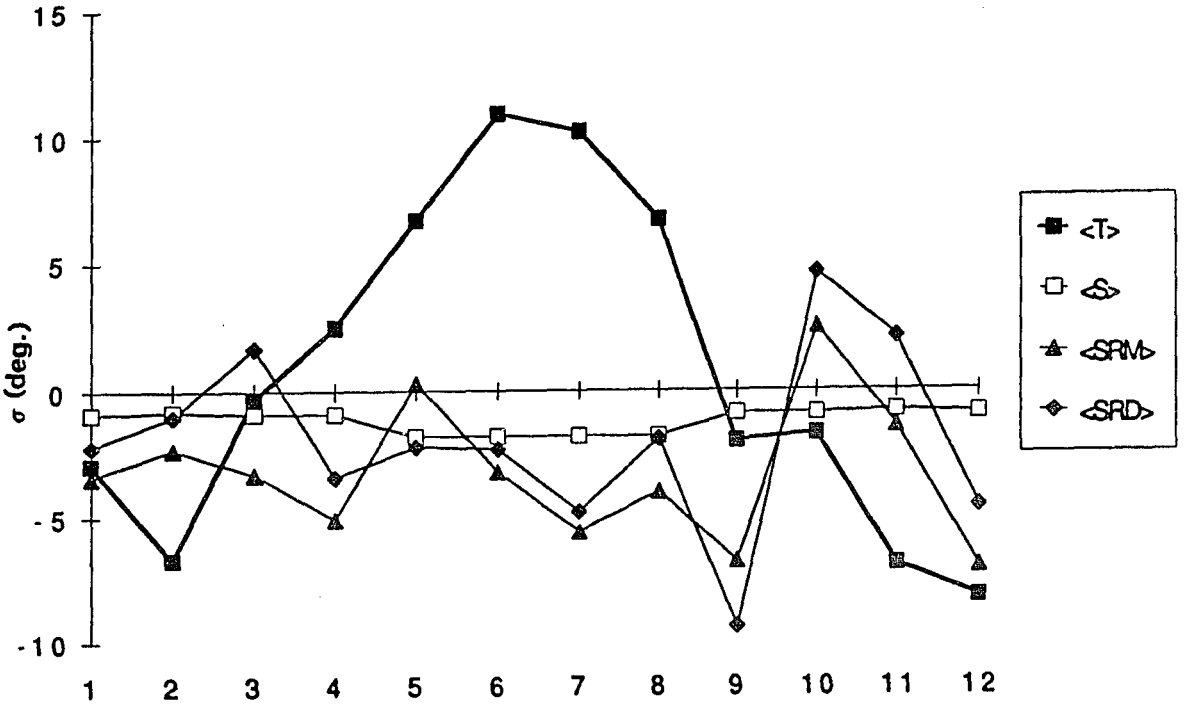


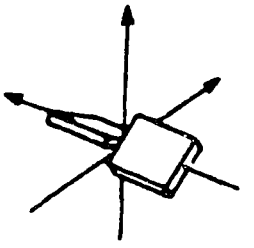
Twist (Ω)



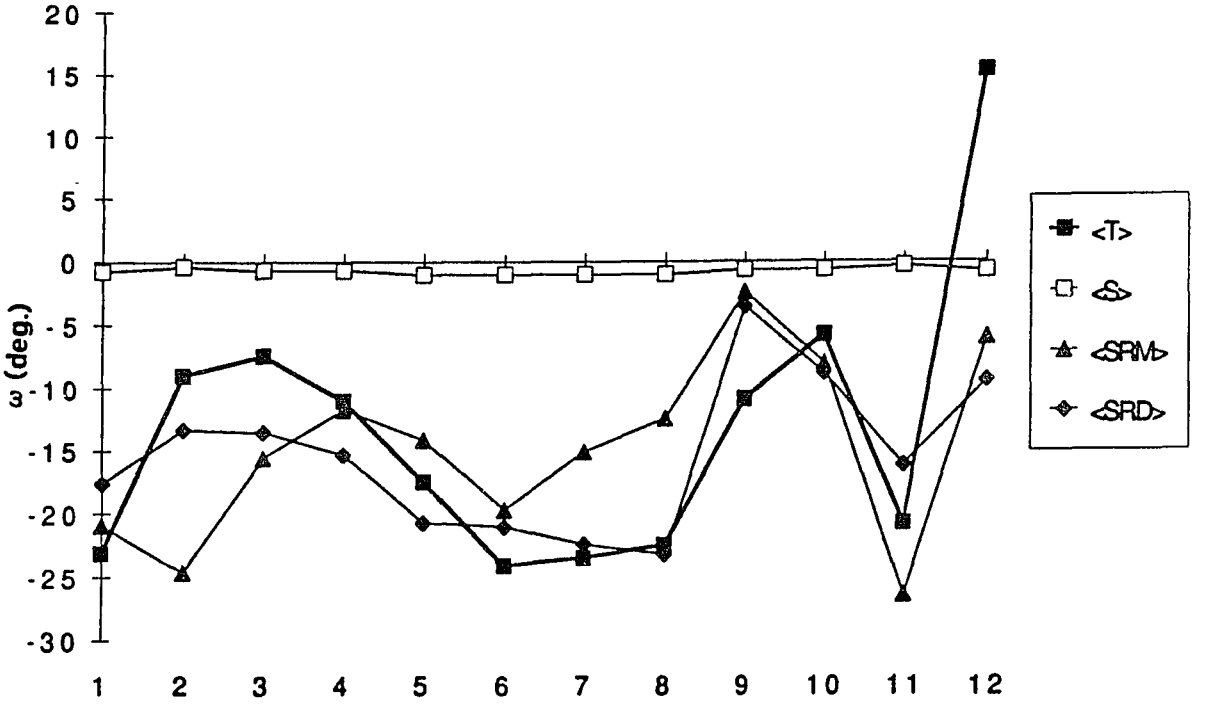


Opening (σ)





Propeller twist (ω)



CHAPTER 5 NMR Results and Discussion

5.1 Chemical Shift Assignments and Stereochemical Assignment of the Phosphotriester

All experiments were performed at 5°C, a temperature at which the self-complementary octamers are double-stranded (see Section 5.1.2). The experiments were performed under conditions of low salt (10 mM sodium phosphate) and neutral pH.

5.1.1 Sequential Assignment of the Nonexchangeable Protons using ^{1}H-{^{1}H} 2D COSY and 2D NOESY Connectivities and Stereochemical Assignment of the Phosphotriester Moiety using ^{1}H-{^{1}H} 2D NOESY Connectivities

Figure 5.1 shows the 5°C, one dimensional 400 MHz ^{1}H spectra of the self-complementary octamer {d[G1pG2pA3pA4pT5pT6pC7pC8]}$_2$ (referred to as the parent) and its two single-site phosphate ethylated analogues, ($\underline{R}_p$,$\underline{R}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ and ($\underline{S}_p$,$\underline{S}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ (referred to as R-A(et)A and S-A(et)A, respectively). Because the samples are dissolved in D_2O, only the nonexchangeable ^{1}H resonances are observed in these spectra. The resonances can be classified by type into the spectral ranges, 7.25-8.25 ppm (purine base H8 and H2 and pyrimidine base H6), 5.40-6.40 ppm (sugar H1' and pyrimidine base H5), 4.50-5.20 ppm (sugar H3'), 3.50-4.70 ppm (sugar H4', H5' and H5" and phosphate ethyl CH_2), 2.00-3.10 ppm (sugar H2' and H2") and 1.20-1.70 ppm (thymidine CH_3 and phosphate ethyl CH_3). The strategy for assigning the nonexchangeable proton resonances in these oligonucleotides is based on 2D NMR techniques that are well-established for the assignment of ^{1}H spectra of DNA oligomers.[15] The general method involves the determination of through-space interproton connectivities of < 5.0 Å from 2D NOESY spectra (Figure 5.2) and through-bond connectivities from 2D COSY spectra (Figure 5.3).[15]

In a right-handed DNA duplex, a direct NOE will be observed between the purine H8 or pyrimidine H6 (H8/H6) and its own H1' (r~3.8 Å for both A-DNA and B-DNA) in addition to the H1' of the (*n-1*) sugar (r~3.0 Å for B-DNA, r~4.0 Å for A-DNA). No NOE will be observed between the H8/H6 and the H1' of the (*n+1*) sugar (r>5.5 Å for A-DNA and B-

DNA). These NOE connectivities enable one to sequentially assign the base protons through their intervening sugar C1' protons as shown in Figure 5.4. Since the sequential through-space connectivities from the H1' to base proton are in the 5' to 3' direction, the 5' terminal base and the 3' terminal H1' are unique in exhibiting only intranucleotide NOEs and thus serve to define the 5' and 3' ends of the sequence.

Once the H1' chemical shift assignments were made, the H2' and H2" protons were assigned by means of their scalar coupling to the adjacent H1' (Figure 5.5). The different COSY contour patterns for the H1'↔H2' and H1'↔H2" connectivities stem from differences in the scalar coupling between the two C2' protons and the C1' proton. The size and shape of a COSY crosspeak depends on both the coupling pattern and the linewidth of each proton giving rise to the COSY connectivity.[78] Because the H1'↔H2' and H1'↔H2" crosspeak intensities are a function of the sugar pucker ($^3J_{1'2'} > {}^3J_{1'2"}$ for a C2'-*endo* conformation and $^3J_{1'2"} > {}^3J_{1'2'}$ for a C3'-*endo* conformation, see Figure 5.6), the distinction between the H2' and H2" signals cannot be determined by their COSY intensity to the H1' proton unless there is prior knowledge regarding the deoxyribose conformation.[88,89] The distinction between H2' and H2" can be resolved by the direct NOE crosspeak intensities between the H1' and its neighboring H2' and H2"; any given H1' is nearer to its H2" than to its H2' for all deoxyribose pseudorotation phase angles and hence will show a stronger NOE crosspeak (Figure 5.7).[15]

The H3' protons could not be assigned by means of a COSY connectivity to the C2' protons because of the lack of crosspeaks in the H3'↔H2',H2" spectral region. This is due to the fact that the components of a COSY crosspeak have anti-phase character and tend to cancel each other when either the resolution is not sufficient to resolve the J separation between them or the linewidths of the components are significantly larger than the J separation.[78] Although the 3-bond coupling constant, $^3J_{3'2'}$, varies within a narrow range of 6-10 Hz (Figure 5.6) the linewidth of the H3' protons for these oligonucleotides at 5°C is apparently large enough to considerably reduce the COSY crosspeak intensities for these connectivities. Hence, the H3' chemical shift was determined by means of a NOESY

connectivity (Figure 5.8) to the adjacent H2' proton (r~2.2-2.4 Å for all pseudorotation phase angles) and the adjacent H2" proton (r~2.6-3.0 Å for all pseudorotation phase angles). This type of connectivity enabled the unambiguous identification of all the H3' chemical shifts.

The H4' chemical shift can be assigned by means of a NOESY connectivity (Figure 5.9) to the adjacent H3' proton (r~2.4-2.9 Å for all pseudorotation phase angles). As shown in Figure 5.9, some of the H3'-H4' connectivities (G2-H3'↔G2-H4' and T6-H3'↔T6-H4' for the parent octamer) are absent, probably resulting from the saturation of signals from protons (G2-H3' and T6-H3') that resonate near the residual HOD solvent signal (the residual HOD signal is suppressed by presaturation). The H3'↔H4' COSY connectivity could not be used to assign the H4' protons because of the lack of crosspeaks in this spectral region. However, the H4' assignments could be confirmed by means of a NOESY connectivity (Figure 5.10) to the adjacent H2' proton (r~3.6 Å for all pseudorotation phase angles) and the adjacent H2" proton (r~3.8 Å for C2'-*endo* sugar pucker, r~2.5 Å for C3'-*endo* sugar pucker).

The H5' and H5" chemical shifts were not assigned. Most of the H5' and H5" protons are poorly resolved and overlap with the chemical shifts of the H4' protons within the 3.5 to 4.7 ppm spectral range (Figure 5.1). The lack of chemical shift dispersion for the H4', H5' and H5" protons prohibits the use of a COSY or a NOESY connectivity between these protons as a means of assigning the C5' protons. 2D TOCSY experiments were performed in an attempt to assign the C5' protons by the multiple transfer of correlations from the C3', C2' or C1' protons. Only a few transferred correlations were observed from the C1', C2' and C3' protons to the C5' protons in the 70 ms spin-lock TOCSY spectrum. Because 2D NOESY crosspeaks to the C5' spectral region were extensively overlapped and could not be accurately quantitated, conditions were not sought (e.g., by varying the length of the spin-lock period in the TOCSY experiment) that would have enabled the complete assignment of the H5' and H5" protons. These protons were thus excluded from further analysis.

The nonexchangeable base protons (A-H2, C-H5 and T-CH_3) were assigned as follows: the C-H5 was assigned by its COSY connectivity to the adjacent C-H6 proton, the T-CH_3 was assigned by its NOESY connectivity to the adjacent T-H6 proton, and the A-H2

was assigned by a NOESY connectivity to its (*n+1*) H1' proton (r~4.5 Å for right-handed helices).[15]

The phosphate ethyl group signals for the S-A(et)A and R-A(et)A analogues were assigned as follows: the phosphate ethyl methyl ($POCH_2\boldsymbol{CH_3}$) signal was assigned as the additional signal appearing in the methyl region within the 1.20-1.70 ppm spectral range and the phosphate ethyl methylene ($PO\boldsymbol{CH_2}CH_3$) was assigned by its COSY connectivity to the adjacent $POCH_2\boldsymbol{CH_3}$ group.

The chemical shifts for all the assigned nonexchangeable protons of the parent, S-A(et)A and R-A(et)A are given in Table 5.1.

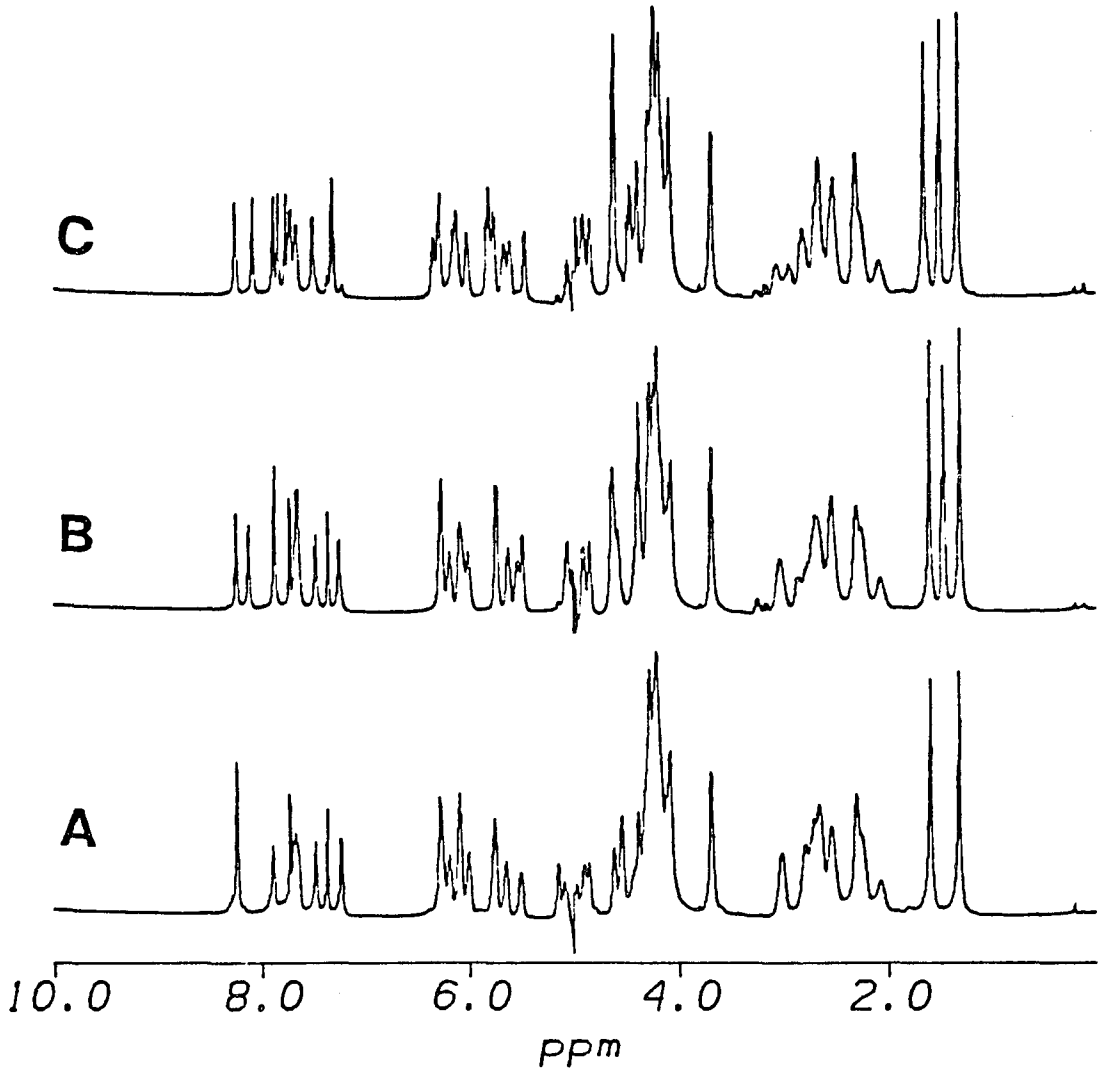


Figure 5.1 400 MHz ^{1}H NMR spectra of the nonexchangeable protons in (A) {d[GGAATTCC]}$_2$, (B) ($\underline{S}_p$,$\underline{S}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ and (C) ($\underline{R}_p$,$\underline{R}_p$)-{d[GGA(ethyl)ATTCC]}$_2$. Sample conditions: 5°C; in D_2O with 10 mM sodium phosphate, pH 7.0. These chemical shifts are referenced relative to TSP and are based on a 5.00 ppm chemical shift for HOD from TSP at 5°C.

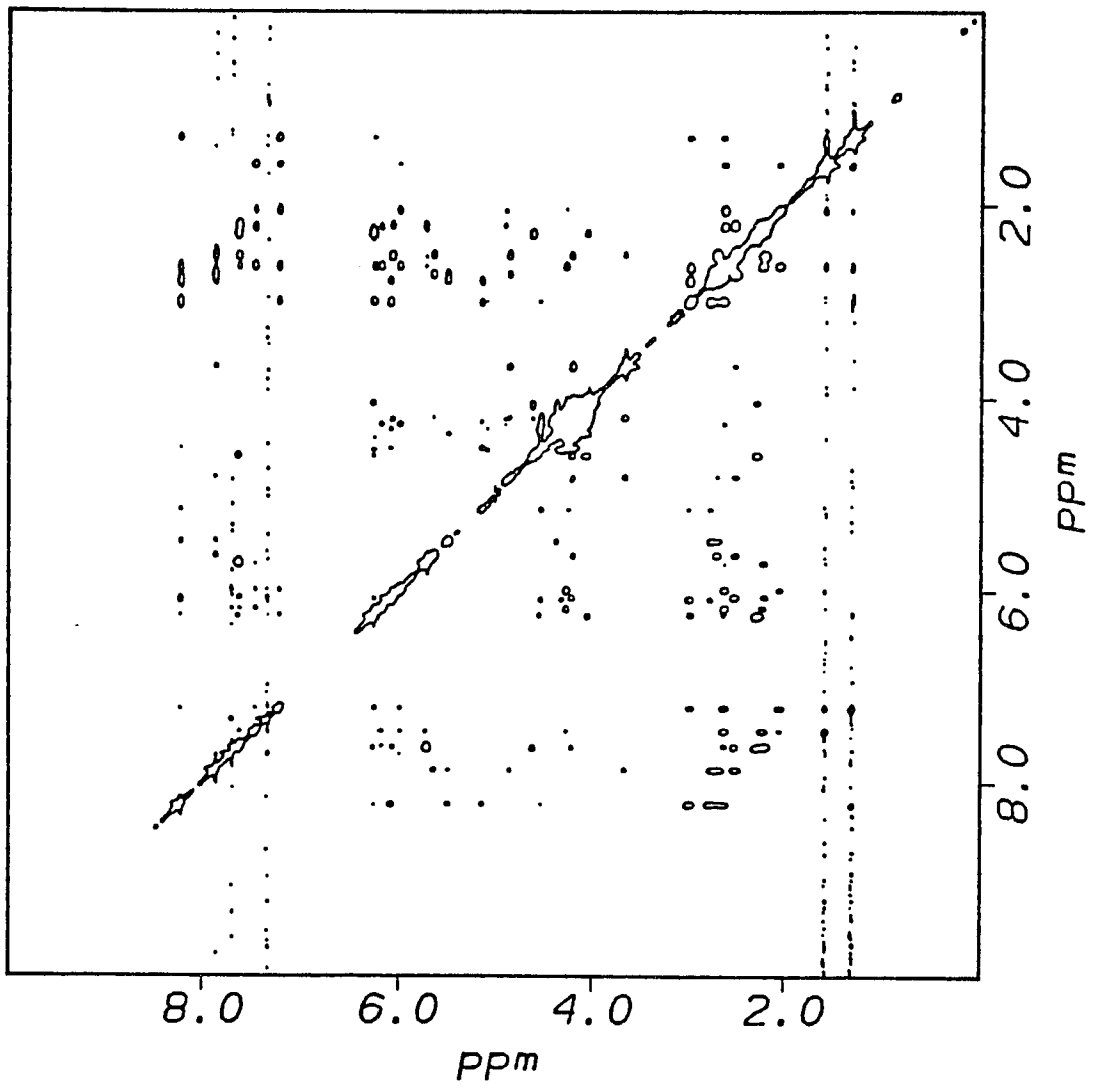


Figure 5.2 A contour plot of the 400 MHz phase-sensitive ^{1}H-{^{1}H} 2D NOESY spectrum (mixing time 200 ms) of (d[GGAATTCC])$_2$ at 5°C.

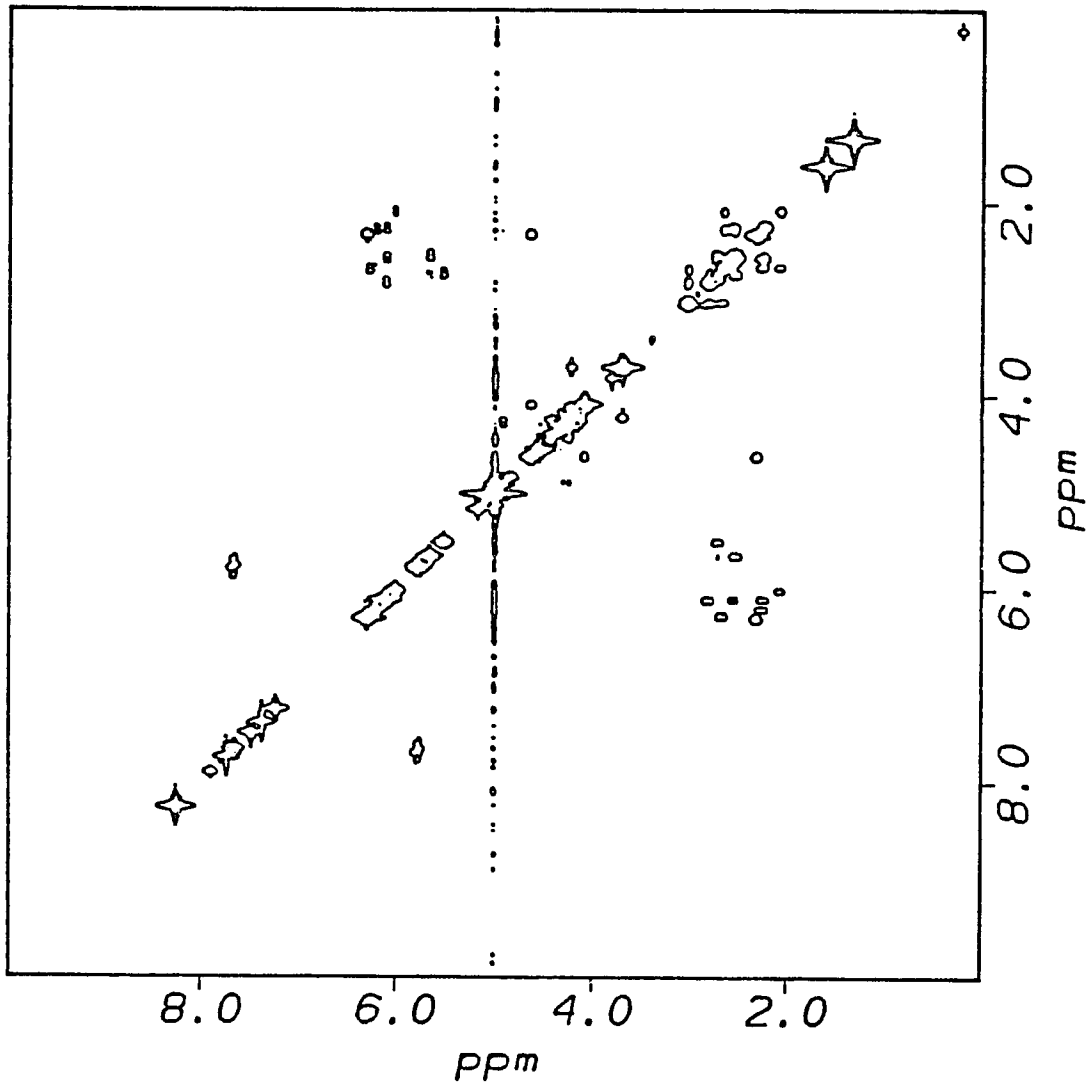


Figure 5.3 A contour plot of the 400 MHz absolute-value ^1H-{^1H} 2D COSY spectrum of {d[GGAATTCC]}$_2$ at 5°C.

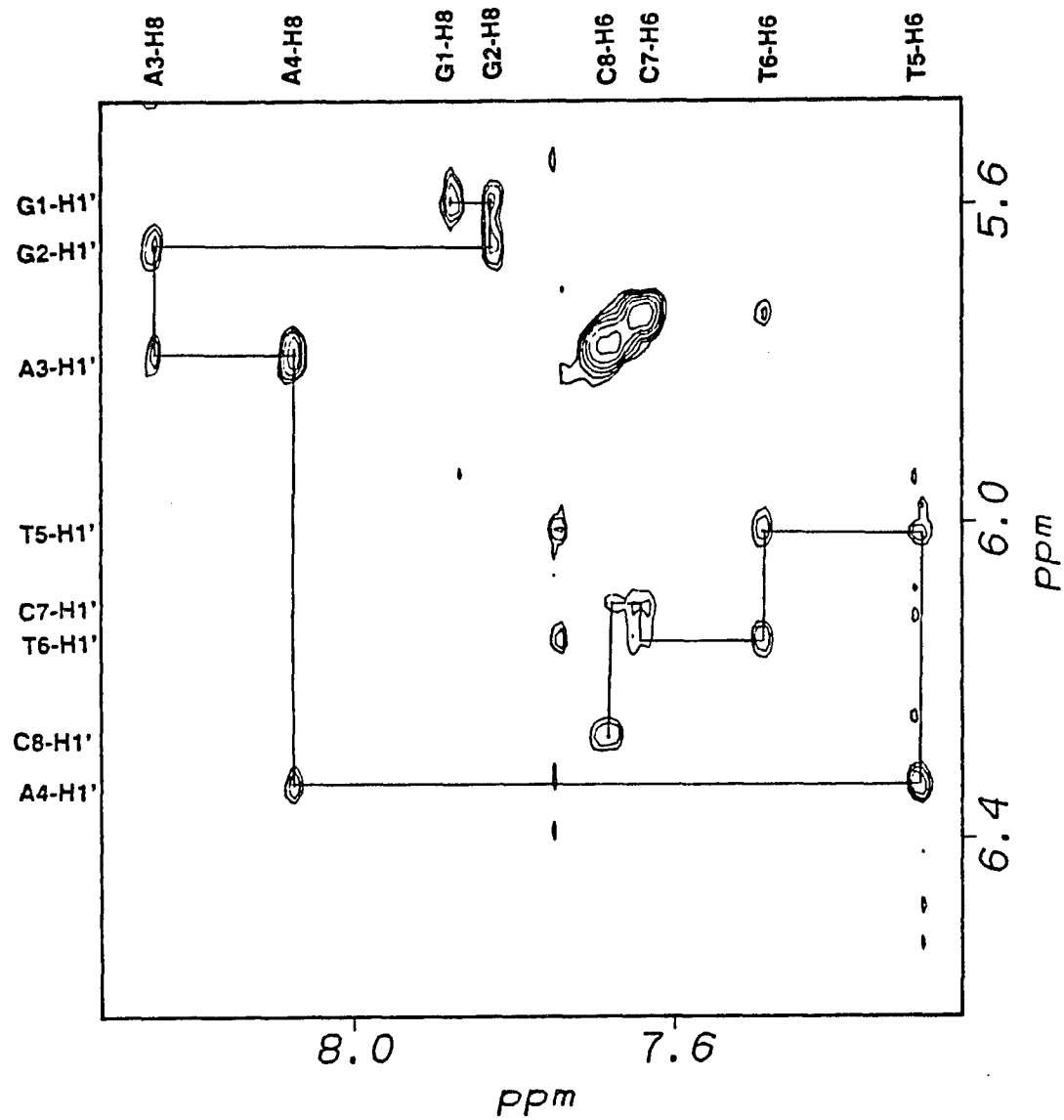


Figure 5.4 An expanded contour plot of the phase-sensitive 2D NOESY spectrum (mixing time 200 ms) of ($\underline{R}_p$,$\underline{R}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ at 5°C correlating the H8/H6 base protons (7.0-8.5 ppm) with the sugar H1' protons (5.1-6.5 ppm). The sequential (*n*)H1'↔(*n*)H8/H6↔(*n+1*)H1' connectivities are shown as solid lines (-).

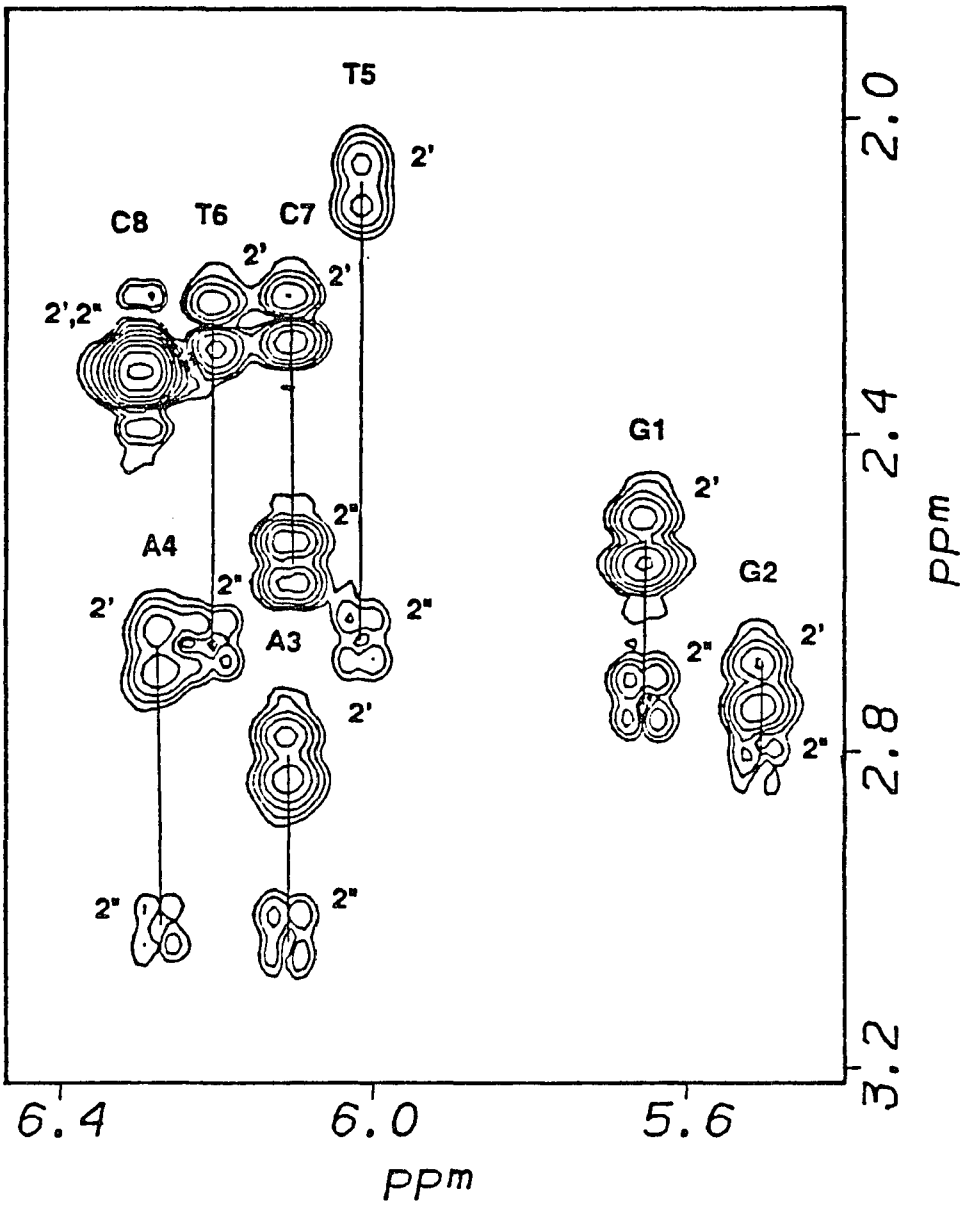


Figure 5.5 An expanded contour plot of the absolute-value 2D COSY spectrum of $\{d[GGAATTCC]\}_2$ at 5°C correlating the sugar H1' protons (5.4-6.4 ppm) with the sugar H2',H2" protons (1.9-3.1 ppm). Each sugar H2' and H2" assignment is given by a vertical line (-) that connects crosspeaks representing H1'↔H2' and H1'↔H2" connectivities within the same residue.

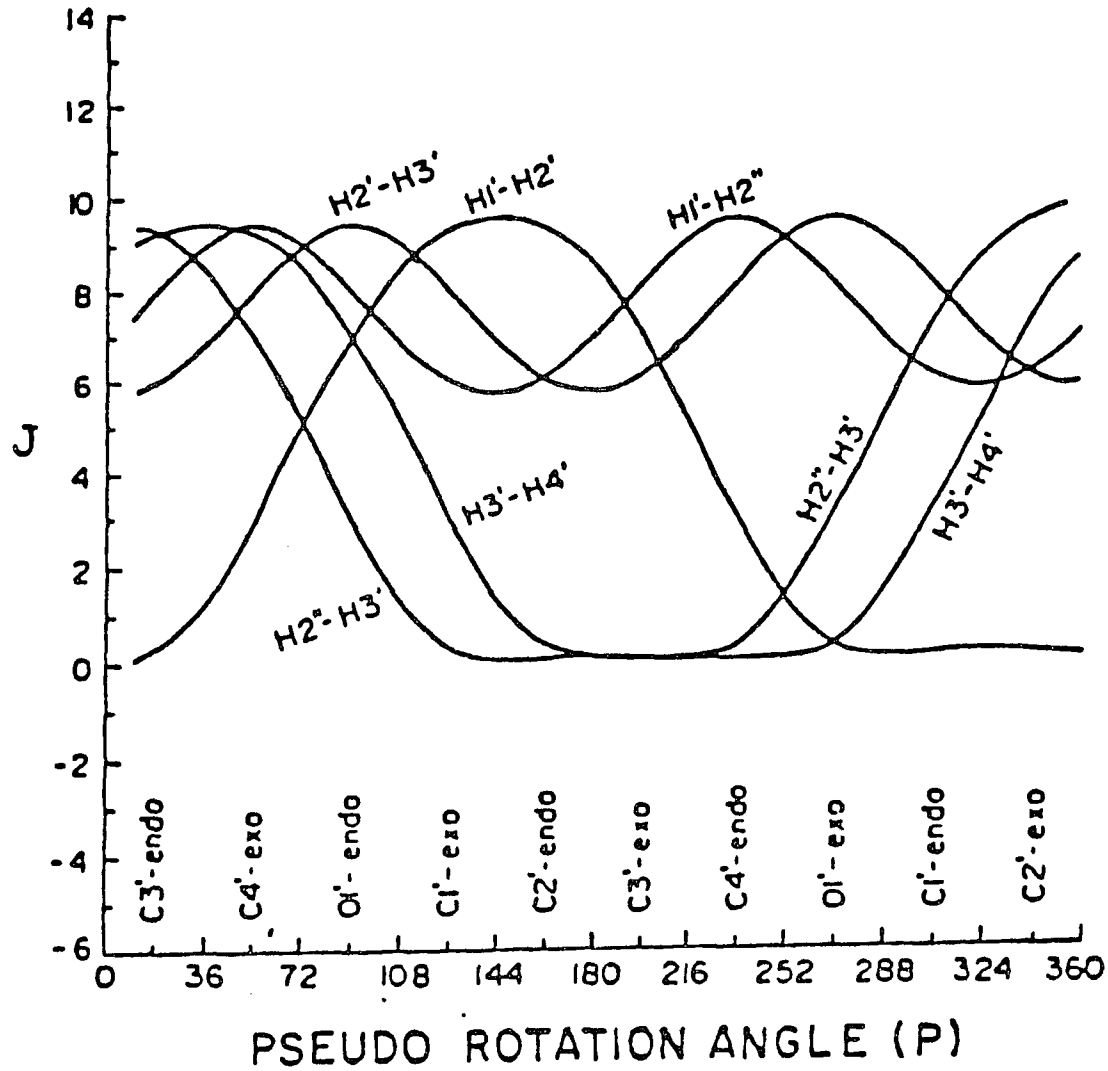


Figure 5.6 Graph showing the dependence of the 3-bond H1'-H2', H1'-H2", H2'-H3', H2"-H3' and H3'-H4' coupling constants on the pseudorotation phase angle, P. Each coupling constant is related to the relevant H-C-C-H dihedral angle ϕ, according to the relation[90]:

$$J = 10.2\cos^2\theta - 0.8\cos\theta$$

The positions of the classical sugar puckers are also indicated in the figure. Figure taken from reference 91.

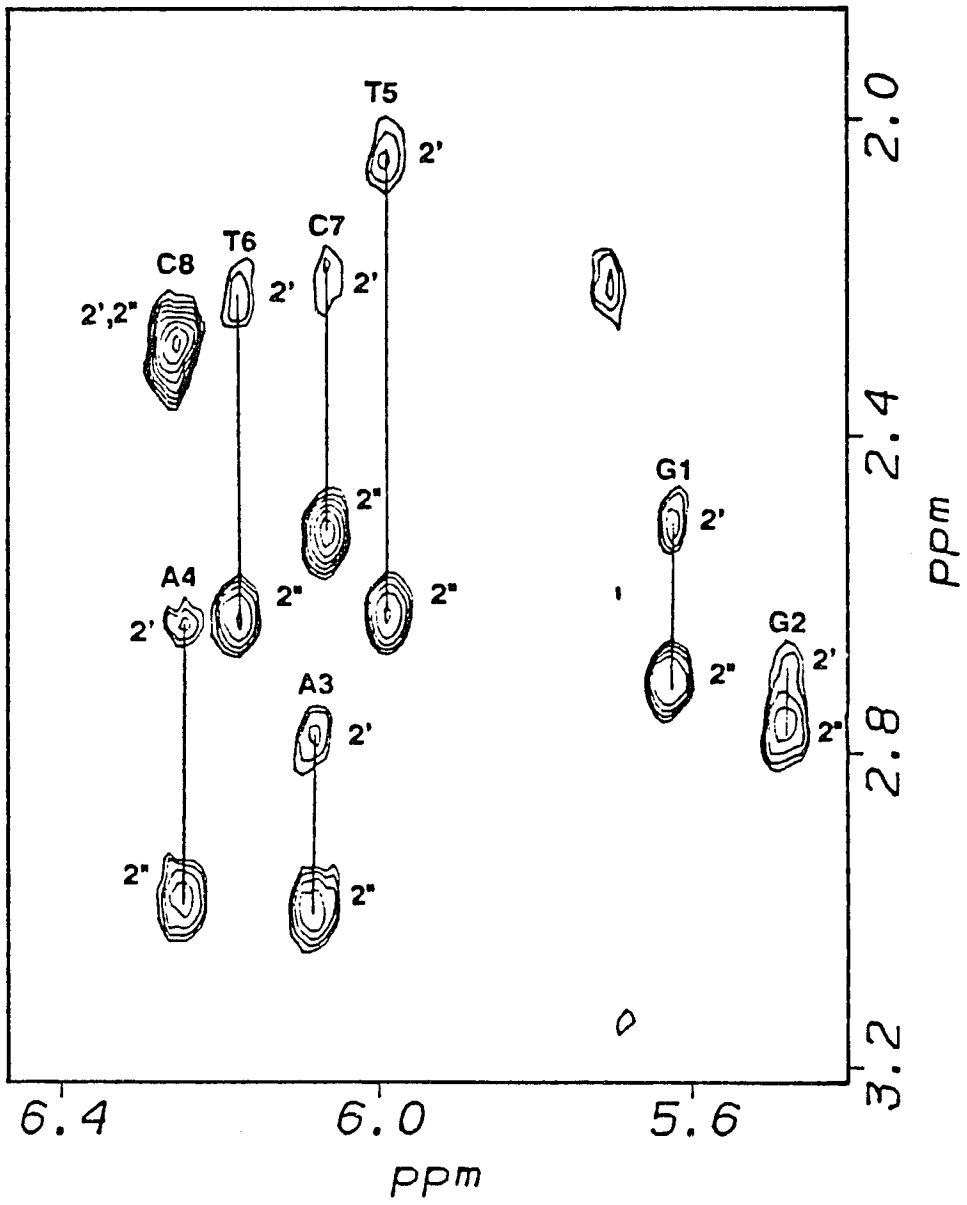


Figure 5.7 An expanded contour plot of the phase-sensitive 2D NOESY spectrum (mixing time 100 ms) of $\{d[GGAATTCC]\}_2$ at 5°C correlating the sugar H1' protons (5.4-6.4 ppm) with the sugar H2',H2" protons (1.9-3.1 ppm). Each sugar H2' and H2" assignment is given by a vertical line (-) that connects crosspeaks representing H1'↔H2' and H1'↔H2" connectivities within the same residue.

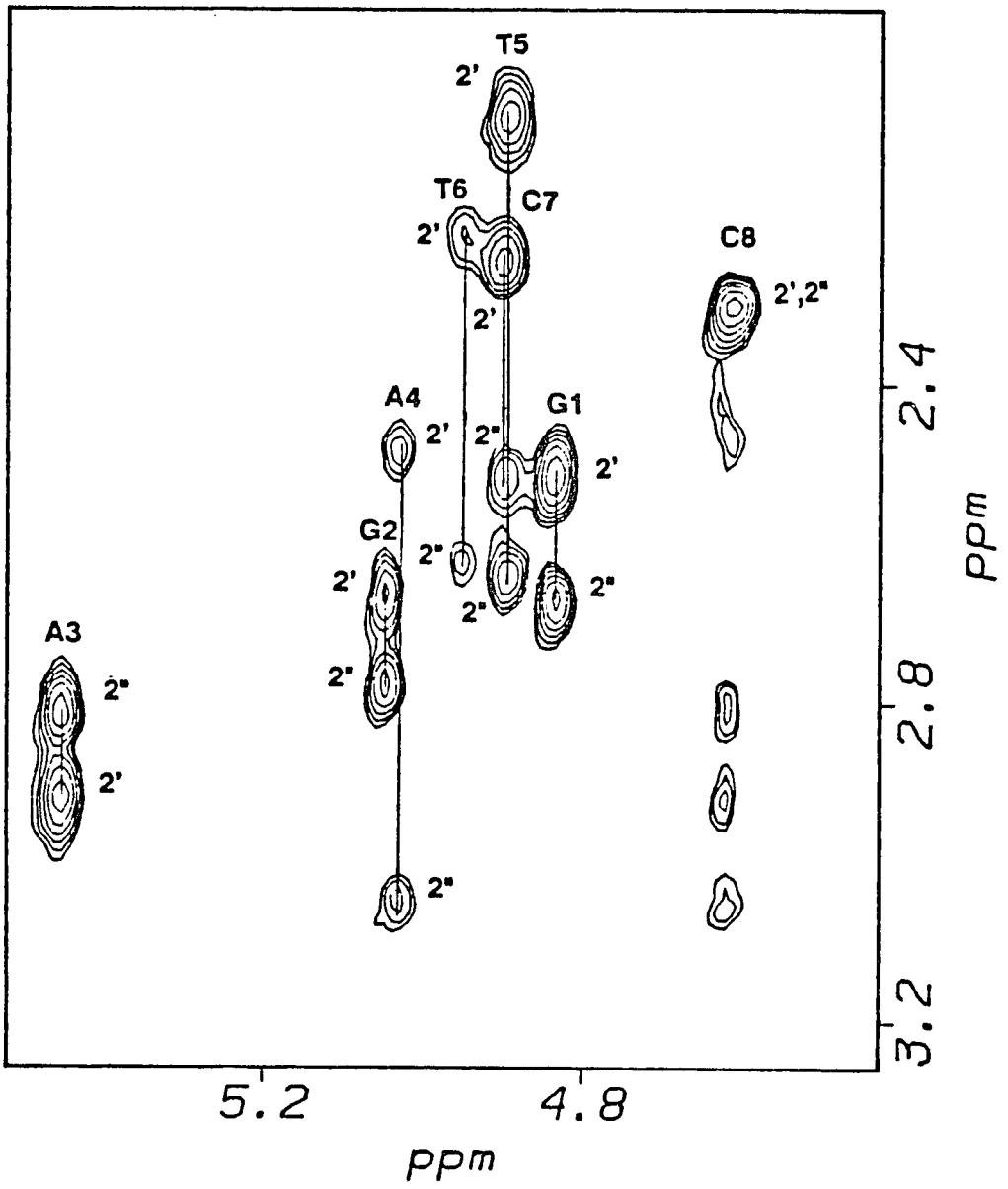


Figure 5.8 An expanded contour plot of the phase-sensitive 2D NOESY spectrum (mixing time 200 ms) of ($\underline{R}_p$,$\underline{R}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ at 5°C correlating the sugar H3' protons (4.5-5.5 ppm) with the sugar H2',H2" protons (1.9-3.1 ppm). Each sugar H3' assignment is given by a vertical line (-) that connects crosspeaks representing H3'↔H2' and H3'↔H2" connectivities within the same residue.

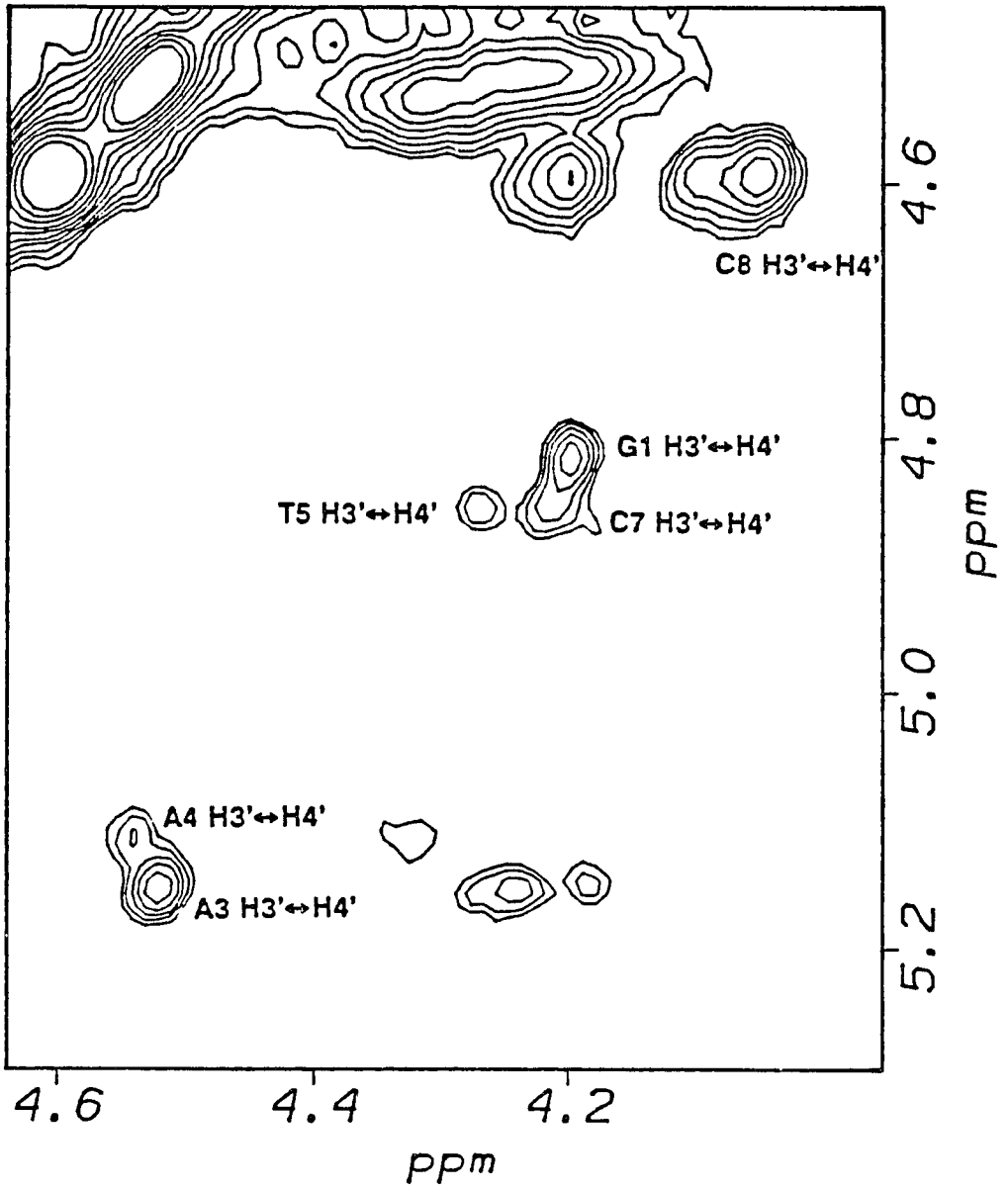


Figure 5.9 An expanded contour plot of the phase-sensitive 2D NOESY spectrum (mixing time 200 ms) of $\{d[GGAATTCC]\}_2$ at 5°C correlating the sugar H4' protons (4.0-4.8 ppm) with the sugar H3' protons (4.5-5.2 ppm). The sugar H3' and H4' assignments are depicted next to the contour peaks.

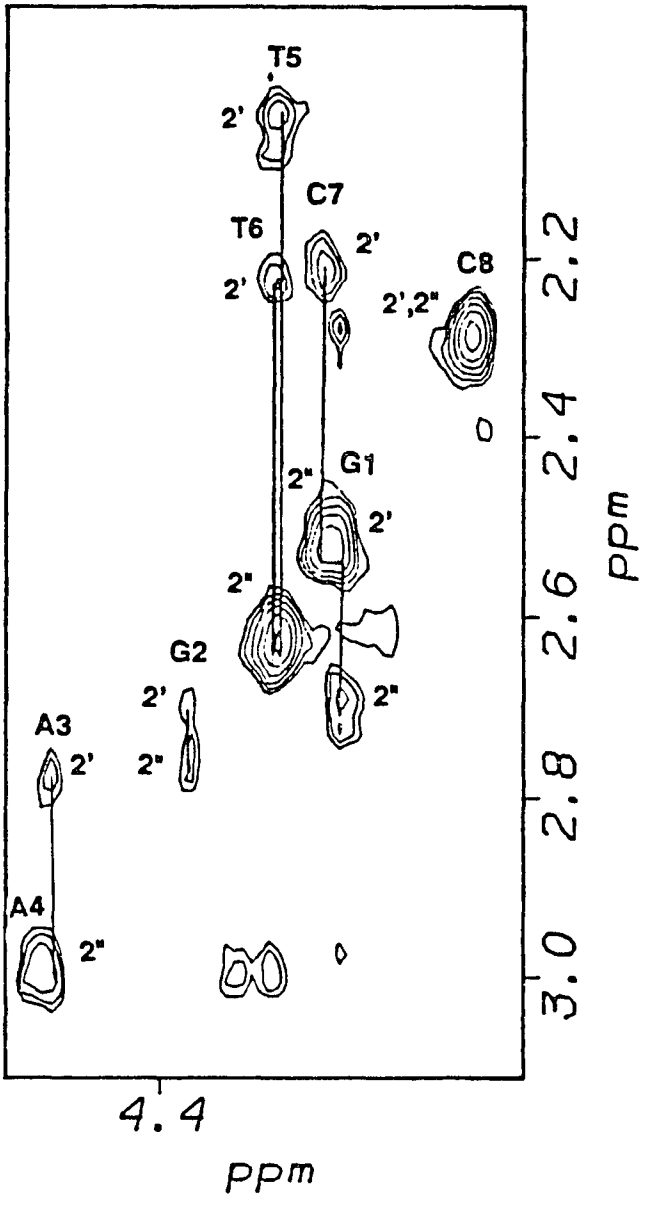


Figure 5.10 An expanded contour plot of the phase-sensitive 2D NOESY spectrum (mixing time 200 ms) of $\{d[GGAATTCC]\}_2$ at 5°C correlating the sugar H4' protons (4.0-4.7 ppm) with the sugar H2',H2" protons (1.9-3.1 ppm). Each sugar H4' assignment is given by a vertical line (-) that connects crosspeaks representing H4'↔H2' and H4'↔H2" connectivities within the same residue.

Table 5.1 Chemical shifts of the nonexchangeable protons of {d[GGAATTCC]}$_2$ (**P**), ($\underline{S}_p$,$\underline{S}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ (**S**) and ($\underline{R}_p$,$\underline{R}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ (**R**). Sample conditions: 5°C; in D_2O with 10 mM sodium phosphate, pH 7.0. These chemical shifts are referenced relative to TSP and are based on a 5.00 ppm chemical shift for HOD from TSP at 5°C.

		8|6	2|5|m	1'	2'	2"	3'	4'	PCH2	PCH3
	P	7.86		5.63	2.51	2.71	4.83	4.19		
G1	S	7.85		5.62	2.52	2.71	4.83	4.20		
	R	7.87		5.61	2.53	2.68	4.83	4.19		
	P	7.86		5.48	2.71	2.78	5.04	4.38		
G2	S	7.85		5.52	2.71	2.76	5.06	4.36		
	R	7.82		5.65	2.68	2.78	5.04	4.40		
	P	8.22	7.32	6.08	2.77	3.00	5.14	4.52		
A3	S	8.22	7.32	6.05	3.00	2.84	5.47	4.64		
	R	8.24	7.30	5.78	2.92	2.81	5.45	4.62		
	P	8.22	7.69	6.24	2.64	3.00	5.08	4.55		
A4	S	8.12	7.70	6.27	2.54	3.03	5.07	4.59		
	R	8.08	7.74	6.33	2.49	3.05	5.06	4.62		
	P	7.21	1.32	5.99	2.04	2.63	4.89	4.28		
T5	S	7.23	1.31	6.01	2.06	2.65	4.89	4.28		
	R	7.29	1.32	6.01	2.07	2.65	4.90	4.29		
	P	7.45	1.59	6.17	2.23	2.63	4.94	4.28		
T6	S	7.45	1.60	6.17	2.22	2.65	4.97	4.28		
	R	7.50	1.65	6.14	2.24	2.64	4.95	4.25		
	P	7.61	5.71	6.06	2.21	2.52	4.89	4.22		
C7	S	7.61	5.71	6.06	2.22	2.52	4.89	4.22		
	R	7.64	5.73	6.10	2.26	2.52	4.90	4.24		
	P	7.63	5.70	6.25	2.29	2.29	4.61	4.05		
C8	S	7.61	5.65	6.24	2.29	2.29	4.61	4.04		
	R	7.67	5.78	6.27	2.30	2.30	4.62	4.06		
	S								4.38	1.48
	R								4.46	1.51

The stereochemistry of the phosphotriester group in each analogue was determined by ^{1}H-{^{1}H} 2D NOESY data and the stereochemical assignment was supported by the HPLC and melting behavior of each analogue. Although NOEs involving the PO*CH*$_2$CH$_3$ protons have served to identify the stereochemistry of other phosphate alkylated analogues,[21,23] the phosphate ethyl methylene group was not used for the stereochemical assignment of the S-A(et)A and R-A(et)A analogues because its chemical shift coincides with some of the C5' proton chemical shifts around 4.4 ppm. Contour sections which display crosspeaks associated with the POCH$_2$*CH*$_3$ protons of the S-A(et)A and R-A(et)A analogues are shown in Figure 5.11. In the R-A(et)A duplex, the POCH$_2$*CH*$_3$ protons display a crosspeak with the A4-H8 proton whereas in the S-A(et)A duplex, this type of crosspeak is absent. The crosspeak for the R-A(et)A analogue was observable only at a mixing time of 500 ms which suggests that either the spatial proximity between the POCH$_2$*CH*$_3$ and A4-H8 protons is at the limit of NOE "detectability" (r~5.0 Å) or that the crosspeak is the result of spin diffusion.

As discussed in Section 4.3, spin diffusion can modify the intensity of a NOESY crosspeak and becomes more significant as the mixing period for the 2D NOESY experiment increases. A NOESY crosspeak due to spin diffusion would be observed between the POCH$_2$*CH*$_3$ and A4-H8 protons as a result of indirect cross-relaxation between the PO*CH*$_2$CH$_3$ and A4-H8 protons if the POCH$_2$*CH*$_3$⋯A4-H8 distance is larger than both the PO*CH*$_2$CH$_3$⋯A4-H8 distance and the POCH$_2$*CH*$_3$⋯PO*CH*$_2$CH$_3$ distance. If these distances are compared from models of the R-A(et)A and S-A(et)A analogues (Figure 5.12), it is possible that the POCH$_2$*CH*$_3$⋯A4-H8 crosspeak observed for the R-A(et)A analogue is the result of spin diffusion from the PO*CH*$_2$CH$_3$⋯A4-H8 spin pair. However, the presence of the POCH$_2$*CH*$_3$⋯A4-H8 crosspeak within one diastereomer but not the other indicates that the phosphotriester ethyl group of this diastereomer can approach the A4-H8 proton whereas this type of approach is not allowed for the other diastereomer. Because the $\underline{R}_p$ diastereomer, with its ethyl group oriented into the major groove, allows for this approach (Figure 5.12), the compound displaying the POCH$_2$*CH*$_3$⋯A4-H8 crosspeak was assigned the $\underline{R}_p$ diastereomer.

The method of spectroscopic assignment of single-site phosphate-alkylated DNA oligomers has been previously advanced by Summers *et al.*[21] and Lawrence *et al.*[23] and the success of this method has been established by comparing the assignment made by spectroscopic means with the assignment made by a chemo-enzymatic method. Furthermore, it has been shown that the HPLC and melting temperature behavior of single-site phosphate alkylated oligomers are correlated with the stereochemistry of the phosphotriester group.[21,23,24] In general, an analogue containing a phosphotriester alkyl group oriented towards the major groove will have a longer HPLC retention time and a decreased melting temperature relative to the analogue of opposite stereochemistry. The stereochemical assignments made for the R-A(et)A and S-A(et)A analogues by ^{1}H-{^{1}H} 2D NOESY measurements are supported by the chromatographic and stability characteristics of these two isomers; the isomer assigned as R-A(et)A has a longer reversed-phase HPLC retention time on a poly(styrene-divinylbenzene) HPLC column eluted with a linear gradient of acetonitrile vs. triethylammonium acetate buffer (G. Zon, private communication) and the isomer assigned as the R-A(et)A analogue also has a lower melting temperature compared to the S-A(et)A analogue (discussed in Section 5.2).

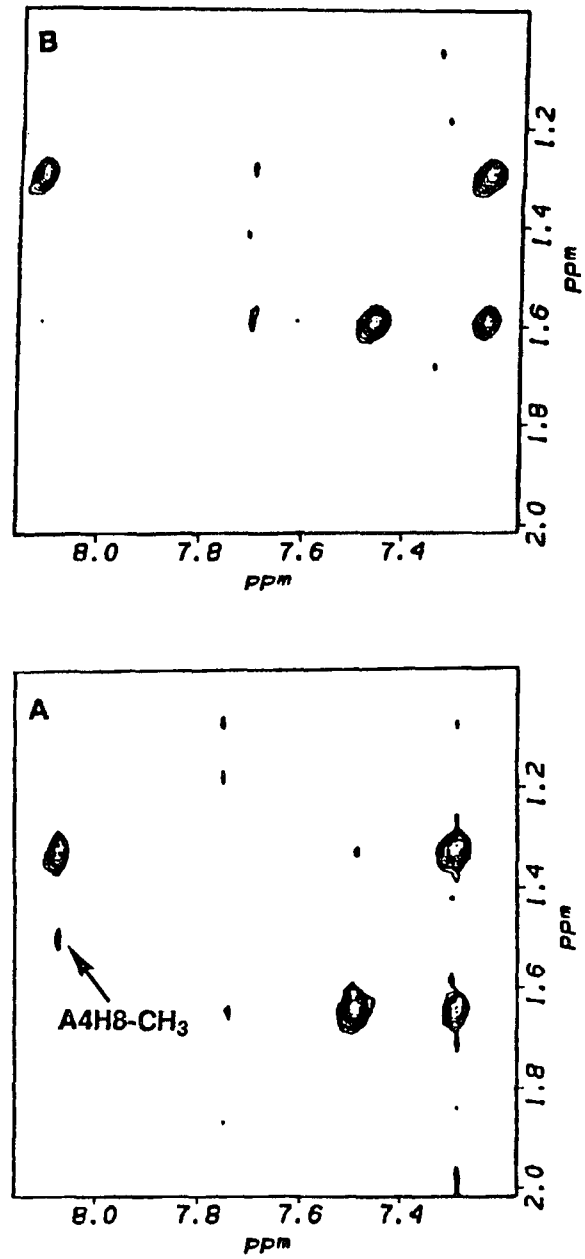


Figure 5.11 Expanded contour plots of the phase-sensitive 2D NOESY spectrum (mixing time 500 ms) of (**A**) (Sp,Sp)-{d[GGA(ethyl)ATTCC]}$_2$ and (**B**) (Rp,Rp)-{d[GGA(ethyl)ATTCC]}$_2$ at 5°C correlating the base protons (7.2 - 8.4 ppm) with the $POCH_2\underline{CH_3}$ protons (~1.5 ppm).

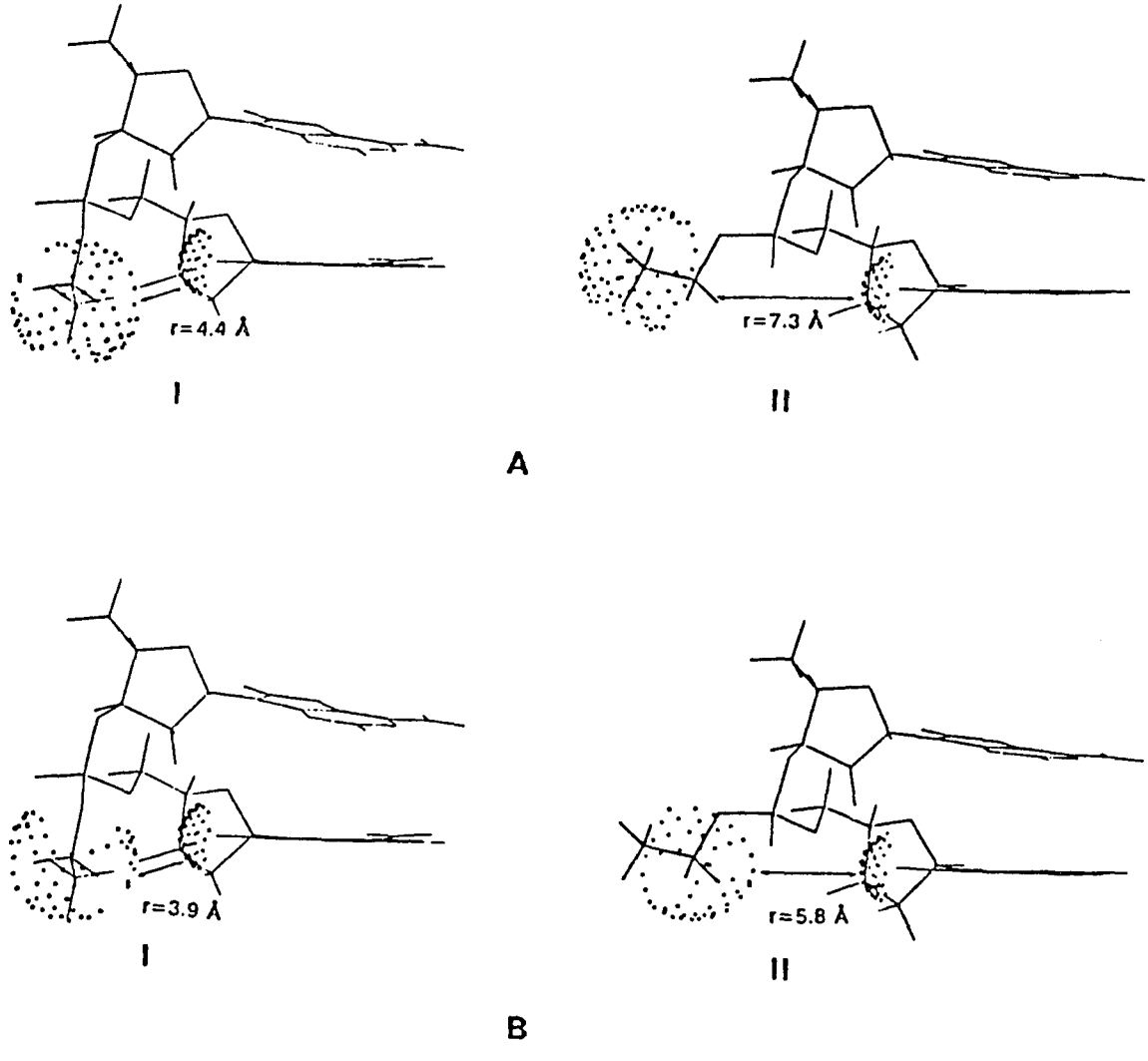


Figure 5.12 Diagram of the d[A(ethyl)A] segment in (I) (Rp,Rp)-{d[GGA(ethyl)ATTCC]}$_2$ and (II) (Sp,Sp)-{d[GGA(ethyl)ATTCC]}$_2$. The distance between the A4-H8 proton and (A) the center of the phosphotriester methyl group and (B) the center of the phosphotriester methylene group for each analogue is shown. The models are taken from reference 25 and represent energy minimized B-form structures of the two analogues.

An analysis of the 2D COSY and NOESY spectra of the parent and its two single-site ethylated analogues, S-A(et)A and R-A(et)A, permitted the assignment of the nonexchangeable base and sugar protons (except for the C5' protons) of these oligomers in D_2O solution at 5°C. Although Pardi *et al.*[92] have described a method using homonuclear ^1H-{^1H} and heteronuclear ^{31}P-{^1H} chemical shift correlations for the sequential assignment of the sugar protons and phosphates without recourse to a specific structural model, this method is limited to systems and conditions that are capable of providing all the necessary scalar coupling connectivities, conditions which were not possible to achieve for the octamers used in this study (the extensive broadening of proton resonances in each octamer at 5°C caused the cancellation of many COSY connectivities within each nucleotide). Therefore, most of the ^1H chemical shift assignments for each octamer were made using NOESY connectivities and assuming that each oligomer was in a right-handed conformation. This assumption is valid for the parent duplex as demonstrated by Broido *et al.*[22] and, indeed, the parent chemical shifts confirm the earlier assignments made by Broido *et al.*[22] and Patel *et al.*[40] on the same duplex at 25°C (after correction is made for the reference chemical shift difference at 25°C and 5°C).

Plotted in Figure 5.13 are the deviations of the ^1H chemical shifts of the R-A(et)A and S-A(et)A analogues from those of the parent molecule. A positive deviation indicates that the resonance of the modified oligomer is shifted upfield relative to that of the parent oligomer. Examination of these plots reveals that significant chemical shift differences (>0.1 ppm) are confined to signals from residues that flank the ethyl phosphotriester moiety (the G2, A3 and A4 residues) and appear to be most significant for the A3 residue. Similar chemical shift differences are observed for the H3', H4', H2', and H2" protons within the A3 residue for both analogues and could indicate a redistribution of charge density in the vicinity of these protons upon phosphate alkylation. For the R-A(et)A analogue, there is a significant change in the chemical shift of the H1' resonance in the G2 (~-0.17 ppm), A3 (~+0.30 ppm) and A4 (~-0.09 ppm) residues. The corresponding chemical shift changes for the S-A(et)A analogue are less than 0.05 ppm. If the chemical shift changes were simply the result of

changes in charge density in the vicinity of the H1' protons upon phosphate alkylation, it would be expected that the H1' protons of both analogues would undergo similar chemical shift changes, contrary to what is observed. A large upfield shift (~0.2 ppm) for the A4-H1' proton was observed in the oligomer, ($\underline{R}_p$,$\underline{R}_p$)-{d[GGAA(ethyl)TTCC]}$_2$ by Summers *et al.*[21] and Lawrence *et al.*[23] observed similar upfield shifts for the A3-H1' and A4-H1' protons in ($\underline{S}_p$,$\underline{S}_p$)-{d[GGA(isopropyl)ATTCC]}$_2$ and ($\underline{S}_p$,$\underline{S}_p$)-{d[GGAA(isopropyl)TTCC]}$_2$, respectively (note that for the isopropylated analogues, $\underline{S}_p$ denotes the diastereomers in which the isopropyl group is oriented toward the major groove). These observations are consistent with the results presented here and suggest that a structural change occurs for oligomers containing a phosphotriester group oriented toward the major groove of the duplex and that such a structural change is localized to the site of phosphate alkylation.

It is well established that the intrinsic proton chemical shifts of the base and sugar protons of DNA fragments can be influenced by

- ring currents arising from delocalized electrons at the purine and pyrimidine rings;
- polarization effects due to an enhancement or depletion of charge via protonation or hydrogen bonding;
- the anisotropy of diamagnetic susceptibility of various proximal groups such as carbonyls.[93]

Such factors could be sensitive to a conformational change in duplex structure. In fact, the H1' resonances of the parent octamer are observed to move downfield with increasing temperature, indicating that these resonances are influenced by shielding aromatic ring currents which are lost during the melting and destacking transitions.[35] However, it is not clear why the (*n-1*)H1' proton should shift significantly upfield when the (*n*)phosphate of an oligomer is ethylated with $\underline{R}_p$ stereochemistry ($\underline{S}_p$ stereochemistry in the case of isopropylation). Although destacking was observed to occur for phosphate alkylated dinucleotides, it is difficult to determine how such stacking changes would affect the ring currents that influence the H1' chemical shift of a polynucleotide. The nature of such

structural changes will become more obvious when spatial information (i.e., distances derived from NOE intensities) for these oligomers is examined (Section 5.5).

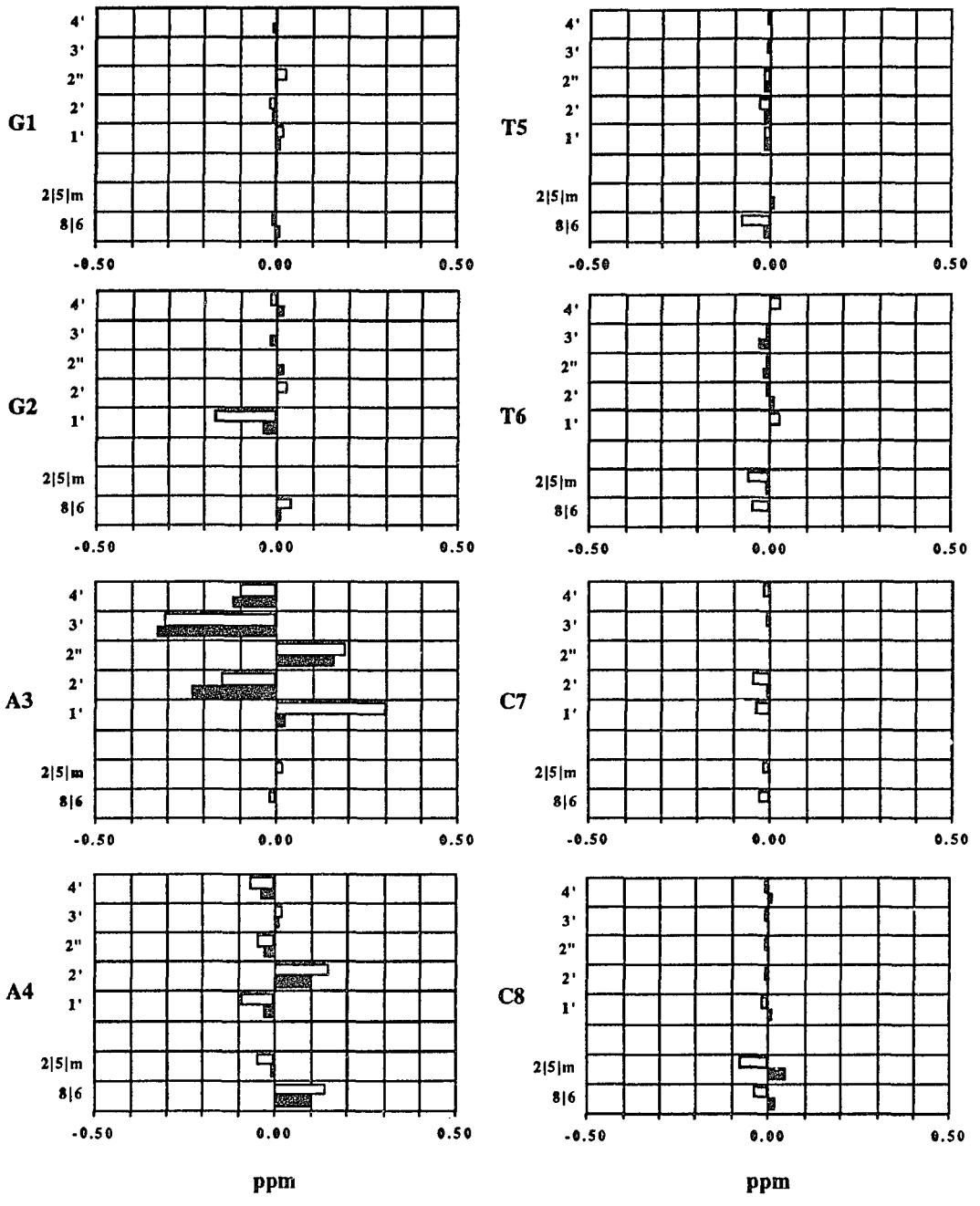


Figure 5.13 ^{1}H chemical shift differences observed between the parent and its two ethylated analogues. The shaded squares (▩) represent the chemical shift difference between $\{d[GGAATTCC]\}_2$ and (Sp,Sp)-$\{d[GGA(ethyl)ATTCC]\}_2$ for the indicated proton; the unshaded squares (□) represent the chemical shift difference between $\{d[GGAATTCC]\}_2$ and (Rp,Rp)-$\{d[GGA(ethyl)ATTCC]\}_2$ for the indicated proton. A positive deviation indicates that the resonance of the modified oligomer is shifted upfield relative to that of the parent oligomer.

5.1.2 Imino Proton Chemical Shift Assignments

The guanine and thymidine bases each contain an imino proton (G-H1 and T-H3) which is capable of forming a hydrogen bond with an appropriate acceptor atom. The imino hydrogen bonds for the A:T and G:C Watson-Crick base pairs are shown in Figure 5.14. An imino proton NMR signal can be observed in aqueous solution when the corresponding imino proton is involved in a hydrogen bond (e.g., Watson-Crick base pairing); otherwise, rapid chemical exchange with water results in the loss of its detectability as the imino proton signal coalesces with the water signal. Figure 5.15 shows the 5°C, one dimensional 400 MHz ^1H spectrum of the imino protons for the parent octamer and the two analogues, R-A(et)A and S-A(et)A. The additional peaks in the R-A(et)A spectrum are most likely due to the presence of small amounts of a heteroduplex in which one strand of the ($\underline{R}_p$,$\underline{R}_p$)-{[GGA(et)ATCC]}$_2$ duplex has lost its phosphotriester ethyl group because of limited hydrolysis. Because there are no additional peaks in the ^1H NMR spectrum of the nonexchangeable protons in R-A(et)A (Figure 5.1), it is likely that limited hydrolysis of the phosphotriester group occurred during the lyophilization procedure necessary to transfer the sample from a D_2O solution to an H_2O:D_2O solution (Section 3.1). However, it should be noted that the exact nature of impurities present in the R-A(et)A sample was not confirmed by other analytical methods such as HPLC.

Four imino proton resonances are observed for each octamer duplex in Figure 5.15 corresponding to the four base pairs related by 2-fold symmetry. Since each octamer is self-complementary, the imino protons can be numbered according to the following scheme:

G1-G2-A3-A4-T$\underline{4}$-T$\underline{3}$-C$\underline{2}$-C$\underline{1}$
C$\underline{1}$-C$\underline{2}$-T$\underline{3}$-T$\underline{4}$-A4-A3-G2-G1

The imino proton resonances can be assigned according to type based on previous studies suggesting that, in general, the intrinsic chemical shift of T-H3 in an A:T Watson-Crick base pair is located approximately 1 ppm downfield from the intrinsic chemical shift of G-H1 in a G:C Watson-Crick base pair.[35] Therefore, the two resonances at ~12.8 ppm in each spectrum

were assigned to the G-H1 protons, whereas the two resonances in the 13.5-14.0 ppm spectral range in each spectrum were assigned to the T-H3 protons. Fraying effects (the rapid opening and closing of a base-pair) coupled with the exposure of one face of the terminal base-pair to solvent leads to an exchange broadening of the terminal base-pair imino proton resonance.[35] Thus, the broadened resonance around 12.8 ppm serves to identify the G1-H1 proton. The remaining resonance around 12.8 ppm in each spectrum was assigned to the G2-H1 proton.

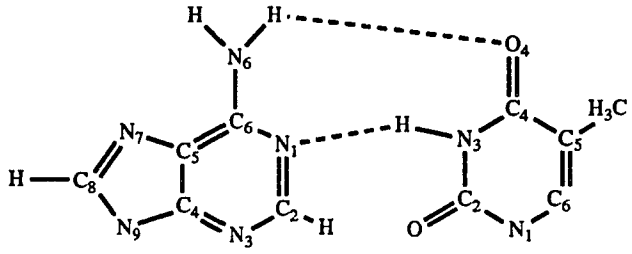
The T3-H3 proton for each octamer was assigned by means of a 1D NOE between this proton and the G2-H1 proton. In an idealized right-handed DNA structure, the distance between sequential imino protons is quite short (r~3.8 Å for A-DNA and r~3.5 Å for B-DNA) and therefore an NOE (a change in signal intensity) will be observed for an imino proton resonance when the resonance of the adjacent imino proton is saturated. This is clearly observed for the G2-H1•T3-H3 connectivity in each duplex as shown in Figure 5.16. The remaining downfield resonance was assigned to the T4-H3 proton. The chemical shift assignments of the imino proton resonances in each of the octamer duplexes are summarized in Table 5.2.

Although it is difficult to obtain a G2-H1•T3-H3 distance using the 1D NOE data obtained for each analogue (due to the lack of a reference distance, Section 2.5), it is clear that the difference spectrum for each duplex displays a similar NOE intensity (~10%) on the T3-H1 signal (Figure 5.16). This observation indicates that single-site phosphate ethylation does not appear to significantly alter the G2-T3 base-pair rise as judged by the G2-H1•T3-H3 NOE interaction (Figure 5.17). It should be noted that NOE connectivities for the sequential imino protons (G1-H1•G2-H1 and T3-H3•T4-H3) could not be determined by 1D NOE experiments due to the presence of decoupler "spillage" artifacts that occurred because the corresponding resonances were not sufficiently separated. Under these conditions, a resonance (e.g., T4-H3) adjacent to the one being saturated (e.g., T3-H3) became partially saturated and a difference signal resulted. Although the difference signal could be due to an actual NOE between the two imino protons, the fact that the difference signal intensity was

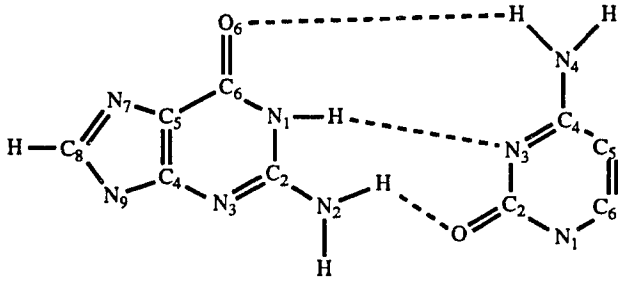
strongly dependent on the strength of the decoupler field suggests that its presence could be due to a decoupler spillage artifact.

The presence of all four imino proton signals (G1-H1, G2-H1, T3-H3 and T4-H3) indicates that duplex formation occurs along the entire length of the helix and is not inhibited by single-site phosphate ethylation of the duplex. However, the G1-H1 and T4-H3 protons shift significantly (~0.2 ppm for both resonances) to high field in the R-A(et)A analogue relative to the parent and the S-A(et)A analogue. The upfield shift of the R-A(et)A G1-H1 imino proton resonance might reflect an increased exchange rate with solvent (H_2O). An increased exchange rate would also lead to an increased broadening of the G1-H1 resonance; however, overlap of the G1-H1 resonance with the G2-H1 resonance prevented an accurate determination of its linewidth. An increased rate of exchange with solvent would be expected for the terminal base-pair of the R-A(et)A analogue because this analogue is destabilized compared to both the parent and the S-A(et)A analogue (see Section 5.2) and such destabilization would lead to increased fraying at the ends of this duplex.

Although the T4-H3 imino proton resonance undergoes an upfield shift in the R-A(et)A analogue, its lack of broadening relative to that observed in the parent and S-A(et)A analogue indicates that the shift is not due to an increased exchange with solvent but probably reflects a structural change instead. In fact, the chemical shift change for the T4-H3 proton in R-A(et)A is consistent with the ^1H chemical shift deviations observed for some of the nonexchangeable protons in R-A(et)A (Figure 5.13) suggesting that the duplex has undergone a structural change in the vicinity of the phosphotriester group.



A:T



G:C

Figure 5.14 The Watson-Crick base pairs A:T and G:C.

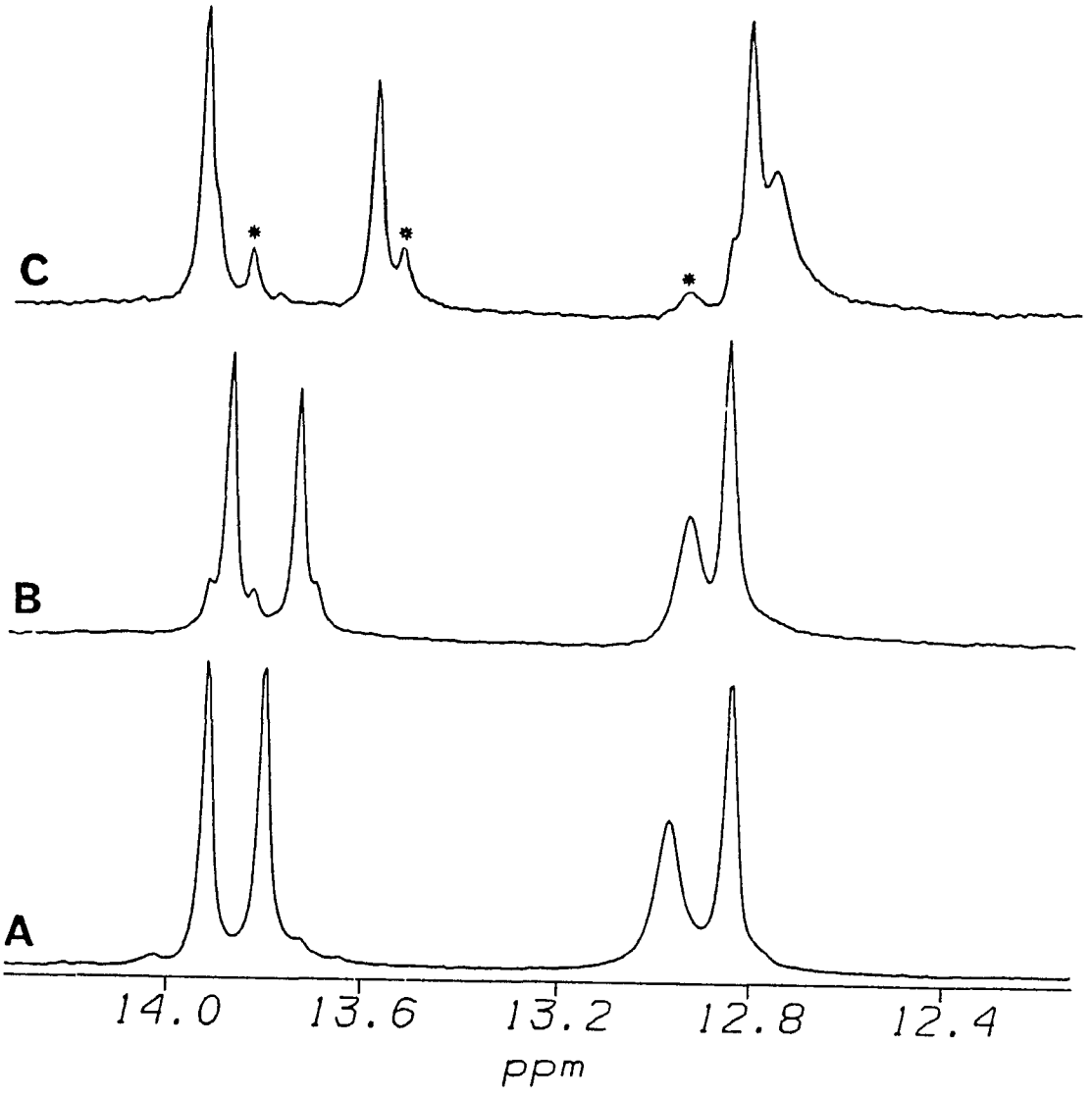


Figure 5.15 400 MHz ^{1}H NMR spectra of the imino protons in (A) {d[GGAATTCC]}$_2$, (B) ($\underline{S}_p$,$\underline{S}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ and (C) ($\underline{R}_p$,$\underline{R}_p$)-{d[GGA(ethyl)ATTCC]}$_2$. Peaks labeled with a (*) are due to impurities and are discussed in text. Sample conditions: 5°C; in H_2O:D_2O (8:2) with 10 mM sodium phosphate, pH 7.0. These chemical shifts are referenced relative to TSP and are based on a 5.00 ppm chemical shift for H_2O from TSP at 5°C.

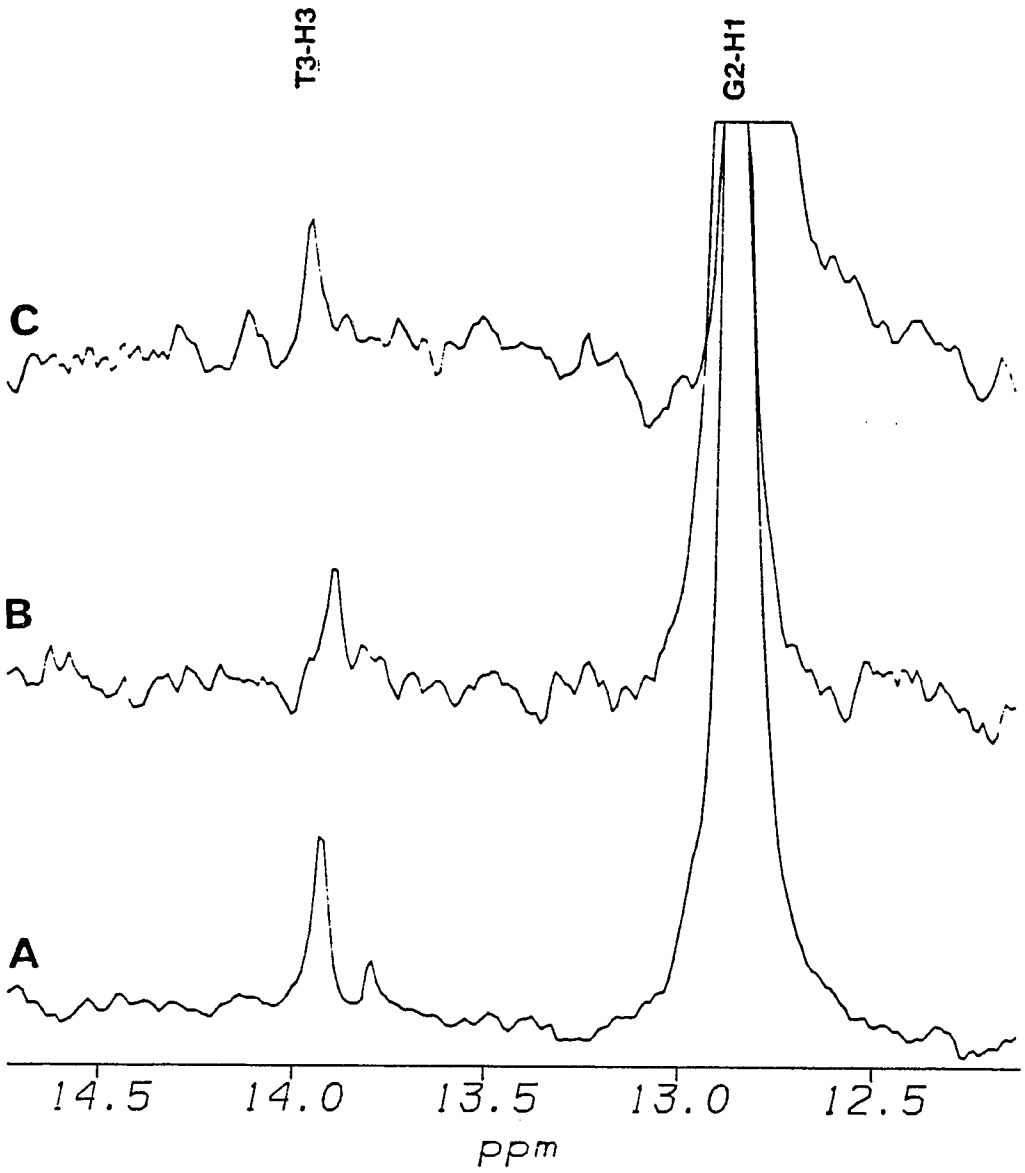


Figure 5.16 5°C, 1D NOE (truncated NOE) difference spectra (irradiation time 1 s) between the reference spectrum (off-resonance irradiation) and the spectrum with irradiation of the G2-H1 resonance for (**A**) $\{d[GGAATTCC]\}_2$, (**B**) $(\underline{S}_p,\underline{S}_p)$-$\{d[GGA(ethyl)ATTCC]\}_2$ and (**C**) $(\underline{R}_p,\underline{R}_p)$-$\{d[GGA(ethyl)ATTCC]\}_2$.

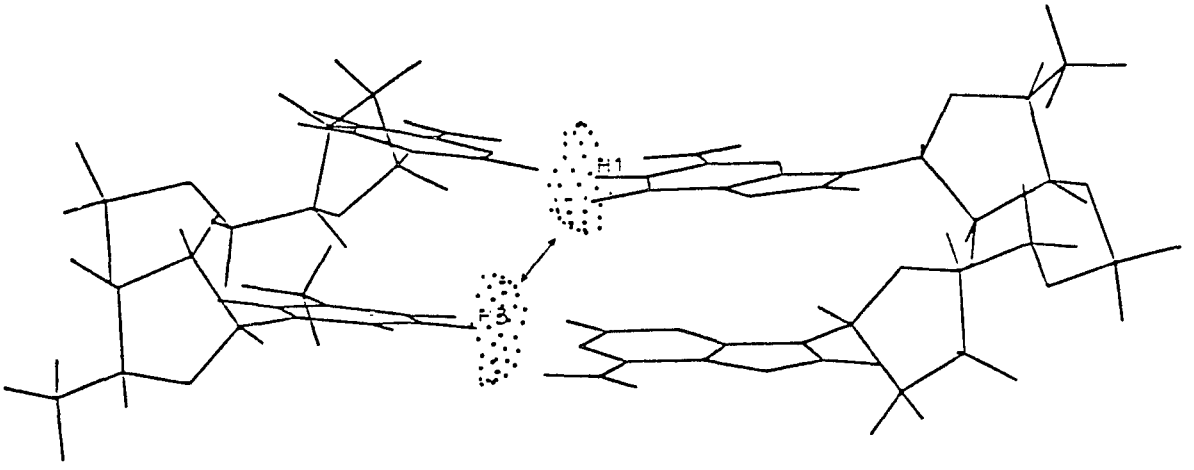


Figure 5.17 Figure of the imino proton interaction between the G2 base-pair and the T3 base-pair of the octamer sequence, $\{d[GGAATTCC]\}_2$. Note that this interaction can be used to determine the base-pair rise helicoidal parameter. The figure was obtained from an energy minimized B-form structure of the octamer sequence.

Table 5.2 ^1H chemical shifts of the imino protons in $\{d[GGAATTCC]\}_2$ (**P**), $(\underline{S}_p,\underline{S}_p)$-$\{d[GGA(ethyl)ATTCC]\}_2$ (**S**) and $(\underline{R}_p,\underline{R}_p)$-$\{d[GGA(ethyl)ATTCC]\}_2$ (**R**) in 10 mM phosphate buffer (pH 7.0), H_2O:D_2O (8:2) at 5°C. These chemical shifts are referenced relative to TSP and are based on a 5.00 ppm chemical shift for H_2O from TSP at 5°C.

	G1-H1	G2-H1	T3-H3	T4-H3
P	12.96	12.83	13.91	13.80
S	12.93	12.86	13.87	13.73
R	12.76	12.81	13.94	13.59

5.1.3 ^{31}P Chemical Shift Assignments

The one dimensional 162 MHz ^{31}P NMR spectra of the parent, S-A(et)A and R-A(et)A are shown in Figure 5.18. Seven distinguishable resonances are observed for the R-A(et)A and S-A(et)A duplexes at 5°C corresponding to two identical sets of seven chemically nonequivalent internucleotide linkages where the seven internucleotide linkages within each strand constitutes a set. Only six distinguishable resonances are observed for the parent duplex because the ^{31}P NMR signals corresponding to the ApT and TpT phosphates are nearly identical (*vide infra*).

The ^{31}P chemical shift assignments for each duplex were determined by the ^{1}H-detected heteronuclear ^{1}H-{^{31}P} chemical shift correlation experiment described in Section 3.2.8. The ^{1}H-{^{31}P} 2D chemical shift correlation spectrum for the S-A(et)A analogue at 5°C is shown in Figure 5.19. Similar spectra were obtained for the parent and the R-A(et)A analogue. The identification of ^{31}P resonances using the ^{1}H-{^{31}P} chemical shift correlation experiment relies on each phosphorus atom having scalar connectivities to protons whose resonances have been previously assigned. The assignment of the ^{31}P resonance corresponding to the (*n-1*)P(*n*) phosphate was accomplished by utilizing the three-bond and four-bond coupling of the ^{31}P spin with the (*n-1*)H3' proton and the (*n*)H4' proton, respectively. For example, the assignment of the ApA ^{31}P resonance was given by the presence of both ApA•A3-H3' and ApA•A4-H4' scalar connectivities. The crosspeaks arising from the three-bond coupling of the ^{31}P spin with the (*n*)H5' and (*n*)H5" protons were not utilized because the C5' protons could not be unambiguously assigned. The ^{31}P chemical shifts of the parent, S-A(et)A and R-A(et)A as determined by the above procedure are given in Table 5.3 along with literature values obtained by ^{17}O-labeling of the parent octamer.[38]

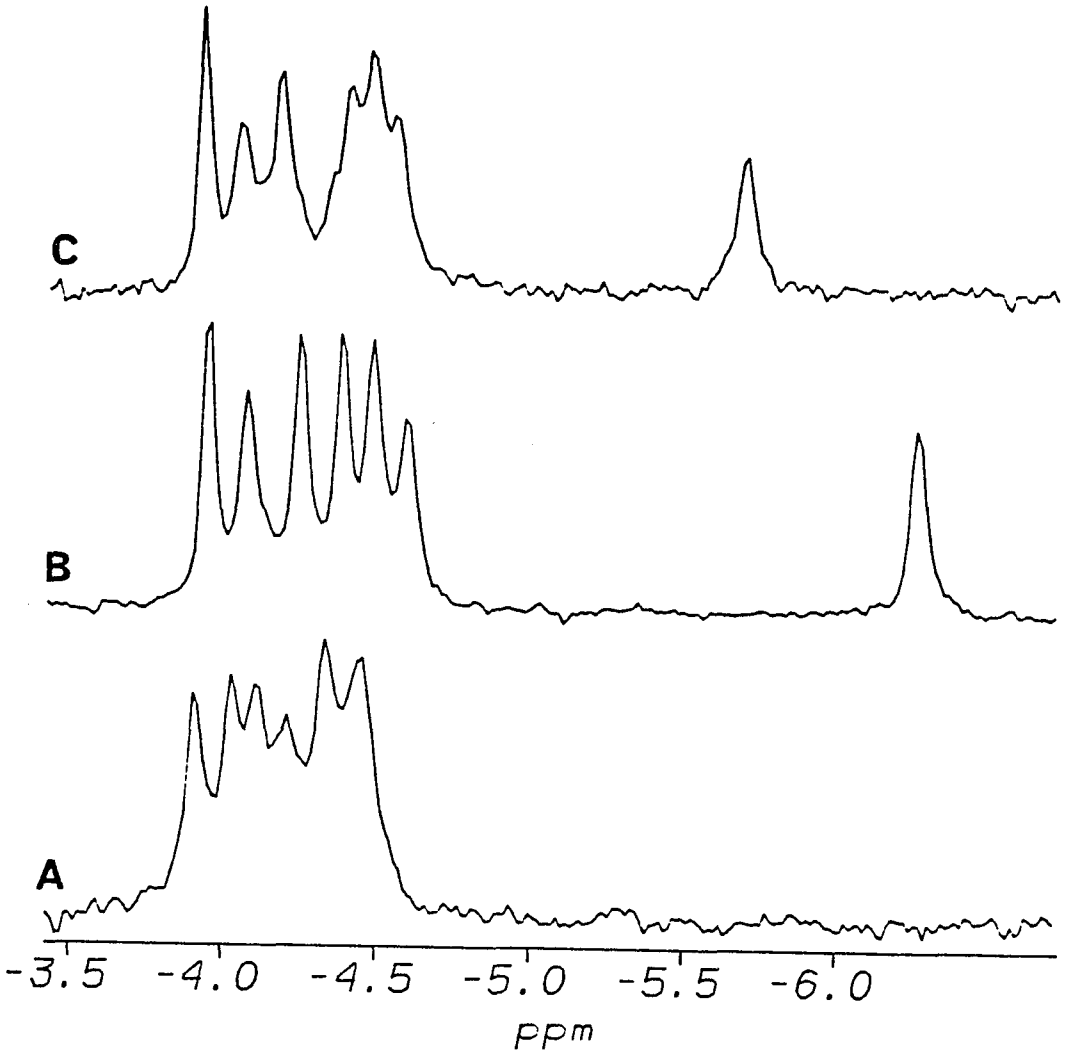


Figure 5.18 162 MHz ^{31}P NMR spectra of (A) {d[GGAATTCC]}$_2$, (B) ($\underline{S}_p$,$\underline{S}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ and (C) ($\underline{R}_p$,$\underline{R}_p$)-{d[GGA(ethyl)ATTCC]}$_2$. Sample conditions: 5°C; in D_2O with 10 mM sodium phosphate, pH 7.0. These chemical shifts are referenced relative to an external 50 mM sodium phosphate standard (phosphate chemical shift at 0.0 ppm).

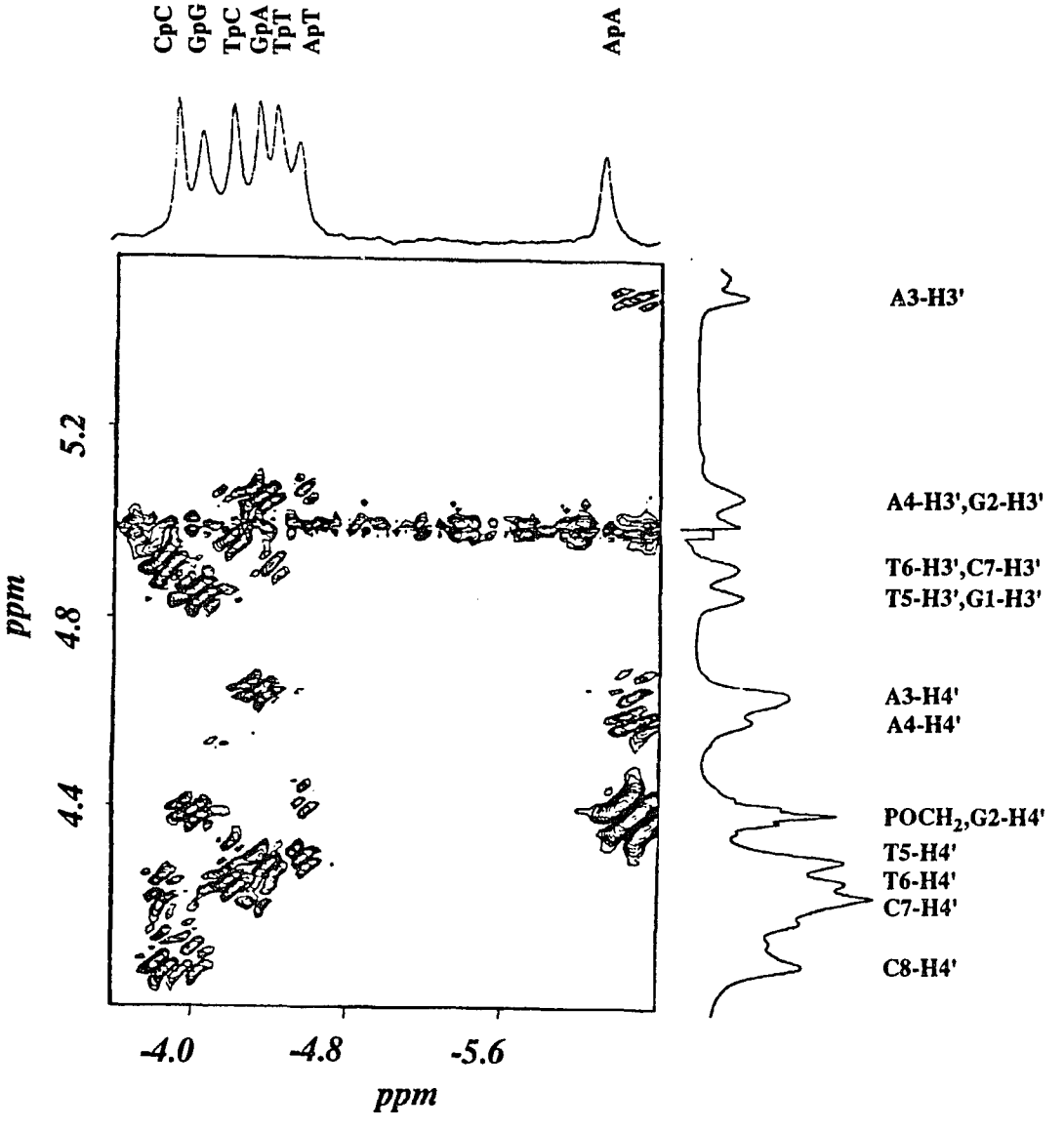


Figure 5.19 ^{1}H-(^{31}P) 2D chemical shift correlation spectrum of ($\underline{S}_p$,$\underline{S}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ at 5°C correlating the H3', H4', H5' and H5" protons (4.0 - 5.6 ppm) with the phosphorus resonances (-6.4 - -3.6 ppm). The 1D ^{1}H and ^{31}P NMR spectra are shown along the two axes of the 2D spectrum. The H3' and H4' resonances are labeled in the ^{1}H spectrum and the phosphorus resonances are labeled in the ^{31}P spectrum.

Table 5.3 ^{31}P chemical shifts for $\{d[GGAATTCC]\}_2$ (**P**), $(\underline{S}_p,\underline{S}_p)$-$\{d[GGA(ethyl)ATTCC]\}_2$ (**S**) and $(\underline{R}_p,\underline{R}_p)$-$\{d[GGA(ethyl)ATTCC]\}_2$ (**R**) in 10 mM phosphate buffer (pH 7.0), D_2O at 5°C. These chemical shifts are referenced relative to an external 50 mM sodium phosphate standard (phosphate chemical shift at 0.0 ppm). Literature values are those reported by Connolly and Eckstein for $\{d[GGAATTCC]\}_2$ at 10°C.[38]

	Lit. Value	P	S	R
GpG	-4.04	-4.06	-4.07	-4.06
GpA	-4.13	-4.14	-4.38	-4.43
ApA	-4.37	-4.37	-6.27	-5.69
ApT	-4.49	-4.48	-4.61	-4.59
TpT	-4.43	-4.45	-4.48	-4.50
TpC	-4.22	-4.23	-4.25	-4.20
CpC	-3.89	-3.93	-3.93	-3.94

Connolly and Eckstein assigned the resonances in the ^{31}P NMR spectrum of $(d[GGAATTCC])_2$ at 10°C by regiospecific labeling with ^{17}O.[38] It is apparent from Table 5.3 that the ^{31}P chemical shifts for the parent as determined by the ^{1}H-$\{^{31}P\}$ heteronuclear chemical shift correlation method are in very good agreement with those determined by Connolly and Eckstein and demonstrates the general utility of the ^{1}H-$\{^{31}P\}$ 2D NMR method for assigning the ^{31}P chemical shifts of deoxyoligonucleotides.

^{31}P chemical shifts, in general, can be influenced by a number of intrinsic effects such as the π bond order, the atomic charge on the phosphorus atom, the σ-bond angle and extrinsic effects such as hydrogen bonding.[94] Recent evidence indicates that variations in ^{31}P chemical shifts for nucleic acids are correlated to changes in the P-O (α and ζ) and C-O (β and ϵ) backbone torsion angles suggesting that ^{31}P chemical shifts can provide a probe of the phosphodiester backbone conformation.[95,96] A comparison of ^{31}P chemical shifts of dinucleoside phosphates and long-chain oligomers indicates that the ^{31}P NMR resonances of oligomers are shifted to higher field relative to the corresponding dinucleoside phosphate ^{31}P NMR resonances.[97] The upfield shifts seen in the ^{31}P chemical shift of oligomers as compared to dinucleoside phosphates have been attributed in large part to the base stacking and subsequent rigidity of oligomers which cause the phosphodiester angles α and ζ to adopt a gauche, gauche conformation. In the more flexible dinucleoside phosphates and at the terminal internucleotide linkages of oligomers, the α and ζ torsion angles tend to adopt a gauche, trans conformation and consequently, the ^{31}P resonances are found at lower field. This type of behavior is found for the parent, the R-A(et)A analogue and S-A(et)A analogue where the ^{31}P resonances corresponding to phosphates at the center of each oligomer are found upfield of the ^{31}P resonances corresponding to phosphates at the ends of each oligomer.

Plotted in Figure 5.20 are the ^{31}P chemical shifts as a function of residue number in the parent and the S-A(et)A and R-A(et)A analogues. Examination of Figure 5.20 reveals that the phosphotriester resonance, ApA, for both analogues is found upfield from the phosphodiester resonances. If charge neutralization of the phosphate in itself was responsible for the upfield shift observed for the phosphotriester resonance, then it would be expected

that the same upfield shift would occur in both the S-A(et)A and R-A(et)A analogues. This, however, is not the case; the phosphotriester resonance in S-A(et)A is further upfield than the phosphotriester resonance in R-A(et)A by approximately 0.6 ppm. In fact, the suggestion has been made in the literature that the absolute stereochemistry at a modified phosphate can be determined by comparing the relative chemical shift positions of the corresponding ^{31}P resonances.[98] Broido *et al.*[24] reported that for the A(iPr)T and A(et)T diastereomeric pairs of the octamer, the resonance of the phosphotriester in which the alkyl group points into the major groove of the helix is shifted further upfield than the phosphotriester resonance of opposite stereochemistry. However, the reverse behavior was found for A(iPr)A analogues of the octamer and the reverse behavior was also found for the A(et)A analogues described here indicating that the relative chemical shift of the phosphotriester ^{31}P resonances cannot be used to assign the stereochemistry of the diastereomers.

Examination of Figure 5.20 also indicates that the phosphodiester resonances, GpA and ApT, for the R-A(et)A and S-A(et)A analogues are shifted upfield, whereas the other phosphodiester resonances are unshifted relative to the corresponding resonances within the parent octamer. Furthermore, the magnitude of the GpA and ApT chemical shift difference is similar for each analogue. The upfield shift observed for the GpA and ApT phosphodiesters in the R-A(et)A and S-A(et)A analogues could be the result of either a structural change at these internucleotide linkages or a change in charge density at these phosphorus spins due to charge neutralization at the ApA phosphate. In order to determine the nature of the structural changes that occur upon phosphate-alkylation as evidenced by the ^{31}P chemical shift differences between the parent and the R-A(et)A and S-A(et)A analogues, it would be necessary to have conformational information on the sugar-phosphate backbone in either distance space or torsion-angle space. Due to the lack of interproton distance data for the C5' protons in these oligomers (the C5' protons were not assigned), there are few distances available that uniquely define the sugar-phosphate backbone.

Conformational information can be obtained in torsion-angle space by using a Karplus

relationship to correlate ^{1}H-^{31}P coupling constants, J_{HP}, with the H-C-O-P dihedral angles of the sugar-phosphate backbone.[41] ^{1}H-^{31}P coupling constants can be extracted from the crosspeak multiplet components in a ^{1}H-{^{31}P} 2D chemical shift correlation spectrum. Because the multiplet components are antiphase for each crosspeak in a ^{1}H-{^{31}P} 2D spectrum, a ^{1}H-^{31}P coupling constant can only be accurately determined from the crosspeak fine structure under conditions where J_{HP} is larger than the linewidth of the ^{1}H and ^{31}P resonances. In DNA fragments, the J_{HP} couplings typically vary between 1-4 Hz for the (*n*)P↔(*n*)H4' interaction and between 2.5-6.5 Hz for the (*n*)P↔(*n-1*)H3' interaction.[41,99] The linewidths of the ^{31}P resonances for the parent, the S-A(et)A analogue and the R-A(et)A analogue are approximately 10-15 Hz as determined from the spectra shown in Figure 5.18. The linewidths of the ^{31}P resonances are clearly larger than the anticipated ^{1}H-^{31}P coupling constants indicating that J_{HP} cannot be accurately determined under the experimental conditions employed for this study. Consequently, there is lack of structural data about the sugar-phosphate backbone for the parent and the S-A(et)A and R-A(et)A analogues in torsion-angle space.

The lack of data capable of providing structural information on the sugar-phosphate backbone makes it difficult to correlate the ^{31}P chemical shift changes observed for the R-A(et)A and the S-A(et)A analogue relative to the parent to a specific conformational change in the phosphodiester backbone induced by phosphate alkylation. However, the ^{31}P chemical shift difference between the parent and the R-A(et)A and S-A(et)A analogues were confined to the phosphodiester groups adjacent to the phosphotriester lesion suggesting that any possible structural changes are localized to the region of phosphate alkylation.

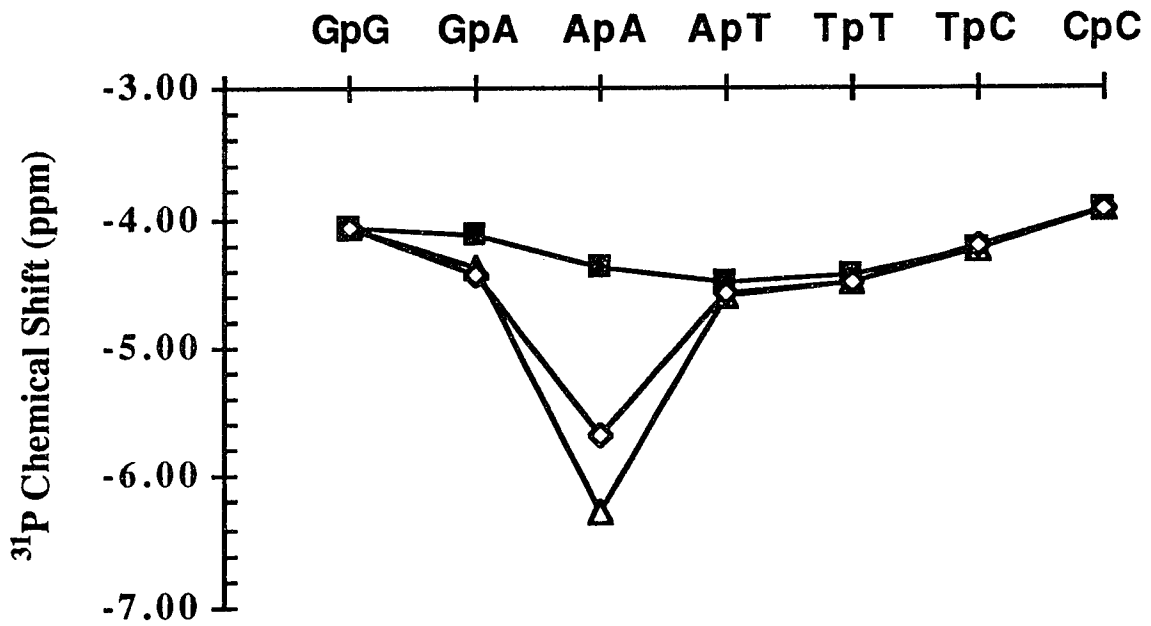


Figure 5.20 The ^{31}P chemical shift as a function of residue number for {d[GGAATTCC]}$_2$ (■), ($\underline{S}_p$,$\underline{S}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ (▵) and ($\underline{R}_p$,$\underline{R}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ (◇).

5.2 Helix-Coil Transition

The thermal dissociation of a nucleic acid duplex into single strands is characterized by a melting temperature, T_m. The melting temperature represents the midpoint of the thermally induced duplex-to-strand (helix-coil) transition and has been extensively used to evaluate the helix stability of DNA duplexes in solution.[14] The melting temperature of a DNA duplex can be measured by a number of techniques which include optical measurements, NMR spectroscopy and differential scanning calorimetry. Although the optical measurement (i.e., UV absorbance at 260 nm) of the loss of hypochromicity from base stacking upon strand separation yields the "classical" melting temperature of a DNA double helix, various NMR measurements can also be used to monitor strand separation.[100,101] In fact, the relative melting temperatures of a series of duplexes determined from NMR measurements is identical to that obtained from optical measurements.[23] The advantage of ^{1}H NMR spectroscopy, unlike most optical methods, is that it can be used to monitor the dissociation of individual base pairs because it permits the resolution and assignment of individual proton markers. The most easily measured property associated with strand separation using NMR spectroscopy is the temperature dependence of imino proton resonance linewidths. An imino proton resonance will broaden beyond detectability when the corresponding imino proton is no longer involved in a hydrogen bond (e.g., base-paired) and undergoes rapid exchange with the bulk solvent (H_2O). The melting temperature of the duplex can be evaluated from the temperature at which an imino proton resonance disappears from the ^{1}H NMR spectrum.

The imino ^{1}H NMR spectra of the parent octamer between 5°C and 40°C are presented in Figure 5.21. Similar melting behavior is observed for the R-A(et)A and S-A(et)A duplexes. In all cases, the G1-H1 resonance exhibits exchange broadening at a lower temperature than do the other resonances, a behavior indicative of fraying at the ends of each duplex.[102] The remaining three resonances (G2-H1,T3-H3 and T4-H3) broaden in a concerted fashion on raising the temperature from 5°C to 32.5°C and broaden beyond

detectability above 32.5°C. The three resonances, G2-H1, T3-H3, and T4-H3, disappear at a common temperature within each octamer studied, demonstrating that the nonterminal base pairs of each duplex melt as a cooperative unit. Therefore, the presence of a phosphotriester lesion does not appear to affect the cooperative melting involved in the two-state helix-to-coil transition of the octamer sequence.

The temperature-dependent linewidths of the T3-H3 and T4-H3 proton resonances for the parent, S-A(et)A and R-A(et)A duplexes are plotted in Figure 5.22. It is readily apparent that the R-A(et)A duplex is thermally less stable than both the S-A(et)A and parent duplexes. Based on the temperature at which the G2-H1, T3-H3, and T4-H3 proton resonances broaden beyond detectibility, the melting temperature (accurate to ±2.5°C) and thus relative duplex stability for each oligomer is estimated to be:

$$\text{R-A(et)A (25°C)} < \text{S-A(et)A (31°C)} < \text{Parent (33°C)}$$

The lower melting temperature for the R-A(et)A analogue is consistent with previous studies demonstrating that, in general, the more destabilized diastereomer is the one containing a phosphate alkyl group pointing into the major groove of the helix. If duplex destabilization were primarily due to an electrostatic effect arising from the partially neutralized backbone, then the R-A(et)A and S-A(et)A analogues would be expected to display similar melting behavior. The lowered T_m and the chemical shift differences (Section 5.1) of the R-A(et)A duplex relative to the S-A(et)A and parent duplexes suggests that a phosphotriester moiety oriented into the major groove has a conformationally induced destabilizing influence on a duplex. Thus, the observed melting behavior supports the hypothesis that the primary source of duplex destabilization by single-site phosphate alkylation is largely steric.

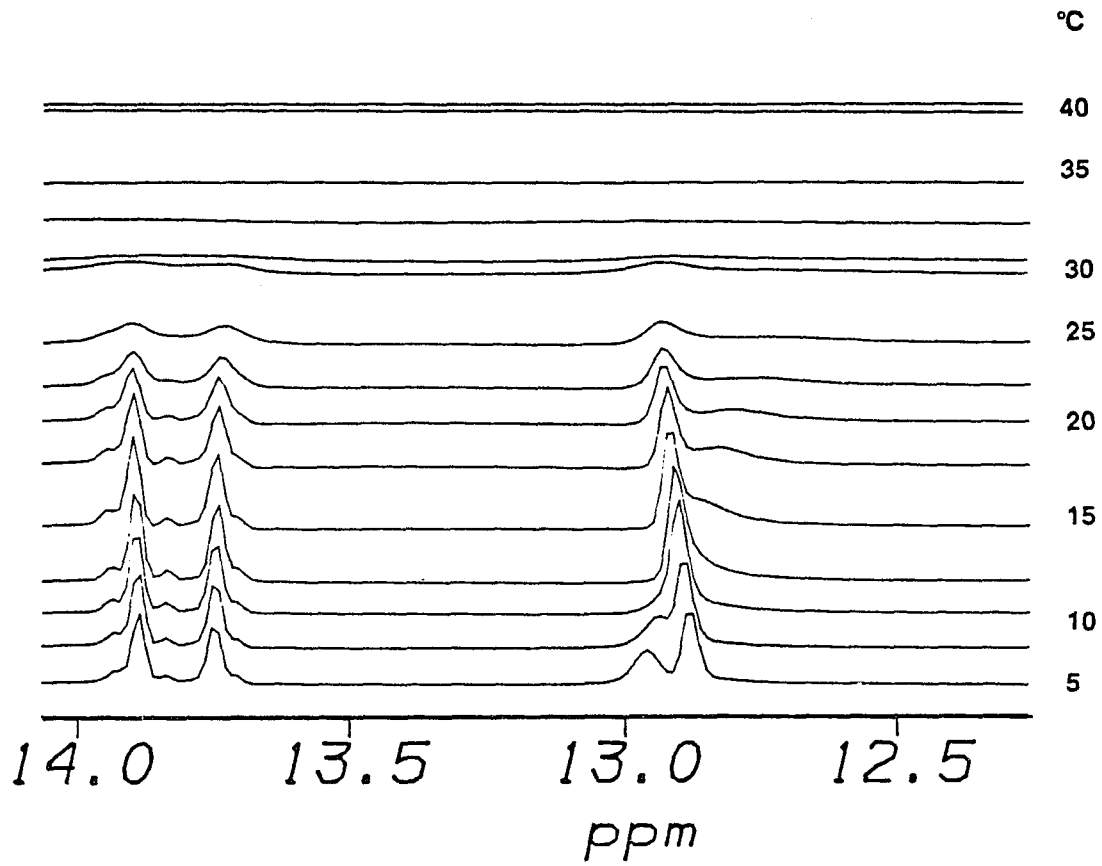


Figure 5.21 400 MHz ^{1}H NMR spectra (12 - 14 ppm) of $(d[GGAATTCC])_2$ between 5°C and 40°C. Sample conditions: in H_2O:D_2O (8:2) with 10 mM sodium phosphate, pH 7.0. These chemical shifts are referenced relative to TSP and are based on a 5.00 ppm chemical shift for H_2O from TSP at 5°C.

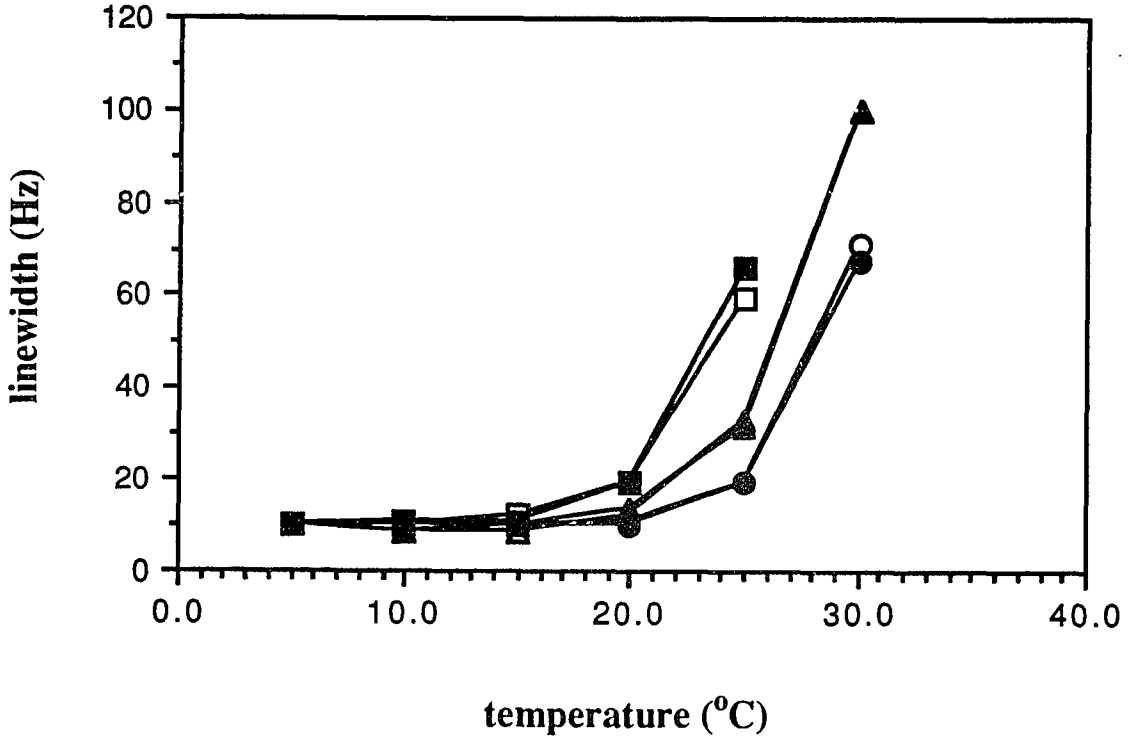


Figure 5.22 The temperature dependence of the linewidths of the T3-H3 (unfilled symbols) and T4-H3 (filled symbols) imino proton resonances in $\{d[GGAATTCC]\}_2$ (○), (Sp,Sp)-$\{d[GGA(ethyl)ATTCC]\}_2$ (△) and (Rp,Rp)-$\{d[GGA(ethyl)ATTCC]\}_2$ (□).

5.3 Base-Pair Lifetime Measurements

The ^{1}H NMR spectrum displaying the imino proton resonances of the parent duplex at 5°C in the presence of three phosphate concentrations is shown in Figure 5.23. As the concentration of phosphate buffer ($[PO_4^{3-}]$) increases, the resonances broaden from base catalyzed exchange of the imino protons as discussed in Section 2.6. Similar exchange behavior is observed for the R-A(et)A and S-A(et)A duplexes. The imino proton linewidths for the parent duplex as a function of the phosphate concentration are shown in Figure 5.24. The exchange time, τ_{ex}, at each catalyst concentration was obtained from the measured imino proton linewidth according to Equation 2.32. The dependence of τ_{ex} on $1/[PO_4^{3-}]$ is shown in Figure 5.25. These plots correspond to concentrations of the phosphate buffer from 200 to 500 mM, a range in which τ_{ex} vs. $1/[PO_4^{3-}]$ is linear as predicted from Equation 2.31. At lower concentrations of phosphate, $[PO_4^{3-}] < 200$ mM, the plots deviate from linearity. This behavior is likely due to the presence of factors such as adjacent base-pairs which are more efficient than phosphate base in catalyzing imino proton exchange at low phosphate concentrations. The base-pair lifetime for the first base pair, G1:C1, is too short to measure accurately at 5°C due to the effects of fraying at the ends of each oligonucleotide. The lifetimes, τ_o, of the G2:C2, A3:T3 and A4:T4 base-pairs in each of the DNA oligomers were determined by linear extrapolation of τ_{ex} to infinite concentration of phosphate base using a least-squares program as described in Section 3.2.3 and they are summarized in Table 5.4. The upper and lower limit of each lifetime was determined graphically from each curve in Figure 5.25. Two lines, representing maximum and minimum values of the base-pair lifetime, were drawn for each curve consistent with the requirement that both lines pass through the x and y error bars for each experimental point. A representative example of the error bars for τ_{ex} and $1/[PO_4^{3-}]$ is given in Figure 5.25A.

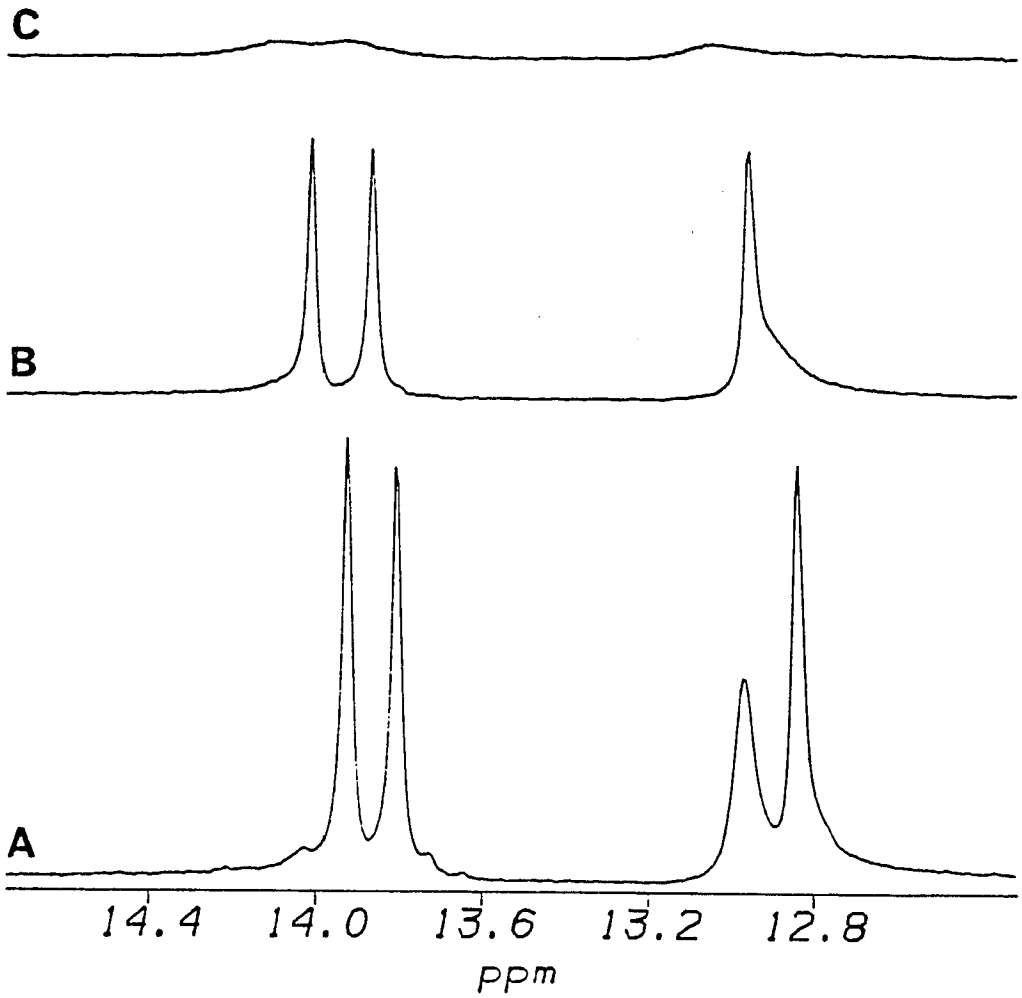


Figure 5.23 Imino proton region of the ^1H NMR spectrum of $\{d[GGAATTCC]\}_2$ containing 0.010 M (**A**), 0.250 M (**B**) and 0.550 M (**C**) sodium phosphate. Sample conditions: 5°C; in H_2O:D_2O (8:2), pH 7.0. These chemical shifts are referenced relative to TSP and are based on a 5.00 ppm chemical shift for H_2O from TSP at 5°C.

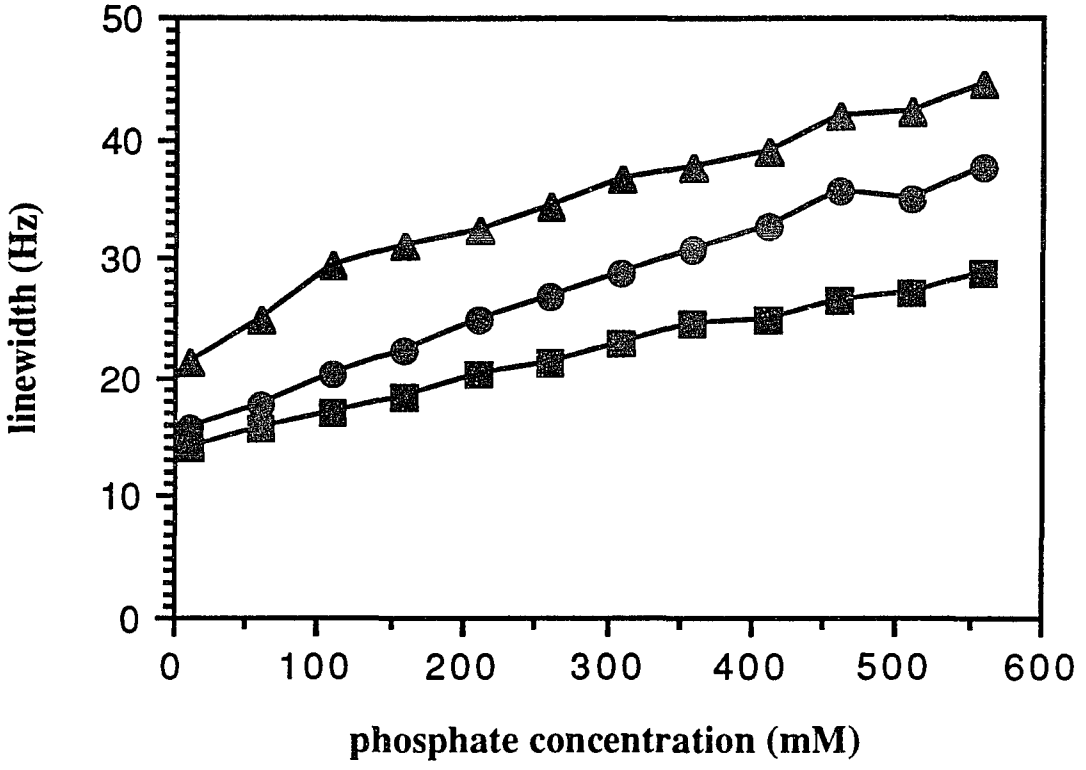
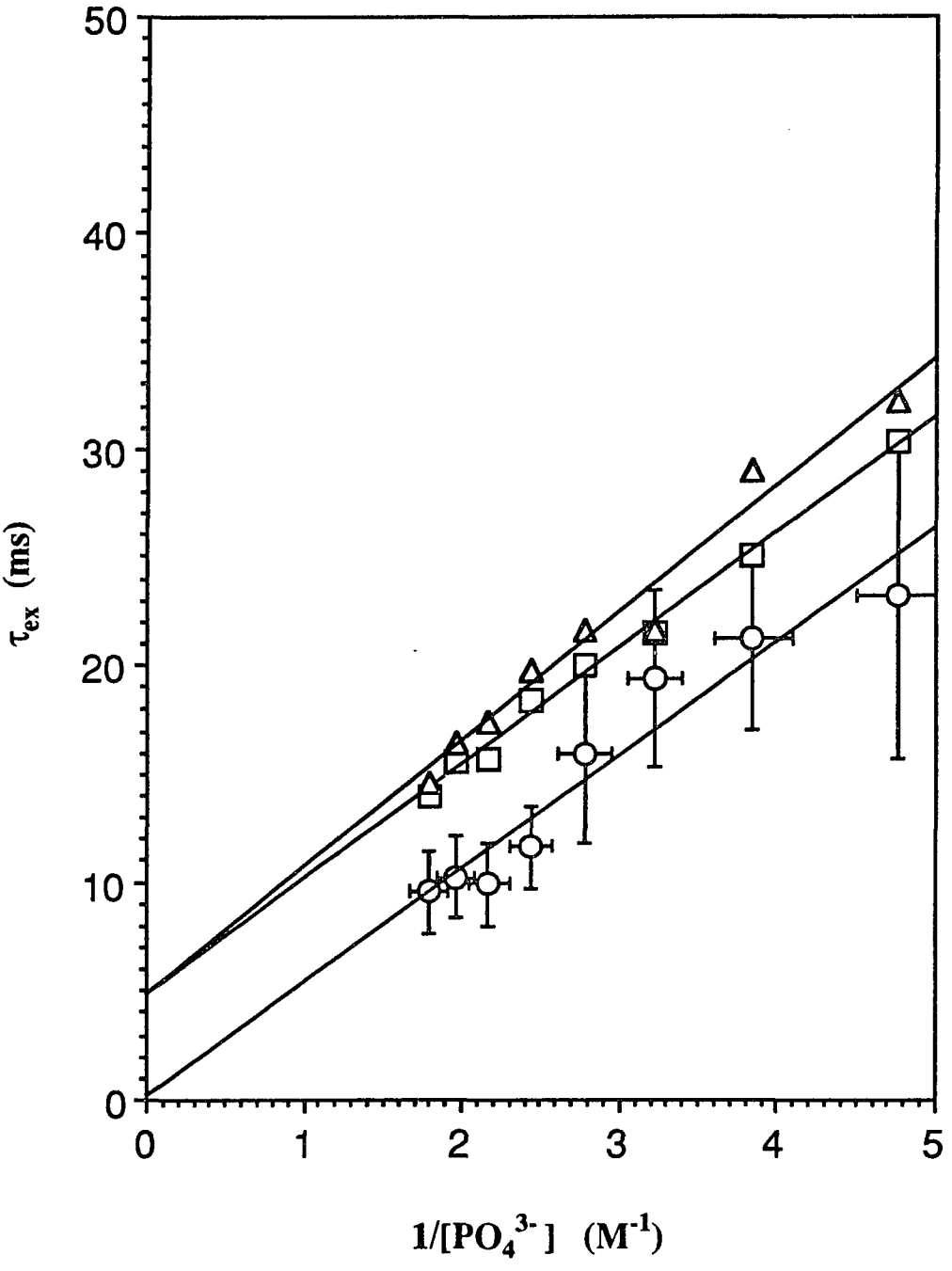


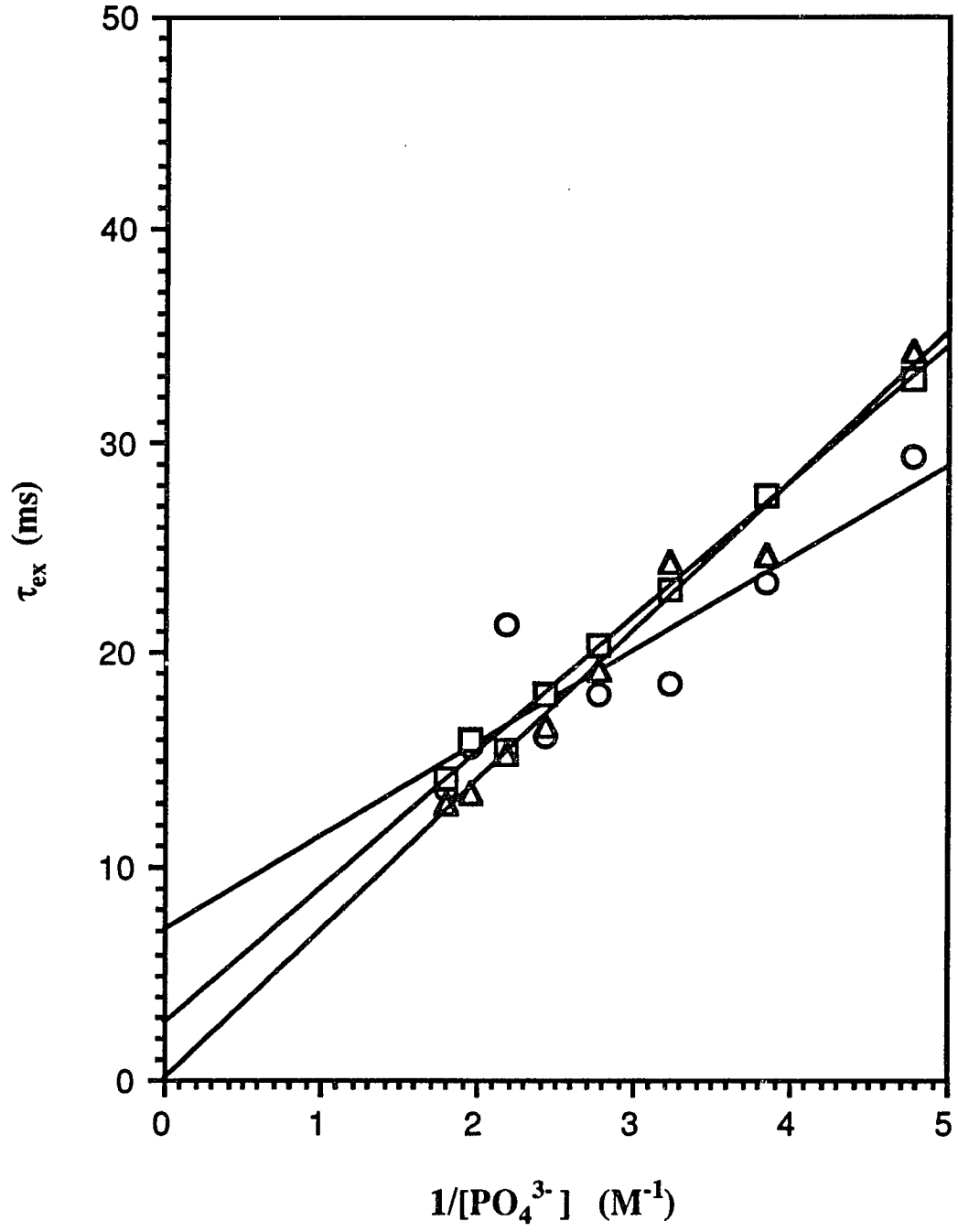
Figure 5.24 The imino proton linewdiths (▲, G2-H1; ●, T3-H3; ■, T4-H3) of (Rp,Rp)-{d[GGA(ethyl)ATTCC]}$_2$ as a function of phosphate concentration.

Figure 5.25 Dependence of τ_{ex} on $1/[PO_4^{3-}]$ for the (A) G2-H1, (B) T3-H3 and (C) T4-H3 imino protons in $\{d[GGAATTCC]\}_2$ (○), (Sp,Sp)-$\{d[GGA(ethyl)ATTCC]\}_2$ (△) and (Rp,Rp)-$\{d[GGA(ethyl)ATTCC]\}_2$ (□). Each line represents a linear least-squares fit of the displayed experimental data. Error bars for τ_{ex} and $1/[PO_4^{3-}]$ are given for the experimental data of the G2-H1 imino proton in $\{d[GGAATTCC]\}_2$.

A. G2-H1



B. T3-H3



C. T4-H3

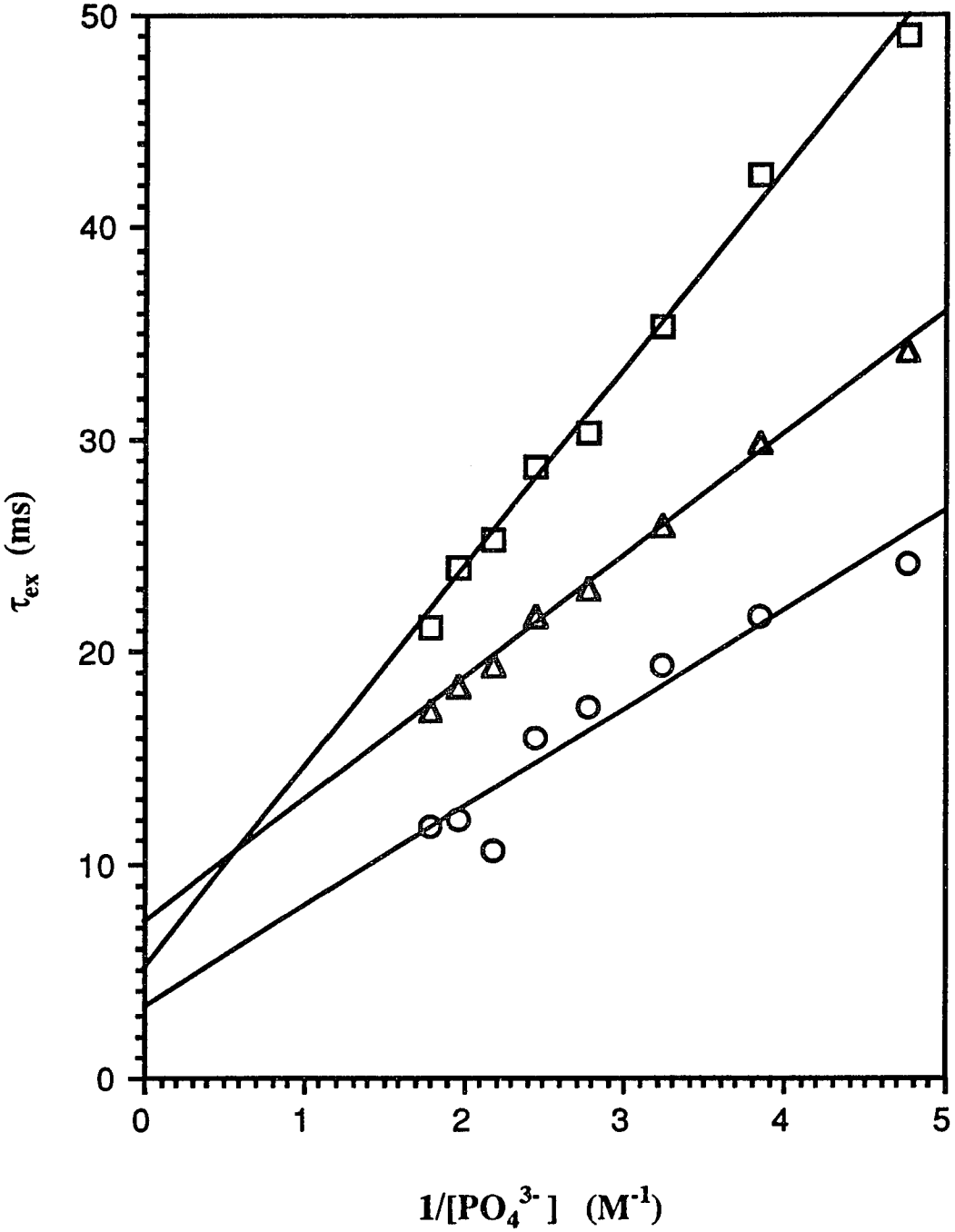


Table 5.4 Base-pair lifetimes, τ_o, for {d[GGAATTCC]}$_2$ (**P**), (Sp,Sp)-{d[GGA(ethyl)ATTCC]}$_2$ (**S**) and (Rp,Rp)-{d[GGA(ethyl)ATTCC]}$_2$ (**R**). The upper and lower limit of each lifetime is given in parentheses.

	G2-H1	T3-H3	T4-H3
P	(6) 0 (0)	(13) 7 (4)	(9) 3 (0)
S	(11) 5 (2)	(6) 0 (0)	(13) 7 (4)
R	(11) 5 (2)	(9) 3 (0)	(11) 5 (2)

Base-pair lifetime determinations on various DNA oligomers suggest a model where imino protons exchange by the opening of individual base-pairs rather than one where the base-pairs open together in a concerted fashion.[103] The base-pair lifetimes are in the range of 1 - 50 milliseconds at room temperature where A:T base-pairs generally have shorter lifetimes (1 - 7 ms) than G:C base-pairs (7 - 40 ms).[66] Previous studies have suggested a correlation between the lifetime of an individual base-pair and the structural rigidity of the duplex at the base-pair site.[42,104] Some base pairs have been found to have extraordinarily long opening lifetimes (e.g., ~10 minutes for base-pairs in the dhU stem of tRNA) and such long lifetimes have been attributed to an unusually tight structure in the region of such base-pairs. On the other hand, some base-pairs are short lived (e.g., <1 ms for base-pairs at the ends of DNA duplexes) suggesting flexibility in the area of such base-pairs.

Examination of Table 5.4 reveals that the lifetime of the central base-pair, A4:T4, for the three duplexes studied here are similar, differing by only a few milliseconds. Although increases in the base-pair lifetime by more than a factor of ten have been correlated to structural features in the vicinity of the corresponding base-pair, it remains unclear what small changes in base-pair lifetimes indicate.[42,66] The small differences in lifetime observed for the A4:T4 base-pair between the parent, S-A(et)A and R-A(et)A duplexes probably indicates that the presence of a phosphotriester lesion does not significantly influence the dynamics of an adjacent base-pair.

The lifetime of the G2:C2 base-pair in the R-A(et)A and S-A(et)A duplexes is longer than in the parent duplex and the lifetime of the A3:T3 base-pair in these analogues is shorter than in the parent duplex. This behavior suggests that phosphate ethylation possibly alters the base-pair dynamics in the region of phosphate ethylation, without regard to stereochemistry. However, because the differences are within experimental error, it cannot be firmly established that phosphate ethylation affects the conformational flexibility of the octamer duplex in the vicinity of the alkylation site. A precise measurement of the base-pair lifetimes using the method described here requires having data at higher phosphate concentrations

where the imino proton linewidths are larger and, consequently, the errors in τ_{ex} are smaller (see Figure 5.25A). The limited solubility of phosphate base at pH 7.0 prevented the collection of data at higher phosphate concentrations and hence restricted the precision of the base-pair lifetimes obtained in this study. In order to obtain base-pair lifetimes with higher precision, it would be necessary to use a base catalyst (i.e., NH_3) that permits the collection of data closer to the limiting value of τ_{ex} (i.e., at higher catalyst concentration). Furthermore, the lifetimes of the G2:C2 and A3:T3 base-pairs are likely to be strongly influenced by the effects of fraying at the ends of the octamer sequence. Therefore, it may be necessary to employ sequences longer than eight base-pairs in order to better separate the effects of fraying at the ends of an oligonucleotide from the effects of a phosphotriester lesion on the base-pair lifetimes within an oligonucleotide. Such a system would provide a more detailed understanding of the effects of single-site phosphate alkylation on the base-pair dynamics and hence, conformational flexibility, of DNA fragments.

5.4 ^{31}P Spin-Lattice Relaxation Measurements

The 162 MHz ^{31}P spin-lattice relaxation time (T_1) of each phosphorus resonance in the parent, S-A(et)A and R-A(et)A duplexes was measured using the inversion-recovery experiment described in Section 3.2.5. An example of an inversion-recovery measurement for the S-A(et)A duplex is shown in Figure 5.26 and a plot of the peak intensity vs. delay time, τ, for each resonance is shown in Figure 5.27. The peak intensities were fitted as a function of the τ delay using a non-linear least-squares fit as described in Section 3.2.5. From these calculations, the spin-lattice relaxation time for each resonance was obtained and the results are given in Table 5.5. To ensure that the spin-lattice relaxation times were not influenced by the presence of trace amounts of paramagnetic metal ion impurities in the samples, T_1 measurements were carried out in the presence of two different concentrations of EDTA (0.1 mM and 0.3 mM). The T_1 data were essentially identical at the two EDTA concentrations indicating that paramagnetic impurities, if present in these samples, do not significantly affect the ^{31}P spin-lattice relaxation times.

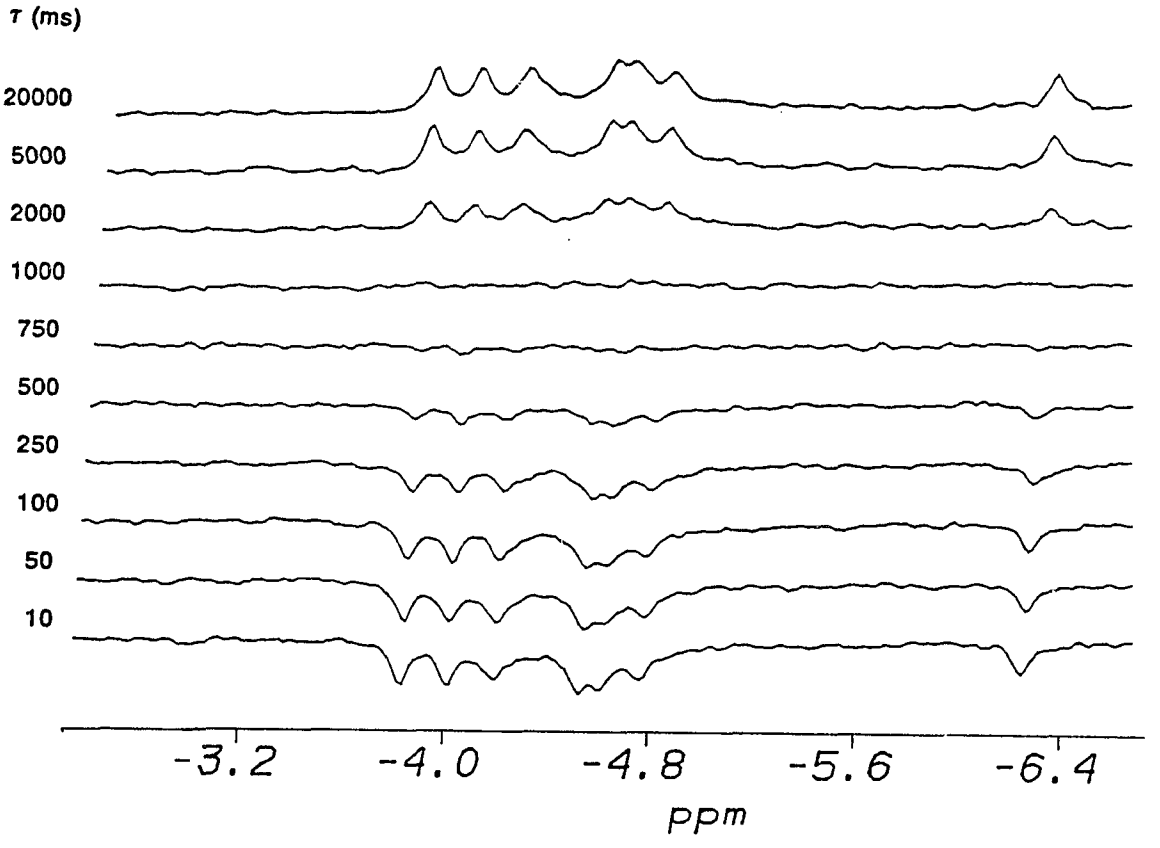


Figure 5.26 162 MHz ^{31}P spectra of (Sp,Sp)-{d[GGA(ethyl)ATTCC]}$_2$ obtained at different delay times, τ, using an inversion-recovery experiment. Sample conditions: 5°C; in D_2O with 10 mM sodium phosphate, pH 7.0.

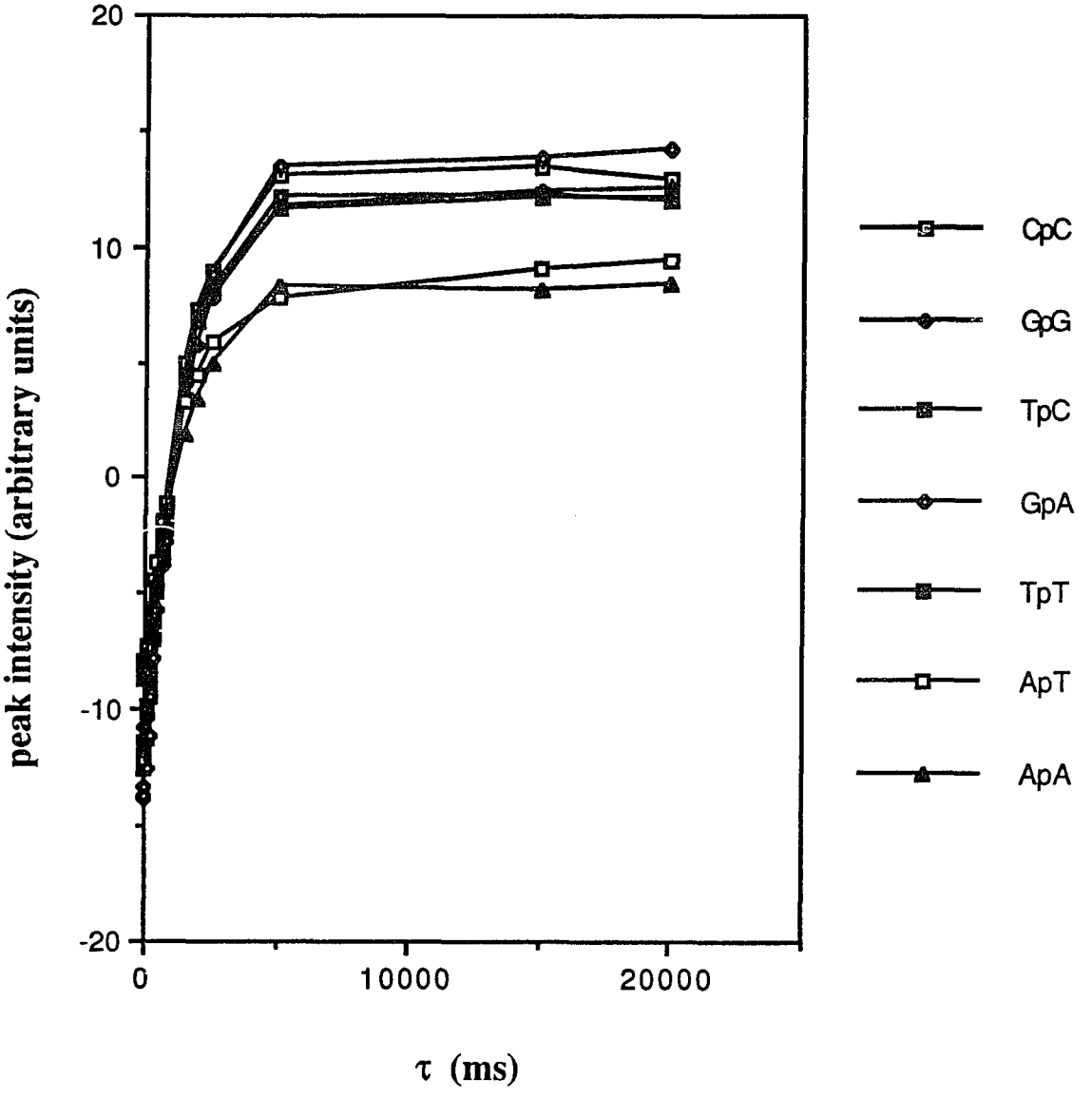


Figure 5.27 Plot of the peak intensity of each phosphorus resonance as a function of the inversion-recovery delay time, τ, for (Sp,Sp)-{d[GGA(ethyl)ATTCC]}$_2$ at 5°C.

Table 5.5 Experimental ^{31}P spin-lattice relaxation times (seconds) for {d[GGAATTCC]}$_2$ (**P**), (Sp,Sp)-{d[GGA(ethyl)ATTCC]}$_2$ (**S**) and (Rp,Rp)-{d[GGA(ethyl)ATTCC]}$_2$ (**R**). Sample conditions: 5°C; in D_2O with 10 mM sodium phosphate, pH 7.0. The error for each T_1 measurement is estimated to be less than 5% based on the standard deviation of the non-linear least squares fit to the experimental data and the uncertainty in the intensity for each phosphorus resonance.

	P	S	R
GpG	1.39	1.46	1.40
GpA	1.41	1.44	1.29
ApA	1.33	1.63	1.34
ApT	1.34	1.37	1.33
TpT	1.35	1.39	1.32
TpC	1.27	1.32	1.31
CpC	1.28	1.29	1.23

^{31}P NMR relaxation techniques have been widely used to study the structure and conformational dynamics of nucleic acids.[94] The availability of a ^{31}P nucleus at each internucleotide linkage in the polynucleotide backbone of a nucleic acid makes ^{31}P relaxation measurements an attractive method for examining the local internal motions and structure of DNA fragments. For example, ^{31}P relaxation measurements indicate that the relaxation parameters for a variety of nucleic acids can be fit very well by a motional model involving long-range bending motions of the helix on a microsecond timescale and internal motions consisting of rotational wobbling about P-O bonds on a nanosecond timescale.[105] In the approach discussed here, ^{31}P NMR spin-lattice relaxation measurements were performed in an effort to characterize the effects of single-site phosphate ethylation on the backbone structure and dynamics of the parent octamer.

NMR spin-lattice relaxation is the process by which a nuclear spin system gives up energy to its surroundings in order to reestablish an equilibrium distribution of spin populations within the available energy states and the spin-lattice relaxation time, T_1, characterizes the time constant for this process. The spin-lattice relaxation rate of a nucleus will, in general, depend on the coupling of the nuclear moment with local magnetic fields that fluctuate at the site of the nucleus as a result of molecular motion.[106] Hence, the spin-lattice relaxation time will depend upon the types and strengths of the local magnetic fields present and the frequencies at which these fields fluctuate. If the fluctuations in the local magnetic fields arise from Brownian motion, then the intensity distribution (the spectral density) of the frequency components is given by

$$J(\omega) \propto \frac{1}{1+\omega^2\tau_c^2} \qquad [5.1]$$

where τ_c is a time (the correlation time) characteristic of the random motion, ω is the Larmor frequency and $J(\omega)$ is the spectral density function.[107] The spin-lattice relaxation rate

$(1/T_1)$ can then be written as

$$\frac{1}{T_1} = R_1 = H^2 J(\omega) \qquad [5.2]$$

where H represents the strength of the particular local magnetic field giving rise to relaxation and $J(\omega)$ is the spectral density function whose exact form depends on the mechanism of spin-lattice relaxation.[46]

For nucleic acids, ^{31}P T_1 relaxation, in the absence of paramagnetic impurities, is dominated by dipolar relaxation with neighboring protons ($T_{1,DD}$) and by chemical shift anisotropy relaxation ($T_{1,CSA}$).[94] The total spin-lattice relaxation rate is a function of the mechanisms giving rise to the relaxation as follows

$$\frac{1}{T_1} = \frac{1}{T_{1,DD}} + \frac{1}{T_{1,CSA}}. \qquad [5.3]$$

The phosphorus relaxation rate $(1/T_{1,DD})$ as a result of dipolar coupling with a proton is given by

$$\frac{1}{T_{1,DD}} = \frac{1}{10}\gamma_H^2\gamma_P^2\hbar^2\frac{\tau_c}{r_{HP}^6}\left[\frac{1}{1+(\omega_H-\omega_P)^2\tau_c^2} + \frac{3}{1+\omega_P^2\tau_c^2} + \frac{6}{1+(\omega_P+\omega_H)^2\tau_c^2}\right] \qquad [5.4]$$

where r_{HP} is the internuclear distance between the phosphorus and the proton, τ_c is the correlation time for fluctuations of the ^{31}P-^1H internuclear vector relative to the reference frame of the magnetic field and ω_P and ω_H are the ^{31}P and ^1H Larmor frequencies.[108] If the phosphorus spin is relaxed by several proton spins then the total phosphorus dipolar relaxation rate will be given by the sum of the individual dipolar relaxation rates. The dipolar contribution to ^{31}P spin-lattice relaxation in nucleic acids arises mainly from the three deoxyribose ring protons, H3', H5' and H5", which are at an average distance of 2.8 Å away

from the phosphorus atom in B-DNA.[94]

Anisotropy of the chemical shift results when the local field experienced by a nucleus depends on the orientation of the molecule in the magnetic field. The orientational dependence of the chemical shift is given by a shielding tensor and the chemical shift anisotropy (CSA) is generally defined by the three principal components of the shielding tensor, σ_{11}, σ_{22} and σ_{33}, as follows:

$$\Delta\sigma = \sigma_{11} - \frac{1}{2}(\sigma_{22} + \sigma_{33}). \qquad [5.5]$$

Chemical shift anisotropy can contribute to the spin-lattice relaxation of a nucleus since the chemical shift and, hence, the local field at the nucleus changes in magnitude as the molecule changes its orientation in the magnetic field. The relaxation rate due to chemical shift anisotropy is given by the equation

$$\frac{1}{T_{1,CSA}} = \frac{1}{15}\omega_P^2 \Delta\sigma^2 \tau_c \left[\frac{2}{1+\omega_P^2\tau_c^2}\right] \qquad [5.6]$$

where $\Delta\sigma$ represents the CSA for the phosphorus nucleus, τ_c is the correlation time for fluctuations of the shielding tensor relative to the reference frame of the magnetic field and ω_P is the ^{31}P Larmor frequency.[108] Linear correlations have been established between the principal values of the shielding tensor for phosphodiesters and the P-O bond lengths and O-P-O bond angles and, therefore, the ^{31}P chemical shift anisotropy is capable of providing conformational information on the sugar-phosphate backbone of nucleic acids.[109]

The functional form of the dipolar and CSA relaxation rates indicates that the ^{31}P spin-lattice relaxation time is sensitive to motional fluctuations that occur at frequencies near

the Larmor frequency (~10^9 s^{-1}) and is thus sensitive to motions on the nanosecond timescale (τ_c~ 10^{-9} s). Because the dipolar and CSA relaxation mechanisms depend on both structural parameters (e.g., the ^{31}P-^{1}H distances for dipolar relaxation and the shielding tensor for CSA relaxation), and dynamical parameters (the effective correlation time), T_1 measurements can potentially yield important information regarding the structure and dynamics of the sugar-phosphate backbone in DNA oligomers.

In order to correlate changes in the ^{31}P spin-lattice relaxation time with either changes in backbone dynamics or changes in the backbone structure, it is important to determine the contribution of each relaxation mechanism to the overall spin-lattice relaxation rate. The spin-lattice relaxation times of the ^{31}P resonances for the parent, R-A(et)A and S-A(et)A duplexes are on the order of 1.5 s (Table 5.5) and suggest that CSA relaxation is the predominant relaxation mechanism for these oligomers at 5°C and a phosphorus Larmor frequency of 162 MHz. This conclusion is supported by comparing the calculated relaxation time for phosphorus under dipolar relaxation and under CSA relaxation. The dipolar and CSA relaxation times can be estimated assuming that reorientation of the dipolar vectors and the shielding tensors relative to the reference frame of the magnetic field coincides with isotropic rotation of the duplex with a correlation time predicted by the Stokes-Einstein-Debye Law for rotational motion. The correlation time, τ_c, for molecular rotation of the octamer duplex can be estimated, using plausible dimensions for the oligomer, by the Stokes-Einstein-Debye Equation

$$\tau_c = \frac{4\pi\eta r^3}{3kT} \qquad [5.7]$$

where r is the mean radius of the molecule and • and T are the viscosity and temperature, respectively, of the solution. Assuming isotropic tumbling of the octamer duplex with a mean

helical diameter of ~24 Å in water at 5°C, the τ_c is estimated to be 3 ns. A correlation of 3 ns is in good agreement with that determined by Broido *et al*[22] who showed that the 2D NOESY intensities of the parent octamer at 25°C could be fit reasonably well by a motional model of the duplex with a rotational correlation time of ~1.7-2.7 ns. Furthermore, a correlation time of 3 ns ($\omega\tau_c$=7.5) is consistent with the sign of the 2D NOESY intensities measured for the octamer duplexes which indicate that the rotational correlation time for these duplexes is in the slow motional regime (see Section 5.5). The chemical shift anisotropy in phosphodiesters has been reported to be approximately 125 ppm,[109] a value which would lead to a ^{31}P spin-lattice relaxation time of 1.6 s using Equation 5.6 and a 3 ns correlation time. Dipolar relaxation of phosphorus by three protons with an internuclear distance (r_{HP}) of 2.8 Å would lead to a ^{31}P spin-lattice relaxation time of approximately 15 s using Equation 5.4 and a 3 ns correlation time. Clearly, the experimental ^{31}P spin-lattice relaxation times given in Table 5.5 are consistent with CSA relaxation as the principal relaxation mechanism when $T_{1,DD}$ and $T_{1,CSA}$ are calculated using the motional model and the structural parameters, $\Delta\sigma$ and r_{HP}, described above. Although the motional and structural models used to calculate the relaxation times were not experimentally validated in this study, two additional experimental observations are consistent with the ^{31}P T_1 being dominated by CSA relaxation: (*i*) the 202 Mz ^{31}P T_1's for the parent duplex were ~10% smaller than the 162 MHz ^{31}P T_1's and (*ii*) a ^{31}P-{^{1}H} NOE was not observed during broadband irradiation of the proton resonances in the parent duplex. If ^{31}P-^{1}H dipolar interactions dominated the ^{31}P spin-lattice relaxation rate obtained for these duplexes, then one would expect to find behavior opposite to that actually observed; namely that either the ^{31}P T_1's would increase with the magnetic field strength (expected when $T_{1,DD}$ dominates T_1 in the slow motional regime) or that a ^{31}P-{^{1}H} NOE would be observed (expected when $T_{1,DD}$ dominates T_1 in the extreme narrowing regime). Therefore, the ^{31}P spin-lattice relaxation times for the parent, S-A(et)A and R-A(et)A duplexes measured at 162 MHz and 5°C are consistent with CSA relaxation as the dominant relaxation mechanism.

Plotted in Figure 5.28 are the experimentally determined ^{31}P T_1's as a function of residue number for the parent, S-A(et)A and R-A(et)A duplexes. Overall, the ^{31}P T_1's within each duplex are quite similar, consistent with CSA relaxation characterized by isotropic reorientation of each duplex with a rotational correlation time in good agreement with that predicted by the Stokes-Einstein-Debye Law for rotational motion. It should be noted that internal motions characterized by subnanosecond correlation times, such as rotational jumps about the phosphodiester bonds, may also contribute to the effective correlation time and therefore be partly responsible for the observed ^{31}P T_1's. Examination of Figure 5.28 indicates that the only ^{31}P T_1 difference between the parent and the two analogues that is larger than the experimental error belongs to the phosphotriester of the S-A(et)A analogue which exhibits a ~25% increase in T_1 relative to the other ^{31}P T_1's. The difference in ^{31}P T_1 observed for this phosphorus could result from differences in the contribution of internal motions to the relaxation rate and/or differences in the CSA for the phosphorus nucleus at this site. However, without prior knowledge regarding the size of the $\Delta\sigma$ term for the phosphorus nucleus, it cannot be determined if a change in CSA and/or a change in the effective correlation time is responsible for this difference. If a change in CSA is solely responsible for the increased ^{31}P T_1 of the S-A(et)A phosphotriester, then a similar increase in ^{31}P T_1 would be expected for the R-A(et)A phosphotriester, contrary to what is actually observed. Although esterification of a phosphotriester results in an upfield shift of the corresponding ^{31}P resonance (Section 5.1.3), its affect on the ^{31}P CSA remains unclear since a direct determination of the ^{31}P shielding tensor of a phosphotriester has not yet been reported. It should be of interest to measure the ^{31}P chemical shift tensors of model phosphotriesters either by measuring the orientational dependence of the ^{31}P chemical shift using single crystals or by recovering the chemical shift tensor elements from the ^{31}P spectra of noncrystalline samples using the magic-angle sample spinning technique.[110] The information obtained from such measurements would allow a direct comparison of the ^{31}P $\Delta\sigma$ term in a phosphodiester with that in a phosphotriester and thus enable a prediction of the

extent to which the ^{31}P spin-lattice relaxation rate of a phosphodiester will change upon phosphate alkylation.

Overall, the similarity of the ^{31}P spin-lattice relaxation rates between the parent, and both the S-A(et)A and R-A(et)A analogues indicates that the presence of a phosphotriester lesion does not appear to strongly influence the structural (i.e., ^{31}P CSA) and the dynamical (i.e., internal motions that would influence the effective correlation time) features of the sugar phosphate backbone as determined by the $T_{1,CSA}$ parameter. If phosphate alkylation alters the structural and/or dynamical features along the sugar-phosphate backbone, then the ^{31}P spin-lattice relaxation rates are insensitive to such changes. However, the increased ^{31}P T_1 of the S-A(et)A phosphotriester may indicate changes to the structural and/or dynamical features of the sugar-phosphate backbone at the phosphotriesterification site. Therefore, it should be of interest to obtain information on the ^{31}P CSA in phosphotriesters so that observed changes in the ^{31}P T_1's of phosphotriesters can be more clearly understood.

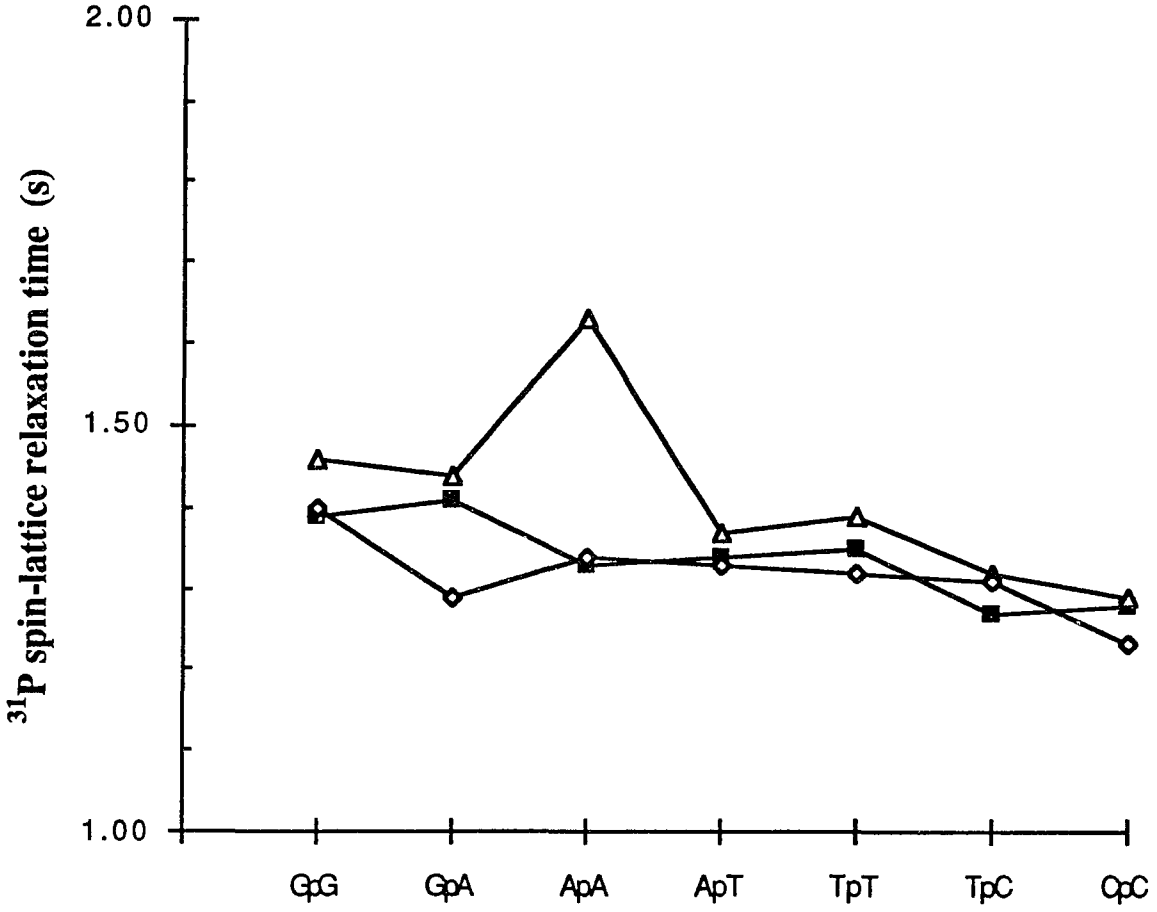


Figure 5.28 The ^{31}P spin-lattice relaxation time as a function of residue number for {d[GGAATTCC]}$_2$ (■), ($\underline{S}_p$,$\underline{S}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ (▵) and ($\underline{R}_p$,$\underline{R}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ (◇). The error for each spin-lattice relaxation time is estimated to be ±0.07 s based on a 5% error in the T_1 measurement (see Figure 5.27).

5.5 NOE Measurements of the Nonexchangeable Protons and Conversion into Distance Data

Two-dimensional NOESY spectra for the parent, S-A(et)A and R-A(et)A duplexes were recorded at five mixing times (2, 50, 100, 200, and 500 ms) and processed as described in Section 3.2.6. Two-dimensional NOESY contour plots for the parent duplex at four mixing times are shown in Figure 5.29. T_1 noise, noise in the ω_1 direction (along the vertical of the 2D contour plots of Figure 5.29), is observed in the spectrum obtained at each mixing time. The t_1 noise artifacts are due to instrument limitations (e.g., less than optimal control of the sample temperature) and are most pronounced for sharp resonances (i.e., T-CH_3 and A-H2).[78] It should be noted from Figure 5.29 that the 2D NOESY diagonal and cross peaks have the same phase. The sign of the NOE changes from positive to negative when the effective correlation time for dipolar relaxation is in the slow motional regime ($\omega\tau_c>1$) and in the 2D NOESY experiment this results in the diagonal and cross peaks having the same phase.[78] Therefore, the effective correlation time for dipolar relaxation in these duplexes must be in the slow motional regime ($\tau_c>0.4$ ns). This observation is consistent with the 3 ns correlation time derived in Section 5.4.

The quantification of crosspeak volumes, not in the regions of t_1 noise, was achieved by two-dimensional integration of each crosspeak using the FTNMR processing program (Section 3.3.1). The volume measurements were used to estimate the interproton distances in each duplex based on Equation 2.24. As discussed in Section 4.3, spin diffusion can modify the intensity of a NOESY crosspeak and becomes more significant as the mixing period for the 2D NOESY experiment increases. The determination of interproton distances using Equation 2.24 is only valid if the τ_m used to acquire the 2D NOESY spectrum is sufficiently short. Therefore, interproton distance determinations using Equation 2.24 require, as a precondition, the estimation of the spin diffusion limits which may be different for different protons pairs. Figure 5.30 shows the buildup of some of the NOE crosspeak volumes as a function of the experimental mixing time. The different slopes of the NOE buildup curves arise from the locally different geometries of the spin pairs (i.e., protons that are closer will

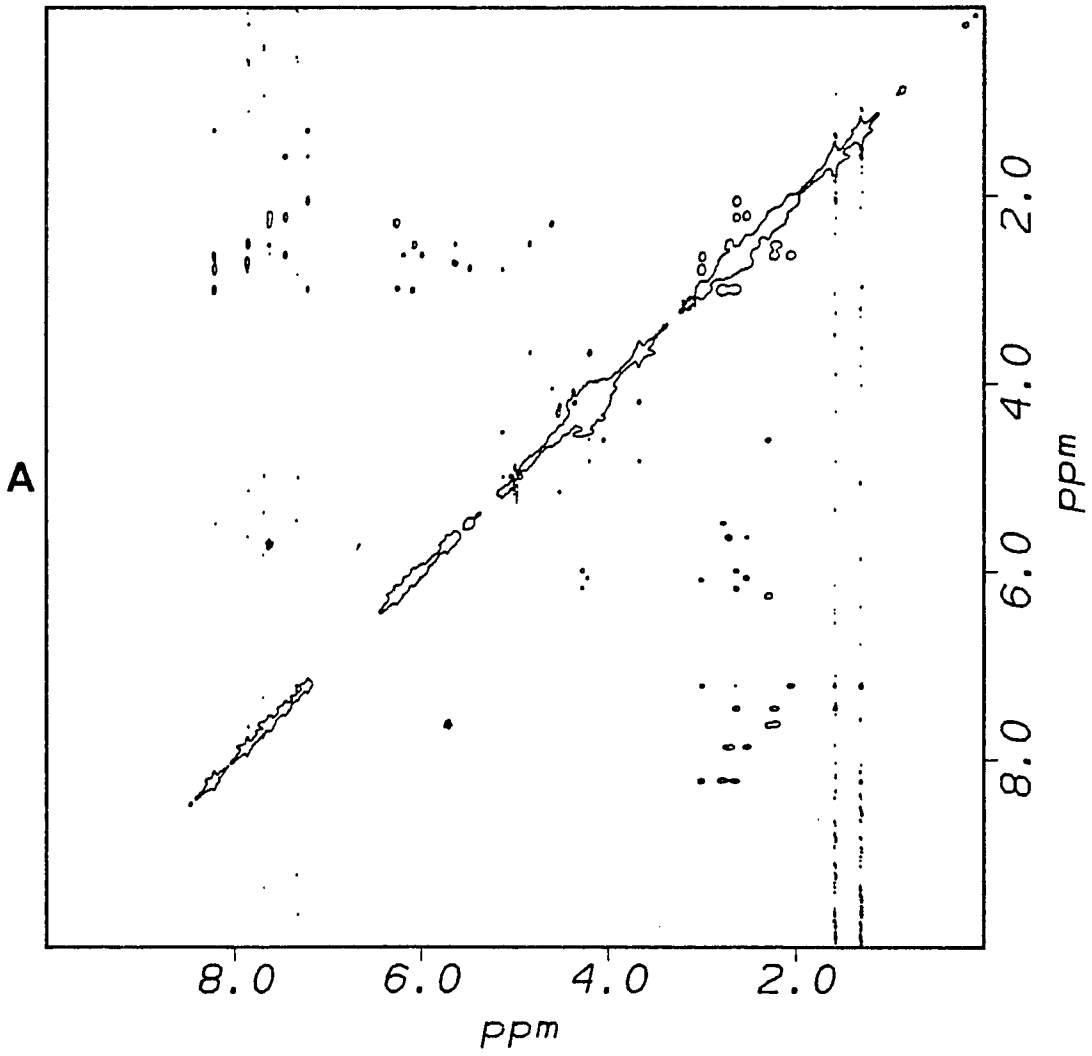
give rise to a steeper slope relative to protons that are farther apart). It is observed that even at 100 ms the curves become nonlinear, indicating contributions from spin diffusion in the system to the measured crosspeak volumes at this mixing time. Therefore, the determination of interproton distances for these oligomers without errors due to spin diffusion would require the measurement of crosspeak volumes at mixing times less than 100 ms. However the very low intensity of most NOESY crosspeaks at short mixing times (<100 ms) prevents their accurate quantitation and therefore most interproton distance determinations were based on crosspeak volumes measured using either a 50 ms, 100 ms or 200 ms mixing time (*vide infra*).

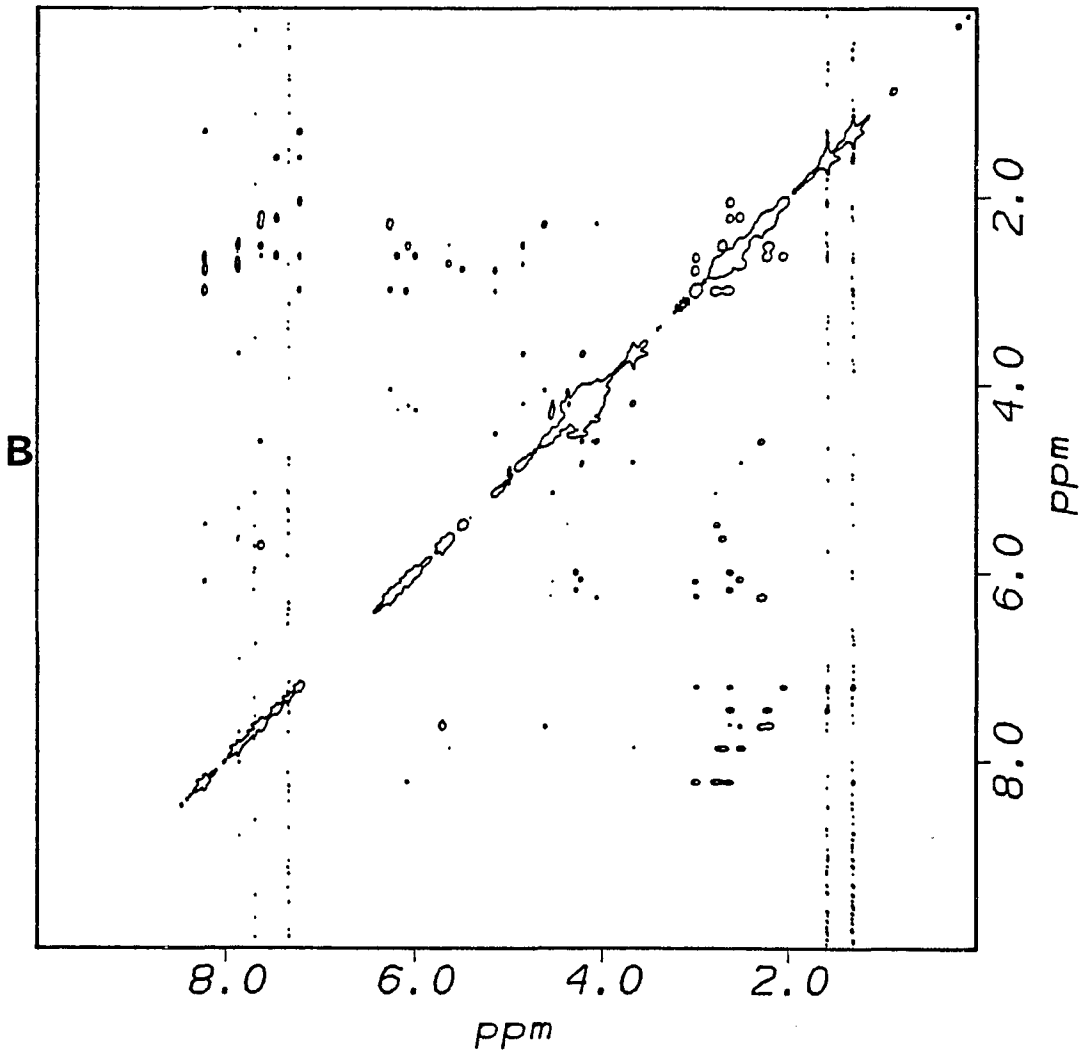
In DNA, there are three fixed distances, T-H6↔T-CH_3 (r=3.0 Å), C-H5↔C-H6 (r=2.5 Å), and the deoxyribose H2'↔H2" (r=1.8 Å) which can be used as standards for distance measurements from crosspeak volumes.[15] The T-H6↔T-CH_3 reference was used as a yardstick for all NOEs involving CH_3 protons because CH_3↔H internuclear vectors generally have a different correlation time than H↔H internuclear vectors due to rotation of the methyl group which is superimposed on the tumbling of the molecule as a whole. The H2'↔H2" reference was used as a yardstick for the few NOEs arising at 50 ms because the C-H5↔C-H6 volumes could not be accurately determined at this mixing time. Only crosspeaks between protons separated by ≤2.8 Å appear at this mixing time (based on an average integrated volume of ~15 for the H2'↔H2" crosspeaks and an integrated volume of ~1 as the noise threshold). The C-H5↔C-H6 reference was used as a yardstick for all NOEs arising at 100 and 200 ms. The H2'↔H2" volume becomes nonlinear after 50 ms and thus was not employed as a reference for NOEs arising at 100 and 200 ms. Only crosspeaks between protons separated by ≤3.8 Å and ≤4.5 Å are seen at mixing times of 100 and 200 ms, respectively (based on an average integrated volume of ~7 and ~15 for the C-H5↔C-H6 crosspeaks at 100 and 200 ms, respectively, and an integrated volume of ~0.5 as the noise threshold at these mixing times).

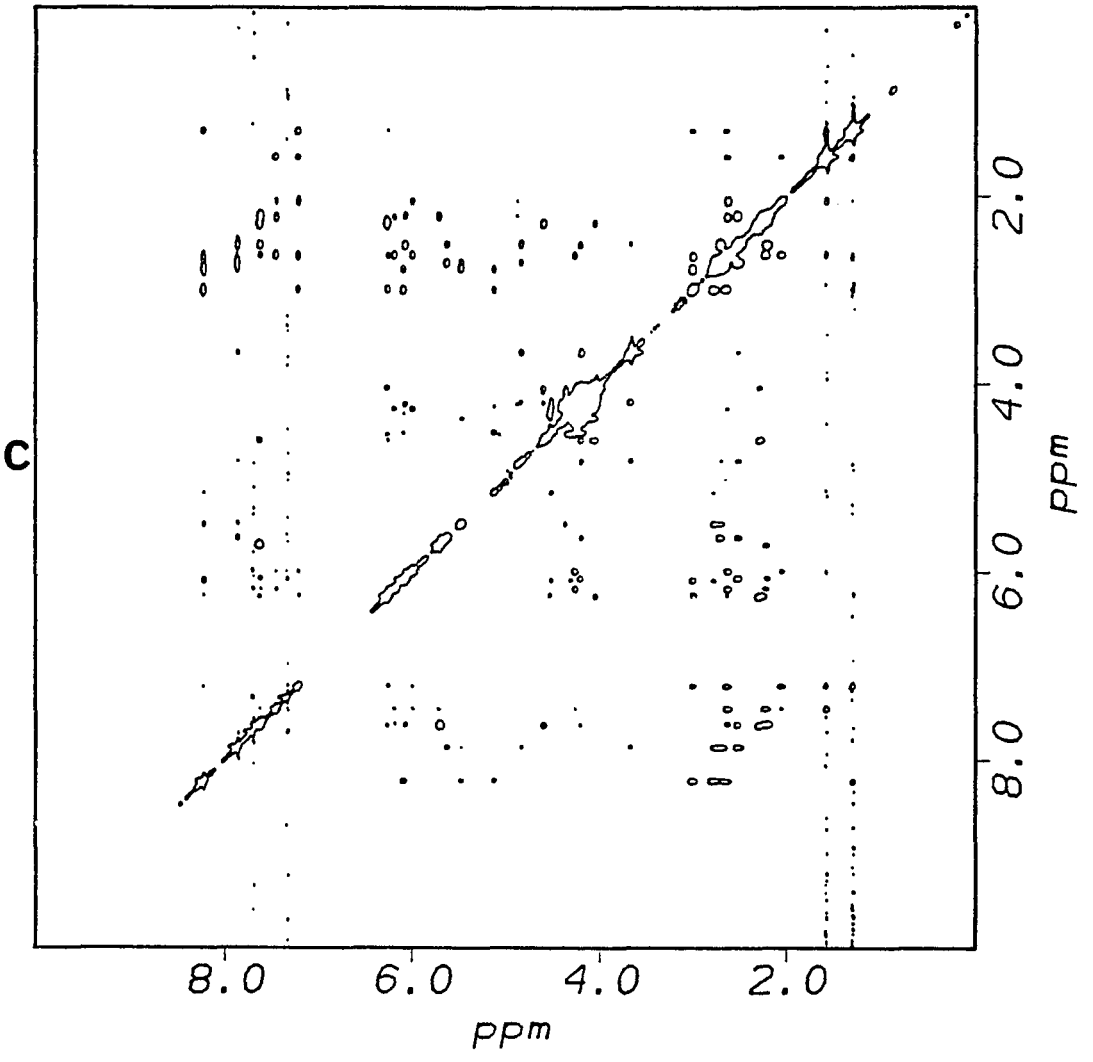
The interproton distances measured, according to the above criteria, between the nonexchangeable protons in each duplex are listed in Table 5.6. The interproton distances within right-handed nucleic acid structures dictate that the majority of ^{1}H-{^{1}H} NOESY crosspeaks will occur exclusively within individual mononucleotides or between sequentially neighboring nucleotides.[15] Therefore, Table 5.6 has been constructed in a convenient format

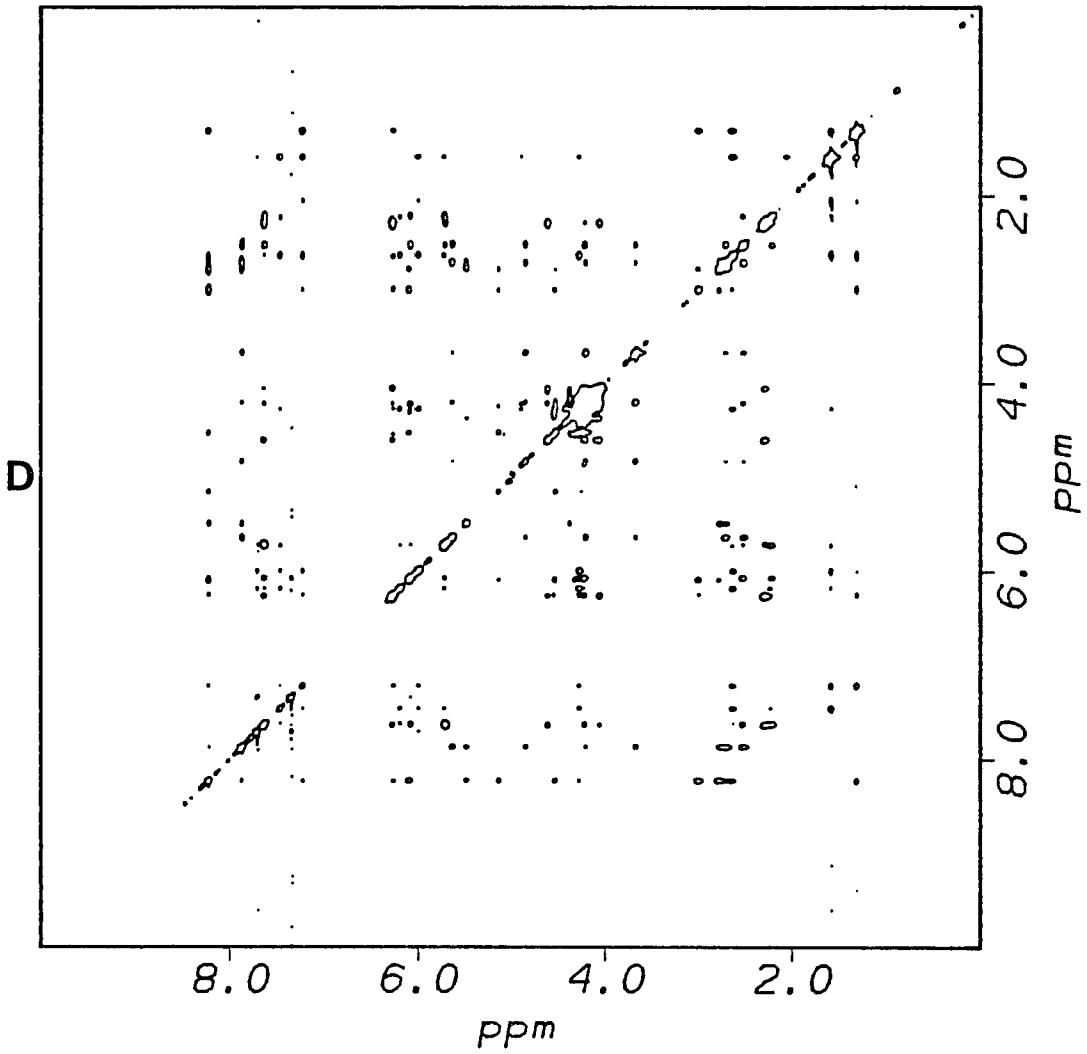
where only sequential neighbors (the (*n-1*) and (*n+1*) residues) to residue *n* are included in the table of measured interproton distances for protons found within residue *n*. Distances that could not be determined because of the lack of a crosspeak are represented by the symbol "-" and those distances that could not be determined because of crosspeak overlap are represented by the symbol "**" in Table 5.6. Distances that were determined but may be inaccurate due to partial crosspeak overlap are represented by the value of the distance followed by the symbol "#". An NOE crosspeak that is observable only at a mixing time of 500 ms (the longest mixing time used in this study) could arise as the result of spin diffusion or could arise because the two protons giving rise to the crosspeak are separated by a distance ≥4.5 Å. Because distances could not be accurately determined under such conditions any crosspeak that exists only at 500 ms is represented by the symbol "5-". The errors associated with the calculated interproton distances given in Table 5.6 will be discussed below.

Figure 5.29 400 MHz ^{1}H-{^{1}H} 2D NOESY spectra of {d[GGAATTCC]}$_2$ acquired at the four mixing times 50 ms (**A**), 100 ms (**B**), 200 ms (**C**) and 500 ms (**D**). Sample conditions: 5°C; in D_2O with 10 mM sodium phosphate, pH 7.0.


A
8.0
6.0
4.0
2.0
ppm
2.0
4.0
6.0
8.0
ppm


B
8.0
6.0
4.0
2.0
ppm
2.0
4.0
6.0
8.0
ppm


C
8.0
6.0
4.0
2.0
ppm
2.0
4.0
6.0
8.0
ppm


D
8.0
6.0
4.0
2.0
ppm
2.0
4.0
6.0
8.0
ppm

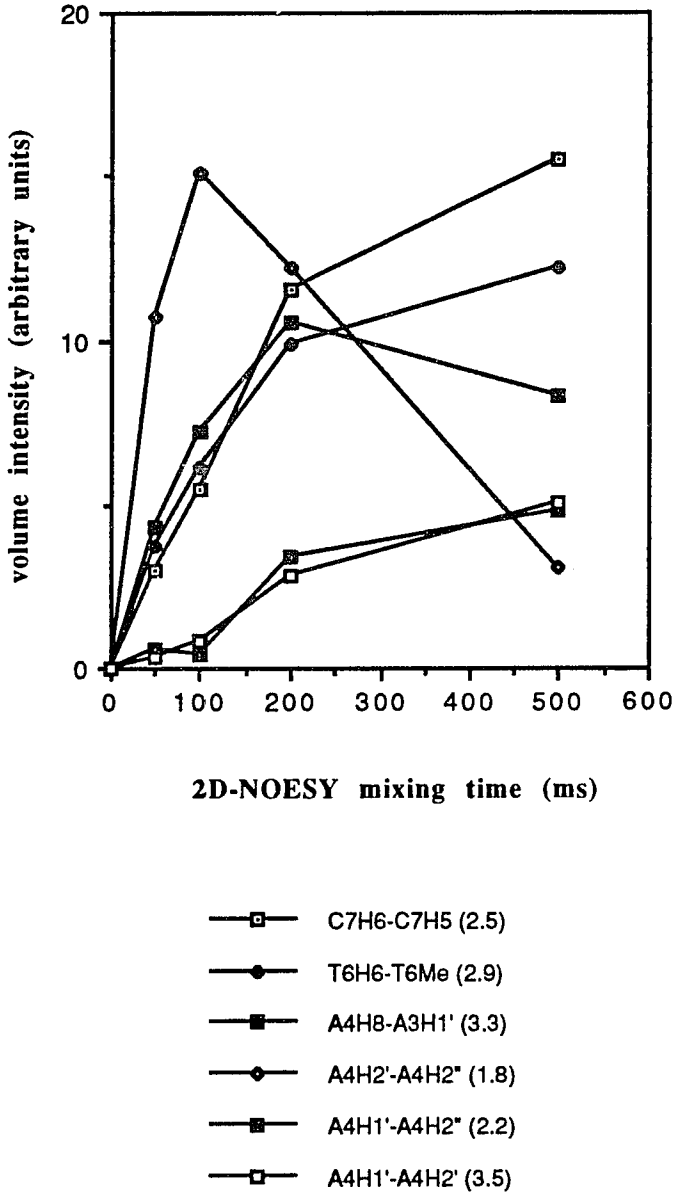


Figure 5.30 Plot of the measured 2D NOESY crosspeak volume as a function of the experimental mixing time for the indicated crosspeaks in ($\underline{R}_p$,$\underline{R}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ at 5°C. The values in parentheses are the calculated interproton distances.

Table 5.6 Calculated interproton distance data for the nonexchangeable protons in $\{d[GGAATTCC]\}_2$ (**P**), $(\underline{S}_p,\underline{S}_p)$-$\{d[GGA(ethyl)ATTCC]\}_2$ (**S**) and $(\underline{R}_p,\underline{R}_p)$-$\{d[GGA(ethyl)ATTCC]\}_2$ (**R**) at 5°C. See text for an explanation of the symbols found in this table.

	G1	1'			2'			2"			3'			4'			8					
G1		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P																					
	S																					
	R																					
2'	P	3.1																				
	S		••																			
	R			3.2																		
2"	P	2.5#			1.8r																	
	S		2.2#			1.8r																
	R			2.6#			1.8r															
3'	P	4.1			3.1			3.5#														
	S		4.1			3.1			3.4#													
	R			3.9			2.5			2.9#												
4'	P	4.3			3.2#			3.8#			4.0											
	S		3.8			4.1#			3.8#			3.4										
	R			3.8			3.5#			4.0#			3.8									
8	P	••			2.3			••			3.9#			5-								
	S		••			2.0			••			3.6#			4.8							
	R			3.4#			2.2			2.8#			3.4			5-						
	P																					
	S																					
	R																					

	G1	1'			2'			2"			3'			4'			8					
G2		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P	4.4			-			••			-			5-			••					
	S		4.4			-			••			-			5-			••				
	R			••			-			••			-			5-			••			
2'	P	••			••			••			••			••			••					
	S		••			••			••			••			••			••				
	R			••			••			••			••			••			••			
2"	P	-			-			••			-			5-			••					
	S		••			-			••			-			••			••				
	R			-			-			••			-			-			5-			
3'	P	-			-			••			-			-			-					
	S		-			-			-			-			-			-				
	R			-			-			••			-			-			-			
4'	P	5-			-			••			-			••			••					
	S		-			-			••			-			••			-				
	R			5-			-			••			-			••			-			
8	P	3.2#			••			2.2#			••			••			••					
	S		2.9#			••			2.1#			••			••			••				
	R			3.5			4.0			2.5#			4.6			5-			••			
	P																					
	S																					
	R																					

	G2	1'			2'			2"			3'			4'			8					
G2		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P																					
	S																					
	R																					
2'	P	3.4																				
	S		2.8#																			
	R			3.1#																		
2"	P	2.3#			1.8r																	
	S		2.5			1.8r																
	R			2.3			1.8r															
3'	P	-			4.8#			4.3#														
	S		-			-			-													
	R			4.2			3.0#			3.1#												
4'	P	3.6			5.1#			4.4#			4.0											
	S		3.7			4.4#			5.1			-										
	R			3.5			4.0			4.7			3.5									
8	P	4.1#			2.2#			2.6#			-			5-								
	S		4.1			2.1#			2.2#			-			5-							
	R			3.6			2.5#			2.8			4.3			5-						
	P																					
	S																					
	R																					

	G2	1'			2'			2"			3'			4'			8					
A3		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P	-			-			••			-			-			-					
	S		5-			-			-			-			-			-				
	R			5-			-			••			-			-			-			
2'	P	••			••			••			••			••			••					
	S		••			-			-			-			-			5-				
	R			5-			-			••			-			-			5-			
2"	P	-			••			••			-			5-			-					
	S		••			-			••			-			-			-				
	R			••			••			••			••			-			5-			
3'	P	-			-			••			-			-			-					
	S		••			-			••			-			••			-				
	R			5-			-			••			-			-			-			
4'	P	5-			-			••			-			••			-					
	S		5-			-			5-			-			••			-				
	R			5-			-			••			••			••			-			
8	P	3.4			••			2.1#			-			5-			4.6					
	S		3.2			4.1#			2.6			-			5-			4.4				
	R			3.6			4.9			2.2#			5-			5-			4.1			
2	P	-			-			-			-			-			-					
	S		-			-			-			-			-			-				
	R			-			••			-			-			-			-			

	A3	1'			2'			2"			3'			4'			8			2		
A3		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P																					
	S																					
	R																					
2'	P	3.2#																				
	S		2.9#																			
	R			2.7#																		
2"	P	2.3#			1.8r																	
	S		2.1			1.8r																
	R			2.4#			1.8r															
3'	P	4.4			2.9#			3.0#														
	S		3.4			2.6#			2.6#													
	R			3.3			2.4			2.5#												
4'	P	3.6#			4.3#			3.8#			3.0											
	S		3.5#			3.8			3.7			3.5										
	R			••			3.8#			3.6#			3.4									
8	P	••			2.1#			••			3.7#			4.2#								
	S		4.2			2.1#			2.4			3.6			4.6							
	R			4.3			2.0			••			3.4			5-						
2	P	••			-			-			-			-			-					
	S		••			-			-			-			-			-				
	R			-			-			-			-			-			-			

	A3	1'			2'			2"			3'			4'			8			2		
A4		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P	••			-			••			-			-			••			5-		
	S		••			••			-			-			••			-			5.4	
	R			5-			-			-			-			••			-			••
2'	P	4.0			••			••			-			-			••			-		
	S		••			••			-			-			5-			-			-	
	R			••			-			-			-			••			-			-
2"	P	••			••			••			••			••			••			-		
	S		••			••			••			••			••			••			-	
	R			-			-			-			-			••			••			-
3'	P	-			-			••			-			••			••			-		
	S		-			-			-			-			-			-			-	
	R			-			-			-			-			••			••			-
4'	P	••			••			••			••			••			••			-		
	S		5-			-			5-			••			••			-			-	
	R			••			••			••			••			••			••			-
8	P	3.5#			••			2.1#			••			••			••			-		
	S		3.2			••			2.2			3.9			4.8			4.0			-	
	R			3.3			3.3			2.5			4.2			••			4.0			••
2	P	-			-			-			-			-			-			3.7		
	S		-			-			-			-			-			-			3.3	
	R			-			-			-			-			-			-			4.2

	A4	1'			2'			2"			3'			4'			8			2		
A4		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P																					
	S																					
	R																					
2'	P	3.7#																				
	S		3.3#																			
	R			3.5																		
2"	P	2.2#			1.8r																	
	S		2.3#			1.8r																
	R			2.2			1.8r															
3'	P	5-			3.5#			3.3#														
	S		-			-			5-													
	R			-			3.8			3.4												
4'	P	3.5			5-			3.8#			3.8											
	S		3.5#			5-			4.1			3.9										
	R			3.7#			3.7#			3.6#			3.4									
8	P	4.4#			2.2#			••			5-			4.9#								
	S		4.2			2.3			••			-			5.0							
	R			3.7			2.4			3.1			5-			4.8						
2	P	5-			-			-			-			-			-					
	S		5-			-			-			-			-			-				
	R			5-			-			-			-			-			-			
	A4	1'			2'			2"			3'			4'			8			2		
T5		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P	4.4#			••			5-			-			-			-			4.6		
	S		5-			-			-			-			-			••			4.7	
	R			5-			-			5-			-			-			-			4.1
2'	P	4.4			••			••			-			-			-			-		
	S		4.9			-			-			-			-			-			-	
	R			5-			-			5-			-			-			-			-
2"	P	••			••			••			••			••			••			-		
	S		-			••			-			-			-			-			-	
	R			-			••			-			••			-			••			-
3'	P	-			••			-			-			-			-			-		
	S		-			••			-			-			-			-			-	
	R			5-			••			-			-			-			-			-
4'	P	5.6			••			3.8			-			••			••			-		
	S		4.4			5-			5-			-			-			-			5-	
	R			5-			-			5-			-			••			-			-
■	P	4.5			5.0#			4.7			5-			5-			4.0			-		
	S		4.8			3.8			3.8			-			5-			3.1			-	
	R			5.5			4.2			4.3			5-			5-			4.4			-
6	P	4.4			3.3#			2.3			-			5-			4.2			5-		
	S		4.2			4.4			2.4			-			-			4.0			5-	
	R			3.5			4.0			2.5			-			5-			4.1			••

	TS	1'			2'			2"			3'			4'			m			6		
T5		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P																					
	S																					
	R																					
2'	P	2.8																				
	S		2.8																			
	R			2.9																		
2"	P	2.3#			1.8r																	
	S		2.2#			1.8r																
	R			2.3#			1.8r															
3'	P	5.2			2.9			3.6#														
	S		-			3.2			3.8#													
	R			4.1			2.5			3.1#												
4'	P	3.2#			3.7#			3.7#			4.2#											
	S		3.0#			3.8#			3.5#			3.7#										
	R			3.0			3.7#			3.8#			3.6#									
m	P	5-			5-			••			-			••								
	S		5-			5-			-			-			••							
	R			5-			5-			5-			5-			-						
6	P	4.1			2.1			3.3#			5-			5-			2.9r					
	S		4.1			2.1			3.4#			-			5-			2.9r				
	R			4.2			2.1			4.0			4.0			5-			2.9r			

	TS	1'			2'			2"			3'			4'			m			6		
T6		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P	4.9			-			••			-			••			-			-		
	S		5-			-			••			-			••			-			-	
	R			4.2			-			••			••			4.5			-			-
2'	P	5-			-			••			••			••			-			-		
	S		-			-			••			••			••			-			-	
	R			-			-			••			••			5-			-			-
2"	P	••			••			••			••			••			••			••		
	S		••			••			••			••			••			-			••	
	R			••			••			••			••			••			-			••
3'	P	-			-			-			-			-			-			-		
	S		-			-			-			-			-			-			-	
	R			-			-			-			-			5-			-			-
4'	P	••			••			••			••			••			••			••		
	S		••			••			••			••			••			••			••	
	R			5-			4.0			••			••			••			-			-
m	P	4.3			5.0			4.6#			4.0#			5-			3.2			3.7		
	S		4.2			4.0			4.0#			5-			5-			3.9			3.7	
	R			5.4			3.8			4.2#			5.2#			5-			3.7			3.4
6	P	4.3			3.1			2.2#			5-			••			5-			4.0		
	S		3.7			2.7			2.1#			-			••			5-			3.8	
	R			3.9			2.8			2.2#			4.5#			5-			5-			4.0

	T6	1'			2'			2"			3'			4'			m			5		
T6		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P																					
	S																					
	R																					
2'	P	3.2#																				
	S		3.3#																			
	R			3.1#																		
2"	P	2.4#			1.8r																	
	S		2.2#			1.8r																
	R			2.4			1.8r															
3'	P	-			-			-														
	S		-			-			-													
	R			4.3			3.4			3.7#												
4'	P	3.5			4.0#			3.7#			-											
	S		3.3#			3.8#			3.5#			-										
	R			3.0			3.8#			3.2#			4.1#									
m	P	5-			5-			••			-			••								
	S		5-			5-			••			-			••							
	R			5-			5-			••			-			••						
6	P	4.6			2.1#			••			-			4.5#			2.9r					
	S		3.7			2.1#			••			-			4.7#			2.9r				
	R			4.0			2.1#			••			4.7			5-			2.9r			
	T6	1'			2'			2"			3'			4'			m			6		
C7		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P	4.1			••			••			-			••			-			-		
	S		3.6			••			5-			-			5-			-			-	
	R			••			••			••			-			••			-			-
2'	P	••			••			••			-			••			••			••		
	S		••			••			••			-			••			••			••	
	R			••			••			••			-			••			••			••
2"	P	5-			••			••			-			-			-			-		
	S		-			••			••			-			••			-			-	
	R			••			••			••			-			••			-			-
3'	P	-			••			••			-			••			••			-		
	S		-			••			••			-			••			••			-	
	R			5-			••			••			-			••			••			••
4'	P	••			••			4.2#			-			••			-			-		
	S		••			••			4.0#			-			••			-			-	
	R			••			••			••			••			••			••			••
5	P	5.0#			••			3.6#			-			-			4.1#			4.4		
	S		4.0			••			3.6			-			-			4.3			3.9	
	R			3.5			••			4.1			-			-			4.7			4.3
6	P	3.6			••			2.7#			-			5-			5-			4.6		
	S		3.7			••			2.4			-			5-			5-			4.2	
	R			3.8			••			2.7			-			••			5-			3.9

	C7	1'			2'			2"			3'			4'			5			6		
C7		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P																					
	S																					
	R																					
2'	P	3.5#																				
	S		3.0#																			
	R			3.1#																		
2"	P	2.2			1.8r																	
	S		2.1			1.8r																
	R			2.4			1.8r															
3'	P	5.4			3.1#			3.6														
	S		-			3.3#			3.7#													
	R			3.7			2.7#			3.2#												
4'	P	3.4#			3.9#			3.2#			3.6#											
	S		3.6			3.5#			4.1#			4.1#										
	R			3.1#			3.8#			3.5#			3.7#									
5	P	••			••			••			••			••								
	S		5-			••			4.9#			-			-							
	R			5-			••			5-			5-			-						
6	P	••			1.8#			••			4.5#			4.0#			2.5r					
	S		••			1.8#			••			5-			3.8#			2.5r				
	R			3.4			1.9			3.2			3.9			5-			2.5r			

	C7	1'			2'			2"			3'			4'			5			6		
C8		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P	••			••			5-			-			••			••			••		
	S		••			••			••			-			5-			-			••	
	R			4.3			••			5-			-			5-			-			-
2'	P	5-			••			••			-			5.0#			••			••		
	S		5-			••			5-			-			4.7#			-			••	
	R			••			••			5-			-			4.3#			••			••
2"	P	••			••			••			-			5.0#			••			••		
	S		••			••			••			-			4.7#			-			••	
	R			••			••			••			-			4.3#			••			••
3'	P	5-			-			-			-			3.8			••			••		
	S		••			-			-			-			••			••			••	
	R			-			-			-			-			••			-			-
4'	P	5-			-			4.6			-			••			-			••		
	S		5-			-			3.9			-			••			-			••	
	R			5-			-			4.1			-			••			-			-
5	P	4.0#			••			3.8#			5-			••			••			••		
	S		5-			3.8#			••			-			••			••			••	
	R			5-			3.4#			3.9#			5-			5-			••			••
6	P	3.7#			••			2.4#			••			••			••			••		
	S		3.5#			••			2.3#			••			••			••			••	
	R			3.9			••			3.3			5-			-			••			••

	CS	1'			2'			2"			3'			4'			5			6		
CS		P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R	P	S	R
1'	P																					
	S																					
	R																					
2'	P	••																				
	S		••																			
	R			••																		
2"	P	••			1.8r																	
	S		••			1.8r																
	R			••			1.8r															
3'	P	3.5			••			••														
	S		3.5			••			••													
	R			3.8			••			••												
4'	P	3.5			3.8#			3.8#			3.1											
	S		3.2			3.6#			3.6#			2.9										
	R			3.1			3.6#			3.6#			3.2									
5	P	5-			••			••			••			-								
	S		5-			••			••			••			••							
	R			••			••			••			4.3			-						
6	P	3.9#			••			••			3.1#			4.2#			2.5r					
	S		3.9#			••			••			3.1#			4.1#			2.5r				
	R			3.7			••			••			3.7			4.1#			2.5r			

The available ^{1}H and ^{31}P chemical shift data on the parent, S-A(et)A and R-A(et)A duplexes indicates that the presence of a phosphotriester lesion induces conformational changes in the octamer duplex and the conformational changes are localized to the site of phosphate alkylation. Furthermore, the ^{1}H chemical shift differences between the parent and its ethylated analogues suggest that the conformational changes are more pronounced for the R-A(et)A analogue which contains a phosphotriester ethyl group oriented into the major groove of the duplex. In an effort to obtain a more detailed understanding of the conformational changes that occur upon single-site phosphate ethylation, spatial data was obtained for each oligomer in the form of NOE-derived interproton distances as described above.

Refinement methods such as restrained energy minimization and restrained molecular dynamics are frequently used to optimize macromolecular structures based on NOE-derived interproton distances.[70,71,72] In Chapter 4, a series of computer simulations were discussed to evaluate the effectiveness of the restrained energy minimization and restrained molecular dynamics methods to refine a nucleic acid structure based on NOE-derived interproton distance data. The set of interproton distance data used in the computer simulations was chosen from a target structure (a dodecamer DNA duplex) and mimicked that derived from typical 2D NOESY experiments on nucleic acids; only interproton distances less than 5 Å were included in the simulations and all interproton distances involving the 5',5" sugar protons and the exocyclic amino and imino protons were excluded from the simulations. It was demonstrated from the simulations that the refinement methods, when used with the set of interproton distance data, generated structures which accurately reproduced some of the conformational features of the target structure such as deoxyribose ring conformations and base-related helicoidal parameters. However, some of the conformational features of the sugar-phosphate backbone were not accurately reproduced in the simulations due to the lack of interproton distance constraints that serve to define the backbone structure (i.e., interproton distances involving the 5',5" sugar protons). The results of the simulations suggest

that the accuracy with which the refinement methods can define the conformational features of a nucleic acid structure depends on both the quality (i.e., error limits for the interproton distances) and the quantity of the distance data used in the refinement procedure. Hence, the results of the simulations discussed in Chapter 4 can be used to judge the effectiveness of a refinement procedure to generate structures from a given set of experimental distance data only if the quality and quantity of the experimental data is similar to that used in the simulations.

Table 5.7 provides some of the interproton distances within mononucleotides and between sequentially neighboring nucleotides for a classical B-DNA structure; only those interproton distances that are accessible from typical 2D NOESY experiments are included in this table. Examination of Table 5.7 indicates that ~15 intraresidue interproton distances and ~8 (*n-1*) and ~8 (*n+1*) interresidue interproton distances would typically be accessible for each nucleotide assuming the structure adopts a B-DNA conformation. The dodecamer duplex used for the computer simulations in Chapter 4 had 717 interproton distance constraints, in good agreement with the ~712 interproton distance constraints expected for a dodecamer duplex using the above criteria. An octamer duplex, assuming it adopts a B-DNA conformation, should have ~240 intraresidue and ~224 interresidue accessible interproton constraints. Examination of Table 5.6 indicates that ~220 intraresidue and ~130 interresidue interproton distances were determined from the available 2D NOESY data on each octamer duplex in this study. Therefore, although most of the accessible intraresidue interproton distances have been determined from the available 2D NOESY data on the octamer duplexes, many of the normally accessible interresidue interproton distances could not be determined, largely because of crosspeak overlap (represented by the symbol "**" in Table 5.6). Furthermore, many of the interresidue interproton distances that could be determined from the 2D NOESY data possess inaccuracies due to partial crosspeak overlap (represented by the symbol "#" in Table 5.6). The accuracy of an octamer structure generated using the refinement methods described in Chapter 4 would be uncertain because the set of interproton distance data for the octamer duplexes is incomplete compared to the set of interproton

distance constraints used in the computer simulations. Because a direct comparison cannot be made with the computer simulations given in Chapter 4, attempts were not made to use the set of interproton distance data obtained on the octamer duplexes in a refinement procedure to generate structures for the parent and its ethylated analogues.

Although incomplete, the interproton distance data obtained on the octamer duplexes can be examined to determine the gross conformational features of each duplex and to determine any structural perturbations that occur to the parent duplex upon phosphate alkylation. Plotted in Figure 5.31 are the (*n*)H4'↔(*n*)H2" and (*n*)H1'↔(*n+1*)H8/H6 interproton distances as a function of residue number for the parent, S-A(et)A and R-A(et)A duplexes. The (*n*)H4'↔(*n*)H2" interproton distance is sensitive to the puckering mode of the deoxyribose ring. Nucleotide structure analysis indicates that two puckering modes, C3'-*endo* and C2'-*endo*, are generally found for deoxyribose rings in nucleic acids; the C3'-*endo* conformation is often preferred in A-DNA and the C2'-*endo* conformation is often preferred in B-DNA.[14] The (*n*)H4'↔(*n*)H2" interproton distance for the C3'-*endo* and C2'-*endo* conformations of the deoxyribose ring is ~2.8 Å and ~4.0 Å, respectively.[15] Examination of Figure 5.31 indicates that the (*n*)H4'↔(*n*)H2" interproton distances for the three octamer duplexes are between 3.6 - 4.6 Å and thus show a reasonable correspondence with the interproton distance expected for a C2'-*endo* deoxyribose ring conformation. However, it is important to consider the error limits for the (*n*)H4'↔(*n*)H2" interproton distances calculated from the 2D NOESY data. As discussed in Chapter 4, spin diffusion can modify the intensity of a 2D NOESY crosspeak leading to errors in the calculated interproton distance. In general, the apparent value of an NOE-derived interproton distance will become closer to that of the reference distance as the extent of spin diffusion increases. Because the C-H5↔C-H6 interproton distance (r=2.5 Å) was used as a reference for calculating the (*n*)H4'↔(*n*)H2" distances, the presence of spin diffusion would result in a systematic underestimation of each (*n*)H4'↔(*n*)H2" interproton distance. The CORMA calculations discussed in Chapter 4 demonstrated that an NOE-derived distance of 3.5-4.5 Å would be underestimated relative to the actual distance by up to 0.5 Å. The

CORMA calculations used conditions (i.e., effective correlation time and 2D NOESY experimental mixing times) that are similar to those found in this study. Because the calculated (*n*)H4'↔(*n*)H2" interproton distances are at the upper limit of that observed for the (*n*)H4'↔(*n*)H2" distance as a function of the pseudorotation phase angle (the (*n*)H4'↔(*n*)H2" distance reaches a maximum of ~4.2 Å for the C2'-*endo*-C3'-*exo* furanose conformation), it is unlikely that the (*n*)H4'↔(*n*)H2" interproton distances were underestimated to the extent predicted from the CORMA calculations. It is likely that internal motions attenuate the effects of spin diffusion at the mixing times employed here (*vide infra*) and, therefore, the 0.5 Å error limit predicted by the CORMA calculations overestimates the error for the NOE-derived (*n*)H4'↔(*n*)H2" interproton distances. However, it is unclear why some of the calculated (*n*)H4'↔(*n*)H2" interproton distances exceed the 4.2 Å upper limit imposed by the geometry of the deoxyribose ring.

The (*n*)H1'↔(*n+1*)H8/H6 interproton distance is sensitive to the global conformation adopted by a nucleic acid (i.e., the A-, B- and Z-DNA conformations) as well as the sequence-dependent local conformation of the duplex structure.[14] The (*n*)H1'↔(*n+1*)H8/H6 interproton distance for A- and B-DNA is 4.5 Å and 3.5 Å, respectively.[15] In general, the experimental (*n*)H1'↔(*n+1*)H8/H6 interproton distances for the three octamer duplexes are between 3.0-4.2 Å and thus show a reasonable correspondence with the interproton distance expected for a standard B-DNA structure. The changes in the (*n*)H1'↔(*n+1*)H8/H6 interproton distance as a function of residue number for each duplex suggest that sequence-dependent variations are present in each structure. The CORMA calculations discussed in Chapter 4 predicted that the NOE-derived (*n*)H1'↔(*n+1*)H8/H6 interproton distances would be underestimated by up to 0.5 Å because of effects due to spin diffusion. If the (*n*)H1'↔(*n+1*)H8/H6 interproton distances are underestimated by 0.5 Å then some of (*n*)H1'↔(*n+1*)H8/H6 distances (i.e., those toward the 3' end of each duplex) show a correspondence with the interproton distance expected for a standard A-DNA structure. However it has been shown by Broido *et al.*[22] that the 2D NOESY intensities for the parent

duplex can be fit using a B-DNA structural model. This result suggests that the (*n*)H1'↔(*n+1*)H8/H6 interproton distances, as was the case for the (*n*)H4'↔(*n*)H2" interproton distances, were not underestimated to the extent determined by the CORMA calculations in Chapter 4. The likely reason that spin-diffusion is not affecting the interproton distances to the extent predicted by the CORMA calculations is that internal motions exist in the duplexes that attenuate the effects of spin-diffusion (the CORMA calculations assumed a rigid molecule with no internal motion). In fact, the results of calculations by Broido *et al.*[22] on the parent molecule indicated that a reasonable fit between the experimental and calculated 2D NOESY intensities were obtained using a motional model for the duplex with nanosecond correlation times (1.5-3 ns) but better fits to some of the 2D NOESY data were obtained by assuming local variations in mobility within the duplex structure. Therefore, a better estimate of the effects of spin-diffusion on interproton distances determined from 2D NOESY experiments could be obtained by performing CORMA calculations assuming a motional model that has internal motions that can influence the interproton cross relaxation rates.

Although the above factors limit the interproton distance data obtained in this study from being used to exactly define the conformational features of each duplex, the interproton distances determined for the parent and its ethylated analogues can be compared to determine if single-site phosphate ethylation affects the structure of the parent duplex. Examination of Figure 5.31 indicates that no significant differences are observed for the (*n*)H4'↔(*n*)H2" interproton distances between the parent and its ethylated analogues. This suggests that single-site phosphate ethylation does not alter, within the resolution limits of the experimental data, the deoxyribose ring conformations in the octamer duplex. Examination of Figure 5.31 indicates that the only (*n*)H1'↔(*n+1*)H8/H6 interproton distance difference observed between the parent and its ethylated analogues is for the A4 residue of the R-A(et)A duplex. The A4-H1'↔T5-H6 interproton distance of the R-A(et)A duplex is ~0.8 Å smaller than the corresponding distance in the parent and S-A(et)A duplexes. Examination of Table 5.6 indicates that the A4-H1'↔T5-CH_3 and A4-H1'↔A4-H8 interproton distances of the R-A(et)A

duplex are ~0.8 Å larger and ~0.6 Å smaller, respectively, than the corresponding distances in the parent and S-A(et)A duplexes. The observed interproton distance differences suggest that the duplex structure in the R-A(et)A analogue adopts a different conformation in the vicinity of the A4 residue than the corresponding region in the parent and the S-A(et)A analogue.

Overall, the interproton distance differences between the parent duplex and the S-A(et)A and R-A(et)A duplexes complement the observed ^{1}H and ^{31}P chemical shift differences discussed in Section 5.1. These data support a model in which a local structural perturbation occurs in the octamer duplex upon single-site phosphate ethylation. This structural perturbation is more pronounced for the $\underline{R}_p$-analogue which contains a phosphotriester ethyl group oriented into the major groove of the helix. In order to resolve the more subtle conformational changes that might occur to the octamer duplex upon single-site phosphate ethylation, it would be necessary to improve both the quality and quantity of interproton distance data. This may be accomplished by decreasing the effects of spin-diffusion and by decreasing the amount of crosspeak overlap in the 2D NOESY data. The effects of spin-diffusion can be minimized by collecting data at shorter mixing times or by using 3D NMR techniques (i.e., 3D NOESY-NOESY) that delineate crosspeaks arising from direct cross relaxation pathways from those that arise from indirect cross relaxation (i.e., spin diffusion) pathways.[111] Crosspeak overlap can be reduced by performing the 2D NOESY experiments at higher field strengths so that the ^{1}H resonances are more clearly resolved or by employing 3D NMR techniques that can provide greatly improved spectral resolution.[112] The additional interproton distance data accessible by these techniques should provide a more detailed understanding of the effects of phosphate alkylation on the structure of DNA duplexes in solution.

Table 5.7 Intraresidue ((*n*)↔(*n*)) and interresidue ((*n*)↔(*n-1*)) interproton distance data for a classical B-DNA structure. Only those interproton distances that are obtained from a typical 2D NOESY experiment are included in this table; all interproton distances larger than 5 Å and all interproton distances involving the 5',5" sugar protons and the exocyclic amino and imino protons were excluded from the table. The distances were obtained from an energy-minimized B-DNA structure.

	n	1'	2'	2"	3'	4'	5'	5"		8	2	5	6	m
n														
1'														
2'		3.0												
2"		2.3	1.8											
3'		3.9	2.4	2.7										
4'		3.6	3.9	4.1	2.7									
5'														
5"														
8		3.9	2.1	3.7	4.2	4.9								
2														
5			4.1											
6		3.7	1.8	3.4	3.8	4.6						2.5		
m		4.8	4.7										2.9	
	n	1'	2'	2"	3'	4'	5'	5"		8	2	5	6	m
n-1														
1'			4.1		4.9	4.2				3.5		4.3	3.5	
2'										3.8		3.6	4.0	3.4
2"			3.7							2.1		2.7	2.2	2.9
3'										4.6			4.6	
4'														
5'														
5"														
8												3.9		3.8
2		4.6									3.6			
5			4.1									4.7		4.7
6										4.8		3.9		3.8
m														

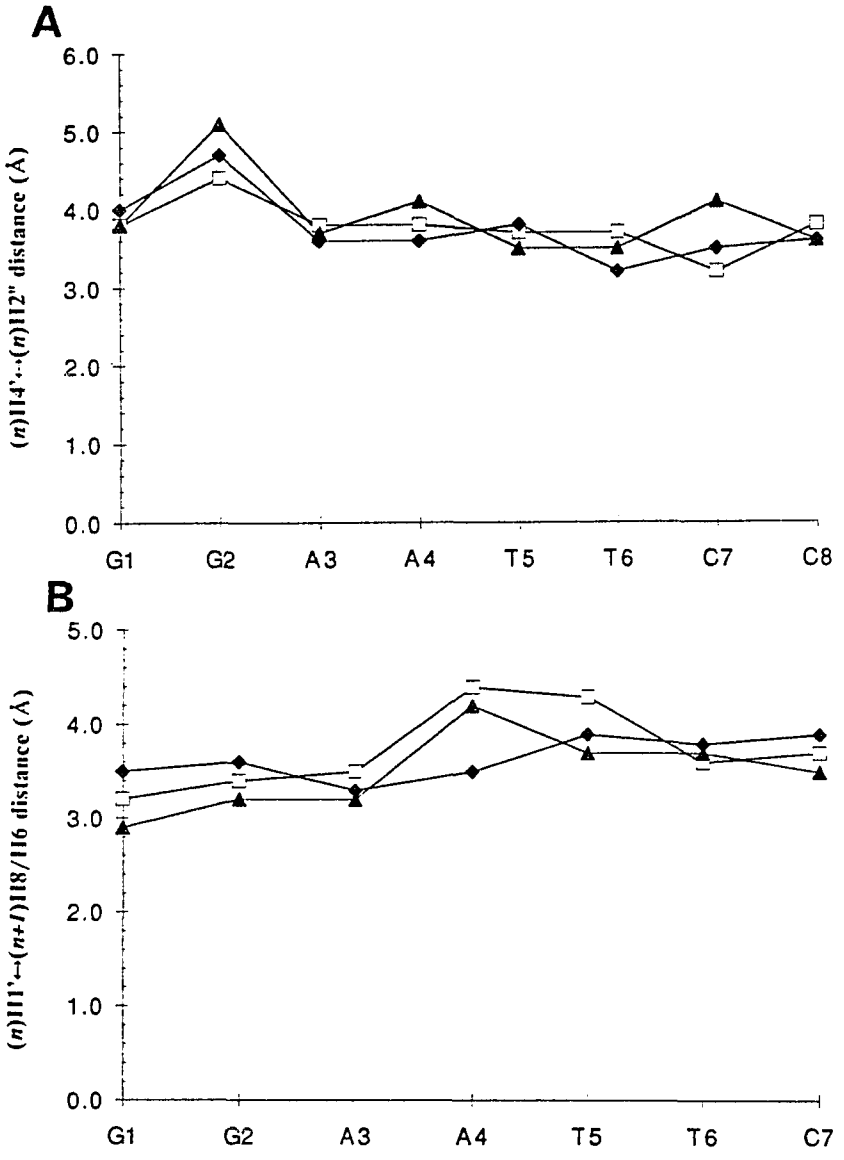


Figure 5.31 Plot of the NOE-derived (*n*)H4'↔(*n*)H2" (A) and (*n*)H1'↔(*n+1*)H8/H6 (B) interproton distance as a function of residue number for {d[GGAATTCC]}$_2$ (□), ($\underline{S}_p$,$\underline{S}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ (▲) and ($\underline{R}_p$,$\underline{R}_p$)-{d[GGA(ethyl)ATTCC]}$_2$ (♦). The interproton distances are taken from Table 5.6.

CONCLUSIONS

A number of NMR measurements have been made on the octamer duplex, $\{d[GGAATTCC]\}_2$, and its two single-site ethylated analogues, S-A(et)A and R-A(et)A. The ^1H chemical shift differences between the native octamer and its ethylated analogues indicate that structural perturbations exist in the S-A(et)A and R-A(et)A duplexes that are localized to the region of phosphate ethylation. The ^1H chemical shift differences, most notably those of the A4-H1' resonance, are more pronounced for the R-A(et)A duplex which contains a phosphotriester group oriented into the major groove of the helix. Interproton distances obtained from 2D NOESY experiments indicate that the A4-H1'↔T5-H6/CH_3 and A4-H1'↔A4-H8 distances in the R-A(et)A duplex are different than those in the parent and S-A(et)A duplexes. The R-A(et)A duplex also has a melting temperature that is ~5°C lower than that found for the parent and S-A(et)A duplexes.

The ^{31}P chemical shift differences between the native octamer and its ethylated analogues were confined to the phosphodiester groups adjacent to the phosphotriester lesion suggesting that any possible structural changes to the sugar-phosphate backbone are localized to the region of the phosphotriester lesion. However, the lack of data capable of providing structural information on the sugar-phosphate backbone makes it difficult to correlate the ^{31}P chemical shift changes observed for the ethylated analogues relative to the native octamer to a specific conformational change in the phosphodiester backbone induced by phosphate ethylation. Furthermore, only small ^{31}P T_1 differences and base-pair lifetime differences were observed between the three octamer duplexes, suggesting that these parameters are largely insensitive to structural perturbations to the parent duplex which result from phosphate ethylation.

The chemical shift, interproton distance, and melting temperature differences observed between the parent and the S-A(et)A and R-A(et)A analogues are consistent with a model in which a phosphotriester lesion induces a local structural change to the duplex in the vicinity of the phosphotriester resulting in destabilization of the duplex. Furthermore,

the structural changes and consequent destabilization are more pronounced for a duplex containing a phosphotriester lesion whose alkyl group is oriented into the major groove of the helix. The similarity of ^{1}H and ^{31}P chemical shifts and interproton distances between the three duplexes in regions that do not flank the phosphotriesterification site suggests that structural perturbations induced by a phosphotriester lesion are not propagated along the backbone of the octamer duplex.

The local, stereochemically dependent structural change induced by single-site phosphate ethylation as evidenced by the NMR data obtained in this study complements the *in vitro* studies of Miller *et al.*[10] The *in vitro* study indicates that a phosphotriester lesion in DNA interferes with template-directed DNA polymerization and that the observed reduction in template activity depends on the stereochemistry of the phosphotriester moiety. Thus it appears that a structural effect is at least partly responsible for the biochemical and biological effects induced by phosphate alkylation. The structural change deduced by the NMR data obtained in this study is consistent with the studies of Miller *et al.*,[17] Weinfeld *et al.*[18] and Kan *et al.*[19] on phosphate ethylated dinucleotides. These studies indicate that a phosphate ethyl group decreases the degree of base-base stacking and that the degree of destacking is more pronounced for an alkyl group oriented toward the bases of the dinucleotide. However, the local, structural change found herein contrast with the NMR studies of Lawrence *et al.*[23] and Broido *et al.*[24], in which single-site isopropyl phosphotriester lesions were found to induce long range structural perturbations to an octamer duplex. This suggests that the nature of the structural perturbation also depends on the steric bulk of the phosphotriester moiety.

Long-range structural perturbations that might exist in the R-A(et)A and S-A(et)A duplexes could not be detected using the interproton distance data obtained here. However, it should be noted that a full set of interproton distances was not obtained on the octamer duplexes due to the lack of resonance assignments for the sugar 5' and 5" protons and because crosspeak overlap in the 2D NOESY spectra precluded the determination of many interresidue interproton distances. Indeed, the results of the computer simulations presented in this study

underscore the necessity of obtaining a full set of high-quality distance data in order to determine the local conformational features of a nucleic acid. Consequently, experimental conditions and methods should be chosen such that the complete set of interproton distance data required for the elucidation of long range structural consequences of DNA phosphate ethylation is obtained.

Appendix I: CURVES-derived conformational parameters for (I) the starting structure (B-form (d[CGCGAATTCGCG])$_2$) and (II) the target structure (the Dickerson dodecamer).

I. CURVES-derived conformational parameters for the starting structure:

```
--------------------------------
|B| Global Base-Axis Parameters |
--------------------------------
```

1st strand	Xdisp (dx)	Ydisp (dy)	Inclin (eta)	Tip (theta)	Bc	Tc
1) CYT 1	-0.59	0.01	-4.54	-0.59	3	0
2) GUA 2	-0.59	0.01	-4.52	-0.61	1	13
3) CYT 3	-0.59	0.01	-4.54	-0.60	3	-13
4) GUA 4	-0.59	0.01	-4.55	-0.59	1	10
5) ADE 5	-0.61	0.01	-4.57	-0.59	2	18
6) ADE 6	-0.61	0.00	-4.58	-0.60	2	24
7) THY 7	-0.58	0.00	-4.55	-0.60	4	-24
8) THY 8	-0.58	0.00	-4.55	-0.62	4	-18
9) CYT 9	-0.59	0.00	-4.56	-0.61	3	-10
10) GUA 10	-0.59	0.01	-4.53	-0.61	1	13
11) CYT 11	-0.59	0.01	-4.53	-0.60	3	-13
12) GUA 12	-0.59	0.01	-4.53	-0.62	1	0

2nd strand	Xdisp (dx)	Ydisp (dy)	Inclin (eta)	Tip (theta)	Bc	Tc
1) GUA 24	-0.59	0.01	-4.53	-0.62	1	0
2) CYT 23	-0.59	0.01	-4.53	-0.60	3	-13
3) GUA 22	-0.59	0.01	-4.53	-0.61	1	13
4) CYT 21	-0.59	0.00	-4.56	-0.61	3	-10
5) THY 20	-0.58	0.00	-4.55	-0.62	4	-18
6) THY 19	-0.58	0.00	-4.55	-0.60	4	-24
7) ADE 18	-0.61	0.00	-4.58	-0.60	2	24
8) ADE 17	-0.61	0.01	-4.57	-0.59	2	18
9) GUA 16	-0.59	0.01	-4.55	-0.59	1	10
10) CYT 15	-0.59	0.01	-4.54	-0.60	3	-13
11) GUA 14	-0.59	0.01	-4.52	-0.61	1	13
12) CYT 13	-0.59	0.01	-4.54	-0.59	3	0

```
-------------------------------------
|C| Global Base pair-Axis Parameters |
-------------------------------------
```

Duplex	Xdisp (dx)	Ydisp (dy)	Inclin (eta)	Tip (theta)	Bc	Tc
1) C 1-G 24	-0.59	0.00	-4.53	0.01	3	0
2) G 2-C 23	-0.59	0.00	-4.53	-0.01	1	13
3) C 3-G 22	-0.59	0.00	-4.54	0.01	3	-13
4) G 4-C 21	-0.59	0.00	-4.55	0.01	1	10
5) A 5-T 20	-0.60	0.00	-4.56	0.02	2	18
6) A 6-T 19	-0.59	0.00	-4.57	0.00	2	24
7) T 7-A 18	-0.59	0.00	-4.57	0.00	4	-24
8) T 8-A 17	-0.60	0.00	-4.56	-0.02	4	-18
9) C 9-G 16	-0.59	0.00	-4.55	-0.01	3	-10
10) G 10-C 15	-0.59	0.00	-4.54	-0.01	1	13
11) C 11-G 14	-0.59	0.00	-4.53	0.01	3	-13
12) G 12-C 13	-0.59	0.00	-4.53	-0.01	1	0
Average:	-0.59	0.00	-4.55	0.00		

```
--------------------------------
|D| Global Base-Base Parameters |
--------------------------------
```

Duplex	Shear (Sx)	Stretch (Sy)	Stagger (Sz)	Buckle (kappa)	Propel (omega)	Opening (sigma)	Bc	Tc
1) C 1-G 24	0.00	0.01	-0.02	-0.01	-1.21	-0.36	3	0

2) G 2-C 23	0.00	0.01	-0.02	0.01	-1.21	-0.36	1 13
3) C 3-G 22	0.00	0.01	-0.02	-0.01	-1.21	-0.34	3 -13
4) G 4-C 21	0.00	0.01	-0.02	0.01	-1.21	-0.35	1 10
5) A 5-T 20	-0.03	0.01	-0.02	-0.02	-1.21	-0.75	2 18
6) A 6-T 19	-0.03	0.01	-0.02	-0.02	-1.21	-0.76	2 24
7) T 7-A 18	0.03	0.01	-0.02	0.02	-1.21	-0.76	4 -24
8) T 8-A 17	0.03	0.01	-0.02	0.02	-1.21	-0.75	4 -18
9) C 9-G 16	0.00	0.01	-0.02	-0.01	-1.21	-0.35	3 -10
10) G 10-C 15	0.00	0.01	-0.02	0.01	-1.21	-0.34	1 13
11) C 11-G 14	0.00	0.01	-0.02	-0.01	-1.21	-0.36	3 -13
12) G 12-C 13	0.00	0.01	-0.02	0.01	-1.21	-0.36	1 0
Average:	0.00	0.01	-0.02	0.00	-1.21	-0.49	

|E| Global Inter-Base Parameters |

1st strand	Shift (Dx)	Slide (Dy)	Rise (Dz)	Tilt (tau)	Roll (rho)	Twist (Omega)	Dc
2) C 1/G 2	0.00	0.00	3.38	0.01	-0.02	35.99	8
3) G 2/C 3	0.00	0.00	3.38	-0.01	0.01	36.02	5
4) C 3/G 4	0.00	0.00	3.38	0.01	-0.01	35.98	8
5) G 4/A 5	-0.03	-0.01	3.38	-0.02	0.01	35.61	2
6) A 5/A 6	0.00	0.00	3.38	0.01	0.00	35.99	4
7) A 6/T 7	0.03	0.00	3.38	0.02	0.02	36.39	7
8) T 7/T 8	0.00	0.00	3.38	-0.01	0.00	36.01	-4
9) T 8/C 9	0.00	0.00	3.38	-0.02	0.00	36.00	-2
10) C 9/G 10	0.00	0.00	3.38	0.00	-0.01	35.99	8
11) G 10/C 11	0.00	0.00	3.38	-0.01	0.01	36.00	5
12) C 11/G 12	0.00	0.00	3.38	0.01	-0.02	35.99	8

2nd strand	Shift (Dx)	Slide (Dy)	Rise (Dz)	Tilt (tau)	Roll (rho)	Twist (Omega)	Dc
2) G 24/C 23	0.00	0.00	3.38	-0.01	0.02	35.99	8
3) C 23/G 22	0.00	0.00	3.38	0.01	-0.01	36.00	5
4) G 22/C 21	0.00	0.00	3.38	0.00	0.01	35.99	8
5) C 21/T 20	0.00	0.00	3.38	0.02	0.00	36.00	-2
6) T 20/T 19	0.00	0.00	3.38	0.01	0.00	36.01	-4
7) T 19/A 18	-0.03	0.00	3.38	-0.02	-0.02	36.39	7
8) A 18/A 17	0.00	0.00	3.38	-0.01	0.00	35.99	4
9) A 17/G 16	0.03	0.01	3.38	0.02	-0.01	35.61	2
10) G 16/C 15	0.00	0.00	3.38	-0.01	0.01	35.98	8
11) C 15/G 14	0.00	0.00	3.38	0.01	-0.01	36.02	5
12) G 14/C 13	0.00	0.00	3.38	-0.01	0.02	35.99	8

|F| Global Inter-Base pair Parameters |

Duplex	Shift (Dx)	Slide (Dy)	Rise (Dz)	Tilt (tau)	Roll (rho)	Twist (Omega)	Dc
2) C 1/G 2	0.00	0.00	3.38	0.00	-0.02	35.99	8
3) G 2/C 3	0.00	0.00	3.38	0.00	0.01	36.01	5
4) C 3/G 4	0.00	0.00	3.38	0.00	-0.01	35.99	8
5) G 4/A 5	-0.01	0.00	3.38	0.00	0.00	35.81	2
6) A 5/A 6	0.00	0.00	3.38	0.01	0.00	36.00	4
7) A 6/T 7	0.00	0.00	3.38	0.00	0.02	36.39	7
8) T 7/T 8	0.00	0.00	3.38	-0.01	0.00	36.00	-4
9) T 8/C 9	0.01	0.00	3.38	0.00	0.00	35.81	-2
10) C 9/G 10	0.00	0.00	3.38	0.00	-0.01	35.99	8
11) G 10/C 11	0.00	0.00	3.38	0.00	0.01	36.01	5
12) C 11/G 12	0.00	0.00	3.38	0.00	-0.02	35.99	8

Average:	0.00	0.00	3.38	0.00	0.00	36.00	

|G| Local Inter-Base Parameters |

1st strand	Shift (Dx)	Slide (Dy)	Rise (Dz)	Tilt (tau)	Roll (rho)	Twist (Omega)	Dc
2) C 1/G 2	0.03	-0.62	3.34	0.38	-2.82	35.88	8
3) G 2/C 3	0.03	-0.62	3.34	0.36	-2.79	35.91	5
4) C 3/G 4	0.03	-0.62	3.34	0.38	-2.82	35.87	8
5) G 4/A 5	0.00	-0.63	3.35	0.35	-2.78	35.50	2
6) A 5/A 6	0.03	-0.64	3.34	0.37	-2.82	35.89	4
7) A 6/T 7	0.06	-0.63	3.34	0.40	-2.83	36.28	7
8) T 7/T 8	0.03	-0.62	3.34	0.37	-2.81	35.90	-4
9) T 8/C 9	0.03	-0.62	3.34	0.36	-2.81	35.89	-2
10) C 9/G 10	0.03	-0.62	3.34	0.38	-2.82	35.88	8
11) G 10/C 11	0.03	-0.62	3.34	0.36	-2.79	35.90	5
12) C 11/G 12	0.03	-0.62	3.34	0.38	-2.82	35.88	8

2nd strand	Shift (Dx)	Slide (Dy)	Rise (Dz)	Tilt (tau)	Roll (rho)	Twist (Omega)	Dc
2) G 24/C 23	-0.03	-0.62	3.34	-0.38	-2.82	35.88	8
3) C 23/G 22	-0.03	-0.62	3.34	-0.36	-2.79	35.90	5
4) G 22/C 21	-0.03	-0.62	3.34	-0.38	-2.82	35.88	8
5) C 21/T 20	-0.03	-0.62	3.34	-0.36	-2.81	35.89	-2
6) T 20/T 19	-0.03	-0.62	3.34	-0.37	-2.81	35.90	-4
7) T 19/A 18	-0.06	-0.63	3.34	-0.40	-2.83	36.28	7
8) A 18/A 17	-0.03	-0.64	3.34	-0.37	-2.82	35.89	4
9) A 17/G 16	0.00	-0.63	3.35	-0.35	-2.78	35.50	2
10) G 16/C 15	-0.03	-0.62	3.34	-0.38	-2.82	35.87	8
11) C 15/G 14	-0.03	-0.62	3.34	-0.36	-2.79	35.91	5
12) G 14/C 13	-0.03	-0.62	3.34	-0.38	-2.82	35.88	8

|H| Local Inter-Base pair Parameters |

Duplex	Shift (Dx)	Slide (Dy)	Rise (Dz)	Tilt (tau)	Roll (rho)	Twist (Omega)	Dc
2) C 1/G 2	0.00	-0.62	3.34	0.00	-2.82	35.88	8
3) G 2/C 3	0.00	-0.62	3.34	0.00	-2.80	35.91	5
4) C 3/G 4	0.00	-0.62	3.34	0.00	-2.82	35.88	8
5) G 4/A 5	-0.01	-0.62	3.34	-0.01	-2.80	35.70	2
6) A 5/A 6	0.00	-0.63	3.34	0.00	-2.82	35.90	4
7) A 6/T 7	0.00	-0.63	3.34	0.00	-2.83	36.28	7
8) T 7/T 8	0.00	-0.63	3.34	0.00	-2.82	35.90	-4
9) T 8/C 9	0.01	-0.62	3.34	0.01	-2.80	35.70	-2
10) C 9/G 10	0.00	-0.62	3.34	0.00	-2.82	35.88	8
11) G 10/C 11	0.00	-0.62	3.34	0.00	-2.80	35.91	5
12) C 11/G 12	0.00	-0.62	3.34	0.00	-2.82	35.88	8
Average:	0.00	-0.62	3.34	0.00	-2.81	35.89	

|I| Global Axis Curvature |

Duplex	Ax	Ay	Ainc	Atip	Adis	Angle	Path	Dc
2) C 1/G 2	0.00	0.00	0.00	0.00	0.00	0.00	3.38	8
3) G 2/C 3	0.00	0.00	0.01	-0.01	0.00	0.01	3.38	5

4)	C 3/G 4	0.00	0.00	0.02	-0.02	0.00	0.02	3.38	8
5)	G 4/A 5	-0.01	0.00	0.01	0.00	0.01	0.01	3.38	2
6)	A 5/A 6	0.00	0.00	0.01	0.02	0.00	0.02	3.38	4
7)	A 6/T 7	0.00	0.00	0.00	0.02	0.00	0.02	3.38	7
8)	T 7/T 8	0.00	0.00	-0.01	0.02	0.00	0.02	3.38	-4
9)	T 8/C 9	0.01	0.00	-0.01	0.00	0.01	0.01	3.38	-2
10)	C 9/G 10	0.00	0.00	-0.02	-0.02	0.00	0.02	3.38	8
11)	G 10/C 11	0.00	0.00	-0.01	-0.01	0.00	0.01	3.38	5
12)	C 11/G 12	0.00	0.00	0.00	0.00	0.00	0.00	3.38	8

Duplex Offset Angle= 0.0 Dirn= 0.0

1)	C 1	0.00
2)	G 2	0.00
3)	C 3	0.01
4)	G 4	0.00
5)	A 5	0.00
6)	A 6	0.01
7)	T 7	0.01
8)	T 8	0.00
9)	C 9	0.00
10)	G 10	0.01
11)	C 11	0.00
12)	G 12	0.00

Path length= 37.18 End-to-end= 37.18 Shortening= 0.00 %

|J| Backbone Parameters |

1st strand		C1'-C2'	C2'-C3'	Phase	Ampli	Pucker	C1'	C2'	C3'
1)CYT	1	24.62	-34.78	192.17	36.27	C3'-exo	107.7	102.0	102.8
2)GUA	2	24.62	-34.78	192.16	36.27	C3'-exo	107.8	102.0	102.8
3)CYT	3	24.65	-34.78	192.13	36.27	C3'-exo	107.8	102.0	102.8
4)GUA	4	24.65	-34.79	192.13	36.27	C3'-exo	107.8	102.0	102.8
5)ADE	5	24.66	-34.78	192.07	36.26	C3'-exo	107.8	102.0	102.8
6)ADE	6	24.62	-34.78	192.17	36.27	C3'-exo	107.7	102.0	102.8
7)THY	7	24.62	-34.78	192.16	36.27	C3'-exo	107.8	102.0	102.8
8)THY	8	24.65	-34.78	192.13	36.27	C3'-exo	107.8	102.0	102.8
9)CYT	9	24.65	-34.79	192.13	36.27	C3'-exo	107.8	102.0	102.8
10)GUA	10	24.66	-34.78	192.07	36.26	C3'-exo	107.8	102.0	102.8
11)CYT	11	24.62	-34.78	192.17	36.27	C3'-exo	107.7	102.0	102.8
12)GUA	12	24.62	-34.78	192.16	36.27	C3'-exo	107.8	102.0	102.8

Torsions		Chi C1'-N	Gamma C5'-C4'	Delta C4'-C3'	Epsil C3'-O3'	Zeta O3'-P	Alpha P-O5'	Beta O5'-C5'
1)CYT	1	-95.45	.....	156.50	159.12	-98.84	-39.25	-151.39
2)GUA	2	-95.45	30.87	156.50	159.14	-98.83	-39.26	-151.41
3)CYT	3	-95.45	30.91	156.49	159.09	-98.83	-39.25	-151.39
4)GUA	4	-95.44	30.88	156.53	159.13	-98.85	-39.25	-151.38
5)ADE	5	-95.41	30.85	156.56	159.15	-98.89	-39.23	-151.43
6)ADE	6	-95.42	30.89	156.50	159.12	-98.84	-39.25	-151.39
7)THY	7	-95.44	30.87	156.50	159.14	-98.83	-39.26	-151.41
8)THY	8	-95.44	30.91	156.49	159.09	-98.83	-39.25	-151.39
9)CYT	9	-95.44	30.88	156.53	159.13	-98.85	-39.25	-151.38
10)GUA	10	-95.43	30.85	156.56	159.15	-98.89	-39.23	-151.43
11)CYT	11	-95.45	30.89	156.50	159.12	-98.84	-39.25	-151.39
12)GUA	12	-95.45	30.87	156.50	.....	.....	.....	.....

2nd strand		C1'-C2'	C2'-C3'	Phase	Ampli	Pucker	C1'	C2'	C3'
1)GUA	24	24.62	-34.78	192.16	36.27	C3'-exo	107.8	102.0	102.8

2) CYT	23	24.62	-34.78	192.17	36.27	C3'-exo	107.7	102.0	102.8
3) GUA	22	24.66	-34.78	192.07	36.26	C3'-exo	107.8	102.0	102.8
4) CYT	21	24.65	-34.79	192.13	36.27	C3'-exo	107.8	102.0	102.8
5) THY	20	24.65	-34.78	192.13	36.27	C3'-exo	107.8	102.0	102.8
6) THY	19	24.62	-34.78	192.16	36.27	C3'-exo	107.8	102.0	102.8
7) ADE	18	24.62	-34.78	192.17	36.27	C3'-exo	107.7	102.0	102.8
8) ADE	17	24.66	-34.78	192.07	36.26	C3'-exo	107.8	102.0	102.8
9) GUA	16	24.65	-34.79	192.13	36.27	C3'-exo	107.8	102.0	102.8
10) CYT	15	24.65	-34.78	192.13	36.27	C3'-exo	107.8	102.0	102.8
11) GUA	14	24.62	-34.78	192.16	36.27	C3'-exo	107.8	102.0	102.8
12) CYT	13	24.62	-34.78	192.17	36.27	C3'-exo	107.7	102.0	102.8

Torsions		Chi C1'-N	Gamma C5'-C4'	Delta C4'-C3'	Epsil C3'-O3'	Zeta O3'-P	Alpha P-O5'	Beta O5'-C5'
1) GUA	24	-95.45	30.87	156.50	.....	.....	.....	.....
2) CYT	23	-95.45	30.89	156.50	159.12	-98.84	-39.25	-151.39
3) GUA	22	-95.43	30.85	156.56	159.15	-98.89	-39.23	-151.43
4) CYT	21	-95.44	30.88	156.53	159.13	-98.85	-39.25	-151.38
5) THY	20	-95.44	30.91	156.49	159.09	-98.83	-39.25	-151.39
6) THY	19	-95.44	30.87	156.50	159.14	-98.83	-39.26	-151.41
7) ADE	18	-95.42	30.89	156.50	159.12	-98.84	-39.25	-151.39
8) ADE	17	-95.41	30.85	156.56	159.15	-98.89	-39.23	-151.43
9) GUA	16	-95.44	30.88	156.53	159.13	-98.85	-39.25	-151.38
10) CYT	15	-95.45	30.91	156.49	159.09	-98.83	-39.25	-151.39
11) GUA	14	-95.45	30.87	156.50	159.14	-98.83	-39.26	-151.41
12) CYT	13	-95.45	.....	156.50	159.12	-98.84	-39.25	-151.39

II. CURVES-derived conformational parameters for the target structure:

```
--------------------------------
|B| Global Base-Axis Parameters |
--------------------------------

  1st strand        Xdisp    Ydisp    Inclin     Tip     Bc   Tc
                    (dx)     (dy)     (eta)    (theta)

  1) CYT    1       -0.42    -0.08     4.79    -14.36    3    0
  2) GUA    2       -0.54     0.02     0.95     -3.39    1   13
  3) CYT    3       -0.23    -0.14    -2.71     -7.92    3  -13
  4) GUA    4       -0.79     0.10     3.69     -2.78    1   10
  5) ADE    5       -0.30    -0.04    -0.14    -10.71    2   18
  6) ADE    6       -0.44     0.10    -3.86    -10.03    2   24
  7) THY    7       -0.22    -0.16    -7.97    -11.95    4  -24
  8) THY    8       -0.26    -0.01    -5.39     -8.40    4  -18
  9) CYT    9       -0.50    -0.14   -15.97     -3.93    3  -10
 10) GUA   10        0.21     0.16     1.22      4.54    1   13
 11) CYT   11       -0.64     0.08    -6.38    -10.24    3  -13
 12) GUA   12        0.11     0.26     0.10      9.54    1    0

  2nd strand        Xdisp    Ydisp    Inclin     Tip     Bc   Tc
                    (dx)     (dy)     (eta)    (theta)

  1) GUA   24        0.01    -0.04    -1.67     -8.93    1    0
  2) CYT   23       -0.74    -0.12    -2.91     -5.76    3  -13
  3) GUA   22       -0.13    -0.02     3.15      1.07    1   13
  4) CYT   21       -0.28    -0.29    -3.15     -8.56    3  -10
  5) THY   20       -0.47    -0.13    -5.53     -7.25    4  -18
  6) THY   19       -0.45    -0.14    -3.88    -14.79    4  -24
  7) ADE   18       -0.59    -0.02    -6.05    -11.88    2   24
  8) ADE   17       -0.26    -0.14    -2.05    -14.73    2   18
  9) GUA   16       -0.38     0.10     5.91     -7.14    1   10
 10) CYT   15        0.29    -0.27    -7.36    -10.37    3  -13
 11) GUA   14       -0.74    -0.29    -6.46    -10.65    1   13
 12) CYT   13        0.56    -0.14    -7.15      5.51    3    0

-------------------------------------
|C| Global Base pair-Axis Parameters |
-------------------------------------

  Duplex            Xdisp    Ydisp    Inclin     Tip     Bc   Tc
                    (dx)     (dy)     (eta)    (theta)

  1) C   1-G  24    -0.21    -0.02     1.56     -2.72    3    0
  2) G   2-C  23    -0.64     0.07    -0.98      1.18    1   13
  3) C   3-G  22    -0.18    -0.06     0.22     -4.49    3  -13
  4) G   4-C  21    -0.53     0.19     0.27      2.89    1   10
  5) A   5-T  20    -0.39     0.04    -2.83     -1.73    2   18
  6) A   6-T  19    -0.44     0.12    -3.87      2.38    2   24
  7) T   7-A  18    -0.41    -0.07    -7.01     -0.04    4  -24
  8) T   8-A  17    -0.26     0.07    -3.72      3.17    4  -18
  9) C   9-G  16    -0.44    -0.12    -5.03      1.60    3  -10
 10) G  10-C  15     0.25     0.22    -3.07      7.45    1   13
 11) C  11-G  14    -0.69     0.18    -6.42      0.20    3  -13
 12) G  12-C  13     0.34     0.20    -3.52      2.02    1    0

  Average:          -0.30     0.07    -2.87      0.99

--------------------------------
|D| Global Base-Base Parameters |
--------------------------------

  Duplex            Shear    Stretch  Stagger  Buckle   Propel   Opening  Bc   Tc
                    (Sx)     (Sy)     (Sz)     (kappa)  (omega)  (sigma)

  1) C   1-G  24    -0.43    -0.12    -0.01      6.46   -23.29    -2.74    3    0
```

2)	G 2-C 23	0.20	-0.09	0.49	3.85	-9.15	-6.77	1 13
3)	C 3-G 22	-0.10	-0.16	0.28	-5.86	-6.85	-0.47	3 -13
4)	G 4-C 21	-0.51	-0.20	-0.14	6.85	-11.34	2.08	1 10
5)	A 5-T 20	0.16	-0.17	-0.14	5.39	-17.97	6.87	2 18
6)	A 6-T 19	0.01	-0.05	-0.39	0.01	-24.82	11.26	2 24
7)	T 7-A 18	0.37	-0.18	-0.10	-1.92	-23.83	10.42	4 -24
8)	T 8-A 17	0.00	-0.16	-0.04	-3.33	-23.13	6.43	4 -18
9)	C 9-G 16	-0.12	-0.04	0.32	-21.88	-11.06	-2.68	3 -10
10)	G 10-C 15	-0.08	-0.11	0.34	8.58	-5.83	-2.15	1 13
11)	C 11-G 14	0.10	-0.21	0.45	0.08	-20.89	-6.76	3 -13
12)	G 12-C 13	-0.44	0.12	0.28	7.24	15.05	-8.16	1 0
Average:		-0.07	-0.11	0.11	0.46	-13.59	0.61	

|E| Global Inter-Base Parameters |

1st strand		Shift (Dx)	Slide (Dy)	Rise (Dz)	Tilt (tau)	Roll (rho)	Twist (Omega)	Dc
2)	C 1/G 2	-0.14	0.04	3.73	-6.07	10.03	39.88	8
3)	G 2/C 3	0.41	-0.10	3.54	-4.05	-3.13	38.50	5
4)	C 3/G 4	-0.65	0.44	2.96	8.04	10.88	26.84	8
5)	G 4/A 5	0.42	-0.07	3.18	-2.52	-4.16	43.96	2
6)	A 5/A 6	-0.06	-0.05	3.25	-3.02	1.09	37.55	4
7)	A 6/T 7	0.28	-0.50	3.37	-3.61	-1.82	35.27	7
8)	T 7/T 8	-0.13	-0.08	3.27	3.41	2.80	31.76	-4
9)	T 8/C 9	-0.11	-0.15	3.85	-10.53	7.73	34.39	-2
10)	C 9/G 10	0.70	0.64	2.83	16.45	7.72	30.69	8
11)	G 10/C 11	-1.13	0.19	3.67	-7.86	-21.95	38.47	5
12)	C 11/G 12	0.83	0.20	3.16	8.29	15.58	32.84	8

2nd strand		Shift (Dx)	Slide (Dy)	Rise (Dz)	Tilt (tau)	Roll (rho)	Twist (Omega)	Dc
2)	G 24/C 23	-0.76	-0.01	3.23	-3.46	4.11	43.90	8
3)	C 23/G 22	0.70	0.04	3.74	5.66	5.42	32.20	5
4)	G 22/C 21	-0.24	-0.48	3.38	-4.56	-15.37	24.29	8
5)	C 21/T 20	-0.26	0.10	3.18	-1.06	-2.46	39.17	-2
6)	T 20/T 19	0.10	0.17	3.50	2.36	-7.94	33.16	-4
7)	T 19/A 18	-0.07	0.37	3.09	-1.68	2.82	36.11	7
8)	A 18/A 17	0.23	0.11	3.21	4.83	-2.10	35.74	4
9)	A 17/G 16	0.01	0.27	3.48	8.02	4.34	43.50	2
10)	G 16/C 15	0.66	-0.71	2.82	-14.02	-2.49	30.16	8
11)	C 15/G 14	-1.32	-0.29	3.56	0.64	6.88	43.08	5
12)	G 14/C 13	1.37	0.13	3.34	1.13	20.37	34.24	8

|F| Global Inter-Base pair Parameters |

Duplex		Shift (Dx)	Slide (Dy)	Rise (Dz)	Tilt (tau)	Roll (rho)	Twist (Omega)	Dc
2)	C 1/G 2	-0.45	0.03	3.48	-4.77	2.96	41.89	8
3)	G 2/C 3	0.55	-0.07	3.64	0.81	-4.27	35.35	5
4)	C 3/G 4	-0.44	0.46	3.17	1.69	13.13	25.57	8
5)	G 4/A 5	0.08	-0.08	3.18	-1.79	-0.85	41.57	2
6)	A 5/A 6	0.02	-0.11	3.37	-0.33	4.51	35.35	4
7)	A 6/T 7	0.10	-0.44	3.23	-2.64	-2.32	35.69	7
8)	T 7/T 8	0.05	-0.10	3.24	4.12	2.45	33.75	-4
9)	T 8/C 9	-0.05	-0.21	3.66	-1.26	1.70	38.95	-2
10)	C 9/G 10	0.68	0.68	2.82	1.21	5.11	30.43	8
11)	G 10/C 11	-1.22	0.24	3.62	-3.61	-14.41	40.78	5
12)	C 11/G 12	1.10	0.04	3.25	4.71	-2.39	33.54	8

Average:	0.04	0.04	3.33	-0.17	0.51	35.71	

|G| Local Inter-Base Parameters |

1st strand	Shift (Dx)	Slide (Dy)	Rise (Dz)	Tilt (tau)	Roll (rho)	Twist (Omega)	Dc
2) C 1/G 2	0.47	-0.23	3.71	0.30	11.22	40.53	8
3) G 2/C 3	0.79	-0.35	3.48	-0.15	-3.46	38.74	5
4) C 3/G 4	-0.38	0.19	3.07	10.44	10.90	26.11	8
5) G 4/A 5	0.75	-0.29	3.17	2.95	-2.32	43.76	2
6) A 5/A 6	0.51	-0.41	3.20	3.89	-0.30	37.45	4
7) A 6/T 7	0.90	-1.01	3.18	3.25	-5.35	35.15	7
8) T 7/T 8	0.43	-0.60	3.20	8.69	-1.03	30.42	-4
9) T 8/C 9	0.43	-1.09	3.71	-6.30	1.17	35.37	-2
10) C 9/G 10	0.53	0.14	2.93	15.70	3.41	30.73	8
11) G 10/C 11	-0.94	0.05	3.79	-5.49	-22.77	39.18	5
12) C 11/G 12	0.65	-0.28	3.18	8.60	13.05	32.79	8

2nd strand	Shift (Dx)	Slide (Dy)	Rise (Dz)	Tilt (tau)	Roll (rho)	Twist (Omega)	Dc
2) G 24/C 23	-1.13	-0.35	3.13	-8.88	-5.48	43.16	8
3) C 23/G 22	0.44	-0.20	3.79	4.17	-5.07	32.45	5
4) G 22/C 21	-0.48	0.33	3.43	-6.23	15.13	24.15	8
5) C 21/T 20	-0.77	-0.56	3.11	-6.37	-0.33	38.85	-2
6) T 20/T 19	-0.61	-0.77	3.40	-4.03	5.00	32.96	-4
7) T 19/A 18	-0.74	-0.91	2.95	-9.74	-5.68	34.77	7
8) A 18/A 17	-0.55	-0.62	3.13	-3.50	-0.46	35.76	4
9) A 17/G 16	-0.71	-0.29	3.44	-0.77	-2.35	44.03	2
10) G 16/C 15	0.26	0.62	2.91	-18.00	1.97	28.02	8
11) C 15/G 14	-2.00	-0.26	3.24	-7.24	-11.88	41.39	5
12) G 14/C 13	1.06	-0.42	3.44	-0.63	-23.74	34.19	8

|H| Local Inter-Base pair Parameters |

Duplex	Shift (Dx)	Slide (Dy)	Rise (Dz)	Tilt (tau)	Roll (rho)	Twist (Omega)	Dc
2) C 1/G 2	-0.36	-0.34	3.48	-3.96	2.64	42.50	8
3) G 2/C 3	0.62	-0.28	3.64	2.04	-4.20	35.67	5
4) C 3/G 4	-0.42	0.26	3.24	2.15	13.14	25.13	8
5) G 4/A 5	-0.01	-0.39	3.21	-1.68	-1.15	41.50	2
6) A 5/A 6	-0.05	-0.54	3.35	0.15	2.66	35.77	4
7) A 6/T 7	0.05	-0.87	3.19	-3.32	-4.87	35.79	7
8) T 7/T 8	-0.07	-0.57	3.20	2.67	-0.31	33.81	-4
9) T 8/C 9	-0.16	-0.66	3.64	-3.61	-0.57	39.68	-2
10) C 9/G 10	0.39	0.38	2.90	-1.14	2.40	29.03	8
11) G 10/C 11	-1.49	-0.13	3.56	-6.64	-17.40	40.25	5
12) C 11/G 12	0.86	-0.34	3.26	5.10	-5.61	33.65	8
Average:	-0.06	-0.32	3.33	-0.75	-1.21	35.71	

|I| Global Axis Curvature |

Duplex	Ax	Ay	Ainc	Atip	Adis	Angle	Path	Dc
2) C 1/G 2	-0.01	-0.06	-2.23	-0.94	0.06	2.42	3.48	8
3) G 2/C 3	0.09	0.06	-0.39	1.40	0.11	1.46	3.64	5

4)	C 3/G 4	-0.09	0.20	1.64	5.75	0.22	5.98	3.18	8
5)	G 4/A 5	-0.06	0.06	1.32	3.77	0.09	3.99	3.18	2
6)	A 5/A 6	0.08	-0.18	0.71	0.40	0.20	0.81	3.38	4
7)	A 6/T 7	0.07	-0.24	0.49	0.10	0.25	0.50	3.24	7
8)	T 7/T 8	-0.10	-0.23	0.83	-0.75	0.25	1.12	3.25	-4
9)	T 8/C 9	0.13	-0.02	0.05	3.26	0.13	3.26	3.67	-2
10)	C 9/G 10	-0.01	0.34	-0.75	-0.75	0.34	1.06	2.84	8
11)	G 10/C 11	-0.29	0.28	-0.26	-7.17	0.40	7.17	3.63	5
12)	C 11/G 12	0.08	0.01	1.82	-4.21	0.08	4.58	3.25	8

Duplex Offset Angle= 17.0 Dirn= 133.4

1)	C 1	0.00
2)	G 2	0.74
3)	C 3	1.24
4)	G 4	1.64
5)	A 5	1.81
6)	A 6	1.56
7)	T 7	1.24
8)	T 8	0.87
9)	C 9	0.92
10)	G 10	0.91
11)	C 11	0.36
12)	G 12	0.00

Path length= 36.73 End-to-end= 36.44 Shortening= 0.80 %

|J| Backbone Parameters |

1st strand		C1'-C2'	C2'-C3'	Phase	Ampli	Pucker	C1'	C2'	C3'
1) CYT	1	48.18	-49.04	166.21	51.96	C2'-endo	104.8	95.3	101.6
2) GUA	2	39.22	-38.39	162.25	41.26	C2'-endo	104.8	100.3	102.2
3) CYT	3	10.47	12.86	70.67	39.72	C4'-exo	106.0	105.0	103.0
4) GUA	4	35.04	-37.28	170.25	38.78	C2'-endo	106.1	101.1	102.1
5) ADE	5	35.62	-32.27	155.43	36.29	C2'-endo	105.0	102.3	102.5
6) ADE	6	32.44	-27.24	149.52	32.45	C2'-endo	105.8	103.4	102.9
7) THY	7	29.99	-10.21	105.19	39.30	O1'-endo	103.7	104.2	103.9
8) THY	8	35.33	-24.03	132.94	36.19	C1'-exo	104.8	102.8	103.2
9) CYT	9	48.62	-39.15	145.97	48.62	C2'-endo	102.7	98.5	101.8
10) GUA	10	49.55	-38.15	142.71	49.83	C1'-exo	104.3	96.0	104.0
11) CYT	11	43.98	-39.01	152.46	44.65	C2'-endo	102.5	100.7	101.9
12) GUA	12	36.25	-15.02	110.19	44.59	C1'-exo	102.7	102.4	104.3

Torsions		Chi C1'-N	Gamma C5'-C4'	Delta C4'-C3'	Epsil C3'-O3'	Zeta O3'-P	Alpha P-O5'	Beta O5'-C5'
1) CYT	1	-107.41	.....	156.10	-140.17	-136.06	-40.97	167.11
2) GUA	2	-95.22	20.48	142.88	-177.18	-113.11	-70.69	159.83
3) CYT	3	-143.56	72.50	88.41	-163.59	-93.28	-39.45	167.45
4) GUA	4	-101.13	47.21	142.17	-172.26	-129.90	-15.80	150.06
5) ADE	5	-110.53	32.11	126.88	178.97	-93.35	-59.30	-173.04
6) ADE	6	-105.83	44.27	126.05	162.56	-80.97	-66.58	-176.86
7) THY	7	-131.68	63.22	103.70	-175.97	-87.81	-47.98	175.82
8) THY	8	-112.62	46.27	117.83	168.89	-92.17	-67.35	178.76
9) CYT	9	-129.15	63.58	136.62	-130.82	-123.56	-41.18	143.20
10) GUA	10	-84.76	34.83	130.06	-101.58	144.50	-67.55	132.94
11) CYT	11	-118.67	44.88	137.71	-151.88	-99.41	-91.93	165.73
12) GUA	12	-116.10	69.66	107.08	.....	.....	.....	.....

2nd strand		C1'-C2'	C2'-C3'	Phase	Ampli	Pucker	C1'	C2'	C3'
1) GUA	24	-18.19	40.78	34.10	50.86	C3'-endo	106.9	101.0	97.9

2) CYT	23	31.61	-10.76	105.29	41.19	O1'-endo	103.5	104.0	104.1
3) GUA	22	45.24	-41.71	157.60	46.64	C2'-endo	106.3	96.4	103.8
4) CYT	21	28.90	-15.76	119.68	32.42	C1'-exo	105.0	103.7	105.4
5) THY	20	37.16	-33.40	155.40	37.77	C2'-endo	105.7	101.3	103.0
6) THY	19	40.24	-25.83	128.37	42.31	C1'-exo	102.6	102.1	104.2
7) ADE	18	34.58	-32.22	159.93	35.50	C2'-endo	106.6	102.1	102.4
8) ADE	17	23.36	-32.11	191.76	33.82	C3'-exo	107.9	103.0	101.9
9) GUA	16	37.03	-32.19	152.13	37.39	C2'-endo	105.1	101.6	103.0
10) CYT	15	8.19	18.90	65.44	46.27	C4'-exo	105.5	103.2	103.2
11) GUA	14	34.53	-24.68	135.06	35.34	C1'-exo	103.8	103.4	104.3
12) CYT	13	39.61	-39.62	164.60	42.09	C2'-endo	104.7	100.5	101.7

Torsions		Chi C1'-N	Gamma C5'-C4'	Delta C4'-C3'	Epsil C3'-O3'	Zeta O3'-P	Alpha P-O5'	Beta O5'-C5'
1) GUA	24	-141.80	56.79	72.07	.....	.....	.....	.....
2) CYT	23	-142.88	63.86	99.35	-167.12	-87.31	-82.07	168.26
3) GUA	22	-84.19	33.06	145.35	-126.71	-177.96	-57.32	117.09
4) CYT	21	-110.44	49.56	115.37	-173.65	-86.13	-46.86	177.86
5) THY	20	-112.39	32.39	130.50	175.09	-96.77	-61.73	-171.88
6) THY	19	-111.91	51.82	118.83	-174.64	-102.54	-31.19	176.30
7) ADE	18	-106.91	31.21	132.34	168.17	-92.03	-57.51	-179.05
8) ADE	17	-104.97	27.99	144.43	172.22	-91.53	-39.50	-175.97
9) GUA	16	-110.24	80.70	136.25	176.12	-105.87	-29.80	-172.19
10) CYT	15	-148.11	68.83	80.78	-166.25	-90.34	-73.45	171.54
11) GUA	14	-123.13	29.26	120.97	-171.86	-99.52	-53.60	151.32
12) CYT	13	-131.63	.....	144.44	-145.23	-156.35	-21.13	135.01

1. Singer, B. and Grunberger, D., *"Molecular Biology of Mutagens and Carcinogens"*, Plenum Press, New York, 1983.

2. Adams, L. P. and Burdon, B. H., *"Molecular Biology of DNA Methylation"*, Springer-Verlag, New York, 1985.

3. Lawley, P. D., *ACS Monogr.* **173**, 83-244 (1976).

4. Pegg, A. E., *Adv. Cancer Res.* **25**, 195-269 (1977).

5. Jensen, D. E., *Biochemistry* **17**, 5108-5113 (1978).

6. Jensen, D. E. and Reed, D. J., *Biochemistry* **17**, 5098-5107 (1978).

7. Shooter, K. V., *Chem.-Biol. Interactions* **13**, 151-163 (1976).

8. Bannon, P. and Verley, W., *Eur. J. Biochem.* **31**, 103-111 (1972).

9. Singer, B., Spengler, S. and Bodell, W. J., *Carcinogenesis* **2**, 1069-1073 (1981).

10. Miller, P. S., Chandrasegaran, S., Dow, D. L., Pulford, S. M. and Kan, L. S., *Biochemistry* **21**, 5468-5474 (1982).

11. Gallo, K. A., Shao, K., Phillips, L. R., Regan, J. B., Koziolkiewicz, M., Uznanski, B., Stec, W. J., Zon, G., *Nucleic Acids Res.* **14**, 7405-7420 (1986).

12. Majumdar, A. and Adhya, S., *J. Mol. Biol.* **208**, 217-223 (1989).

13. Weinfeld, M. and Livingston, D. C., *Biochemistry* **25**, 5083-5091 (1986).

14. Saenger, W., *"Principles of Nucleic Acid Structure"*, Springer-Verlag, New York, 1984.

15. Wüthrich, K., *"NMR of Proteins and Nucleic Acids"*, John Wiley & Sons, New York, 1986.

16. Brennan, R. G., Kondo, N. S. and Sundaralingam, M., *J. Am. Chem. Soc.* **106**, 5671-5676 (1984).

17. Miller, P. S., Fang, K. N., Kondo, N. S. and Ts'o P. O. P., *J. Am. Chem. Soc.* **93**, 6657-6665 (1971).

18. Weinfeld, M., Drake, A. F., Kuroda, R. and Livingston, D. C., *Analyt. Biochem.* **178**, 93-101 (1989).

19. Kan, L. -S., Cheng, D. M., Chandrasegaran, S., Pramanik, P. and Miller, P. S., *J. Biomol. Struct. Dyn.* **4**, 785-796 (1987).

20. Pramanik, P. and Kan, L. -S., *Biochemistry* **26**, 3807-3812 (1987).

21. Summers, M. F., Powell, C., Egan, W., Byrd, R. A., Wilson, W. D. and Zon, G., *Nucleic Acids Res.* **14**, 7421-7436 (1986).

22. Broido, M. S., James, T. L., Zon, G. and Keepers, J., *Eur. J. Biochem.* **150**, 117-128 (1985).

23. Lawrence, D. P., Wenqiao, C., Zon, G., Stec, W. J., Uznanski, B. and Broido, M. S., *J. Biomol. Struct. Dyn.* **4**, 757-783 (1987).

24. Broido, M. S., Lawrence, D. P., Wenqiao, C. and Zon, G., in *"Structure and Expression"*, Vol. 2: DNA and Its Drug Complexes, pp. 167-179, R. H. Sarma and M. H. Sarma, Eds., Adenine Press, New York, 1987.

25. Broido, M. S. and Mezei, M., *Biopolymers* **29**, 597-607 (1990).

26. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta Jr., S. and Weiner, P. K., *J. Am. Chem. Soc.* **106**, 765-784 (1984).

27. Weiner, P. K. and Kollman, P. A., *J. Comp. Chem.* **2**, 287-303 (1981).

28. Weiner, S. J., Kollman, P. A., Nguyen, D. T. and Case, D. A., *J. Comp. Chem.* **7**, 7163-7172 (1986).

29. Zubay, G., *"Biochemistry"*, Addison-Wesley, Massachusetts, 1984.

30. McClarin, J. A., Frederick, C. A., Wang, B.-C., Greene, P., Boyer, H. W., Grable, J. and Rosenberg, J. M., *Science* **234**, 1526-1541 (1986).

31. Dickerson, R. E., Drew, H. R., Conner, B. N., Wing, R. M., Fratini, A. V. and Kopka, M. L., *Science* **216**, 475-485 (1982).

32. Drew, H. R. and Dickerson, R. E., *J. Mol. Biol.* **151**, 535-556 (1981).

33. Drew, H. R., Samson, S. and Dickerson, R. E., *Proc. Natl. Acad. Sci., USA* **79**, 4040-4044 (1981).

34. Dickerson, R. E. and Drew, H. R., *J. Mol. Biol.* **149**, 761-786 (1981).

35. Patel, D. J. and Canuel, L. L., *Eur. J. Biochem.* **96**, 267-276 (1979).

36. Patel, D. J., Kozlowski, S. A., Marky, L. A., Broka, C., Rice, J. A., Itakura, K. and Breslauer, K. J., *Biochemistry* **21**, 428-436 (1982).

37. Hare, D. R., Wemmer, D. E., Chou, S. -H., Drobny, G. and Reid, B. R., *J. Mol. Biol.* **171**, 319-336 (1983).

38. Connolly, B. A. and Eckstein, F., *Biochemistry* **23**, 5523-5527 (1984).

39. Oh, J. and Eckstein, F., *Biochemistry* **24**, 2530-2535 (1985).

40. Patel, D. J., Shapiro, L. and Hare, D., *J. Biol. Chem.* **261**, 1223-1229 (1986).

41. Sklenar, V. and Bax, A., *J. Am. Chem. Soc.* **109**, 7525-7526 (1987).

42. Moe, J. G. and Russu, I. M., *Nucleic Acids Res.* **18**, 821-827 (1990).

43. Bloch, F., Hansen, W. W. and Packard, M. E., *Phys. Rev.* **69**, 127 (1946).

44. Purcell, E. M., Torrey, H. C. and Pound, R. V., *Phys. Rev.* **69**, 37 (1946).

45. Macomber, J. D., *"The Dynamics of Spectroscopic Transitions"*, John Wiley, New York, 1976.

46. Farrar, T. C. and Becker, E. D. *"Pulse and Fourier Transform NMR. Introduction to Theory and Methods"*, Academic Press, New York, 1971.

47. Abragam, A., *"The Principles of Nuclear Magnetism"*, Clarendon Press, Oxford, UK, 1961.

48. Slichter, C. P., *"Principles of Magnetic Resonance"*, Springer-Verlag, New York, 1980.

49. Ernst, R. R., Bodenhausen, G. and Wokaun, A., *"Principles of Nuclear Magnetic Resonance in One and Two Dimensions"*, Oxford University Press, New York, 1987.

50. Fukushima, E. and Roeder, S. B., *"Experimental Pulse NMR. A Nuts and Bolts Approach"*, Addison Wesley Publishing Company, Massachussets, 1981.

51. Ernst, R. R. and Anderson, W. A., *Rev. Sci. Instr.* **37**, 93 (1966).

52. Ernst, R. R., *Adv. Mag. Res.* **2**, 1-135 (1966).

53. Bax, A., *"Two-Dimensional Nuclear Magnetic Resonance in Liquids"*, Delft University Press, Delft, Holland, 1982.

54. Jeener, J., *"Ampere International Summer School"*, Basko Polje, Yugoslavia, 1971.

55. Aue, W. P., Bartholdi, E. and Ernst, R. R., *J. Chem. Phys.* **64**, 2229-2247 (1976).

56. Braunschweiler, L. and Ernst, R. R., *J. Magn. Reson.* **53**, 521-528 (1983).

57. Neuhaus, D. and Williamson, M. P., *"The Nuclear Overhauser Effect in Structural and Conformational Analysis"*, Verlag Chemie, New York, 1989.

58. Noggle, J. H. and Schirmer, R. E., *"The Nuclear Overhauser Effect. Chemical Applications"*, Academic Press, New York, 1971.

59. Macura, S. and Ernst, R. R., *Mol. Phys.* **41**, 95-117 (1980).

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

61. Borgias, B. A., Thomas, P. D. and James, T. L., *COmplete Relaxation Matrix Analysis (CORMA)*, University of California, San Fransisco (1987,1989).

62. Borgias, B. A., Gochin, M., Kerwood, D. J. and James, T. L., *Prog. NMR Spectrsc.* **22**, 83-100 (1990).

63. Solomon, I., *Phys. Rev.* **99**, 559 (1955).

64. Goldman, M., *"Quantum Description of High-Resolution NMR in Liquids"*, Clarendon Press, Oxford, 1988.

65. Gueron, M., Kochoyan, M. and Leroy, J. L., *Nature* **328**, 69-92 (1987).

66. Gueron, M., Kockoyan, M. and Leroy, J. L., "Proc. 1st EBSA Workshop: Structure, Dynamics and Function of Biomolecules", Ehrenberg, A., Rigler, G., Graslund, A. and Nilsson, L., Eds., pp 196-200, Springer-Verlag, Berlin, 1987.

67. Leroy, J. L., Kochoyan, M., Huynh-Dinh, T., and Gueron, M., *J. Mol. Biol.* **200**, 223-238 (1988)

68. Kochoyan, M., Leroy, J. L. and Gueron, M., *J. Mol. Biol.* **196**, 599-609 (1987).

69. Eigen, M., *Angew. Chem. Int. Ed. Engl.* **3**, 1-19 (1964).

70. Brooks III, C. L., Karplus, M. and Pettitt, M., *"Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics"*, John Wiley, New York, 1988.

71. McCammon, J. A. and Harvey, S. C., *"Dynamics of Proteins and Nucleic Acids"*, Cambridge University Press, Cambridge, 1987.

72. *"Molecular Dynamics: Applications in Molecular Biology"*, J.M. Goodfellow, Ed., CRC Press, 1990.

73. Verlet, L., *Phys. Rev.* **159**, 98 (1967).

74. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J. , Swaninathan, S. and Karplus, M., *J. Comp. Chem.* **4**, 187-217 (1983).

75. Lavery, R. and Sklenar, H., *J. Biolmol. Str. Dyn.* **6**, 63-91 (1988).

76. Lavery, R. and Sklenar, H., *J. Biomol. Str. Dyn.* **6**, 655-667 (1989).

77. Plateau, P. and Gueron, M., *J. Am. Chem. Soc.* **104**, 7310-7311 (1982).

78. Freeman, R., *"A Handbook of Nuclear Magnetic Resonance"*, John Wiley & Sons, New York, 1988.

79. States, D. J., Haberkorn, R. A. and Ruben, D. J., *J. Magn. Reson.* **48**, 286-292 (1982).

80. Sklenar, V., Miyashiro, H., Zon, G., Miles, H. T. and Bax, A., *FEBS Lett.* **208**, 94-98 (1986).

81. Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R. and van Gunsteren, W. F., *J. Mol. Biol.* **182**, 179-182 (1985).

82. Nilsson, L., Clore, G. M., Gronenborn, A. M., Brunger, A. T. and Karplus, M., *J. Mol. Biol.* **188**, 455-475 (1986).

83. Crippen, G. M., *J. Comput. Phys.* **24**, 96-107 (1977).

84. Havel, T. F., Crippen, G. M. and Kuntz, I. D., *Biopolymers* **18**, 73-81 (1979).

85. Clore, G. M. and Gronenborn, A. M., *J. Magn. Reson.* **61**, 158-164 (1985).

86. Clore, G. M. and Gronenborn, A. M., *FEBS Lett.* **172**, 219-225 (1984).

87. Ryckert, J. P., Cicotti, G. and Berendsen, H. J. C., *J. Comput. Phys.* **23**, 327-337 (1977).

88. Bax, A. and Lerner, L., *J. Magn. Reson.* **79**, 429-438 (1988).

89. Celda, B., Widmer, H., Leupin, W., Chazin, W. J., Denny, W. A. and Wuthrich, K., *Biochemistry* **28**, 1462-1471 (1989).

90. Davies, D. B., *Prog. Nucl. Magn. Reson. Spectrosc.* **12**, 135-225 (1985).

91. Hosur, R. V., Ravikumar, M., Chary, K. V. R., Sheth, A., Govil, G., Zu-Kun, T. and Todd Miles, H., *FEBS Lett.* **205**, 71-76 (1986).

92. Pardi, A., Walker, R., Rapoport, H., Wider, G. and Wuthrich, K., *J. Am. Chem. Soc.* **105**, 1652-1653 (1983).

93. *"Biological Applications of Magnetic Resonance"*, R.G. Shulman, Ed., Academic Press, New York, 1979.

94. *"Phosphorus-31 NMR: Principles and Applications"*, D.G. Gorenstein, Ed., Academic Press, Orlando, 1984.

95. Gorenstein, D. G., Schroeder, S. A., Fu, J. M., Metz, J. T., Roongta, V. and Jones, C. R., *Biochemistry* **27**, 7223-7237 (1988).

96. Schroeder, S. A., Fu, J. M., Jones, C. R. and Gorenstein, D. G., *Biochemistry* **26**, 3812-3821 (1987).

97. Gorenstein, D. G., *Annu. Rev. Biophys. Bioeng.* **10**, 355-386 (1981).

98. Herderling, W. and Seela, F., J. Org. Chem. **50**, 5314-5323 (1985).

99. Mellema, J. R., Haasnoot, C. A. G., Van Boom, J. H. and Altona, C., *Biochem. Biophys. Acta* **655**, 256-264 (1981).

100. Kearns, D. R., *Ann. Rev. Biophys. Bioeng.* **6**, 477-523 (1977).

101. Kallenbach, N. R. and Berman, H. M., *Q. Rev. Biophys.* **10**, 138-236 (1977).

102. Patel, D. J. and Hilbers, C. W., *Biochemistry* **14**, 2651-2660 (1975).

103. Leroy, J. L., Broseta, D., Bolo, N. and Gueron, M., *NATO ASI Ser.* **110**, 31-42 (1985).

104. Leroy, J. L., Charretier, E., Kochoyan, M. and Gueron, M., *Biochemistry* **27**, 8894-8898 (1988).

105. Keepers, J. W. and James, T. L., *J. Am. Chem. Soc.* **104**, 929-939 (1982).

106. Bloembergen, N., Purcell, E. M., and Pound, R. V., *Phys. Rev.* **73**, 679 (1948).

107. Woessner, D. E., *J. Chem. Phys.* **37**, 647-655 (1962).

108. Sudmeier, J. L., Anderson, S. E. and Frye, J. S., *Concepts Magn. Reson.* **2**, 197-212 (1990).

109. Un, S. and Klein, M. P., *J. Am. Chem. Soc.* **111**, 5119-5124 (1989).

110. Fyfe, C. A., *"Solid State NMR For Chemists"*, C.F.C Press, Ontario, Canada, 1983.

111. Boelens, R., Vuister, G. W., Koning, T. M. G. and Kaptein, R., *J. Am. Chem. Soc.* **111**, 8525-8526 (1989).

112. Griesinger, C., Sorensen, O. W. and Ernst, R. R., *J. Magn. Reson.* **84**, 14-63 (1989).