Title
DYNAMIC AND (INFORMATIONAL EFFECTS OF STRUCTURAL PERTURBATIONS IN DNA AND RNA STUDIED BY NUCLEAR MAGNETIC RESONANCE AND CHEMICALLY INDUCED DYNAMIC NUCLEAR POLARIZATION

Permalink
https://escholarship.org/uc/item/1jk6r2br

Author
Morden, K.M.

Publication Date
1983-09-01
DYNAMIC AND CONFORMATIONAL EFFECTS OF STRUCTURAL PERTURBATIONS IN DNA AND RNA STUDIED BY NUCLEAR MAGNETIC RESONANCE AND CHEMICALLY INDUCED DYNAMIC NUCLEAR POLARIZATION

K.M. Morden
(Ph.D. Thesis)

September 1983

TWO-WEEK LOAN COPY

This is a Library Circulating Copy which may be borrowed for two weeks. For a personal retention copy, call Tech. Info. Division, Ext. 6782.

Prepared for the U.S. Department of Energy under Contract DE-AC03-76SF00098
DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.
DYNAMIC AND CONFORMATIONAL EFFECTS OF STRUCTURAL PERTURBATIONS IN DNA AND RNA STUDIED BY NUCLEAR MAGNETIC RESONANCE AND CHEMICALLY INDUCED DYNAMIC NUCLEAR POLARIZATION

Kathleen Margaret Morden
Ph.D. Thesis

September 1983

Chemical Biodynamics Division
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

This work was supported by the U.S. Department of Energy under Contract No. DE-AC03-76SF0098 and NIEHS Training Grant No. ES07075.
Dynamic and Conformational Effects of Structural Perturbations in DNA and RNA Studied by Nuclear Magnetic Resonance and Chemically Induced Dynamic Nuclear Polarization

by

Kathleen Margaret Morden

Abstract

The conformation and dynamics of oligonucleotides in solution were studied using several different NMR techniques. The oligonucleotides contain a variety of perturbations; I) an extra cytosine: $dCA_3CA_2G + dCT_6G$, II) an extra adenine: $d(CGCAGAATTCGCG)_2$, and III) a G·T mismatch: $d(CGTGAATTCGCG)_2$. These are compared to the unperturbed duplexes IV) $dCA_6G + dCT_6G$ and V) $d(CGCGAATTCGCG)_2$. Through the use of nuclear Overhauser effects, the unpaired cytosine in duplex I was found to be extrahelical. The temperature dependence of the chemical shifts for the aromatic base protons supports this result. Thermodynamic parameters were determined for duplexes I and IV using UV absorption melting curves. The extrahelical cytosine causes a 15°C decrease in the duplex melting temperature (at 200μM per single strand). The standard free energy for duplex formation at 25°C is 2.9 kcal mol$^{-1}$ more positive for the duplex with the extra cytosine.

The kinetics for exchange of the imino protons in these oligonucleotide duplexes were studied by saturation recovery techniques. In duplex I, the extra cytosine causes the exchange lifetime for one of the neighboring A·T base pairs to decrease by a
factor of two over the other A·T base pairs in the duplex. The activation energy for exchange of this base pair is \( \sim 20 \text{ kcal mol}^{-1} \), while, the activation energy for the other base pairs averages 55 kcal mol\(^{-1}\). In duplex IV, the activation energies average \( \sim 20 \text{ kcal mol}^{-1} \). The extra adenine in duplex II affects the exchange lifetimes throughout the duplex, however, the G·T mismatch in duplex III has a more local effect on the lifetimes. The unperturbed duplexes, IV and V, undergo exchange by local base pair opening, whereas the perturbed duplexes, I, II, and III, involve a higher energy mechanism indicating the involvement of whole helix opening.

Also presented are solvent accessibility studies of oligonucleotides using Chemically Induced Dynamic Nuclear Polarization (CIDNP). CIDNP was observed for guanine(H8) protons and for guanine(H1') protons in dinucleotides and longer oligonucleotides. Several mechanisms for this polarization are discussed. This technique was also used to investigate the solvent accessibility of guanines in yeast tRNA\(^{phe}\).
Dedicated to my mother Margaret

and

in memory of my father Donald

for teaching me to live life to its fullest
ACKNOWLEDGEMENTS

First and foremost, I would like to thank my research advisor Nacho Tinoco for his guidance and encouragement during my time at Berkeley. On occasion, I also had the pleasure of seeing Nacho in a non-science setting where he was good company and very entertaining, especially his alpine accident reports.

Arthur Pardi was my mentor and as such taught me a lot of what I know about NMR. He has continued to be a source of advice about such things.

Frank Martin was influential in my early years as a graduate student. He provided a wealth of information about nucleic acids and was always provoking interesting discussions in the lab. He also provided many of the oligonucleotides that I used in my research.

Jeff Nelson was always willing to listen to problems, especially concerning kinetics and optical studies of nucleic acids. His extensive knowledge about computers often helped me to deal with them. During the process of writing this thesis (and doing the research) he provided much help and lessened the agony with his continuous support and encouragement.

I would like to thank Gloria Meng for doing the optical studies presented here. It was a pleasure to work with her and she always brightened my day.

David Keller and Emil Scoffone also provided much computer advice. David was always willing to discuss my seemingly endless
questions about ring current shielding parameters. Emil took much kidding in the lab but yes Emil I do like you and by the way where are we going for dinner?

Ti-Sheng Young was my source of all around general science advice. I could always count on him to get excited about my latest result or be willing to help me solve my latest problem.

As my office mates during the last push, Kathi Hall and Phil (Easy Mark) Cruz were wonderful. They were very supportive during my moments of total panic and my sometimes manic-depressive moods. Phil also read and gave advice on much of this thesis.

Betsy McCord was the main instigator of the CIDNP studies on nucleic acids. I am very glad I had the opportunity to get to know Betsy as she is not only a very good scientist but has become a valued friend.

Joe Kao has provided continuous support since the beginning of my graduate career. He is always willing to discuss any scientific problem that I might have, but the most important thing he provides is friendship. He has helped me through several rough spots and has always lent a shoulder for me to cry on. Joe has also contributed to many sessions of uncontrollable laughter.

A large portion of my research was done at institutes other than Berkeley and several of the people that ran these NMR facilities deserve mention. Dave Wemmer provided much advice during my many trips to Stanford and also obtained some of the NOE data in my thesis. Jerry Matson helped me get started on the UC Davis NMR and helped set up the double quantum experiments on that instrument.
Karen Elliot taught me how to use the word processor, thus allowing me to write my thesis without the aid of a Swiss bank account. She also kept me calm during the ordeal of writing.

There are three people in the lab who I would like to give special thanks to as they have become like family to me. Helen Lok was always cheery and concerned and she made working in the lab a real pleasure. Barbara Dengler provided support both scientifically and personally. David Koh, who I can't thank enough, is one of the most considerate people I know, especially when it comes to late rent payments.

Beth Schwehr, Will Hollingsworth, Paul Kobrin, Bruce Schardt, John van Curen, Jan Cole, Mike Mayberry and Vicki Mayberry all provided friendship that made my stay in Berkeley a memorable one.

Finally I would like to thank my family for all their love and support.

This work was supported in part by NIEHS training grant ES07075 and the U.S. Department of Energy under Contract Number DE-AC03-76SF00098.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................. ii
TABLE OF CONTENTS ................................................................. v
LIST OF TABLES ................................................................. ix
LIST OF FIGURES ................................................................. xi
CHAPTER I. Introduction ......................................................... 1
1. Background ................................................................. 1
2. The Scope of This Thesis ................................................ 14
References ................................................................. 15
CHAPTER II. The Conformation and Stability of an Oligonucleotide Duplex Containing a Cytosine Bulge ... 17
1. Synopsis ................................................................. 17
2. Introduction ................................................................. 18
3. Materials and Methods .................................................. 19
   A. Sample Preparation ................................................... 19
   B. Experimental Techniques ............................................. 20
4. NMR Results and Discussion ......................................... 28
   A. Imino-Imino NOE ................................................... 28
   B. Imino-A(H2) NOE ................................................... 39
   C. Conclusions from Observed NOE ................................... 42
   D. Non-Exchangeable Protons ........................................... 45
5. Thermodynamic Results .................................................. 63
6. Discussion of the Thermodynamics .................................... 69
7. Conclusion ................................................................. 71
References ................................................................. 73
CHAPTER III. Kinetics of Exchange in Oligonucleotides

1. Introduction .................................................. 75

2. Theory of Imino Proton Exchange .......................... 76

3. Effects of an Extra Cytosine on the Kinetics of Exchange .......................... 85
   A. Materials and Methods .................................. 85
   B. Results ..................................................... 86
   C. Discussion .................................................. 99

4. Effects of a G·T Mismatch and an Extra Adenine on the Exchange .......................... 118
   A. Materials and Methods .................................. 118
   B. Results ..................................................... 119
      i. 12-mer: d(C-G-C-G-A-A-T-T-C-G-C-G)₂ ............. 119
      ii. 12-mer G·T: d(C-G-T-G-A-A-T-T-C-G-C-G)₂ .... 127
   C. Discussion .................................................. 127
      i. T₁ vs. Chemical Exchange ............................... 127
      ii. Are These Imino Protons in the Open-Limited Region for Exchange? ....... 141
      iii. Effects of Perturbations on the Lifetimes of Exchange ..................... 142
      iv. Activation Energies and Exchange Mechanisms ............................... 143

5. Conclusions .................................................... 150

References .......................................................... 154
CHAPTER IV. CIDNP Studies of Nucleic Acids

1. CIDNP Theory .................................. 156
2. Introduction .................................. 163
3. Experimental ................................. 168
4. Results ..................................... 173
   A. Bases, Mononucleosides, and Mononucleotides .... 173
   B. Dinucleotides and Trinucleotide .................. 186
   C. Tetranucleotide .............................. 191
   D. tRNA^phe .................................... 191
5. Discussion ................................... 204
   A. Mechanism .................................. 204
   B. Adenine Polarization .......................... 207
   C. Sugar Polarization ............................ 208
      i. Mechanisms ................................ 208
      ii. The Verdict .............................. 212
   D. Tetranucleotide .............................. 216
   E. tRNA^phe .................................... 217
6. Conclusions .................................. 223
References .................................... 224

APPENDIX A. Ethidium Ion Binding Studied By NMR ....... 229
1. Materials and Methods .......................... 232
2. Discussion .................................. 235

APPENDIX B. Distances Between Protons in B-form dCA_6G + dCT_6G .................................... 237
APPENDIX C. Theoretical Calculations of Chemical Shifts ...... 251

1. Theoretical Shielding Parameters ......................... 251

2. Empirically Determined High Temperature Chemical
   Shifts .................................................. 261

References .................................................. 265
List of Tables

2.1 Chemical Shifts of Imino Protons for dCA₆G + dCT₆G and dCA₃CA₃G + dCT₆G ........................................ 44

2.2 Changes in Chemical Shift Upon Double-Strand Formation ... 61

2.3 Thermodynamic Parameters for dCA₆G + dCT₆G and dCA₃CA₃G + dCT₆G .................................................. 68

3.1 Lifetimes for Exchange of Imino Protons at pH 7 in dCA₃CA₃G + dCT₆G ............................................. 95

3.2 Lifetimes for Exchange of Imino Protons at pH 7 in dCA₆G + dCT₆G ..................................................... 96

3.3 pH Dependence of Lifetimes in dCA₃CA₃G + dCT₆G ............. 97

3.4 pH Dependence of Lifetimes in dCA₆G + dCT₆G .................. 98

3.5 Activation Energies for Exchange in dCA₆G + dCT₆G and dCA₃CA₃G + dCT₆G .......................................... 112

3.6 Lifetimes for Exchange of Imino Protons in 12-mer at pH 6 ................................................................. 122

3.7 pH Dependence of Lifetimes in 12-mer ............................. 128

3.8 Lifetimes for Exchange of Imino Protons in 12-mer G·T at pH 6 ................................................................. 131

3.9 pH Dependence of Lifetimes in 12-mer G·T ......................... 134

3.10 Lifetimes for Exchange of Imino Protons in 13-mer at pH 8 ................................................................. 137

3.11 Activation Energies for Exchange in 12-mer, 12-mer G·T and 13-mer ......................................................... 144
4.1 Summary of Spin Polarization of Bases, Mononucleosides and Mononucleotides ..................................................... 183

4.2 Summary of Spin Polarization of Dinucleotides and Oligonucleotides .............................................................. 188

B.1 X, Y, Z Coordinates for B-form dCA₆G + dCT₆G ................. 240

B.2 Cylindrical Polar and Cartesian Coordinates for the Base Protons ................................................................. 242

B.3 Calculated Distances Between Protons in B-form dCA₆G + dCT₆G ................................................................. 243

C.1 Shielding Parameters for Protons in dCA₆G + dCT₆G............. 252

C.2 Shielding Parameters for Protons in dCA₃CA₃G + dCT₆G....... 253

C.3 Comparison of Empirically Calculated and Experimental Chemical Shifts for CA₆G ...................................................... 262

C.4 Comparison of Empirically Calculated and Experimental Chemical Shifts for CA₃CA₃G ................................................ 263
List of Figures

1.1 The chemical structure of a single strand of DNA ............ 4
1.2 Watson-Crick base pairs in DNA ........................................ 6
1.3 A schematic of the ring current effect ................................. 9
1.4 A schematic of the helix-to-coil transition ........................... 11

2.1 The NOE difference spectra of dCA₆G + dCT₆G ..................... 23
2.2 The NOE difference spectra of dCA₃CA₃G + dCT₆G .................. 25
2.3 The pulse sequence and parameters for NOE spectra ............. 27
2.4 ¹H NMR double quantum coherence spectrum of
    dCA₃CA₃G + dCT₆G .................................................. 30
2.5 The double quantum coherence pulse sequence ..................... 32
2.6 ¹H NMR spectra of the imino protons in dCA₅G + dCT₅G,
    dCA₃CA₃G + dCT₆G, and dCA₆G + dCT₆G ............................. 34
2.7 ¹H NMR spectrum of the imino protons in dGCA₂CA₂GC +
    dGCT₄GC .......................................................... 47
2.8 ¹H NMR spectrum of the non-exchangeable protons in
    dCA₃CA₃G + dCT₆G as a function of temperature ................ 50
2.9 Chemical shift vs. temperature for the C(H6) protons in
    dCA₆G + dCT₆G and dCA₃CA₃G + dCT₆G ............................ 53
2.10 Chemical shift vs. temperature for C(H6) protons in
    dCA₃CA₃G + dCT₆G: double strand vs. single strand ........... 55
2.11 Chemical shift vs. temperature for A(H2) protons in
\[ dCA_6G + dCT_6G \] ................................. 58

2.12 Chemical shift vs. temperature for A(H2) protons in
\[ dCA_3CA_3G + dCT_6G \] ............................... 60

2.13 Optical melting curves for \[ dCA_3CA_3G + dCT_6G \] ............................. 65

2.14 \( l/T_m \) vs. \( C_0 \) plot for \( dCA_3CA_3G + dCT_6G \) and \( dCA_6G + dCT_6G \) ...................................................... 67

3.1 A mechanism for chemical exchange of imino protons ........ 79

3.2 Kinetic analysis for the exchange mechanism............... 82

3.3 Temperature dependence of the NMR spectra of the imino
protons in \[ dCA_3CA_3G + dCT_6G \] ............................... 88

3.4 Temperature dependence of the NMR spectra of the imino
protons in \[ dCA_6G + dCT_6G \] ................................. 90

3.5 Spectra from a saturation recovery experiment on
\[ dCA_3CA_3G + dCT_6G \] ................................. 92

3.6 Spectra from a saturation recovery experiment on
\[ dCA_6G + dCT_6G \] ................................. 94

3.7 Arrhenius plots for \( dCA_3CA_3G + dCT_6G \) .................. 101

3.8 Arrhenius plots for \( dCA_3CA_3G + dCT_6G \) .................. 103

3.9 Arrhenius plots for \( dCA_3CA_3G + dCT_6G \) .................. 105

3.10 Arrhenius plots for \( dCA_6G + dCT_6G \) .................. 107

3.11 Arrhenius plots for \( dCA_6G + dCT_6G \) .................. 109

3.12 Arrhenius plots for \( dCA_6G + dCT_6G \) .................. 111

xii
4.11 The temperature dependence of the $^1$H NMR CIDNP spectra of yeast tRNA$^{\text{phe}}$ ........................................... 200

4.12 $^1$H NMR CIDNP spectra of tRNA$^{\text{phe}}$ and tRNA$^{\text{phe-Y}}$ .................. 203

A.1 The chemical structure of the ethidium ion .................. 231
A.2 $^1$H NMR spectra of the titration of rCA$_6$G + dCT$_6$G with ethidium bromide .......................... 234

B.1 The duplex dCA$_6$G + dCT$_6$G in B-form geometry ............. 239
B.2 The three base stack dCAA + dTTG .............................. 246
B.3 The three base stack dAAA + dTTT .............................. 248
B.4 The three base stack dAAG + dCTT .............................. 250

C.1 A view of d$^{\text{AA}}_{\text{TT}}$ down the helix axis ................. 256
C.2 A view of d$^{\text{CA}}_{\text{AT}}$ down the helix axis ................. 258
C.3 A view of d$^{\text{AC}}_{\text{AT}}$ down the helix axis ................. 260
Chapter I
Introduction

1. BACKGROUND

Over the past few years there has been a renewed interest in the conformation and dynamics of deoxyribonucleic acids (DNA). Several breakthroughs have contributed to this interest. Because of the quantities needed to do physical studies on nucleic acids, in the past, studies were restricted to either large pieces of DNA, synthetic polymers and random sequences such as calf thymus DNA, or to small oligonucleotides of known sequence. With recent improvements in the techniques of chemical synthesis, it is now possible to obtain substantial quantities of large oligonucleotides with known sequence. This has led to several important discoveries in the area of nucleic acid structure. One such discovery was made by Wang et al. (1979); they determined, using X-ray crystallography, that d(C-G)₃ forms a left-handed helix. The possible existence of a left-handed helix had been proposed earlier by the circular dichroism studies of Pohl and Jovin (1972). Several structures have been determined for right-handed oligonucleotide helices which also indicate that nucleic acids have great structural diversity (Dickerson & Drew, 1981; Shakked et al., 1981; Wang et al., 1982; Conner et al., 1982). If one considers that DNA must be versatile enough to perform many biological functions then it is perhaps not so surprising that the conformation must also be somewhat versatile.

The discovery of these structures by X-ray crystallography, led to the question of whether these conformations also existed in
solution. One of the most powerful techniques for studying the solution structure of nucleic acids is Nuclear Magnetic Resonance (NMR). In the past, using NMR to obtain detailed information about structure has been limited to small oligonucleotides because of the problems with resolution and assignment of the resonances. However, the advancement of NMR technology in recent years has allowed larger oligonucleotides to be studied. Some of the more important of these advancements are the development of higher field magnets and the use of new techniques, especially nuclear Overhauser effects, for assigning resonances.

One of the unique properties of NMR, as a tool to study nucleic acid conformation in solution, is the ability to obtain information about specific areas of a molecule. The basic structure of a single strand of DNA is shown in Figure 1.1. Of the nuclei found in this molecule, \(^1\text{H}\) and \(^3\text{P}\) are the most easily studied by NMR. Of these two nuclei, protons are most often used for studying conformation in solution. There are two types of protons, the sugar protons and the base protons. I will restrict this discussion to the base protons, however much work has been done on the sugar protons as well (see Sarma, 1980). When DNA forms a double strand, or helix, the bases hydrogen bond to form base pairs as shown in Figure 1.2. There are two types of base protons I will discuss, non-exchangeable and exchangeable.

The non-exchangeable protons are the \(\text{H}_8\) and \(\text{H}_2\) of adenine, the \(\text{H}_6\) of thymine, the \(\text{H}_8\) of guanine, and the \(\text{H}_5\) and \(\text{H}_6\) of cytosine. These protons resonate between 6 and 9 ppm, with the exception of
Figure 1.1 The chemical structure of a single strand of DNA.
Figure 1.2  Watson-Crick base pairs in DNA.
WATSON-CRICK BASE PAIRS

ADENINE

H-8
N
H
N
H-2
deoxyribose

THYMINE

H-6
N
H-3
N
H-6
O
CH3
deoxyribose

GUANINE

H-8
N
H
N
H-1
deoxyribose

CYTOSINE

H-5
H-6
N
H
N
H-1
deoxyribose
the cytosine H5 which resonates at higher field (lower ppm) than the other protons. These protons are also referred to as the aromatic protons because they are attached to the heteroaromatic ring system of the bases. The methyl protons of thymine are also non-exchangeable protons and they resonate at ~1-2 ppm. The main technique for studying conformation has been to study the temperature dependence of the chemical shifts of the protons in solution. The chemical shift, that is the frequency at which a proton resonates, is determined by the magnetic field experienced by the proton. In nucleic acids, the biggest effect on chemical shift comes from ring currents induced in the heteroaromatic ring systems of the bases. When placed in a field, the electrons in the aromatic ring begin to circulate. This induces a local magnetic field which is opposite to the applied magnetic field. The induced field is such that a proton directly above or below the aromatic ring will be shielded from the applied field while a proton in the plane of the ring will be deshielded. This effect is shown in Figure 1.3 and is called the ring current effect. An oligonucleotide duplex made up of complementary strands will form a double strand at low temperatures and dissociate into single strands at higher temperatures. This equilibrium is illustrated in Figure 1.4. In terms of the environment of a proton, one can think of this equilibrium as being between a state where the base pairs are stacked (the double strand) and a state where they are unstacked (the single strand at high temperature). Thus the chemical shift will change as the temperature changes and this can be used to study
Figure 1.3  A schematic of an aromatic ring in a magnetic field. In (a) $H_0$ is the applied field, the circle in the plane of the ring represents the current induced by the applied field and the lines coming out of the ring represent the magnetic field lines for the induced field. In (b) the (+) represents a region where a proton would be shielded and a (-) represents a region of deshielding.
Figure 1.4  A schematic representation of the single strand to double strand equilibrium (helix-to-coil transition).
Helix ↔ Coil
the change in environment of a proton. The local information which NMR can provide depends on being able to assign resonances to specific protons in the molecule. In the past, the method of incremental analysis has been used to assign resonances (Borer et al., 1975). This method requires NMR spectra of small oligonucleotide fragments from the molecule being studied. This can be a tedious task and requires the synthesis of many nucleotides. Recently several groups have demonstrated that observation of nuclear Overhauser effects (NOE) provides an alternative method for assigning the non-exchangeable resonances. The nuclear Overhauser effect allows for magnetization transfer between protons which are close in space without the necessity of being connected covalently. This method has been used to assign both the aromatic and the sugar protons (Kan et al., 1982; Petersheim & Turner, 1983; Reid et al., 1983). Because the nuclear Overhauser effect is extremely sensitive to distance, the intensity is inversely related to $r^6$, the technique can also be used to study conformation. Patel at al. (1982) have used NOE between the non-exchangeable protons to differentiate between syn and anti conformations of the bases in poly(dG–dC).

The exchangeable protons, are the protons involved in base pairing. These protons include the exocyclic amino protons as well as the imino protons [thymine(H3) and guanine(H1)]. The imino protons resonate from 12 to 15 ppm when they are hydrogen-bonded. The amino protons resonate in the same region as the aromatic protons (7–9 ppm). The exchangeable protons have been studied to
obtain information about the dynamics of the helix (Early et al., 1981; Pardi & Tinoco, 1982) but they can also be used to obtain information about conformation. These protons exchange with other protons or deuterons and thus must be studied in H₂O. The observation of a small concentration of protons from the nucleotides in a sea of protons from the H₂O requires the use of special pulse sequences which suppress the water signal. The technique I have used is the Redfield 2-1-4-1-2 pulse sequence (Redfield et al., 1975). However, several other pulse sequences for suppressing the water signal have recently been demonstrated (Wright et al., 1981; Haasnoot & Hilbers, 1983). The observation of nuclear Overhauser effects has also been used to assign the resonances of the imino protons in both tRNA (Roy & Redfield, 1981; Hare & Reid, 1982; Heerschap et al., 1982) and oligonucleotides (Kan et al., 1982; Chou et al., 1983). The sensitivity of NOE to distance has been used with imino protons to study the conformation of a G•T mismatch (Patel et al., 1982) and a G•A mismatch (Kan et al., 1983).

Recently, a very powerful NMR technique, two-dimensional Fourier transform NMR, has shown great promise for studying the conformation and dynamics of nucleic acids in solution (Feigon et al., 1982; Haasnoot & Hilbers, 1983). This technique has been applied to proteins with great success and has shown the potential for obtaining, in solution, the same type of information obtainable for crystals by X-ray crystallography. The technique of two-dimensional Fourier transform NMR may revolutionize the use of NMR to study nucleic acids in solution.
2. THE SCOPE OF THIS THESIS

In this thesis, I have used several different techniques to study the conformation and dynamics of oligonucleotides in solution.

In Chapter II, I discuss the conformation of an oligonucleotide duplex containing an extra cytosine, dCA$_3$CA$\cdot$G + dCT$_6$G. The conformation was determined through the observation of NOE and through the temperature dependence of the chemical shifts. The resonances were assigned using double quantum coherence techniques and NOE. UV absorption as a function of temperature was used to study the thermodynamic properties of the perturbed duplex.

In Chapter III, the kinetics for exchange of the imino protons with water is presented for the duplex containing the extra cytosine, and compared with the kinetics for exchange of the duplex without the perturbation (dCA$_6$G + dCT$_6$G). The method used was saturation recovery. The effects of an extra adenine and a G·T mismatch on the kinetics of exchange are also discussed and compared with the effects of the extra cytosine.

In Chapter IV, the solvent accessibility of oligonucleotides in solution was studied through the use of Chemically Induced Dynamic Nuclear Polarization (CIDNP). The mechanisms for this polarization were investigated by studying mononucleosides, mononucleotides, dinucleotides and a trimer. These results were applied to a double stranded tetranucleotide and to yeast tRNA$^{\text{Phe}}$. 
References


CHAPTER II

The Conformation and Stability of an Oligonucleotide Duplex Containing a Cytosine Bulge

1. SYNOPSIS

The conformation and thermodynamics of duplex formation are determined for the oligonucleotide duplex dCA₃CA₃G + dCT₆G which contains an extra, unpaired cytosine residue. The conformation is investigated using nuclear magnetic resonance, observing both the aromatic base protons and base pairing imino protons. The nuclear Overhauser effects between imino protons on either side of the unpaired cytosine show that the cytosine is extrahelical; the base is not stacked in the helix. The temperature dependence of the chemical shifts of the aromatic base protons support this conclusion. The observed nuclear Overhauser effects and a large upfield shift in the imino proton resonance of A·T base pair 4, adjacent to the extrahelical cytosine indicate there is a perturbation in the hydrogen bonding of this A·T base pair as well as a change in the overlap with adjacent base pairs. The nuclear Overhauser effect is also used to assign both the imino and A(H2) resonances in the dCA₃CA₃G + dCT₆G duplex as well as the dCA₆G + dCT₆G duplex. Thermodynamic parameters are determined for both duplexes. The extrahelical cytosine causes a decrease in the duplex melting temperature of 15°C, at a concentration of 200µM per single strand. The standard free energy for duplex formation at 25°C is 2.9 kcal mol⁻¹ more positive for the duplex with the extra cytosine. This
decrease in stability for the perturbed helix is caused by an unfavorable change in enthalpy and a favorable change in entropy for duplex formation.

2. INTRODUCTION

The ability to predict the structure of RNA or DNA is based on an understanding of the energetics involved in forming that structure. Although there are data available on the thermodynamics of double strand formation in RNA (Borer et al., 1974) there are limited data for DNA. Predictions for the thermodynamics of duplex formation in DNA helices must, presently, be done using the RNA values. Relatively little is known about the thermodynamics for the formation of secondary structures in RNA, such as hairpins (Uhlenbeck et al., 1973; Gralla & Crothers, 1973), mismatched bases, or base bulges (Fink & Crothers, 1972), and even less is known for DNA. Previous investigations have probed the effects of base pair mismatches in synthetic polynucleotide helices (for a review, see Lomant and Fresco, 1974). These studies used optical properties to determine if the polynucleotides formed helices containing non-complementary base oppositions, or if the bases, either stacked or not stacked in the helix, had no base opposition on the cross strand. Evidence for looped out bases in polynucleotides has been obtained using circular dichroism (Gray et al., 1980) and photodimerization (Evans and Morgan, 1982). The free energy for a bulge defect has been measured for poly(A,A*) + poly(U), where A* is the N-1 oxide of adenine (Fink and Crothers, 1972). The energetics for base bulges may be valuable in investigating mechanisms of muta-
genesis, in particular frameshift mutagenesis (Streisinger et al., 1966).

In dC-G-C-G-A-A-T-T-C-G-C-G the extra adenes were found to be stacked in the helix upon double strand formation (Patel et al., 1982). In this chapter, I will demonstrate the use of nuclear magnetic resonance to study the deoxyoligonucleotide duplex, dCA₃CA₃G + dCT₆G, in comparison with the parent duplex, dCA₆G + dCT₆G. By analyzing the temperature dependence of the chemical shifts of aromatic base protons and the nuclear Overhauser effects (NOE) between imino protons, I have determined that the unpaired cytosine is not stacked in the helix. UV absorption melting curves have been measured and analyzed, for these helices, to obtain the standard free energies, enthalpies and entropies of duplex formation.

3. MATERIALS AND METHODS

A. Sample Preparation

All of the oligonucleotides were synthesized by one of two techniques; dCT₆G, dCA₅G and dCT₅G were synthesized using diester solution techniques (Khorana, 1968); dCA₃CA₃G, dCA₆G, dGCT₄GC, and dGCA₂CA₂GC were synthesized using phosphoramidite solid support techniques (Beaucage & Caruthers, 1981; Matteucci & Caruthers, 1981). They were desalted several times using Biogel P-2 columns and then lyophilized for storage. Extinction coefficients were estimated from the coefficients for mononucleotides and dinucleoside phosphates (Handbook of Biochem. & Mol. Biol., 1975). The extinction coefficients per mole of strand at 260 nm and 25°C for dCA₃CA₃G, dCA₆G, dCA₅G, dCT₅G, dCT₆G, dGCT₄GC, and dGCA₂CA₂GC are
$98 \times 10^3$, $91 \times 10^3$, $79 \times 10^3$, $58 \times 10^3$, $66 \times 10^3$, $66 \times 10^3$, and $91 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$, respectively.

NMR samples were prepared with buffer made up of 28 mM phosphate, 0.1 mM EDTA and 0.2 M NaCl. D$_2$O samples containing oligonucleotide and buffer were lyophilized 3 or 4 times from 99.98% D$_2$O (Biorad) and finally dissolved in 100% D$_2$O (Biorad) in a glove bag under N$_2$ atmosphere. H$_2$O samples were prepared following the same procedure, but were lyophilized from H$_2$O and prepared outside the glove bag. The water samples contained 10-15% D$_2$O. All H$_2$O samples were adjusted to pH = 7 ± 0.1 and D$_2$O samples were adjusted to pD = 7 ± 0.1 (measured as 6.6 on a pH meter). All samples have sodium 3-(trimethylsilyl)propionate-2,2,3,3-d$_4$, TSP, added as an internal standard.

Samples used for thermodynamic studies were lyophilized from H$_2$O and then dissolved in a 0.01M phosphate buffer, pH=7, containing 0.1mM EDTA and 1M NaCl. The optical melting curves were obtained using a procedure described previously (Nelson et al., 1981).

B. Experimental Techniques

The spectra taken in D$_2$O were all obtained using the 360 MHz instrument at the Stanford Magnetic Resonance Laboratory. Typically, spectra were the sum of 100 accumulations, each accumulation being collected with 16K points and a sweep width of ±1800 Hz. The temperature for spectra obtained using the 360 MHz instrument was regulated to ±1°C using a B-ST 100/700 Bruker temperature controller. The accuracy of the controller was monitored periodically by placing a copper-constantan thermocouple
into an NMR tube containing \( \text{H}_2\text{O} \) and then placing the tube into the magnet. These checks were made every 5-10°C.

The spectra taken in \( \text{H}_2\text{O} \) for the temperature and kinetic studies of the base pairing imino protons were obtained using the 360 MHz instrument at the Stanford Magnetic Resonance Laboratory. These spectra typically were obtained using an 8K block size, sweep width of ±5000 Hz and the sum of 500-1000 accumulations. All of these spectra were obtained using a Redfield 214 pulse sequence (Redfield et al., 1975).

The NOE spectra in Figures 2.1 and 2.2 were also obtained in \( \text{H}_2\text{O} \) using a Bruker WM500 instrument (500 MHz) at the University of Washington, Seattle for the dCA\(_3\)CA\(_3\)G + dCT\(_6\)G and the 500 MHz Nicolet instrument at the UC Davis NMR facility for the dCA\(_6\)G + dCT\(_6\)G. Both sets of spectra were obtained using the pulse sequence shown in Figure 2.3. The difference spectra for dCA\(_3\)CA\(_3\)G + dCT\(_6\)G represent 12,000 accumulations collected in a cycle of 16 scans with irradiation on resonance followed by 16 scans with irradiation off resonance. The block of on resonance acquisitions is subtracted from the block of off resonance acquisitions during the course of the experiment so that the Fourier transform of the final free induction decay (FID) gives the NOE difference spectrum. The difference spectra for dCA\(_6\)G + dCT\(_6\)G, obtained on the UC Davis instrument represent 12,000 scans collected in a cycle of 40 acquisitions with irradiation on resonance followed by 40 acquisitions with irradiation off resonance. The major difference with these spectra, however, is that on and off resonance acquisitions are collected in separate
Figure 2.1. The 500 MHz $^1$H NOE difference spectra from dCA$_6$G + dCT$_6$G. Spectra b–h are the difference of 6,000 scans irradiating off resonance and 6,000 scans irradiating on resonance. Spectrum a was obtained without irradiation using a Redfield pulse sequence (with 90° phase shifting of the second and fourth pulse intervals) for observation. The sample was 1mM per single strand in 90% H$_2$O.
Figure 2.2. The 500 MHz $^1$H NOE difference spectra from dCA$_3$CA$_3$G + dCT$_6$G. Spectra b–g are the difference of 6,000 scans irradiating off resonance and 6,000 scans irradiating on resonance. Spectrum a was obtained using a Redfield pulse with no irradiation. The sample was 1mM per single strand in 90% H$_2$O.
Figure 2.3. The pulse sequence and parameters used to obtain NOE difference spectra.
D6 = 500 msec, decoupler on
D4 = 1 msec
P2 = 0.056 msec
P1 = 0.028 msec
P4 = 0.120 msec

Acquisition Time = 279 msec
D5 = 10 msec

Offset from H2O = 3570 Hz
Sweep width = ± 7142
Block size = 8 K

* The Redfield pulse sequence contains P1 intervals which are 180° phase shifted. The modified Redfield pulse sequence contains P1 intervals which are 90° phase shifted.
blocks of memory. After the experiment is completed these blocks are Fourier transformed individually and then subtracted to give the difference spectra as shown in Figure 2.1.

The double quantum coherence spectrum shown in Figure 2.4 was obtained using the 500 MHz instrument at the UC Davis NMR facility. The pulse sequence and the pulse program used are shown in Figure 2.5. This pulse sequence has been discussed by Hore et al. (1982a,b).

4. NMR RESULTS AND DISCUSSION

Before discussing the thermodynamic results it is helpful to know more about the conformation of the oligonucleotide duplex with the extra base. The conformation has been studied by NMR, both by using NOE to probe proton-proton distances in the duplex and by investigating the temperature dependence of the chemical shifts of aromatic protons.

A. Imino - Imino NOE

Spectra of the base pairing imino protons from dCA_5G + dCT_5G, dCA_3CA_3G + dCT_6G, and dCA_6G + dCT_6G, along with the numbering convention for each, are shown in Figure 2.6. The assignments of the imino resonances for dCA_6G + dCT_6G were determined by imino-imino NOE, as discussed below. The assignments of the resonances from base pairs 1, 7, and 8 in the dCA_3CA_3G + dCT_6G duplex were made in analogy to the assignments for dCA_6G + dCT_6G and the assignments for the dCA_5G + dCT_5G duplex determined by Pardi et al. (1981).
Figure 2.4. The 500 MHz $^1$H NMR spectrum of dCA$_3$CA$_3$G + dCT$_6$G. The top spectrum is the sum of 10,400 scans using the double quantum pulse sequence shown in Figure 2.5. The bottom spectrum is the single quantum spectrum of the same sample. The sample was 0.5mM per single strand in D$_2$O.
Figure 2.5. The double quantum coherence pulse sequence and the pulse programs for the Nicolet 1280 to carry out the 16 step cycle and the 32 step cycle experiments, as described in Hore et. al. (1982 a & b). The symbols after the "/" define the phase of the transmitter pulse, relative to the x axis of the rotating frame. Using the "@" allows the use of phase expressions, as given at the bottom of the figure. In step #8 the "@" defines the receiver phase. The phase expressions define the phase as a function of the scan count S, where S = 0, 1, 2, etc. Both the overall expression and any expression in parentheses are evaluated in modulo 4. The final phase values will therefore be 0, 1, 2, or 3; the definition of these values is 0 = 0° of phase shift, 1 = 90°, 2 = 180°, 3 = 270° (or -90°). The phase expression @S used during acquisition will result in standard quadrature phase detection. (See Nicolet 1280 manual for further discussion.)
16 STEP CYCLE

#1. P2
#2. D3
#3. P1/1
#4. D3
#5. P2
#6. D8
#7. P2/αA+0
#8. AαB
#9. D2
#10. D5 JUMP TO #1

PHASE A = S
PHASE B = -S

32 STEP CYCLE

#1. P2/αD + 0
#2. D3
#3. P1/αB + 1
#4. D3
#5. P2/αB + 0
#6. D8
#7. P2/αC + 0
#8. AαA
#9. D2
#10. D5 JUMP TO #1

PHASE A = (S/4)+((S/16)×2)-S
PHASE B = S/4
PHASE C = S+(S/4)
PHASE D = (S/4)+((S/16)×2)
Figure 2.6. The 360 MHz $^1$H NMR spectra of the base pairing imino protons from dCA$_5$G + dCT$_5$G, dCA$_3$CA$_3$G + dCT$_6$G, and dCA$_6$G + dCT$_6$G. Samples were 1mM per single strand in H$_2$O. Spectra are the sum of 400-600 accumulations.
The nuclear Overhauser effect is a process where magnetization can be transferred from one nucleus to another through space (Noggle and Schirmer, 1971). The process depends inversely on the sixth power of the distances between the nuclei. The NOE is also affected by several other factors. The intensity is dependent on the correlation time of the molecule, which depends on the length of the duplex. The NOE can also be in competition with other possible relaxation processes, such as exchange with water. I have used this technique to assign resonances and to obtain information about the conformation around the extra cytosine. In a regular helix the distance between imino protons on neighboring base pairs is less than 4Å, a reasonable distance to see NOE's. In A·T base pairs there is also an NOE from the imino proton to the adenine(H2) on the same base pair which is characterized by a sharp resonance in the aromatic region (6-9 ppm) of the difference spectrum. This NOE has been used to assign imino resonances to A·T (or A·U) vs. G·C base pairs (Sanchez et al., 1980; Hare and Reid, 1982). NOE can also be observed from imino protons to amino protons or adenine (H2)'s on the neighboring base pairs. The interproton distances for the dCA₆G + dCT₆G duplex, assuming a B-form geometry, have been calculated and are presented in Appendix B.

The NOE difference spectra of dCA₆G + dCT₆G are shown in Figure 2.1b-h, as well as a spectrum showing all the imino and aromatic resonances for reference (Figure 2.1a). In spectra c,d,e,f and g, NOE from an imino proton to a sharp line in the aromatic region upfield indicates that the saturated resonances are due to A·T base
pairs. The two imino resonances furthest upfield remain to be assigned to G·C base pairs 1 or 8. Ring current calculations done for the duplex dCA₅G + dCT₅G (Pardi, 1980) predict that the imino resonances can be assigned as shown in Figure 2.6a. The assignments of the terminal G·C base pairs in dCA₆G + dCT₆G can be predicted by the same ring current calculations and are given in Figure 2.6c. By starting with the assignment of one of the imino resonances one can "walk" down the helix by observing NOE from one imino proton to the proton or protons on neighboring base pairs. This procedure is illustrated in Figure 2.1b-h and is discussed below.

Experimentally one observes NOE by saturating the resonance of interest and looking for an effect on the other resonances. The effect will be either an increase or a decrease in the intensity of the other resonances. In the case of oligonucleotide helices 7-9 base pairs in length a negative NOE, or a decrease in the signal intensity is expected (Noggle and Schirmer, 1971; Brothner-By, 1979). As mentioned before, the NOE is inversely proportional to the sixth power of the interproton distance. This severely limits the protons to which NOE will be observable when saturating an imino resonance (see Appendix B). However, we do expect to see an effect between an imino proton and the imino proton on the neighboring base pairs. For examples of this we need only look at the difference spectra in the region between 12.5 and 15 ppm. Thus in Figure 2.1h we see NOE from base pair 8 to the imino resonance from A·T base pair 7. In Figure 2.1g we observe NOE to A·T base pair 6 and a small effect to G·C base pair 8 when saturating the imino resonance
from base pair 7. The imino protons on the terminal G·C base pairs exchange more quickly with water than the imino protons from the other base pairs, as will be discussed in Chapter III. This exchange competes with the NOE, therefore it is more difficult to observe NOE to and from the terminal base pairs. From the imino proton in base pair 6 we observe NOE to imino protons in base pairs 5 and 7 as shown in Figure 2.1f. In Figure 2.1e, when saturating the resonance from A·T base pair 5 NOE is observed to the resonances due to base pairs 4 and 6. Analysis of the difference spectra in Figures 2.1d and c is complicated by having two imino protons resonating at 14.09 ppm. Therefore, in Figure 2.1d, observation of NOE from the imino proton from base pair 3 to only one other imino resonance indicates that the imino proton on base pair 3 is between the two imino protons which both resonate at 14.09 ppm. Therefore this resonance has been assigned to A·T base pairs 2 and 4. This assignment is also supported by the observed NOE's from base pairs 1 and 5, as shown in Figures 2.1b and e, respectively. The assignments of the double resonance at 14.09 ppm can also be supported by the difference spectrum in Figure 2.1c. When saturating this resonance, which has been assigned to base pairs 2 and 4, we indeed observe NOE to base pair 5 and to base pair 3. The NOE to base pair 3 is significantly larger in intensity than the NOE to base pair 5 and therefore it is reasonable to believe it may be due to saturating both the resonances at 14.09 ppm. That is, the imino proton on base pair 3 is getting NOE from imino protons on both base pairs 2 and 4. There is no NOE observed to base pair 1, however this is
probably due to the rapid exchange of the imino proton in base pair 1 with water (see Chapter III). However, in Figure 2.1b we can observe the results of the reverse experiment. By saturating the imino resonance from base pair 1 with slightly more power we are able to observe an NOE to base pair 2. Thus we have been able to observe NOE from one imino proton to the imino proton on a neighboring base pair for the entire length of the duplex. This procedure results in the assignment of the imino resonances in dCA₆G + dCT₆G as shown in Figure 2.1a.

The NOE difference spectra for dCA₃CA₃G + dCT₆G are shown in Figure 2.2 along with a Redfield spectrum showing all the imino and aromatic resonances for reference. In spectra 2.2b, c, d, e and f, NOE to a sharp line in the aromatic region, upfield from the imino region, indicates the saturated resonances are due to AT base pairs. The NOE difference spectra can be analyzed using the same procedure as that used for dCA₆G + dCT₆G.

In Figure 2.2b saturating the imino resonance from base pair 2 should show an NOE to resonances from base pairs 1 and 3. Base pair 1 is one of the terminal G C base pairs and is in rapid exchange with water (see Chapter III). Thus, observation of NOE to this resonance is not likely. As expected, only a peak from base pair 3 is observed in the difference spectrum in Figure 2.2b. It should be noted that this resonance could also be due to partial saturation as it is close to the resonance for base pairs 2 and 5. Saturation of the resonance from base pair 5 shows NOE to base pairs 4 and 6, as expected. NOE from base pair 3 to base pair 4 and possibly 2 is
shown in Figure 2.2c. The peak from base pairs 2 or 5 may again be due to partial saturation. NOE is observed from the imino resonance for base pair 4 to base pairs 3 and 5 in Figure 2.2d. In Figure 2.2e NOE is observed to base pairs 5 and 7 when saturating the imino resonance from base pair 6. In Figure 2.2f NOE is observed from base pair 7 to base pair 6. No NOE is observed to base pair 8, because this resonance has a rapid relaxation rate. This is also a problem when trying to observe NOE from the imino resonance of base pair 8 and, as shown in Figure 2.2g, none are observed in the lowfield region. From the observation of imino-imino NOE, described above, we have assigned the lowfield resonances as given in Figure 2.2a.

B. Imino - A(H2) NOE

As discussed previously and as calculated in Appendix B, there are protons other than the neighboring imino protons which are close enough to show NOE when saturating an imino resonance. Another observable NOE (and the only other one I will discuss) is from imino protons, T(H3)'s, to A(H2)'s. As shown in Appendix B, in an A·T base pair the closest proton to the imino proton is the A(H2) on the same base pair; NOE between these protons will be referred to as intra-base-pair NOE. As discussed previously this will result in a sharp resonance in the aromatic region (6-9 ppm). NOE can also be observed from an imino proton to the A(H2)'s on neighboring base pairs; this will be referred to as inter-base-pair NOE. In the block of A·T base pairs in dCA₆G + dCT₆G only the A(H2) on the neighboring base pair in the 3' direction, relative to the dCA₆G
strand, is close enough to show inter-base-pair NOE. The inter-base-pair NOE involves larger interproton distances and thus is expected to be weaker than the intra-base-pair NOE. Observation of NOE to the A(H2) resonances can be used to assign these resonances and to confirm the assignments of the imino protons. This is illustrated in Figures 2.1 and 2.2 and is discussed below.

The experimental NOE's from imino protons to A(H2)'s in dCA6G + dCT6G are shown in Figures 2.1b-h. These results are in good agreement with the expected NOE discussed previously. In Figure 2.1g saturation of the imino resonance from base pair 7 results in an intra-base-pair NOE to the A7(H2). In the 3' direction from base pair 7 is base pair 8 which is a G·C base pair. Therefore, there is no A(H2) in the 3' direction and the intra-base-pair NOE is the only imino-A(H2) NOE observed from base pair 7. Figure 2.1f shows both the intra- and inter-base-pair NOE. When the imino resonance from base pair 6 is saturated, intra-base-pair NOE is observed to A6(H2) and inter-base-pair NOE is observed to A7(H2). Figures 2.1e and d show this same pattern; one resonance from the intra-base-pair NOE and another weaker resonance from the inter-base-pair NOE. In Figure 2.1c the double resonance for base pairs 2 and 4 is being saturated. Thus, there is an intra-base-pair NOE to A2(H2) and A4(H2) and an inter-base-pair NOE to A3(H2) and presumably to A5(H2). The inter-base-pair NOE to A5(H2) is difficult to assign unambiguously because A4(H2) and A5(H2) have the same chemical shifts. In Figure 2.1b the terminal G·C base pair is saturated. There is no intra-base-pair H2 in a G·C base pair, however there is
an inter-base-pair NOE to the A2(H2), as predicted by distances calculated in Appendix B. The broad resonances in the region from 6-9 ppm are due to NOE between imino and amino protons. The only inter-base-pair NOE which is observed from an imino proton to an A(H2) on the 5' side is shown in Figure 2.1h. This NOE was also predicted by the distances calculated in Appendix B.

The experimental NOE's from imino protons to A(H2)'s in dCA3CA3G + dCT6G are shown in Figures 2.2b-g. These NOE are very similar to the NOE's from dCA6G + dCT6G, discussed above. Figures 2.2f and e show the same inter- and intra-base-pair NOE's as observed in Figures 2.1g and f, respectively, and the discussion need not be repeated. However, the other NOE are complicated by the presence of overlapping peaks and therefore will be discussed. In Figure 2.2b, the imino resonances from both base pairs 2 and 5 are saturated and result in intra-base-pair NOE to A2(H2) and A5(H2) and inter-base-pair NOE to A3(H2) and A6(H2), respectively. In Figure 2.2c saturation of the imino resonance from base pair 3 gives intra-base-pair NOE to A3(H2) and inter-base-pair NOE to A4(H2), as expected. We also observe resonances for A2(H2) and A5(H2), however, this is probably due to partial saturation of these resonances because they are close to the resonance for base pair 3. The only example of an NOE in the block of A-T base pairs that does not follow the trend is shown in Figure 2.2d. Saturation of the imino resonance from base pair 4, one of the base pairs adjacent to the unpaired C, shows intra-base-pair NOE to A4(H2) and inter-base-pair NOE to A5(H2), as expected. In addition, we observe another inter-
base-pair NOE to the A3(H2) which is the neighboring base pair on the 5' side. This implies that the distance between the imino proton from base pair 4 and the H2 proton on base pair 3 is less than the analogous distance in other parts of the duplex, or in the parent duplex dCA6G + dCT6G, where similar NOE is not observed. This decrease in distance could be achieved by a decrease in winding angle between these two base pairs (3 and 4) or a translation of one base pair relative to the other.

C. Conclusions from the Observed NOE

The observation of NOE, both imino-imino NOE and imino-A(H2) NOE, from base pair 4 to base pair 5 characterizes the conformation of the unpaired cytosine. If the extra cytosine were stacked in the helix, the distance between base pairs 4 and 5 would be almost 7 Å and no imino-imino NOE would be observed. However, as shown in Figure 2.2d we do observe NOE from the imino proton on base pair 4 to the imino proton on base pair 5. Thus, the unpaired cytosine cannot be stacked in the helix. This is corroborated by the observation of NOE from the imino proton on base pair 4 to the A5(H2), as shown in Figure 2.2d.

There are several other potential assignment schemes for the imino protons. However, these can be ruled out, either because of inconsistencies in the imino-imino NOE's or inconsistencies in the imino-A(H2) NOE's. There is one assignment scheme for the imino protons, other than the one presented, that can not be ruled out by any obvious inconsistencies. This alternate assignment consists of exchanging the assignments for the imino resonances from base
pairs 3 and 4, compared to those shown in Figure 2.2. Assignment of the A(H2) resonances can be made consistent with this imino assignment by exchanging the A(H2) assignments for base pair 3 and 4 as well as base pair 2 and 5. With this alternative assignment scheme, NOE is observed from the imino protons of base pair 3, 4, and 5 to the A(H2) on the cross-strand and to the A(H2)'s on both of the neighboring base pairs. This is not consistent with the expected NOE from a regular DNA helix of this sequence, but the assignment can not be ruled out completely on this basis. There is NOE observed from the imino proton of base pair 4 to the A5(H2) and from the imino proton of base pair 5 to the A4(H2), so this alternate assignment scheme is also consistent with the extra cytosine being out of the helix. However, this alternate assignment does not agree as well with the trends found for the NOE in dCA₆G + dCT₆G. Therefore, I favor the original assignments.

With the imino proton resonances assigned, the spectra of the imino protons for dCA₃CA₃G + dCT₆G and dCA₆G + dCT₆G can be compared as shown in Figure 2.6 and summarized in Table 2.1. The largest change in chemical shift is for base pair 4 which has shifted upfield by 0.6 ppm. Some of this shift could be due to the increased overlap with A·T base pair 3, as discussed previously. The maximum shielding value for adenine is ~1.6 ppm for a proton 3.4 Å above the center of the base. However, most of the ring current effect due to this A·T base pair is already present in dCA₆G + dCT₆G so 0.6 ppm is a large shift to be explained by only an improved overlap. I propose that some of the upfield shift is due to a
Table 2.1 Chemical Shifts of the Imino Protons (ppm), 0°C

<table>
<thead>
<tr>
<th>Base Pair #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>δ₁ - δ₁I</td>
<td>0.11</td>
<td>-0.05</td>
<td>-0.10</td>
<td>0.64</td>
<td>0.11</td>
<td>0.01</td>
<td>-0.01</td>
<td>-0.02</td>
</tr>
</tbody>
</table>
lengthened and weakened hydrogen bond in A·T base pair 4. Work done by Wagner et al. (1983) on proteins indicates a strong correlation between H-bond length and chemical shift. It is unusual that the imino proton from an A·T base pair will resonate so far upfield; in this case it is more upfield than G·C base pair 8. I have taken the NMR spectrum of another duplex containing an extra cytosine which provides more evidence that the imino proton from an A·T base pair can resonate in this upfield region. Figure 2.7 shows the spectrum of the imino resonances for the duplex dGCA<sub>2</sub>CA<sub>2</sub>GC + dGCT<sub>4</sub>GC. In this spectrum there are three resonances between 14-14.7 ppm presumably due to imino protons in A·T base pairs. There are four resonances between 12.8-13.1 ppm presumably due to G·C base pairs. This leaves the resonance at ~13.5 ppm to be assigned to an imino proton in A·T base pair. This is very close to the chemical shift of 13.45 ppm for the imino proton from A·T base pair 4 in dCA<sub>3</sub>CA<sub>3</sub>G + dCT<sub>6</sub>G. It should also be noted that the resonance from A·T base pair 4 in dCA<sub>3</sub>CA<sub>3</sub>G + dCT<sub>6</sub>G is broader than the other A·T resonances and may indicate faster exchange with water. This increase in exchange rate would be expected if the hydrogen bond were weaker. These exchange rates are discussed in Chapter III.

D. Non-Exchangeable Protons

More evidence for the cytosine being extrahelical is obtained by monitoring the chemical shifts of the aromatic proton resonances and their temperature dependence. When the aromatic bases are placed in a magnetic field there is a local field induced, thus, a proton directly above or below the base will be shielded from the
Figure 2.7. The 360 MHz $^1$H NMR spectrum of the base pairing imino protons of $dGCA_2CA_2GC + dGCT_4GC$, at $1^\circ C$. The sample was 0.5mM per single strand in 90% H$_2$O. The spectrum is the sum of 1000 scans.
applied field and a proton in the plane of the base will be de-shielded. For an aromatic proton in a double strand, such ring current effects will be mainly due to the base pairs above and below the proton and from the base it is base paired to on the cross strand. The equilibrium between double-strand and single-strand formation changes as a function of the temperature, thus the environment around a proton changes as the temperature changes. This is observed in the NMR spectrum as a change in the chemical shift and can be seen in Figure 2.8.

In order to monitor the temperature dependence of the chemical shift of any individual proton the resonances must be assigned. At low temperature the resonances are broad due to exchange between the single strand and double strand environment. This broadening makes assignment of the resonances more difficult. Assignment of the cytosine doublets at low temperature was accomplished using a double quantum pulse sequence (spectrum shown in top of Figure 2.4). This pulse sequence creates double quantum coherence and allows observation of only those resonances which have double quantum transitions (Hore et al., 1982a,b). The region of the spectrum between 6.5 - 8.5 ppm contains only singlets and doublets and becomes simplified to contain only doublets when the double quantum pulse sequence is used. It is characteristic of this pulse sequence that the two peaks in the doublet will be 180° out of phase, as seen in the top of Figure 2.4. The low temperature chemical shifts of the cytosine doublets were obtained from the spectrum in Figure 2.4. The high temperature chemical shifts of the cytosine doublets in dCA₃CA₃G +
Figure 2.8. The 360 MHz $^1$H NMR spectrum of dCA$_3$CA$_3$G + dCT$_6$G as a function of temperature. The sample is 1mM per single strand in D$_2$O.
dCT₆G can be assigned by comparison with dCA₅G + dCT₂G (Pardi et al., 1981) and with dCA₆G + dCT₆G. In Figure 2.9 the chemical shifts as a function of temperature are compared for dCA₃CA₃G + dCT₆G and dCA₆G + dCT₆G. The only major difference between the curves for the two terminal cytosines is that the curve for the dCA₆G + dCT₆G resonances shift to higher temperature due to the increased melting temperature of this duplex. The curve at the bottom of Figure 2.9 can now be assigned to the extra cytosine in dCA₃CA₃G + dCT₆G. Because this duplex has non-self-complementary stands, the single-strand chemical shifts as a function of temperature can also be obtained. The chemical shift versus temperature of dCA₃CA₃G, dCT₆G, and the double strand are shown in Figure 2.10. As duplex formation occurs, with decreasing temperature, both the Cl(H₆) and the C₈(H₆) show an increase in chemical shift, a downfield shift, due to the deshielding effect of the bases on the cross strand. This deshielding will not be present for the extra cytosine, whether or not it is stacked in the helix, because there is no base directly across from it on the cross strand. Chemical shift calculations done for the extra cytosine, assuming it is stacked in the helix and all bases are stacked in a B-form geometry around it, predict an upfield shift on going from high to low temperature (see Appendix C). If the cytosine were stacked in the helix, it would be shielded by the A·T base pairs around it. However, what is observed in Figure 2.10 is a large (0.3 ppm) increase in the chemical shift for this resonance, relative to the chemical shift in the single strand. Also note that the chemical
Figure 2.9. Chemical Shift vs. Temperature curves for the C(H6) protons in (x) dCA₃CA₃G + dCT₆G and (Δ) dCA₆G + dCT₆G. The samples were 1mM per single strand in D₂O.
Figure 2.10. Chemical Shift vs. Temperature curves for the C(H6) protons in dCA₃CA₃G + dCT₆G. Both the (x) double strand and the (o) single strand curves are shown. The sample was 1mM per single strand in D₂O.
shift of the extra cytosine at high temperature, where the bases are partially unstacked, is almost identical to the shift at low temperature, where the bases should be stacked. If the unpaired cytosine were extrahelical then the environment around that base could be similar to the environment in the unstacked or single strand state. This chemical shift behavior suggests that the extra cytosine is not stacked in the helix.

The problem of assigning resonances at low temperature has been solved for the A(H2) resonances by observing NOE from imino protons to the A(H2) protons. These assignments were discussed previously and are shown in Figures 2.1 and 2.2. At high temperature the A(H2) resonances can be identified because the H2 protons have a longer T1 than the other protons (Ts'o et. al., 1973). The only problem that remains is to assign the resonances to specific bases in the strand. This has been done by comparing experimental values of the A(H2) chemical shifts with values calculated empirically by the method of Hader et. al. (1982). These calculated values for the chemical shift are given in the second part of Appendix C. The experimental curves of chemical shift versus temperature are shown in Figure 2.11 for dCA6G + dCT6G and in Figure 2.12 for dCA3CA3G + dCT6G. Table 2.2 summarizes the changes in chemical shift upon double strand formation for both the duplexes. The value for Δδ' is calculated by subtracting the chemical shift in the double strand at 1°C from the chemical shift in the single strand at 60°C. This can be thought of as the change in chemical shift due to stacking and double-strand formation. The value for Δδ'' is calculated by
Figure 2.11. Chemical Shift vs. Temperature curves for the A(H2) protons in dCA₆G + dCT₆G. The top plot (a) shows the single strand dCA₆G and the bottom plot (b) shows the helix-to-coil behavior for the duplex. The sample was 1mM per single strand in D₂O.
Figure 2.12. Chemical Shift vs. Temperature curves for the A(H2) protons in dCA₃CA₃G + dCT₆G. The top plot (a) shows the single strand dCA₃CA₃G and the bottom plot (b) shows the helix-to-coil behavior for the duplex. The sample was 1mM per single strand in D₂O.
```
1 2 3 4 5 6 7 8
G-T-T-T-T-T-T-Cd
```

**Diagram a:**
- Graph showing temperature (TEMP C) vs. pH
- Data points for different samples labeled A2(H2), A3(H2), A7(H2), A6(H2), A5(H2), A4(H2)

**Diagram b:**
- Graph showing temperature (TEMP C) vs. PPM
- Data points for different samples labeled A2(H2), A7(H2), A3(H2), A6(H2), A5(H2), A4(H2)
Table 2.2 Changes in Chemical Shift Upon Double-Strand Formation

\[ \delta_{ss,60°} - \delta_{ds,1°} = \Delta \delta' \]

\[ \delta^*_{ss,1°} - \delta_{ds,1°} = \Delta \delta'' \]

<table>
<thead>
<tr>
<th></th>
<th>( \delta_{ss,60°} )</th>
<th>( \delta_{ds,1°} )</th>
<th>( \Delta \delta' )</th>
<th>( \delta^*_{ss,1°} )</th>
<th>( \delta_{ds,1°} )</th>
<th>( \Delta \delta'' )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A2(H2)</strong></td>
<td>7.97</td>
<td>7.30</td>
<td>0.67</td>
<td>7.73</td>
<td>7.30</td>
<td>0.44</td>
</tr>
<tr>
<td>A3</td>
<td>7.73</td>
<td>2.04</td>
<td>0.69</td>
<td>7.37</td>
<td>7.04</td>
<td>0.34</td>
</tr>
<tr>
<td>A4</td>
<td>7.65</td>
<td>6.92</td>
<td>0.73</td>
<td>7.29</td>
<td>6.92</td>
<td>0.37</td>
</tr>
<tr>
<td>A5</td>
<td>7.66</td>
<td>6.91</td>
<td>0.75</td>
<td>7.30</td>
<td>6.91</td>
<td>0.39</td>
</tr>
<tr>
<td>A6</td>
<td>7.70</td>
<td>7.04</td>
<td>0.66</td>
<td>7.37</td>
<td>7.04</td>
<td>0.34</td>
</tr>
<tr>
<td>A7</td>
<td>7.84</td>
<td>7.59</td>
<td>0.25</td>
<td>7.59</td>
<td>7.59</td>
<td>0.003</td>
</tr>
</tbody>
</table>

\[ \delta_{ss,60°} - \delta_{ds,1°} = \Delta \delta' \]

\[ \delta^*_{ss,1°} - \delta_{ds,1°} = \Delta \delta'' \]

<table>
<thead>
<tr>
<th></th>
<th>( \delta_{ss,60°} )</th>
<th>( \delta_{ds,1°} )</th>
<th>( \Delta \delta' )</th>
<th>( \delta^*_{ss,1°} )</th>
<th>( \delta_{ds,1°} )</th>
<th>( \Delta \delta'' )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A2(H2)</strong></td>
<td>7.97</td>
<td>7.55</td>
<td>0.42</td>
<td>7.78</td>
<td>7.55</td>
<td>0.23</td>
</tr>
<tr>
<td>A3</td>
<td>7.85</td>
<td>7.38</td>
<td>0.47</td>
<td>7.77</td>
<td>7.38</td>
<td>0.39</td>
</tr>
<tr>
<td>A4</td>
<td>7.60</td>
<td>7.12</td>
<td>0.64</td>
<td>7.50</td>
<td>7.12</td>
<td>0.37</td>
</tr>
<tr>
<td>A5</td>
<td>7.74</td>
<td>6.87</td>
<td>0.87</td>
<td>7.55</td>
<td>6.87</td>
<td>0.68</td>
</tr>
<tr>
<td>A6</td>
<td>7.83</td>
<td>7.20</td>
<td>0.63</td>
<td>7.61</td>
<td>7.2</td>
<td>0.41</td>
</tr>
<tr>
<td>A7</td>
<td>7.88</td>
<td>7.62</td>
<td>0.26</td>
<td>7.34</td>
<td>7.62</td>
<td>0.12</td>
</tr>
</tbody>
</table>

* single strand value at low temperature is calculated by extrapolating data at other temperatures linearly to 1°C.
subtracting the chemical shift in the double strand at 1°C from the chemical shift in the single strand at the same temperature which means the single-strand stacking contribution is subtracted. Both of these $\Delta \delta$ values show similar trends so I will just focus on the $\Delta \delta'$ values. For the purposes of the following discussion I will only be concerned with the internal bases that are the nearest and the next nearest neighbors to the extra cytosine; that is $A_4(H2)$, $A_5(H2)$, $A_3(H2)$ and $A_6(H2)$. For the perfect duplex, $dCA_6G + dCT_6G$, all the internal base pair $A(H2)'s$ have similar values for $\Delta \delta'$ while in $dCA_3CA_3G + dCT_6G$ there are much larger differences in the values of $\Delta \delta'$. If the cytosine were stacked in the helix one would expect the $\Delta \delta'$ value to be less for the $dCA_3CA_3G + dCT_6G$ duplex than for the $dCA_6G + dCT_6G$ duplex due to a decrease in shielding of the cytosine over the adenine. This qualitative conclusion agrees with calculations of the shielding values assuming the cytosine is stacked in the helix. These calculations are shown in Appendix C. However, the $\Delta \delta$ values given in Table 2.2 show a slight decrease in shielding of the $A_4(H2)$ and an increase in shielding of the $A_5(H2)$. This implies that the cytosine is not stacked in the helix. The calculations and to a certain extent the intuitive arguments are based on the helix having a regular B-form geometry. However, with the perturbation of the extra cytosine this is probably not true. Therefore, trying to predict the changes in chemical shifts is complicated by not knowing the actual geometry of this duplex.
5. THERMODYNAMIC RESULTS

The optical melting curves of dCA3CA3G + dCT6G taken at seven different concentrations are shown in Figure 2.13. The curves are normalized at 60°C to an absorbance of 1. The upper curve is the single-strand melting curve, determined from a weighted average of melting curves from the two individual strands. The curves show an increase in hypochromicity as the concentration increases; this has been attributed to aggregation of the double strands (Nelson et al., 1981). The lower baseline of the double strand melt is difficult to determine, because this duplex melts at approximately room temperature. A non-linear least squares program based on a two-state model was used to fit \( \Delta H^\circ \), \( \Delta S^\circ \), and the lower baseline slope to the normalized melting curve. The slope for the most concentrated sample was determined to be \( 1.04 \times 10^{-3} \) per °C. This slope was used to determine low temperature baselines for the other concentrations. The melting curves were analyzed to determine the melting temperature using the slope of the lower baseline previously discussed and the experimental single strand melting curve. Figure 2.14 shows the \( 1/T_m \) vs. \( \log C_0 \) plot used to determine \( \Delta H^\circ \) and \( \Delta S^\circ \), where \( C_0 \) is the total concentration of each of the oligomer strands (Nelson et al., 1981). These results are given in Table 2.3. The same procedure was used for dCA6G + dCT6G with the exception that the slope of the low temperature baseline for the second most concentrated sample was used to obtain low temperature baselines for the other curves. This slope was used because it gave baselines that fit the curves better than those determined from the slope of the most concentrated
Figure 2.13. Melting curves of dCA$_3$CA$_3$G + dCT$_6$G in 1M NaCl with concentrations ranging from 0.01mM-0.9mM in single strand. The curves are normalized to 1 at 60°C. The upper curve is the single strand melting curve as described in the text.
Figure 2.14. The $1/T_m$ vs. log $C_0$ plot for $\circlearrowleft$ dCA$_3$CA$_3$G + dCT$_6$G and $\blacktriangle$ dCA$_6$G + dCT$_6$G.
$\frac{1}{T_E} \times 10^3 \left(^\circ K^{-1}\right)$

$\log($concentration$)$
Table 2.3 Thermodynamic Parameters from 1/T vs. log C₀ Plot

<table>
<thead>
<tr>
<th></th>
<th>Tₘ (°C, 200μM)</th>
<th>ΔG° (kcal mol⁻¹)</th>
<th>ΔH° (kcal mol⁻¹)</th>
<th>S° (e.u.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dCA₂CA₂G + dCT₆G</td>
<td>22 ± 2</td>
<td>-4.9 ± 0.3</td>
<td>-53 ± 2</td>
<td>-161 ± 5</td>
</tr>
<tr>
<td>dCA₆G + dCT₆G</td>
<td>37 ± 3</td>
<td>-7.8 ± 0.6</td>
<td>-59 ± 3</td>
<td>-172 ± 8</td>
</tr>
<tr>
<td>dCA₅G + dCT₅Gᵃ</td>
<td>32</td>
<td>-6.6</td>
<td>-47 ± 2</td>
<td>-136 ± 5</td>
</tr>
</tbody>
</table>

a. Data from Nelson et al., 1981.
sample. The slope used was $1.55 \times 10^{-3}$ per °C. Values obtained using the low temperature baseline slope of the most concentrated sample were within 3% of the values reported here. The $1/T_m$ vs. log $C_o$ plot from these data is also shown in Figure 2.14. Thermodynamic parameters are reported in Table 2.3 for dCA$_6$G + dCT$_6$G and for dCA$_5$G + dCT$_5$G (Nelson et al., 1981).

6. DISCUSSION OF THE THERMODYNAMICS

The thermodynamic results can be discussed in view of the result that the cytosine is extrahelical. It should be emphasized that the thermodynamic analysis is based on a two-state model. Table 2.3 compares the thermodynamic parameters for all three helices. The melting temperature, at 200μM per single strand, decreases by 15°C when we introduce the perturbation of the extra cytosine. It is decreased 10°C over the duplex with one less A·T base pair and no perturbation. Another indication of the destabilization is that the $\Delta G^o(25°C)$ of double strand formation is not as energetically favorable for the duplex with the unpaired cytosine: a 2.9 kcal mol$^{-1}$ destabilization from the parent duplex and a 1.7 kcal mol$^{-1}$ destabilization compared to the duplex with one less A·T base pair. In RNA there is a favorable enthalpy and unfavorable entropy change on going from a single-strand to a double-strand state (Borer et al., 1974). This trend is observed when comparing dCA$_6$G + dCT$_6$G to dCA$_5$G + dCT$_5$G. In this case we have one more A·T base pair, which decreases the enthalpy by 12 ± 5 kcal mol$^{-1}$ and decreases entropy by 36 ± 13 e.u. When comparing dCA$_3$CA$_3$G + dCT$_6$G with dCA$_6$G + dCT$_6$G there is an unfavorable change in enthalpy of
6 ± 5 kcal mol⁻¹. This change depends on differences in the single strands \((dCA₂,CA₂,G \text{ vs. } dCA₂,G)\) and in the two double strand helices. The increase in enthalpy is consistent with a loss of stacking of the cytosine in the double strand relative to the single strand, and the deformation of the double helix around the unpaired cytosine. There is a small, probably favorable, change in the entropy, 11 ± 13 e.u., presumably explained by the freedom of the extra cytosine which is not stacked in the helix.

Several other groups have studied the effects of an extra base on the conformation and thermodynamics (Lomant & Fresco, 1974; Evans & Morgan, 1982; Gray et al., 1980; Fink & Crothers, 1972; Patel et al., 1982). Several of these studies have shown that these bases are extrahelical. Work done by Fink and Crothers on the polyribo-nucleotide \(\text{poly}(A,A*) + \text{poly}(U)\), where \(A*\) is the N-1 oxide of adenine, shows a destabilization in the standard free energy at 25°C of 2.8 kcal mol⁻¹ due to this perturbation. This is in good agreement with the value of 2.9 kcal mol⁻¹ obtained in this paper for a deoxyribo-oligonucleotide. Work done by Patel et al. (1982) on the nearly self-complementary tridecamer \(dC-G-C-A-G-A-T-T-C-G-C-G\) indicates that the extra adenine is stacked in the helix. There is no difference in enthalpy, within experimental error, for forming the duplex with the extra adenine compared to the completely self-complementary dodecamer, \(dC-G-C-G-A-A-T-T-C-G-C-G\). The extra base and the sequence of the duplex around that base apparently determine whether the base remains in the helix or is bulged out. As adenine is the base with the strongest stacking tendency, it is reasonable.
that an extra adenine remains in the helix, but that a cytosine surrounded by adenines is forced out of the helix. Our understanding of the sequence dependence of conformation will become clearer as different sequences containing perturbations are studied.

7. CONCLUSION

I have studied the conformation of an extra cytosine in the deoxyribonucleotide duplex dCA₃CA₃G + dCT₆G. Using the nuclear Overhauser effect, which depends on the inverse sixth power of the interproton distances, we find that the A·T base pairs on either side of the extra cytosine are stacked on each other. Thus the cytosine must be extrahelical; that is, it is not stacked with neighboring base pairs in the helix. These experiments also demonstrate that the extrahelical cytosine causes local perturbations in the conformation of these adjacent A·T base pairs. The temperature dependence of the chemical shifts of the cytosine H₆ doublets in this duplex also indicates that the unpaired cytosine is extrahelical.

The thermodynamic parameters were determined for this duplex and compared to parameters for dCA₆G + dCT₆G. The extrahelical cytosine causes a 2.9 kcal mol⁻¹ decrease in the standard free energy at 25°C and a decrease in the enthalpy of double strand formation. Comparing the results we obtain for this sequence with results obtained on other sequences (Patel et al., 1982) indicates that stacking interactions may determine whether a base will be stacked or not stacked in the helix.
The effects of an extra base on conformation and energetics of the duplex has implication on the mechanisms of mutagenesis. Models have been proposed for stacking interactions between frameshift mutagens and an extrahelical base (Streisinger et al, 1966; Drake & Baltz, 1976). Several studies have demonstrated the possibility of this type of interaction (Lee & Tinoco, 1978; Helfgott & Kallenbach, 1979). The work presented here has shown how dCA₃CA₃G + dCT₆G in solution forms a stable duplex with the unpaired cytosine not stacked on the neighboring base pairs in the helix and thus this duplex would be ideal for studying the interaction between frameshift mutagens and extrahelical bases.
REFERENCES


Wagner, G., Pardi, A., & Wüthrich, K., to be published.
CHAPTER III
Kinetics of Exchange in Oligonucleotides

1. INTRODUCTION

Relaxation rates of the base-paired imino protons have been measured by nuclear magnetic resonance experiments in several nucleic acid systems (Crothers et al., 1974; Johnston and Redfield, 1977; Hurd & Reid, 1980; Early et al., 1981a,b; Pardi & Tinoco, 1982). The saturation recovery technique for aqueous solutions developed by Redfield (Johnston & Redfield, 1977) was used in the present studies. The basic theory for interpretation of the exchange behavior of imino protons measured by NMR will be presented here; it has also been discussed by Johnston & Redfield (1981) and Pardi & Tinoco (1982).

In the previous chapter, I discussed the effects of an extra cytosine on the conformation of the duplex dCA₃CA₃G + dCT₆G. In this chapter, I will discuss the kinetic effects of this perturbation as well as several other perturbations. The helices studied in this chapter are: I) dCA₃CA₃G + dCT₆G, II) dCA₅G + dCT₆G, III) d(CGCGAATTCGGC-G), referred to as the 12-mer, IV) d(CGTGATAATTGGC-G), referred to as the 12-mer C-G-T, and V) d(CGCAAGAATTGGC-G), referred to as the 13-mer. In Chapter II, I demonstrated that the extra cytosine in duplex I was not stacked in the helix and that this perturbation caused a destabilization of this duplex compared to its totally base paired counterpart, duplex II (see Table 2.3). The last three duplexes are used to study
several other perturbations. Duplex IV contains a G•T wobble base pair (Patel et al., 1982b). Duplex V contains an extra adenine which was shown using NMR to be stacked in the helix (Patel et al., 1982c). The "perfect" helix counterpart to duplexes IV and V is duplex III. The conformation and thermodynamics of duplex III have been previously studied (Patel et al., 1982a). The perturbations in duplexes IV and V cause destabilization of the double strand relative to the single strand. This is reflected in their melting temperatures which were determined by calorimetry (Patel et al., 1982a, b, c). Duplex III, IV, and V have melting temperatures of 71°C, 51°C, and 52°C, respectively (sample concentration of ~0.7mM per strand in 10mM phosphate, 0.1M NaCl). The conformation studies and preliminary kinetic studies have been presented previously (Pardi et al., 1982; Patel et al., 1982d).

In this chapter, I will discuss the results of saturation recovery experiments which measure the relaxation lifetimes of imino protons in duplexes in solution. The pH and temperature dependence of these lifetimes will also be discussed. These results can be interpreted in terms of mechanisms of chemical exchange in oligonucleotide duplexes. The effects of a variety of perturbations on exchange along with the effect of helix length on the mechanism for exchange will be discussed.

2. THEORY OF IMINO PROTON EXCHANGE

The theory of measuring chemical exchange using saturation recovery has been presented previously (Johnston & Redfield, 1981). Therefore, I will present only what is necessary for the
reader to understand the results and discussion which follow.

The following scheme has been proposed (Crothers et al., 1974; Teitelbaum & Englander, 1975; Kallenbach et al., 1976) for the exchange of base pairing protons in nucleic acids:

\[
\text{closed} \xrightarrow{k_{\text{cl}}} \text{open} \xrightarrow{k_{\text{ex[cat]}}} \text{exchanged}
\]

where closed represents the base paired duplex state and open is some state capable of exchange. This scheme defines the closed state of the nucleic acid as the state from which there is no exchange. The reaction for opening is characterized by a rate constant \(k_{\text{op}}\) and the closing by a rate constant \(k_{\text{cl}}\). The exchange step is base catalyzed and is characterized by a rate constant \(k_{\text{ex[cat]}}\). An example of this mechanism as it would be applied to an NMR saturation recovery experiment is shown in Figure 3.1. Saturation of the resonance is used to label the proton, \(H^*\), which will then be exchanged with unlabeled protons, in this case \(H_2O\) protons. In the NMR experiment, we start by saturating a resonance, species A in Figure 3.1 which is not observable in the NMR experiment, and then monitor the rate at which it becomes unlabeled, species D in Figure 3.1 which is observable. The relaxation, or the rate at which the labeled proton becomes visible, is a function of time as follows:

\[
I(t) = I_{eq} \left(1 - e^{-t/\tau_{\text{obs}}}\right)
\]

where \(I(t)\) is the intensity at time \(t\), \(I_{eq}\) is the intensity at
Figure 3.1. A schematic representation of a mechanism for chemical exchange of an imino proton in an oligonucleotide duplex. Species A is not observable in the NMR experiment whereas species D is its observable counterpart.
\[ A \xrightarrow{k_{op}} B \xleftarrow{k_{cl}} A \]
\[ C \xrightarrow{k_{ex}[cat]} B \xleftarrow{k_{cl}} C \]

Diagram:

equilibrium and $\tau_{\text{obs}}$ is the observed lifetime for the proton. How $\tau_{\text{obs}}$ relates to the kinetic processes of opening, closing, and exchange is shown in Figure 3.2. The result of this kinetic analysis is that

$$\bar{k}_{\text{ex}} = \frac{k_{\text{op}} k_{\text{ex [cat]}}}{k_{\text{cl}} + k_{\text{ex [cat]}}}$$

(2)

where $k_{\text{ex}}$ is the overall exchange rate for the process shown in Figure 3.1. This expression can be simplified by considering several limiting cases (Crothers et al., 1974; Hilbers, 1979; Kallenbach et al., 1976; Teitelbaum & Englander, 1975). First we assume $k_{\text{cl}} \ll k_{\text{ex [cat]}}$. When the closing rate is slow compared to the rate of exchange, the helix will exchange every time it is in the open state. This assumption defines the open-limited region and leads to the following simplification of equation (2):

$$\bar{k}_{\text{ex}} = k_{\text{op}}$$

(3)

If however, $k_{\text{cl}} \gg k_{\text{ex [cat]}}$ then equation (2) simplifies to

$$\bar{k}_{\text{ex}} = \frac{k_{\text{op}}}{k_{\text{cl}}} k_{\text{ex [cat]}}$$

(4)

This limit is called the pre-equilibrium limit. In this case the base pair can open and close many times before any exchange occurs. The $\tau_{\text{obs}}$ measured in the NMR experiment will be $1/k_{\text{ex}}$ if chemical exchange is the only source of relaxation. So in the
Figure 3.2. The kinetic analysis for the mechanism shown in the previous figure. The analysis assumes the step from \( B \rightarrow C \) has no back reaction because the \( H^+ \) concentration will be very small in solution. The steady state approximation is made for species \( B \) and \( C \). \( K_{ex} \) is the overall rate constant for exchange by this mechanism.
\[
\begin{align*}
\frac{d[B]}{dt} &= 0 = k_{op}[A] - k_{c1}[B] - k_{ex[cat]}[B] \\
\frac{d[C]}{dt} &= 0 = k_{ex[cat]}[B] + k_{op}[D] - k_{c1}[C]
\end{align*}
\] (1) (2)

Observe \( \frac{d[D]}{dt} \).

\[
\frac{d[D]}{dt} = -k_{op}[D] + k_{c1}[C]
\]

\[
= k_{ex[cat]}[B] \quad \text{from eqn. (2)}
\]

\[
= \frac{k_{ex[cat]} k_{op}[A]}{k_{c1} + k_{ex[cat]}} \quad \text{from eqn. (1)}
\]

\[
K_{ex} = \frac{k_{ex[cat]} k_{op}}{k_{c1} + k_{ex[cat]}}
\]
open-limited case, equation (3), the $\tau_{obs}$ will be directly measuring $1/k_{op}$ whereas, in the pre-equilibrium case, equation (4), the $\tau_{obs}$ is measuring a more complex combination of the rates. Whether $\tau_{obs}$ for a proton is in the open-limited region can be determined by varying the concentration of the catalyst. This is done experimentally by changing the pH and thus the concentration of OH$^-$ or HPO$_4^{2-}$, from a phosphate buffer.

The next problem is determining the nature of the open state. By measuring the temperature dependence of the lifetimes for exchange one can determine an activation energy for the exchange process. There are two processes which I will consider. One process involves the "opening" of one or perhaps a few base pairs (process I). Opening is used loosely to indicate going to a state which is able to exchange its base-pairing imino proton. The other process involves the double helix going to single strands and then exchanging (process II). These two possibilities have been discussed by Pardi & Tinoco (1982). Because process II involves the breaking of many hydrogen bonds and possibly loss of base stacking, one might expect this process to have a higher activation energy. Indeed the activation energy determined for dCA$_5$G + dCT$_5$G, which was attributed to whole helix opening, process II, was $\sim$47 kcal mol$^{-1}$ (Pardi & Tinoco, 1982). The rate of whole helix opening can also be determined by temperature-jump kinetics. This was done for dCA$_5$G + dCT$_5$G and the activation energy was found to be 43 kcal mol$^{-1}$ (Nelson & Tinoco, 1982). This value for the activation energy is in good agreement with the value determined using NMR. Of course the
activation energy for whole helix opening will be dependent on the length and the sequence of the oligonucleotide. The activation energy for whole helix opening of dCA5G + dCT5G determined by NMR can be compared to activation energies of 14-16 kcal mol⁻¹ which have been attributed to local base-pair opening, process I, (Early et al., 1981a,b). It is not clear if the activation energies for local base-pair opening are due to a single base pair exchanging or to several base pairs exchanging. In any case, if both processes are available the observed activation energy will have a dependence on the activation energies for both processes, as given below:

\[
E_{a}^{\text{obs}} = \frac{1}{k_{\text{op}}^{I} + k_{\text{op}}^{II}} \left( \frac{k_{\text{op}}^{I} E_{a}^{I} + k_{\text{op}}^{II} E_{a}^{II}}{k_{\text{op}}^{I} + k_{\text{op}}^{II}} \right)
\]  

(5)

where \(k_{\text{op}}^{I}\) and \(k_{\text{op}}^{II}\) are rate constants and \(E_{a}^{I}\) and \(E_{a}^{II}\) are activation energies for (I) local base-pair opening and (II) whole helix opening processes, respectively. So by determining the activation energy we can begin to understand the mechanisms for exchange of the base-pairing imino protons.

The preceding discussion assumes that chemical exchange is the only, or at least the dominant, pathway for relaxation of the imino protons. However, all protons are capable of relaxing, without chemical exchange, through a mechanism known as spin-lattice relaxation. This process is characterized by a relaxation lifetime \(T_1\). As will be discussed later, the extent to which magnetic relaxation, \(T_1\), contributes to the \(\tau_{\text{obs}}\) must be considered before these lifetimes can be interpreted in terms of chemical exchange.
3. EFFECTS OF AN EXTRA CYTOSINE ON THE KINETICS OF EXCHANGE

A. Materials and Methods

The syntheses of dCA₃CA₃G, dCA₆G, and dCT₆G are described in Chapter II. The sample preparation has also been discussed in Chapter II. Samples were adjusted from pH 7 to pH 8 by the addition of NaOH; this resulted in a change in sample concentration of not more than 5%. NMR experiments were done using the saturation recovery technique. This involves application of a 200-300 msec rf pulse at the resonance frequency for the desired proton, followed by a variable delay, and then the observation pulse. The Redfield 214 pulse sequence is used for observation, as discussed in Chapter II. The initial saturating pulse was applied either at a specific frequency or over a wider frequency band through the use of a noise modulator. There was no significant difference in the results obtained by these two methods. Typically 10-15 delay times were used to determine each lifetime value. The spectrum at each delay time was the sum of 500-600 scans. The lifetimes were determined by fitting the intensities at various delays to equation (1). In the fitting routine each data point is weighted by I(t)/Iₑq, where Iₑq is determined experimentally by taking a spectrum with the saturating power off. Several of the proton resonances overlap and for these resonances areas were used rather than using intensities. These areas were analyzed using the multi-exponential fit procedures described by Drobnies (1979) and Nelson (1982). None of the overlapping resonances show any significant multi-exponential behavior, thus the values for the single exponential fit are reported in all
these cases. The lifetimes are estimated to be accurate to within ±30%.

B. Results

The imino resonances for dCA₃CA₃G + dCT₆G are shown at several temperatures in Figure 3.3. The resonances broaden as the temperature is increased due to an increase in the exchange rate with H₂O. At 15°C the resonances have broadened so significantly that they are difficult to distinguish from the noise. A similar trend is observed for dCA₆G + dCT₆G, as shown in Figure 3.4. The broadening occurs at higher temperatures because this duplex is more stable than the dCA₃CA₃G + dCT₆G duplex. Examples of spectra from a saturation recovery experiment on dCA₃CA₃G + dCT₆G, at pH 7 and 5°C, are shown in Figure 3.5. In this example, a saturating rf pulse is applied at the resonance frequencies for base pairs 4 and 8. An example of spectra from a saturation recovery experiment on dCA₆G + dCT₆G, at pH 7 and 5°C, is shown in Figure 3.6. Here the saturating pulse is applied to the frequencies for base pairs 2, 4, and 7 (note: base pairs 2 and 4 resonate at the same frequency). From this type of data, peak intensities or areas are measured and used to determine the relaxation lifetimes. The lifetimes for exchange of the imino protons in dCA₃CA₃G + dCT₆G at pH 7 are summarized in Table 3.1 and for dCA₆G + dCT₆G at pH 7 in Table 3.2. To determine if these rates are in the open-limited region, the experiments have also been done at pH 8. The values for the relaxation lifetimes at pH 8 are given, and compared to the values at pH 7, in Table 3.3 for dCA₃CA₃G + dCT₆G and in Table 3.4 for dCA₆G + dCT₆G. The values for
Figure 3.3. The temperature dependence of the 360 MHz $^1$H NMR spectra of the base-pairing imino protons from dCA$_3$CA$_3$G + dCT$_6$G. Spectra are shown for 1°C, 5°C, 10°C, and 15°C. Each spectrum is the sum of 600-1000 accumulations. The sample was 1mM per single strand in H$_2$O.
1 2 3 4 5 6 7 8
G-T-T-T-T-T-T-T-Cd

15°C

10°C

5°C

1°C
Figure 3.4. The temperature dependence of the 360 MHz $^1$H NMR spectra of the base-pairing imino protons from dCA$_6$G + dCT$_6$G. Spectra are shown for 5°C, 8°C, 15°C, 20°C, and 25°C. Each spectrum is the sum of 500-600 accumulations. The sample was 1mM per single strand in H$_2$O.
G-T-T-T-T-T-T-Cd
Figure 3.5. The 360 MHz \(^1\)H NMR spectra from a saturation recovery experiment on dCA\(_3\)CA\(_3\)G + dCT\(_6\)G at 5°C (only partial data shown). The arrows indicate the frequencies which were saturated, corresponding to base pairs 4 and 8. The time in milliseconds refers to the delay between the saturation pulse and the observation pulse. The top spectrum is done under the same conditions as the other spectra however there is no power on for the saturation pulse. Each spectrum is the sum of 500 accumulations.
1 2 3 4 5 6 7 8
\text{G-T-T-T-T-T-T-Cd}

\text{infinity}

30 msec

15 msec

0.2 msec

\text{15.0 14.5 14.0 13.5 13.0 12.5 PPM}
Figure 3.6. The 360 MHz $^1$H NMR spectra from a saturation recovery experiment on dCA$_6$G + dCT$_6$G at 5°C (only partial data shown). The arrows indicate the frequencies which were saturated, corresponding to base pairs 7 and the double resonance for base pairs 2 and 4. The time in milliseconds refers to the delay between the saturation pulse and the observation pulse. The top spectrum is done under the same conditions as the other spectra however there is no power on for the saturation pulse. Each spectrum is the sum of 500 accumulations.
G-T-T-T-T-T-T-Cd

infinity

200 msec

40 msec

0.2 msec

15.0 14.5 14.0 13.5 13.0 12.5 PPM

1 2 3 4 5 6 7 8
Table 3.1. Lifetimes for Exchange of Imino Protons at pH 7 in

\[
\begin{array}{cccccccc}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
G-T-T-T----T-T-T-dC \\
\end{array}
\]

<table>
<thead>
<tr>
<th>Imino Proton (Base Pair #)</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
<th>#7</th>
<th>#8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp (°C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>121</td>
<td>162</td>
<td>77</td>
<td>121</td>
<td>163</td>
<td>150</td>
<td>70</td>
</tr>
<tr>
<td>5</td>
<td>89</td>
<td>123</td>
<td>41</td>
<td>89</td>
<td>99</td>
<td>70</td>
<td>34</td>
</tr>
<tr>
<td>7</td>
<td>46</td>
<td>61</td>
<td>25</td>
<td>46</td>
<td>69</td>
<td>53</td>
<td>25</td>
</tr>
<tr>
<td>9</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>29</td>
<td>14</td>
<td>24</td>
</tr>
</tbody>
</table>


### Table 3.2. Lifetimes for Exchange of Imino Protons at pH 7 in

$$12345678$$


G-T-T-T-T-T-T-Cd

<table>
<thead>
<tr>
<th>Imino Proton (Base Pair #)</th>
<th>#1</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
<th>#7</th>
<th>#8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp ($^\circ$C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>19</td>
<td>228</td>
<td>187</td>
<td>228</td>
<td>223</td>
<td>184</td>
<td>144</td>
<td>57</td>
</tr>
<tr>
<td>5</td>
<td>22</td>
<td>224</td>
<td>290</td>
<td>224</td>
<td>246</td>
<td>247</td>
<td>192</td>
<td>44</td>
</tr>
<tr>
<td>10</td>
<td>--</td>
<td>149</td>
<td>253</td>
<td>276</td>
<td>245</td>
<td>318</td>
<td>86</td>
<td>24</td>
</tr>
<tr>
<td>15</td>
<td>--</td>
<td>58</td>
<td>196</td>
<td>191</td>
<td>236</td>
<td>218</td>
<td>59</td>
<td>--</td>
</tr>
<tr>
<td>20</td>
<td>--</td>
<td>22</td>
<td>72</td>
<td>73</td>
<td>118</td>
<td>121</td>
<td>30</td>
<td>--</td>
</tr>
</tbody>
</table>
Table 3.3. Comparison of Lifetimes at pH 8 and pH 7 in

\[
\begin{array}{cccccccc}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
\end{array}
\]

<table>
<thead>
<tr>
<th>Imino Proton (Base Pair #)</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
<th>#7</th>
<th>#8</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Temp (°C)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5°C, pH 8</td>
<td>45</td>
<td>66</td>
<td>27</td>
<td>45</td>
<td>66</td>
<td>44</td>
<td>27</td>
</tr>
<tr>
<td>5°C, pH 7</td>
<td>89</td>
<td>123</td>
<td>41</td>
<td>89</td>
<td>99</td>
<td>70</td>
<td>34</td>
</tr>
<tr>
<td>9°C, pH 8</td>
<td>20</td>
<td>20</td>
<td>16</td>
<td>20</td>
<td>21</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>9°C, pH 7</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>29</td>
<td>14</td>
<td>24</td>
</tr>
</tbody>
</table>
Table 3.4. Comparison of Exchange Lifetimes at pH 8 and pH 7 in

\[ G-T-T-T-T-T-T-Cd \]

<table>
<thead>
<tr>
<th>Temp (°C)</th>
<th>Imino Proton (Base Pair #)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#2</td>
</tr>
<tr>
<td>5°C, pH 8</td>
<td>117</td>
</tr>
<tr>
<td>5°C, pH 7</td>
<td>224</td>
</tr>
<tr>
<td>15°C, pH 8</td>
<td>27</td>
</tr>
<tr>
<td>15°C, pH 7</td>
<td>58</td>
</tr>
</tbody>
</table>
the lifetimes can be plotted as an Arrhenius plot, $-\ln \tau_{\text{obs}}$ vs. $1/T$, to determine activation energies for the relaxation processes. These plots are shown in Figures 3.7, 3.8, and 3.9 for dCA$_3$CA$_3$G + dCT$_6$G and in Figures 3.10, 3.11, and 3.12 for dCA$_6$G + dCT$_6$G. The activation energies determined from these plots are summarized in Table 3.5 for both duplexes.

C. Discussion

By measuring the relaxation lifetimes as a function of temperature, I can probe the dynamics of oligonucleotides in solution. NMR is a very good technique for probing dynamics because it gives information from each base pair. However, there are limitations and experimental difficulties with this technique. One of these difficulties is that the technique involves monitoring protons which are exchangeable. This means the experiments must be done in H$_2$O if these protons are to be observed. The potential dynamic range problem (1mM sample as compared to 110M protons from H$_2$O) has been overcome by using a special pulse sequence to observe these protons (Redfield et al., 1975). The next problem occurs when the duplex approaches its melting temperature. As the equilibrium shifts towards the single strand the exchange rate for these protons is faster and eventually is so fast that the imino resonance broadens and becomes indistinguishable from the noise. This usually occurs ~10-15°C below the melting temperature of the duplex and is called kinetic melting. This can be seen in Figures 3.3 and 3.4. If we wish to study these duplexes in solution, then we are also limited at the low temperature extreme by freezing. For a destabilized
Figure 3.7. Arrhenius plots for the observed lifetimes of imino protons from dCA$_2$CA$_3$G + dCT$_6$G. a) A·T base pairs 2 and 5 (double resonance), b) A·T base pair 3.
G-T-T-T-T-T-T-Cd

**a**

Base pair 2 & 5
\[ E_a = 51 \text{ kcal/mol} \]

\[
\begin{array}{c}
\text{-ln } T_{\text{obs}} \\
3.5 & 3.6 & 3.7
\end{array}
\]

\[
\begin{array}{c}
1/T \times 10^3 (^\circ \text{K}^{-1})
\end{array}
\]

**b**

Base pair 3
\[ E_a = 64 \text{ kcal/mol} \]

\[
\begin{array}{c}
\text{-ln } T_{\text{obs}} \\
3.5 & 3.6 & 3.7
\end{array}
\]

\[
\begin{array}{c}
1/T \times 10^3 (^\circ \text{K}^{-1})
\end{array}
\]
Figure 3.8. Arrhenius plots for the observed lifetimes of imino protons from $dCA_3CA_3G + dCT_6G$. a) A·T base pair 6, b) A·T base pair 7.
$4.0$
$3.0$
$2.0$
$1.0$
$0.0$
$-\ln T_{\text{obs}}$

$E_a = 48 \text{ kcal/mol}$

$1/ T \times 10^3 (\circ \text{K}^{-1})$

$12345678$

G-T-T-T-T-T-T-Cd

$3.6$
$3.7$

$4.0$
$3.0$
$2.0$
$1.0$
$0.0$
$-\ln T_{\text{obs}}$

$E_a = 63 \text{ kcal/mol}$

$1/ T \times 10^3 (\circ \text{K}^{-1})$
Figure 3.9. Arrhenius plots for the observed lifetimes of imino protons from dCA$_3$CA$_3$G + dCT$_6$G. a) A·T base pair 4, b) G·C base pair 8.
![Graph](image)

**Base Pair 4**

- **Formula:** G-T-T-T-T-T-T-Cd

**Graph**

- **Equation:** $E_a = 21$ kcal/mol

---

**Base Pair 8**

- **Formula:**
- **Formula:**

**Graph**

- **Equation:** $E_a = 22$ kcal/mol
Figure 3.10. Arrhenius plots for the observed lifetimes of imino protons from dCA₆G + dCT₆G. a) A·T base pair 2, b) A·T base pair 3.
**Base Pair 2**

\[ \text{E}_a = 30 \text{ kcal/mol} \]

**Base Pair 3**

\[ \text{E}_a = 21 \text{ kcal/mol} \]
Figure 3.11. Arrhenius plots for the observed lifetimes of imino protons from dCA₆G + dCT₆G. a) A•T base pair 4, b) A•T base pair 5.
base pair 4

\( E_a = 22 \text{ kcal/mol} \)

base pair 5

\( E_a = 23 \text{ kcal/mol} \)
Figure 3.12. Arrhenius plots for the observed lifetimes of imino protons from dCA₆G + dCT₆G. a) A•T base pair 6, b) A•T base pair 7.

(a) base pair 6
\[ E_a = 16 \text{ kcal/mol} \]

(b) base pair 7
\[ E_a = 17 \text{ kcal/mol} \]
Table 3.5. Activation Energies (kcal mol\(^{-1}\)) for Exchange of Imino Protons

<table>
<thead>
<tr>
<th>Molecule</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
<th>#7</th>
<th>#8</th>
</tr>
</thead>
<tbody>
<tr>
<td>dCA(_3)CA(_3)G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ dCT(_6)G</td>
<td>51 ± 10</td>
<td>64 ± 13</td>
<td>21 ± 10</td>
<td>51 ± 10</td>
<td>48 ± 12</td>
<td>63 ± 23</td>
<td>22 ± 10</td>
</tr>
<tr>
<td>dCA(_6)G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+ dCT(_6)G</td>
<td>31 ± 6</td>
<td>21 ± 8</td>
<td>22 ± 10</td>
<td>23 ± 11(^a)</td>
<td>16 ± 4</td>
<td>17 ± 4</td>
<td>12 ± 8</td>
</tr>
</tbody>
</table>

\(^a\) This value was determined from the slope between two points.
duplex like $\text{dCA}_3\text{CA}_3\text{G} + \text{dCT}_6\text{G}$ these limits severely restrict the
temperature range in which experiments can be done. Notice in Table
3.1 lifetimes could only be obtained from $1^\circ$ to $9^\circ\text{C}$, whereas for
$d\text{CA}_6\text{G} + d\text{CT}_6\text{G}$ (Table 3.2) data was obtained from $-1^\circ$ to $20^\circ\text{C}$.

A more general problem with using the technique of saturation
recovery to measure chemical exchange is that we are actually
measuring relaxation. This relaxation can occur by several
processes, not solely chemical exchange, and it is hard to distin­
guish these in the NMR experiment. Magnetic spin-lattice relaxation
is one of these processes. A possible way to distinguish this
process from chemical exchange will be discussed later, however, the
data presented here have not been corrected for any contribution due
to spin-lattice relaxation.

A problem that is specific to the two duplexes discussed here,
but potentially a problem in other systems, is how to interpret
relaxation for distinct imino protons with overlapping resonance
frequencies. In the case of $d\text{CA}_6\text{G} + d\text{CT}_6\text{G}$, imino protons from base
pairs 2 and 4 have overlapping resonances. For $d\text{CA}_3\text{CA}_3\text{G} + d\text{CT}_6\text{G}$,
the resonances for imino protons from base pairs 2 and 5 overlap, as
do those from base pairs 4 and 8. The data from these resonances
was analyzed using a multi-exponential function. However, the
single exponentials were always the best fit. Perhaps with better
signal-to-noise, multi-exponential solutions could be obtained. The
overlap of resonances causes problems with the imino proton from
base pair 4, as its lifetime is probably greater than that for base
pair 8. This is suggested at $1^\circ\text{C}$ and $5^\circ\text{C}$ where the resonances are
separated sufficiently to allow determination of individual lifetimes (see Figure 3.3). However, above 5°C it is difficult to separate the lifetimes and therefore the determined lifetime has been assigned to both imino protons.

Keeping all of these problems in mind, I can now discuss the results. In Table 3.2 the general trend observed for dCA₆G + dCT₆G is an increase in the lifetimes as you go from the ends of the helix into the middle. At -1°C and 5°C the lifetimes for imino protons from the internal base pairs 3-6 are experimentally the same. However, for dCA₃CA₃G + dCT₆G (shown in Table 3.1) this trend is not strictly followed. At 1°C the lifetimes for the imino protons from base pairs 3, 5 and 6 are experimentally the same, but the lifetime for base pair 4 is 40% less than the lifetime for base pair 5 and more than a factor of two less than the lifetimes for base pairs 3 and 6. This is also observed for the lifetimes at 5°C and 7°C. So this confirms my belief from Chapter II that the extra cytosine perturbs base pair 4 and causes faster exchange with H₂O. At 9°C all the internal imino protons (base pairs 3-6) have the same lifetimes.

As discussed in the introduction of this chapter, we can test whether the observed lifetimes correspond to opening by measuring the lifetimes at different concentrations of catalyst. This was done by measuring the lifetimes at pH 8. The results of these experiments are shown in Tables 3.3 and 3.4. Lifetimes at pH 7 are also shown in these tables for comparison. Values at pH 7 that are within 30% of the values at pH 8 are considered to be the same and
therefore open-limited. Values for lifetimes that are within 40% have been indicated as borderline open-limited. At 5°C the imino protons from base pairs 4, 6, and 7 are borderline; however, at 9°C all of the A·T base pairs are open limited, with the exception of base pair 4 which is again borderline. The lifetimes determined at 5°C, in dCA₆G + dCT₆G (Table 3.4) lifetimes for base pairs 3, 5, and 6 show open-limited behavior and at 15°C base pairs 3, 4, and 5 are open-limited. So I will consider the internal base pairs in this duplex (base pairs 3-6) to be open limited above 5°C. The observed lifetimes for protons in base pairs that are not open-limited are upper limits to the lifetimes for opening.

Information about the mechanism for exchange is obtained by measuring the lifetimes as a function of temperature. These values are then plotted on an Arrhenius plot in order to determine activation energies. The Arrhenius plots for dCA₃CA₃G + dCT₆G are shown in Figures 3.7-3.9 and for dCA₆G + dCT₆G in Figures 3.10-3.12. The activation energies determined from these plots are summarized in Table 3.5. These values were determined using only the higher temperature lifetimes: 5°C-9°C for dCA₃CA₃G + dCT₆G and 10°C-20°C for dCA₆G + dCT₆G. The high temperature lifetimes were used in order to minimize the effect of spin lattice relaxation, which becomes important at lower temperatures. When determining activation energies, the problem of working in a limited temperature range is very
evident. Activation energies for base pairs that were not open limited are also included in Table 3.5 for comparison. The activation energies for dCA₆G + dCT₆G are a factor of 2-3 smaller than the values for dCA₃CA₃G + dCT₆G, excluding base pair 4. The values of 12-20 kcal mol⁻¹ obtained for dCA₆G + dCT₆G are similar to the values obtained by Early et al. (1981a) for a dodecamer and to the values obtained for the dodecamer dC-G-C-G-A-A-T-T-C-G-C-G and its related duplexes, as will be discussed later in this chapter. These values have been attributed to local base-pair opening.

It is reasonable that as one looks at longer oligonucleotide duplexes the local base-pair opening mechanism will be favored over the higher energy whole helix opening process. Studies done on dCA₅G + dCT₅G (Pardi & Tinoco, 1982) show large activation energies for exchange of all the internal A•T base pairs. This was attributed to whole helix opening in this duplex. So by adding one more A•T base pair on going from dCA₅G + dCT₅G to dCA₆G + dCT₆G we may have reached a turning point from whole helix opening to local base-pair opening. With the exception of base pair 4, the activation energies in dCA₃CA₃G + dCT₆G are large, averaging 55 kcal mol⁻¹, which indicates this process could be whole helix opening. The value of 55 kcal mol⁻¹ is in good agreement with the ΔH for helix opening (53 kcal mol⁻¹) which was given in Table 2.3 of Chapter II. There is a possibility that both whole helix opening and local base-pair opening are contributing to this activation energy as given by equation (5). Again base pair 4 in dCA₃CA₃G + dCT₆G shows anomalous behavior. As shown in Table 3.5, the
activation energy for base pair 4 is more than a factor of two less than the values for base pairs 6 and 7. Evidence for base pair 4 having a smaller activation energy is illustrated by the melting behavior of the imino proton resonances (Figure 3.3). At low temperature the rate for exchange of the imino proton from base pair 4 is faster than the rate for protons from base pairs 2 and 5. If the activation energies were the same for all the base pairs, one would expect that at high temperature the exchange rate for base pair 4 would still be significantly higher than the rates for base pairs 2 and 5. Contrary to this, the resonances from base pairs 2 and 5 and base pair 4 have similar linewidths at high temperature (Figure 3.3) which indicates that their rates are similar. As discussed in Chapter II, there is evidence that the conformation around base pair 4 is different than expected for a regular helix. Perhaps there is some type of pocket that allows base pair 4 to be accessible to H$_2$O and to exchange with a minimal perturbation in structure. Or another possibility, which was suggested in Chapter II, is that the hydrogen bonding in base pair 4 is weaker. So, perhaps it is easier to open this base pair without having to open the whole helix.

The last question I will address is how much of a contribution there is from magnetic spin-lattice relaxation, $T_1$, in the observed lifetimes. The Arrhenius plots give some idea as to the answer to this question. In Figures 3.7a and b, one can see that at low temperature the lifetimes begin to change less drastically. This has been observed in other systems as well (Early et al., 1981; Pardi & Tinoco, 1982) and has been attributed to $T_1$. In the duplexes
studied here, there are not many data points where this effect is present, so it is unreasonable to attempt to quantify the contribution due to $T_1$. However, by determining activation energies using the highest temperatures, I have attempted to minimize this effect. It is possible that the effect of $T_1$ could be determined experimentally by doing a saturation recovery experiment on water in the presence of the oligonucleotide duplex. In this experiment the water resonance would be saturated, thus labeling water, and the labeled water would then be transferred through exchange to the imino resonances. This means exchange with water would cause the imino resonances to become non-observable in the NMR experiment. The only process that would allow the resonances to relax or become observable is the magnetic spin-lattice relaxation process. So if a signal was not observed after waiting a length of time equal to several times the observed lifetime determined by saturation recovery for the imino proton, it could be assumed that the contribution of magnetic spin-lattice relaxation was negligible.

4. EFFECTS OF A G·T MISMATCH AND AN EXTRA ADENINE ON THE EXCHANGE
A. Materials and Methods

The oligonucleotides were prepared by a modified triester method followed by deprotection and purification (Hirose et al., 1978; Patel et al., 1982a-c). The NMR experiments were performed on the HXS-360 MHz instrument at the Stanford Magnetic Resonance Laboratory, with the experimental methods described in the previous section. The NMR spectra were all run in 0.1M phosphate buffer, 2.5mM EDTA. The chemical shifts were referenced to the internal
standard sodium 4,4-dimethyl-4-silapentane-1-sulfonate (DSS). The lifetimes were calculated as previously discussed without any significant double exponential behavior seen in the data. The lifetimes reported here are estimated to be accurate to within ±20%. Typically, 10-15 different delay times were taken with 220-250 accumulations for each delay time.

Some of the experiments discussed here were repeated on a 200 MHz instrument using methods similar to the long pulse inversion-recovery technique described by Early et al. (1980). The values for the lifetimes calculated from these experiments are not reported here but were found to be within experimental error of the reported values.

B. Results

i. 12-mer: d(C-G-C-G-A-A-T-T-G-G-C-G)₂. Figure 3.13 shows a saturation recovery experiment performed on the 12-mer at 15°C, pH 8. As discussed by Patel et al. (1982a), the terminal G·C base pair imino proton was observed only at very low temperatures and so was not seen in the temperature range used in this study. The measured lifetimes of the other five imino resonances for the 12-mer at pH 6 are given as a function of temperature in Table 3.6. Lifetimes are determined by fitting the data to equation (1) as shown in Figure 3.14. Arrhenius plots for the lifetimes of base pairs 3, 4, 5, and 6 are shown in Figure 3.15.

The pH dependence of the lifetimes of the imino protons in the 12-mer was also measured to test for open-limited behavior in these protons. The measured lifetimes of the imino protons in the 12-mer
Figure 3.13. Spectra from a saturation recovery experiment (only partial data shown) on the imino protons in the 12-mer double strand at 15°C and pH 8. The peaks at 13.15, 12.98 and 12.78 ppm, corresponding to imino protons on base pairs 2, 3 and 4, respectively, were saturated. The times in milliseconds correspond to the delay times between the saturation and the detection pulses. The spectrum at infinite time was taken under the same conditions with no power in the saturation pulse.
Table 3.6. Lifetimes (msec) of Imino Protons in 12-mer at pH 6

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>150</td>
<td>260</td>
<td>230</td>
<td>250</td>
<td>270</td>
</tr>
<tr>
<td>15</td>
<td>140</td>
<td>205</td>
<td>255</td>
<td>360</td>
<td>320</td>
</tr>
<tr>
<td>25</td>
<td>100</td>
<td>140</td>
<td>190</td>
<td>310</td>
<td>350</td>
</tr>
<tr>
<td>30</td>
<td>90</td>
<td>175</td>
<td>240</td>
<td>230</td>
<td>280</td>
</tr>
<tr>
<td>35</td>
<td>40</td>
<td>130</td>
<td>170</td>
<td>140</td>
<td>235</td>
</tr>
<tr>
<td>40</td>
<td>30</td>
<td>80</td>
<td>160</td>
<td>125</td>
<td>200</td>
</tr>
<tr>
<td>45</td>
<td>18</td>
<td>60</td>
<td>150</td>
<td>90</td>
<td>130</td>
</tr>
<tr>
<td>50</td>
<td>--</td>
<td>50</td>
<td>80</td>
<td>65</td>
<td>90</td>
</tr>
<tr>
<td>55</td>
<td>--</td>
<td>35</td>
<td>80</td>
<td>35</td>
<td>55</td>
</tr>
</tbody>
</table>
Figure 3.14. Saturation recovery experiment on an interior A·T imino resonance (base pair 5) in the 12-mer. The triangles are experimental data and the solid line is the best exponential fit to the data. This exponential fit results in a $\tau_{\text{obs}}$ of 125 msec.
A·T base pair 5, 40°C

\[ \tau_{\text{obs}} = 125 \text{ msec} \]
Figure 3.15. Arrhenius plots for the observed lifetimes of the 12-mer double helix for a) the G·C base pairs and b) the A·T base pairs. The line used to determine the activation energies was calculated from the first five points (temperatures of 35°C and above).
\begin{align*}
\text{dC-G-C-G-A-A-T-T-C-G-C-G} \\
\text{G-C-G-C-T-T-A-A-G-C-G-Cd}
\end{align*}

\begin{align*}
\text{\textbullet{} base pair 3} \\
\text{\textsquare{} base pair 4}
\end{align*}

\begin{align*}
\text{\textcircled{}} \text{ base pair 5, } E_a = 14 \text{ kcal/mol} \\
\text{\texttriangle{} base pair 6, } E_a = 15 \text{ kcal/mol}
\end{align*}
are given at pH 6 and pH 8 for several temperatures in Table 3.7.

ii. 12-mer G·T: d(C-G-T-G-A-A-T-T-C-G-C-G)_2. Figure 3.16 shows an example of a saturation recovery experiment on the 12-mer G·T at 15°C, pH 8. The measured lifetimes of the imino protons on base pairs 2-6 at pH 6 are given as a function of temperature in Table 3.8. Arrhenius plots for the lifetimes of base pairs 4-6 are shown in Figure 3.17. The pH dependence of the lifetimes of the imino protons in the 12-mer G·T for several temperatures is shown in Table 3.9.

iii. 13-mer: d(C-G-C-A-G-A-A-T-T-C-G-C-G)_2. Figure 3.18 shows a saturation recovery experiment on the 13-mer at 15°C, pH 8. The measured lifetimes of the imino protons at pH 8 are given as a function of temperature in Table 3.10. Arrhenius plots for the lifetimes of base pairs 4-6 are shown in Figure 3.19.

C. Discussion

i. T₁ vs. Chemical Exchange. I will again consider two relaxation processes which contribute to the relaxation lifetimes: chemical exchange and magnetic spin-lattice relaxation. At temperatures well below the melting temperature the lifetimes measured by NMR are dominated by the spin-lattice relaxation time, T₁, of the imino protons. As the temperature approaches the melting temperature, Tₘ, chemical exchange with water dominates the observed lifetimes (Johnston & Redfield, 1978; Early et al., 1981a,b; Pardi & Tinoco, 1982).

In the 12-mer the T₁ becomes important below 30°C. This can be seen in Table 3.6 where the lifetimes of the imino protons increase
<table>
<thead>
<tr>
<th>Temp.</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH 6  15°C</td>
<td>140</td>
<td>205</td>
<td>255</td>
<td>360</td>
<td>320</td>
</tr>
<tr>
<td>pH 8</td>
<td>160</td>
<td>230</td>
<td>260</td>
<td>275</td>
<td>280</td>
</tr>
<tr>
<td>pH 6  35°C</td>
<td>40</td>
<td>130</td>
<td>170</td>
<td>140</td>
<td>235</td>
</tr>
<tr>
<td>pH 8</td>
<td>23</td>
<td>170</td>
<td>280</td>
<td>180</td>
<td>235</td>
</tr>
<tr>
<td>pH 6  45°C</td>
<td>18</td>
<td>60</td>
<td>150</td>
<td>90</td>
<td>130</td>
</tr>
<tr>
<td>pH 8</td>
<td>a</td>
<td>b</td>
<td>180</td>
<td>70</td>
<td>105</td>
</tr>
</tbody>
</table>

*a Too fast to measure, <5 msec.*

*b Lifetime is difficult to measure due to overlapping peaks at this temperature.*
Figure 3.16. Spectra from a saturation recovery experiment (only partial data shown) on the imino protons in the 12-mer G·T helix at 15°C and pH 8. The peaks at 13.16, 11.65 and 10.6 ppm, corresponding to base pairs 2 and the imino protons from the G·T base pair, respectively, were saturated. Assignment of the low field resonance to the T and the high field resonance to the G in the G·T base pair was made by Patel et al. (1982c). The times in milliseconds correspond to the delay times between the saturation and the detection pulses. The spectrum at infinite time was taken under the same conditions with no power in the saturation pulse. The curvature of the baseline in these spectra is due to the Redfield 214 pulse.
dC-G-T-G-A-A-T-T-C-G-C-G

(GTmer)

infinity
80 msec
30 msec
0.1 msec

PPM
Table 3.8. Lifetimes (msec) of Imino Protons in the 12-mer G·T at pH 6

dC-G-T-G-A-A-T-T-C-G-C-G

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>210</td>
<td>210</td>
<td>160</td>
<td>225</td>
<td>255</td>
</tr>
<tr>
<td>15</td>
<td>170</td>
<td>150</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>40</td>
<td>90</td>
<td>110</td>
<td>185</td>
<td>200</td>
</tr>
<tr>
<td>25</td>
<td>30</td>
<td>55</td>
<td>55</td>
<td>150</td>
<td>235</td>
</tr>
<tr>
<td>30</td>
<td>35</td>
<td>35</td>
<td>90</td>
<td>205</td>
<td>255</td>
</tr>
<tr>
<td>35</td>
<td>9</td>
<td>12</td>
<td>12</td>
<td>85</td>
<td>135</td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td>40</td>
<td>75</td>
</tr>
<tr>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9</td>
</tr>
</tbody>
</table>
Figure 3.17. Arrhenius plots for the observed lifetimes of the 12-mer G·T helix for a) the G·C base pairs and b) the A·T base pairs. The activation energy was calculated from the first four points in (a) and the first three points in (b) (temperatures of 35°C and above).
\[
\begin{align*}
\text{dC-G-T-G-A-A-T-T-C-G-C-G} \\
\text{G-C-G-C-T-T-A-A-G-T-G-Cd}
\end{align*}
\]

\[\begin{array}{cccccc}
1 & 2 & 3 & 4 & 5 & 6 \\
\hline
\end{array}\]

- In T (b)

\[\frac{1}{T} \times 10^3 (\text{K}^{-1})\]

- In T (b)

\[\frac{1}{T} \times 10^3 (\text{K}^{-1})\]

XBL 825-10192
Table 3.9. pH Dependence of Lifetimes (msec) of Imino Protons in the 12-mer G·T

<table>
<thead>
<tr>
<th>Temp. 20°C</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH 6</td>
<td>40</td>
<td>90</td>
<td>110</td>
<td>190</td>
<td>200</td>
</tr>
<tr>
<td>pH 8</td>
<td>20</td>
<td>32</td>
<td>30</td>
<td>220</td>
<td>260</td>
</tr>
<tr>
<td>pH 6</td>
<td></td>
<td>35</td>
<td>35</td>
<td>--</td>
<td>200</td>
</tr>
<tr>
<td>pH 8</td>
<td></td>
<td>--</td>
<td>10</td>
<td>--</td>
<td>200</td>
</tr>
<tr>
<td>pH 6</td>
<td>40</td>
<td>75</td>
<td>95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pH 8</td>
<td>60</td>
<td>85</td>
<td>130</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 2 3 4 5 6 6 5 4 3 2 1
dC-G-T-G-A-A-T-T-C-G-C-G

Imino Proton (Base Pair #)
Figure 3.18. Spectra from a saturation recovery experiment (only partial data shown) on the imino protons in the 13-mer helix at 15°C and pH 8. The peaks at 13.14, 12.93, and 12.34 ppm, corresponding to base pairs 2, 3 and 4, respectively, were saturated. The times in milliseconds correspond to the delay times between the saturation and the detection pulses. The spectrum at infinite time was taken under the same conditions with no power in the saturation pulse.

(13mer)
Table 3.10. Lifetimes (msec) of Imino Protons in 13-mer at pH 8

<table>
<thead>
<tr>
<th>Temperature(°C)</th>
<th>Imino Proton (Base Pair #)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1  2  3  4  5  6  5  4  3  2  1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Temperature(°C)</th>
<th>#2</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>32</td>
<td>45</td>
<td>160</td>
<td>250</td>
<td>290</td>
</tr>
<tr>
<td>20</td>
<td>195</td>
<td>300</td>
<td>270</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
<td>155</td>
<td>220</td>
<td>265</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
<td>100</td>
<td>150</td>
<td>225</td>
</tr>
<tr>
<td>35</td>
<td></td>
<td></td>
<td>60</td>
<td>85</td>
<td>165</td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
<td>32</td>
<td>45</td>
<td>60</td>
</tr>
<tr>
<td>45</td>
<td></td>
<td>15</td>
<td>14</td>
<td>22</td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.19. Arrhenius plots for the observed lifetimes of the 13-mer helix for a) the G·C base pairs and b) the A·T base pairs. The activation energy was calculated from the first three points (temperatures of 35°C and above).
\[ \text{dC-G-C-A-G-A-T-T-C-G-C-G} \]
\[ \text{G-C-G-C-T-T-A-A-G-A-C-G-Cd} \]

\[ \begin{align*}
\text{1} & \quad \text{2} & \quad \text{3} & \quad \text{4} & \quad \text{5} & \quad \text{6} & \quad \text{6} & \quad \text{5} & \quad \text{5} & \quad \text{4} & \quad \text{3} & \quad \text{2} & \quad \text{1} \\
\text{dC-G-C-A-G-A-T-T-C-G-C-G} \\
\end{align*} \]

\[ \begin{align*}
\text{-ln } T_{\text{obs}} & \quad \text{vs} \quad \frac{1}{T} \times 10^3 (\text{oK}^{-1}) \\
\text{O} - \text{base pair \#4, } E_a = 27 \text{ kcal/mol} \\
\text{A} - \text{base pair \#6, } E_a = 39 \text{ kcal/mol} \\
\end{align*} \]

\[ \begin{align*}
\text{3.00} & \quad \text{3.20} & \quad \text{3.40} & \quad \text{3.60} \\
\text{I/T \times 10^3 (oK^{-1})} \\
\end{align*} \]
sharply when the temperature changes from 55°C to 25-30°C. For base pairs 4, 5 and 6 the lifetimes then level off until 5°C where they begin to decrease. The contribution of $T_1$ at low temperatures becomes more apparent in the Arrhenius plots for the lifetimes of these protons (Figure 3.15). The observed lifetimes for temperatures above 30°C is dominated by chemical exchange of the imino protons with water, which is consistent with studies on other systems (Johnston & Redfield, 1978; Early et al., 1981a,b; Pardi & Tinoco, 1982). Similar behavior can be seen for the imino protons on the 12-mer G·T and 13-mer. This is illustrated for the 12-mer G·T in Table 3.8 and Figure 3.17 and for the 13-mer in Table 3.10 and Figure 3.19. Therefore, I will concentrate on the lifetimes above 30°C, which measure the chemical exchange of the imino protons in all three helices.

The lifetimes at the high temperatures will have some contribution due to the spin-lattice relaxation, $T_1$. A contribution from $T_1$ will result in a lifetime for chemical exchange which is greater than the observed lifetime. Because the low-temperature data, where $T_1$ is dominant, are limited to temperatures above 0°C, it is difficult to determine the effect of $T_1$ accurately. I have attempted to estimate the contribution from $T_1$ by using the low temperature points on an Arrhenius plot to estimate the activation energy (i.e. the slope on the Arrhenius plot) for the $T_1$ process. The slope of the low temperature points can then be used to determine the value of $T_1$ at higher temperatures. The actual value for the exchange lifetime can then be determined by subtracting $1/T_1$ from $1/\tau_{obs}$. 
Using this method, I have estimated that accounting for the contribution from $T_1$ will increase the lifetimes for chemical exchange and the activation energies by not more than 50%. This will not change any of the qualitative conclusions made in the following sections.

ii. Are These Imino Protons in the Open-Limited Region for Exchange? The exchange rates are easily interpreted in terms of base-pair opening rates if exchange is open-limited. The exchange of the imino protons has been found to be in the open-limited region in tRNA (Hurd & Reid, 1980) and also for the interior base pairs in the double helix dCA$_2$G + dCT$_2$G (Pardi & Tinoco, 1982). To determine if exchange is open-limited, the concentration of catalyst is varied and any changes in the measured lifetimes are observed. If the system is open-limited the lifetimes will be independent of catalyst concentration. In these studies the catalysts are OH$^-$ and HPO$_4^{2-}$ and the concentrations of these catalysts are changed by varying the pH.

The measured lifetimes for the imino protons in the 12-mer at pH 6 and pH 8 are given in Table 3.7. The lifetimes should decrease as the concentration of base increases, if the exchange is not in the open-limited region (Crothers et al., 1974; Hilbers, 1979; Pardi & Tinoco, 1982). At 35°C and 45°C only the imino proton from base pair 2 has a lifetime that shows this decrease on going from pH 6 to pH 8. The measured lifetimes for base pairs 3-6 above 30°C therefore measure the rates for opening of these base pairs. The measured lifetimes for base pair 2 represent only an upper limit for the lifetime for base-pair opening. In the 12-mer G·T, again the
chemical exchange process dominates the lifetimes above 30°C in base pairs 4, 5 and 6 (Table 3.8). The pH dependence of the imino protons in the 12-mer G•T is given in Table 3.9. For base pairs 4-6, exchange is independent of pH and therefore is in the open-limited region. At 20°C the lifetimes of the imino protons on base pairs 2 and 3 decrease by a factor of two or more on going from pH 6 to pH 8. The dependence of these protons on pH indicates that opening the base pair is not the rate-limiting step in the exchange of these imino protons.

In the 13-mer, I will again consider only points above 30°C and will thus only be measuring chemical exchange (Table 3.10). The pH dependence of the lifetimes of the 13-mer was not measured. I will assume that base pairs 4-6 are analogous to those base pairs in the 12-mer and 12-mer G•T and thus are in the open-limited region.

iii. Effects of Perturbations on the Lifetimes of Exchange.
By comparing relaxation lifetimes of the three related dodecamer oligonucleotide duplexes we can determine the effects of a G•T base pair or an extra nucleotide on a double-stranded helix. For the 12-mer at 35°C and pH 6 the lifetimes for base pairs 4, 5 and 6 are 170, 140 and 235 msec, respectively (Table 3.6).

In the 12-mer G•T at 35°C and pH 6 the lifetimes are 85, 135 and 225 msec for base pairs 4, 5 and 6, respectively (Table 3.8). Lifetimes for base pairs 5 and 6 are comparable to those in the 12-mer. The lifetime for base pair 4, which is next to the G•T base pair, decreases by a factor of two. Although the G•T containing helix is destabilized relative to the 12-mer (a decrease in $T_m$ of
these lifetimes show that the perturbation does not affect the lifetimes of the interior A·T base pairs and therefore is a very local effect. The lifetimes of both the G and T imino protons in the G·T base pair are the same at all temperatures which means that when this base pair opens both protons exchange at the same rate. At 35°C the G·T base pair (#3) has a lifetime of 12 msec compared to 130 msec for the analogous G·C base pair in the 12-mer.

For the 13-mer at 35°C and pH 8, lifetimes for base pairs 4, 5 and 6 are 60, 85 and 165 msec, respectively (Table 3.10), compared to 280, 180 and 235 msec in the 12-mer at pH 8 (Table 3.7). Thus all of the lifetimes have decreased in the 13-mer. The perturbation due to the extra adenine is not localized as in the G·T case but affects the whole molecule.

iv. Activation Energies and Exchange Mechanisms. By investigating the temperature dependence of the exchange rates of the imino protons it is possible to obtain activation energies for the exchange process. The magnitudes of these activation energies can give information on the mechanisms of exchange (Pardi & Tinoco, 1982).

The activation energies for exchange of the imino protons of base pairs 5 and 6 of the 12-mer are shown in Table 3.11 and were calculated from the plots in Figure 3.15. Points from 35-55°C were used to calculate these activation energies. As discussed previously, there may be several mechanisms for chemical exchange of the imino protons in these molecules, including local base-pair opening and whole helix opening. The activation energy for helix
Table 3.11. Activation Energies (kcal mol\(^{-1}\)) for Exchange of Imino Protons

<table>
<thead>
<tr>
<th>Molecule</th>
<th>#3</th>
<th>#4</th>
<th>#5</th>
<th>#6</th>
</tr>
</thead>
<tbody>
<tr>
<td>12-mer</td>
<td>---</td>
<td>---</td>
<td>14 ± 2</td>
<td>15 ± 2 (pH 6)</td>
</tr>
<tr>
<td>12-mer G·T</td>
<td>---</td>
<td>30 ± 5</td>
<td>37 ± 8</td>
<td>48 ± 9 (pH 6)</td>
</tr>
<tr>
<td>13-mer</td>
<td>---</td>
<td>27 ± 5</td>
<td>35 ± 8</td>
<td>39 ± 8 (pH 8)</td>
</tr>
</tbody>
</table>
opening of a duplex the size of the 12-mer can be approximated by
the enthalpy for double-strand formation, 102 kcal mol\(^{-1}\) as
determined by calorimetry (Patel et al., 1982a). The values
measured for the exchange of the imino protons in the 12-mer are 14-
15 kcal mol\(^{-1}\). These low values indicate that helix opening is not
a dominant process in the exchange of these protons and suggests
that local base-pair opening may be the important pathway for
exchange. Other evidence for local base-pair opening comes from the
lifetimes of the individual base pairs. At 45°C exchange lifetimes
of base pairs 4, 5 and 6 are 150, 90 and 130 msec, respectively
(Table 3.6). Because the lifetime of base pair 5 is much shorter
than either of its neighbors this imino proton must, to some extent,
exchange independently of the other two. Our results are in
agreement with Early et al. (1981a,b) who found activation energies
of around 15 kcal mol\(^{-1}\) for the A-T imino protons in another
dodecamer. The difference in lifetimes (or rates) of the two A-T
base pairs could be due to small differences in their activation
energies or differences in the pre-exponential factor, \(A\), for each
base pair as given in the Arrhenius equation:

\[
-k = A e^{\frac{-E}{RT}}
\]

A difference in the activation entropies for the two A-T imino
protons in the 12-mer might reflect differences in the flexibility
of the base pairs. The activation energies for base pairs 3 and 4
were not calculated because it is not clear where chemical exchange
becomes the dominant mechanism for the relaxation lifetimes.
For the 12-mer G·T the imino protons broaden and disappear at a lower temperature than in the 12-mer. Therefore only 3 or 4 temperatures above 30°C were used for calculation of the activation energies, and thus these values are less reliable than the values obtained for the 12-mer. Table 3.11 gives the activation energies for the 12-mer G·T. The Arrhenius plot is shown in Figure 3.17. It is clear that the activation energies for exchange of the imino protons in the 12-mer G·T are much larger than those in the 12-mer. This indicates local base-pair opening is no longer the dominant pathway for exchange in the 12-mer G·T and that helix opening is probably contributing to the exchange. The activation energies for base pairs 4, 5 and 6 in the 12-mer G·T are 30, 37 and 48 kcal mol⁻¹, respectively (Table 3.11). If exchange were taking place by helix opening we would expect the activation energies and lifetimes to be the same for all the imino protons. This result was found in the heptamer dCA₅G + dCT₅G (Pardi & Tinoco, 1982). Thus the measured activation energies for the 12-mer G·T are probably due to contributions from both base-pair opening and helix opening.

The activation energies for the imino protons of base pairs 4-6 of the 13-mer are given in Table 3.11 and were calculated from the Arrhenius plots in Figure 3.19. Again only points above 30°C were used. The activation energies for base pairs 4, 5 and 6 are 27, 35 and 39 kcal mol⁻¹, respectively. This is very similar to the 12-mer G·T in that the high activation energies probably indicate contributions from both local base-pair opening and helix opening.

In order to better understand how contributions from both local
base-pair opening and helix opening will affect the rates of exchange and the activation energies, I have calculated rates for the system at different temperatures using the model developed previously (Pardi & Tinoco, 1982). For a system where two pathways contribute to the exchange of an imino proton, the observed activation energy, \( E_a^{\text{obs}} \), is given by equation (5). Values for \( E_a^{\text{II}} \) and \( k_{\text{op}}^{\text{II}} \) can be obtained by measuring the kinetics of the helix-to-coil transition using temperature-jump techniques (Porschke & Eigen, 1971). Values were measured for the 12-mer G·T and the 13-mer in the same buffer as the NMR experiments. The activation energies for dissociation of the double helix (\( E_a^{\text{II}} \)) were 68 and 74 kcal mol\(^{-1}\) for the 12-mer G·T and 13-mer, respectively. The pre-exponential factors, \( A \), for this process were found to be \( 1.2 \times 10^{48} \) and \( 2.6 \times 10^{52} \) sec\(^{-1}\) for the 12-mer G·T and 13-mer, respectively (Chu & Tinoco, 1983). Data from base pair 6 of the 12-mer is used to get approximate values for local base-pair opening: \( E_a^{\text{I}} = 14 \) kcal mol\(^{-1}\), \( A = 8 \times 10^{10} \) sec\(^{-1}\). Using these numbers and the Arrhenius equation (eqn. 6) we can calculate the observed rate constant for this two-pathway process at any temperature. The observed rate will be the sum of the individual rates for each pathway. Figure 3.20 shows an Arrhenius plot calculated for the 12-mer G·T using this model. At high temperature the steeper slope indicates a larger activation energy (68 kcal mol\(^{-1}\)) which is the activation energy for helix opening. At lower temperature base-pair opening becomes dominant and there is a smaller activation energy (14 kcal mol\(^{-1}\)). This model predicts that the helix opening pathway
Figure 3.20. An Arrhenius plot calculated for A·T base pair 6 of the 12-mer G·T helix, assuming only helix opening and base-pair opening for exchange of the imino proton. Discussion of these exchange mechanisms and values for the rates are given in the text.
becomes dominant above 40°C, that below 30°C, base-pair opening is the main exchange pathway, and between 30-40°C both pathways are important. This is consistent with the interpretation of the measured activation energies for the 12-mer G·T and the 13-mer. It is clear from these results that NMR experiments alone will not always give enough information to explain the kinetics of exchange in oligonucleotides, and may actually be misleading by themselves. Thus, complementing the NMR kinetic measurements with temperature-jump kinetic measurements leads to a much better understanding of the dynamics of the helices in solution.

5. CONCLUSIONS

This chapter describes experiments which probe the relaxation lifetimes of imino protons of base pairs in different environments. The effects of destabilizing perturbations in a helix on the exchange rates have been studied. The saturation recovery NMR experiments performed here allow the lifetimes for opening of individual base pairs in a helix to be determined. Thus the kinetic stability of each base pair is studied, instead of the overall stability of the whole helix. The helices used in this chapter, dCA_3CA_3G + dCT_6G, dCA_6G + dCT_6G, the 12-mer: d(C-G-C-G-A-A-T-T-C-G-C-G)_2, the 12-mer G·T: d(C-G-T-G-A-A-T-T-C-G-C-G)_2, and the 13-mer: d(C-G-C-A-G-A-A-T-T-C-G-C-G)_2 represent excellent models for probing the effects of a perturbation (an extra cytosine, a G·T mismatch and an extra adenine) on the stability of individual base pairs in the helix. In Chapter II, I have reported on the temperature dependence of the chemical shifts of protons in the first two duplexes and
obtained information on the conformations and the relative stabilities of these molecules; this has been done for the last three duplexes by Patel et al. (1982a,b,c).

The observed lifetimes measured for the helices studied in this chapter were shown to correspond to lifetimes for chemical exchange of the imino protons for temperatures above 5°C for dCA₃CA₃G + dCT₆G and dCA₆G + dCT₆G and temperatures above 30°C for the 12-mer and related helices. These lifetimes were also shown to be in the open-limited region (every time the base pair opens the imino proton exchanges with solvent water) for the internal base pairs in all these duplexes.

The lifetimes from dCA₃CA₃G + dCT₆G indicate that the extra cytosine causes a two fold increase in the exchange rate of the imino proton from base pair 4 compared to the other internal base pairs. Comparison of the lifetimes in the 12-mer helices shows that the G∙T base pair in the 12-mer G∙T causes over an order of magnitude decrease in the relaxation lifetime of its imino protons relative to the G·C base pair in the 12-mer. The destabilization of the G·T base pair also has a large effect on the neighboring base pair (#4) but does not affect the lifetime of the next-nearest-neighbor A·T base pair in the helix. Thus the effect of the G·T base pair is very localized in the helix. The destabilization due to the unpaired adenine in the 13-mer has a larger effect on the helix, with the lifetimes of all the imino protons decreasing, relative to the 12-mer. This overall decrease in the lifetimes is also observed in the duplex containing the extra cytosine.
Activation energies for exchange of the imino protons were obtained by observation of the temperature dependencies of the exchange rates. In \( \text{dCA}_6 \text{G} + \text{dCT}_6 \text{G} \) the imino protons from the internal A\( \cdot \)T base pairs (3-6) have activation energies of 16-23 kcal mol\(^{-1}\); similarly in the 12-mer, both A\( \cdot \)T imino protons have activation energies of 14-15 kcal mol\(^{-1}\). The values indicate that the exchange mechanism of the imino protons is local base-pair opening. Early et al. (1981b) found the same results for the activation energy for exchange of A\( \cdot \)T imino protons in \( \text{dC-C-G-C-A-C-T-G-A-T-G-G + dC-C-A-T-C-A-G-T-G-C-G-G} \). The activation energies for the imino protons in the \( \text{dCA}_3 \text{CA}_3 \text{G} + \text{dCT}_6 \text{G} \), the 12-mer G\( \cdot \)T, and the 13-mer are much larger in magnitude than those in the \( \text{dCA}_6 \text{G} + \text{dCT}_6 \text{G} \) and the 12-mer indicating a helix opening mechanism is involved in the exchange of the imino protons in these three helices. The helix opening pathway was shown to be the dominant exchange process of the double strand \( \text{dCA}_5 \text{G} + \text{dCT}_5 \text{G} \) studied by Pardi \\& Tinoco (1982).

It is clear from these studies that the relaxation lifetimes of an imino proton in a double helix are dependent upon many factors. Whether the imino proton is in an A\( \cdot \)T or G\( \cdot \)C base pair will affect its lifetime, as will the sequence of its neighboring base pairs. A very important factor in the exchange of an imino proton in a double helix is its distance from a terminal base pair. Terminal base pairs are known to kinetically fray (Patel \\& Hilbers, 1975; Pardi \\& Tinoco, 1982), which means that they open and close at very fast rates. This kinetic fraying of the ends of a helix can affect the rates of exchange for the penultimate base pairs, and possibly the third base pair in from the end, depending upon the sequence and the temperature. A perturbation in a helix will also have a large
effect upon the exchange lifetimes of the imino protons in a double helix. These perturbations could be destabilizing factors such as internal loops, bulges, or non-standard base pairs such as the G·T base pair studied in this work. Many drugs are known to bind to nucleic acids and to stabilize the double helix. These drugs would be expected to have a large effect on the exchange rates of the imino protons in the helix (see Patel et al., 1982d for a preliminary report of such studies). The lifetimes for exchange of imino protons from individual base pairs in a helix give extremely valuable information on the extent of such perturbations throughout the helix (Pardi et al., 1983).
REFERENCES


Chapter IV
CIDNP Studies of Nucleic Acids

1. CIDNP THEORY

The phenomenon of chemically induced dynamic nuclear polarization (CIDNP) was first observed in 1967 as a decrease or an increase in the NMR signal intensities of spectra taken during a chemical reaction (Bargon et al., 1967; Ward & Lawler, 1967). The role of free radical reactions in this process is explained by the Radical Pair Mechanism (Closs, 1969; Kaptein & Oosterhoff, 1969) which I will briefly explain here. For a more in depth discussion of the CIDNP phenomenon see Closs (1974) and Kaptein (1975).

CIDNP has been used to study biological molecules by creating the radical pair through the interaction of the biological molecule and a photo-excited dye. This technique involves a cyclic reaction in which the dye is photo-excited, interacts with the biological molecule, and forms a radical pair all without any overall chemical change. This process is shown schematically in Figure 4.1. In order to understand the process which takes place when CIDNP is created we must take a closer look at the pair of radicals. In Figure 4.1, the species under the bar schematically represents a "caged" radical pair; the radicals are constrained to be close to each other unless they "escape" from the cage. Several things happen during the "caged" pair step. First, the excited dye molecule interacts with the biological molecule to create two unpaired electrons, then the two radicals diffuse apart and finally the radicals can come together again and back react to form the
Figure 4.1. The cyclic reaction scheme for the creation of CIDNP in biological molecules. D represents the dye, where $^1D$ is the photo-excited singlet of the dye which undergoes intersystem crossing (ISC) to form $^3D$, the triplet state of the dye. BH represents the biological molecule containing the proton being observed. BH* is the biological molecule containing polarized protons. The species under the bar are caged radical pairs.
original species. When the two electrons, one on each radical, are in the cage they form a radical pair similar to an electron pair in a molecule. In a magnetic field the spins of the two electrons can be in one of four states. These are represented vectorially in Figure 4.2. The environment of the electron will affect its precession rate, or more quantitatively, the precession rate is characterized by its g-factor. If the environments or g-factors of the two electrons are different, the electrons will precess at different rates. This can result in S-T<sub>0</sub> mixing as shown in Figure 4.2. The back reaction required for the cyclic mechanism, shown in Figure 4.1, is kinetically favored when the radical pair is in the singlet state. So, the environment around the electrons will affect their relative precession rates in the radical pair. This will affect the S-T<sub>0</sub> mixing rate and will, therefore, determine if the back transfer of the electron will occur.

How does all this relate to an increased or decreased NMR signal? The nucleus with which an unpaired electron is associated also has a spin. In a magnetic field the nuclear spin will be aligned with or against the applied field. The two states are referred to as the β and α states, respectively. These two states will be populated according to the Boltzmann distribution. This means that there are slightly more protons in the lower energy, or β, state than in the α state and results in the net absorption observed as an NMR signal. During the NMR experiment, energy is put into the system and causes β to α transitions to occur; through relaxation, α to β transitions will also occur.
Figure 4.2. The vector representation of the possible spin states for electrons in a radical pair when placed in a magnetic field $H_0$. $S_1$ and $S_2$ are the spin vectors for the two electrons in the radical pair.
TRIPLETS

SINGLET
Now, consider the CIDNP experiment. The state of the nucleus affects the environment of the unpaired electron and will therefore influence its precession rate, as discussed above. Let us suppose the $\beta$ state of the proton affects the electron spin such that it promotes the $S-T_0$ mixing. In this case the back reaction will be more likely to occur between the dye and molecules with protons in the $\beta$ state than between the dye and molecules with protons in the $\alpha$ state. The molecules with protons in the $\alpha$ state will then escape from the cage and go into solution as free radicals. Since the nuclear relaxation rate in the presence of an unpaired electron is very fast (at least 100 times faster than in the absence of the electron) these protons will relax back into a normal distribution of $\alpha$ and $\beta$ states. This will occur faster than the time scale of the NMR experiment. The back transfer has therefore selected for the $\beta$ nuclei (assuming their relaxation rate, $T_1$, is slower than the time scale of the experiment) and the proton population becomes $\beta$ polarized. This will result in a net absorption over that expected from a normal distribution of spins. If the $\alpha$ state of the proton had increased the $S-T_0$ mixing rate then there would be a decrease in the absorption compared to that expected for a normal distribution.

From this simple explanation of the CIDNP effect, it is clear that there will be no polarization if the dye can not get close enough to react with the biomolecule. Also the reaction scheme shown in Figure 4.1 involves reaction with a native biomolecule and gives the native biomolecule back as a product. So, the technique
is theoretically nondestructive. Of course, photo-exciting the dye can cause some photo-degradation of the sample.

2. INTRODUCTION

The conformation of nucleic acids and the forces responsible for maintaining the structures of oligonucleotides provide a basis for understanding their biological function. NMR and circular dichroism (CD) have been the primary methods used to determine the conformation of nucleotides, oligonucleotides and nucleic acids in solution (Sarma, 1980; Davies, 1978a; Johnson, 1978). Measurements of vicinal coupling constants can provide values (or a range of values for species equilibrating between conformations) for the backbone torsion angles $\beta$, $\gamma$, $\delta$ and $\epsilon$, but not for the glycosidic torsion angle, $\chi$, or for the phosphodiester bond angles, $\alpha$ and $\zeta$ (see Figure 4.3 for notation). At low temperatures the phosphodiester bond angles are assumed to be predominantly in the $\alpha^{-}\zeta^{-}$ conformation, based on chemical shift calculations and molecular models (Sarma, 1980). This is the only conformation, out of the nine possible conformations around the phosphodiester bonds, that allows both base stacking and a reasonable distance between the bases.

It is generally assumed that in neutral aqueous solution all naturally occurring nucleic acids exist as an equilibrium mixture of syn ($-90^\circ < \chi < 90^\circ$) and anti ($90^\circ < \chi < 270^\circ$) conformations, with a preference for the anti conformation (Lee, et al., 1976). Indirect NMR parameters can be used to estimate $\chi$, such as the chemical shifts of the base and ribose protons as a function of temperature,
Figure 4.3. The chemical structure and numbering system for ApGpCpU. The backbone and glycosidic torsion angles are illustrated for G.
pH, and various substituents, changes in chemical shifts and relaxation times upon metal binding, and $^{13}\text{C}-^{1}\text{H}$ and long range $^{1}\text{H}-^{1}\text{H}$ coupling constants. In some cases these values are in conflict with those derived from $T_1$ and nuclear Overhauser enhancement (NOE) measurements (Sarma, 1980; Davies, 1978). In dinucleotides, $\chi$ has been estimated from proton chemical shift values as a function of pH, temperature, and composition, by using molecular models, and by $T_1$ measurements (Sarma, 1980; Davies, 1978a; Neuman, et al., 1979; Chachaty, et al., 1977; Tran-Dinh, et al., 1981; Chachaty, et al., 1980).

Photo-CIDNP has been used extensively by Kaptein and his co-workers to study the surface accessibility of photo-excited flavin dyes to histidine, tyrosine and tryptophan residues in proteins (Kaptein, 1980). Kaptein has also shown that CIDNP is observed for adenine and guanine bases, nucleotides, and nucleosides, for the pyrimidines, thymine, 5-methyl cytidine, and 3-methyl cytidine (Kaptein, et al., 1979), and for the dinucleotides $d(ApA)$ and $d(pApA)$ (Kaptein, et al., 1981) following reaction with a photo-excited flavin. Solutions of flavins and the oligonucleotides $d(pCpGpCpG)$ (Garssen, et al., 1978), $d(pCpGpCpGpCpG)$ (Hilbers, et al., 1978), oligo[d(AT)] (Buck, et al., 1980), and poly(A) (Garssen, et al., 1978), however showed no CIDNP following laser irradiation.

In this chapter, I will discuss the CIDNP we have observed in other nucleotides, oligonucleotides and tRNA, with particular emphasis on the G polarization. This work was done in collaboration with Dr. E. F. McCord and Prof. S. G. Boxer at Stanford University.
and is also discussed in McCord (1983). The CIDNP experiments provide a convenient method for assigning NMR resonances belonging to G nucleosides in oligonucleotides, including the H1' proton. The polarization intensity of the latter varies with the temperature and the oligonucleotide length, suggesting that it may be a monitor of the glycosidic torsion angle. Because CIDNP requires the target residue to be accessible to the photo-excited dye, polarization is expected only for single-stranded solvent-accessible regions of the oligonucleotides. This is confirmed by observing CIDNP while following the melting transition of the self-complementary tetramer ApGpCpU.

This work represents the first use of CIDNP as a probe of the solvent accessibility of guanine residues in tRNA^{phe}. The observed CIDNP in tRNA is the result of non-destructive, cyclic electron transfer reactions and therefore reflects the accessibility of the G bases in the native tRNA molecule.

There has been a great deal of interest in the conformation of tRNA in solution and in the effects of various perturbations such as aminoacylation, pH, Mg$^{2+}$ concentration, and anticodon-codon base pairing on this conformation. Many methods, including NMR, CD, and chemical modification have been used to study the conformation of yeast tRNA in solution. The results of these studies generally agree with the conformation determined for the yeast tRNA^{phe} crystal by X-ray crystallography (Holbrook & Kim, 1983).

The CIDNP results presented here show that at 20°C both aromatic and aliphatic resonances of yeast tRNA^{phe} show
polarization. The variations in this polarization are studied as a function of temperature, phosphate concentration, dye, and excision of the Y base. Two of the polarized aliphatic resonances are assigned to the methyl groups of m^2_G26 and m^2_G10. Assignments are suggested for the polarized aromatic resonances based on chemical modification studies (Rhodes, 1975), NOE results, and observation of CIDNP due to dipolar cross-relaxation between resonances.

3. EXPERIMENTAL

Nucleic acid mononucleotides, dinucleotides and ApGpU were purchased from Sigma, Aldrich, or P. L. Biochemicals. ApGpCpU was a generous gift from Dr. P. Borer of Syracuse University and was also synthesized by Mr. D. Koh in our laboratory using enzymatic methods (Martin, et al., 1971; Uhlenbeck, et al., 1971). The chemical structure and numbering system of ApGpCpU are shown in Figure 4.3 for reference. N^2-acetyl-2',3'-'O-isopropylidene-N^3,5'-cycloguanosine (V1) was a generous gift from Dr. J. P. H. Verheyden of Syntex Corp., Palo Alto. 8,5'-Imino-1-methoxymethylene-9-(5'-deoxy-ß-D-ribofuranosyl)guanine (S1), 8,5'-aminoimino-1-methoxymethylene-9-(5'-deoxy-2',3'-O-isopropylidene-ß-D-ribofuranosyl)guanine (S2), and 8,2'-(N^2-methylhydrazino)-9-(2'-deoxy-ß-D-arabinofuranosyl)-guanine (S3) were generous gifts from Professor Sasaki, Nagoya University, Japan. These compounds will be referred to as V1, S1, S2, and S3, as indicated. Their chemical structures are illustrated in Figure 4.4. 7-methyl-8-bromo-10-(1'-D-ribityl)isoalloxazine (8BrF) was a generous gift from Dr. J. Lambooy at the University of Maryland.
Figure 4.4. The structure and numbering systems of:

VI: $N^2$-acetyl-2',3'-O-isopropylidene-\(N^3,5'\)-cycloguanosine

S1: 8,5'-Imino-1-methoxymethylene-9-(5'-deoxy-\(\beta\)-D-ribo-
    furanosyl)guanine

S2: 8,5'-aminoimino-1-methoxymethylene-9-(5'-deoxy-2',3'-O-
    isopropylidene-\(\beta\)-D-ribofuranosyl)guanine

S3: 8,2'-(\(N^2\)-methylhydrazino)-9-(2'-deoxy-\(\beta\)-D-arabino-
    furanosyl)guanine

note: the 2',3'-O-isopropylidene group of S2 is not shown in
the figure.
$V_1$

$S_1, S_2^*$

$S_3$

$S_1 \ R = H$

$S_2 \ R = NH_2$
Typically, samples were prepared by combining the nucleotide dissolved in D$_2$O and 0.5 ml of deuterated pH 7 Sorensen's phosphate buffer (Handbook of Biochem. & Mol. Biol., 1970), usually 66mM phosphate, saturated with riboflavin (0.4mM), followed by lyophilization and dissolution in D$_2$O (100% D$_2$O, Aldrich). The pH values of deuterated solutions were measured with a glass electrode and were not corrected for the isotope effect. The H$_3$ protons of the purines were exchanged for deuterons by heating the samples in an 80°C water bath in the sealed NMR tube for 3 hours.

The 8BrF used in some of the tRNA experiments contained approximately 20% of the corresponding chloro compound, 8ClF. Yeast tRNA$^{\text{phe}}$ was purchased from Boehringer-Mannheim. Sample concentrations for the tRNA were determined optically using the extinction coefficient at 260nm of $5.1 \times 10^5$ M$^{-1}$ cm$^{-1}$ (Yoon, 1976). The tRNA samples were prepared by repeated dialysis against a buffer containing 0.15M NaCl, 5mM phosphate, and 1mM EDTA at pH 7. Samples were then dialyzed repeatedly against the same buffer without the EDTA, lyophilized, and dissolved in D$_2$O to give a final nucleic acid concentration of ~1.5mM. The sample was then added to the lyophilized dye. The Y base was excised from tRNA$^{\text{phe}}$ (the resulting molecule will be referred to as tRNA$^{\text{phe}}$-Y) by the method of Thiebe & Zachau (1968). The tRNA$^{\text{phe}}$ sample (containing dye) was diluted to 2.5 ml, the pH was adjusted to 2.9 with 1N HCl, and the sample was heated for four hours in the dark at 37°C. The sample was then neutralized to pH 7, lyophilized, and dialyzed against a buffer containing 5mM phosphate and 0.15M NaCl at pH 7. Additional 8BrF
was added to the sample which was then lyophilized and dissolved in D$_2$O. As mentioned previously, the pH of deuterated buffers refers to the uncorrected pH meter readings.

NMR data were obtained on the 360 MHz NMR spectrometer at the Stanford Magnetic Resonance Laboratory. The chemical shifts are referenced to the internal standard 3-(trimethylsilyl)-1-propane sulfonic acid, DSS, except for the ApGpCpU samples which were referenced to 3-(trimethylsilyl)propionate, TSP. For the tRNA samples the DSS was added to the samples after the CIDNP experiments. The signal from the residual water was suppressed by presaturating the HDO resonance. The temperature was 25±1°C, unless otherwise indicated, for all samples except the tRNA samples. The tRNA samples were 20±1°C, unless otherwise indicated. T$_1$ data were collected on the same samples used for the CIDNP experiments. The data were collected using a standard inversion-recovery pulse sequence and the values are accurate to ±0.2 seconds. The photo-CIDNP apparatus and the data collection system have been previously described (McCord, et al., 1981). CIDNP spectra were obtained using a 90° rf pulse. CIDNP spectra are presented as the difference between spectra obtained with photo-excitation (light spectra) and those obtained without photo-excitation (dark spectra). The photo-excitation is provided by an Argon-ion laser irradiating at 480 nm. Enhancement factors are defined as the difference between the integrated intensity of a polarized peak and its unenhanced dark intensity, divided by the latter. These values should be used for comparison only as the absolute CIDNP intensity depends on sample
geometry, dye concentration, etc. The $T_m$ value of ApGpCpU was determined by measuring the temperature dependence of the absorbance at 260nm on a Gilford Model 250 Spectrophotometer (Nelson, et al., 1981).

4. RESULTS

A. Bases, Mononucleosides, and Mononucleotides

Solutions of riboflavin (0.2mM) and the following molecules were studied: uracil (Ura); thymine (Thy), cytosine (Cyt), inosine (I), guanosine (G), uridine (U), adenosine (A), 3'-GMP, 3'-UMP, 3'-AMP, 3'-CMP, 3'-IMP, 3'-dTMP, 5'-GMP, 5'-CMP, 5'-AMP, 5'-UMP, 5'-dAMP, 5'-dTMP, 9-methylguanine (9MeGua), and 5-methylcytosine (5MeCyt). Solutions were either 33mM or 66mM phosphate buffer, 2.5mM in base, except for 9MeGua, which was a saturated solution (~1.5mM). All samples were pH 7, except for 5'-CMP, 5'-AMP, 5'-UMP, and 5'-dAMP, which were pH 5. Samples of 5'-GMP were examined at both pH 5 and 7. Samples of I (2.5mM) and G (2.5mM) containing various phosphate concentrations (0-750mM) and 0.4mM riboflavin at pH 7 were examined. Samples of G were also studied as a function of pH (pH 5-10, phosphate or borate buffer), and at pH 7 (66mM phosphate buffer) as a function of temperature.

The I(H8) and G(H8) enhancements decrease monotonically as the phosphate concentration increases; at 750mM phosphate the enhancement is about half that at zero phosphate concentration. The sign of the polarization does not change as the phosphate concentration is varied. The G polarization is strongest at pH 7-8 and decreases at both lower and higher pH. The polarization decreases
as the temperature is increased; the enhancement at 55°C is approximately half that at 20°C. The polarization also decreases ~15% as the temperature is lowered to 5°C. The sign and magnitude of the A(H8) and A(H2) polarization are highly variable, as previously discussed (Scheek, et al., 1981). Polarization is only observed for U, T, and C derivatives if a proton or deuteron is present in the 1 position of the pyrimidine ring. Thus, H5 shows negative polarization in Ura and Cyt (see Figure 4.5), while the 5-methyl protons in 5MeCyt and Thy show positive polarization. There is no polarization observed for U, 3'-UMP, 3'-CMP, 3'-dTMP, 5'-UMP, 5'-CMP or 5'-dTMP.

The G(H8) resonance shows a positive CIDNP in all cases (see Figure 4.6). The measured G(H8) enhancement factors for most of the monomers are in the range of 4-8, while that of 9MeGua is 11.5. Polarization of this proton can easily be observed in a single scan with solutions of G at concentrations as low as 0.06mM. Polarization of the 1'-sugar proton in G (see Figure 4.7) and 5'-GMP or of the methyl protons in 9MeGua is extremely weak and barely detectable at 6°C. No sugar polarization is observed in 9MeGua, 5'-GMP or G when the H8 proton is either saturated with rf during the light pulse or replaced by deuterium. Because relaxation causes the return of the spins to their equilibrium populations, and thus a loss of polarization, it is important to characterize the relaxation lifetimes, T1. The T1 values measured for H8 and H1' in G are 1.8 s and 2.4 s, while those for HB and the methyl protons in 9MeGua are 5.2 s and 1.3 s, respectively. In solutions containing both A and G, the relative amount of A polarization as compared to G
Figure 4.5. The 360 MHz $^1$H NMR light-minus-dark difference spectra (1 accumulation each) in 66mM phosphate buffer, pH 7, 0.4mM riboflavin at 25°C. The dotted lines indicate resonances which are not spin polarized and therefore cancel in the difference spectrum. The two spectra are A) Cytosine (2.5mM) and B) Uracil (2.5mM).
Figure 4.6. The 360 MHz $^1H$ NMR spectra (1 accumulation each) of guanosine (2.5mM) in 66mM phosphate buffer, pH 7, 0.4mM riboflavin at 25°C. The structure illustrates the anti conformation of the nucleoside. The two spectra are A) the light-minus-dark spectrum and B) the dark spectrum. The R indicates resonances due to the riboflavin.
GUanosine
Figure 4.7. The 360 MHz $^1$H NMR spectra of ApGpU (1.2 mM) in 66 mM phosphate buffer, pH 7, 0.4 mM riboflavin at 22°C. R indicates riboflavin resonances. The two spectra are A) the light-minus-dark spectrum (4 accumulations) and B) the dark spectrum (100 accumulations).
polarization varies depending on the A and G concentrations. When both nucleosides are 2.5mM (0.4mM riboflavin), the relative G(H8) to A(H8) polarization is 18 to 1, whereas, when both nucleosides are 0.1mM (0.4mM riboflavin), the relative G(H8) to A(H8) polarization is 0.9 to 1.

Solutions of the modified or rare bases 1-methylguanosine (m1G), N2-dimethylguanosine (m2G), N2-methylguanosine (m2G), dihydrouridine (D), 2'-O-methylguanosine (Gm), pseudouridine (Ψ), 7-methylguanosine (m7G), 1-methylinosine (m1I), 2-aminopurine, 8-methylaminoadenine (8-me-amino-A), and purine were studied. The m1G shows weak negative polarization of the G(H8) resonance (enhancement factor about -2). The m2G is nearly insoluble and shows very weak negative polarization of the methyl protons at 25°C. The solubility of this compound is very temperature dependent. At 40°C and above, strong CIDNP is observed for m2G(H8) and for the methyl protons. At 40°C the ratio of the enhancements of the H8 to the methyl resonance is 1 to -2.5. When observed separately, m2G and G have almost the same enhancement factor for the H8 proton (0.3mM base, 0.25mM riboflavin, 66mM phosphate, pH 7, 40°C). In an equimolar mixture of these two compounds (1mM each) under the same conditions, the H8 protons again showed almost identical enhancements. The m2G is increasingly soluble at higher temperatures. It, too, shows positive m2G(H8) polarization and negative methyl proton polarization. Dihydouridine shows no polarization and pseudouridine shows weak positive polarization of the H1' proton. The Gm shows only positive H8 polarization. In m7G, H8 is not observed due to the
fast exchange of this proton with deuterium, however, the H1' shows weak positive polarization. The m^1A shows positive polarization of the H8 proton. For both m^1I and purine there is no observable polarization, while 2-aminopurine shows positive polarization of H8 and H6. 8-me-amino-A shows positive polarization of H2 and negative polarization of the methyl resonance. The results for all the mononucleotides are summarized in Table 4.1.

A series of synthetic purine mononucleoside derivatives with fixed glycosidic angle, \( \chi \), was also studied in a solution of 66mM phosphate buffer, 0.4mM riboflavin, and pH 7 (see Figure 4.4 for the structures of these compounds). The cyclic derivative V1 gives strong positive polarization for H8 and weak negative polarization for H1', H5', H5'', and the methyl proton resonances of the acetyl group. The polarization of H1' disappears when H8 is either deuterated or saturated with rf during the light pulse. S3 shows positive polarization of the 8-aminomethyl protons and of H2', and weak negative polarization of H1'. The weak negative polarization of H1' was not affected by rf irradiation of H2' during the light pulse. S1 and S2 show very weak negative polarization of H1' and of the N1 methylene protons. It should be cautioned that weak negative polarization could well be an artifact in light-minus-dark spectra arising from the broadening and loss of signal intensity due to minor heating effects of the laser. S2 shows weak positive polarization of the 5' and 5'' protons, while S1 shows weak negative polarization of these protons. Upon raising the pH to 9-10 by the addition of NaOD, S2 shows positive polarization of H1', H5' and
Table 4.1  Summary of Spin Polarization of Mononucleotides Upon Reaction with Photo-excited Riboflavin$^a$

<table>
<thead>
<tr>
<th>Monomer</th>
<th>H2</th>
<th>H8</th>
<th>H5</th>
<th>H6</th>
<th>H1'</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uracil</td>
<td></td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>n</td>
</tr>
<tr>
<td>Uridine</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td></td>
<td>n</td>
</tr>
<tr>
<td>3'-UMP</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td></td>
<td>n</td>
</tr>
<tr>
<td>5'-UMP</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td></td>
<td>n</td>
</tr>
<tr>
<td>Thymine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+(_Me)</td>
<td>n</td>
</tr>
<tr>
<td>3'-dTMP</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td></td>
<td>n</td>
</tr>
<tr>
<td>5'-dTMP</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td></td>
<td>n</td>
</tr>
<tr>
<td>Cytosine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-</td>
<td>n</td>
</tr>
<tr>
<td>3'-CMP</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td></td>
<td>n</td>
</tr>
<tr>
<td>5'-CMP</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td></td>
<td>n</td>
</tr>
<tr>
<td>5'-methylcytosine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+(_Me)</td>
<td>n</td>
</tr>
<tr>
<td>Adenosine$^b$</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td></td>
<td>n</td>
</tr>
<tr>
<td>3'-AMP$^b$</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td></td>
<td>n</td>
</tr>
<tr>
<td>5'-AMP$^b$</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td></td>
<td>n</td>
</tr>
<tr>
<td>5'-dAMP$^b$</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td></td>
<td>n</td>
</tr>
<tr>
<td>Guanosine$^c$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>w+</td>
</tr>
<tr>
<td>3'-GMP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>n</td>
</tr>
<tr>
<td>5'-GMP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>w+</td>
</tr>
<tr>
<td>9-methylguanine</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>w(_Me)</td>
</tr>
<tr>
<td>Inosine$^d$</td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>n</td>
</tr>
<tr>
<td>3'-IMP$^d$</td>
<td></td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td>n</td>
</tr>
</tbody>
</table>

$^a$ Reaction with photo-excited riboflavin.

$^b$ Data from Ref. 1.

$^c$ Data from Ref. 2.

$^d$ Data from Ref. 3.
Table 4.1 continued.

<table>
<thead>
<tr>
<th>Monomer</th>
<th>H2</th>
<th>H8</th>
<th>H5</th>
<th>H6</th>
<th>H1'</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-methylguanosine</td>
<td>-</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N^2-dimethylguanosine</td>
<td>+</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td>-(Me)_e</td>
</tr>
<tr>
<td>Dihyouridine</td>
<td>n</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2'-O-methylguanosine</td>
<td>+</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td>n(Me)</td>
</tr>
<tr>
<td>Pseudouridine</td>
<td>w_+</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7-methylguanosine</td>
<td>n_f</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td>+(Me)</td>
</tr>
<tr>
<td>1-methylinosine</td>
<td>n</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-aminopurine</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3'5' cyclic-8-methyladenine</td>
<td>+</td>
<td>-(Me)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>purine</td>
<td>n</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9-methylguanine</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>n(Me)</td>
</tr>
</tbody>
</table>

Symbols used in the table are: +, positive polarization; -, negative polarization; w, weak polarization; n, no observed polarization; Me, polarization of methyl protons.

The intensity and sign of the polarization are dependent on temperature, buffer concentration, and on the presence of ionized phosphate groups in the nucleotide (Scheek, et al., 1981).
Table 4.1 continued.

c Polarization is greatest at pH 7-8 and decreases when the pH is either lower or higher than this range. The polarization is temperature dependent at pH 7 (see text).

d The polarization decreases as buffer phosphate concentration increases; however, no change in the sign of the polarization is observed.

e The solubility of these molecules is very temperature dependent (see text).

f This proton exchanges rapidly with \(^2\text{H}\).
H5''. This polarization disappears when the pH is subsequently lowered to pH 5. S1 also shows increasing positive polarization of H1', H5' and H5'' as the pH is increased to 10. S2 decomposes readily in aqueous solution to a compound which shows only weak positive polarization of the 5' and 5'' protons at any pH. Deoxyribo-N1-isobutrylguanosine-5'-cyanoethylphosphate shows negative polarization of H8 only. This reversal of sign is probably due to an increase in the magnitude of the g-factor caused by the substituent at N1 (see discussion).

Solutions of 8BrF (about 0.3mM) and either I, m1I, G, m1G, or A (2.5mM) in 66mM phosphate buffer, pH 7 were examined. The enhancement factors for H8 are: G (+18), m1G (+8), I (+2), m1I (0), A (+5), and A with no phosphate (+6). By contrast the enhancements using riboflavin as the dye are: G (+9), m1G (-3), I (-8), m1I (0), A (-7). As the extinction coefficient, intersystem crossing yield, and reactivities of the two flavins may well be different, the absolute magnitudes of these enhancements are not particularly meaningful; however, it should be noted that in three cases the sign of the polarization changes.

B. Dinucleotides and Trinucleotide

Solutions of the following were studied: CpG, GpA, ApG, GpC, GpU, UpG, GpG, d(CpG), d(TpG), d(pCpG), d(pGpG), d(pGpA), CpG-d8, IpI, ApC, CpA, UpA, ApA, d(pApT), d(CpA), d(ApT), G(2'-5')A, A(2'-5')G, and ApGpU. Solutions of 5'-phosphate containing molecules and d(CpG) were pH 5; all other solutions were pH 7. Solutions were 33 or 66mM phosphate buffer; the phosphate
concentration did not affect the relative $G(H_8)$ to $G(H_{1'})$ polarization. Nucleic acid concentrations were 1-2.5mM. The results for these molecules are summarized in Table 4.2.

$A(H_8)$ and $A(H_2)$ show polarization unless the molecule contains $G$; in molecules containing both $G$ and $A$, no polarization is observed for the $A$ protons. This was true for $ApG$ at either 0 or 66mM phosphate concentration. $A$ shows very weak polarization in a pH 7 solution (66mM phosphate buffer) of $UpG$ (2.5mM), $A$ (2.5mM), and riboflavin (0.2mM), whereas $A$ alone (no $UpG$) shows strong polarization under the same conditions. The $A(H_8)$ enhancement factor is $-0.3$ in the presence of $UpG$ and $-6.5$ in the absence of $UpG$. $I(H_8)$ and $I(H_2)$ both show negative polarization.

$G(H_8)$ shows positive polarization in all the dinucleotides. The polarization is much weaker in $GpG$ and $d(GpG)$ but this is probably due to aggregation. Except for these two dinucleotides, the enhancement factors of $G(H_8)$ are in the range 11-16 for all $G$-containing dinucleotides. The $G(H_8)$ polarization of $GpG$ increases with increasing temperature ($20^\circ-50^\circ C$), but the enhancement factor is always less than 0.5.

In sharp contrast to the monomers, positive polarization is observed for the $G(H_{1'})$ sugar proton in all the $G$-containing dinucleotides (except $GpG$ and $d(GpG)$), in the trimer $ApGpU$ (see Figure 4.7) and in the tetramer $ApGpCpU$ (see Figure 4.8). The $A(H_{1'})$ or $I(H_{1'})$ protons do not show polarization in the molecules we have studied. Assignment of the positively polarized resonances as $G(H_{1'})$ follows Lee, et al., (1976) and Ezra, et al., (1977) for
Summary of Spin Polarization of Dinucleotides and Oligonucleotides Upon Reaction With Photo-excited Riboflavin $^a,b$

<table>
<thead>
<tr>
<th></th>
<th>A(H2)</th>
<th>A(H8)</th>
<th>G(H8)</th>
<th>G(H1')</th>
</tr>
</thead>
<tbody>
<tr>
<td>CpG</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GpA</td>
<td>n</td>
<td>n</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ApG</td>
<td>n</td>
<td>n</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>GpC</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GpU</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UpG</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GpG$^c$</td>
<td>+</td>
<td></td>
<td>n</td>
<td></td>
</tr>
<tr>
<td>d(CpG)</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d(TpG)</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d(pCpG)</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d(pGpG)$^c$</td>
<td>+</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>d(pGpA)</td>
<td>n</td>
<td>n</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>IpI</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ApCd</td>
<td>+/-</td>
<td></td>
<td>+/-</td>
<td></td>
</tr>
<tr>
<td>CpAd</td>
<td>+/-</td>
<td></td>
<td>+/-</td>
<td></td>
</tr>
<tr>
<td>UpAd</td>
<td>+/-</td>
<td></td>
<td>+/-</td>
<td></td>
</tr>
<tr>
<td>ApAd</td>
<td>+/-</td>
<td></td>
<td>+/-</td>
<td></td>
</tr>
<tr>
<td>d(pApT)$^d$</td>
<td>+/-</td>
<td></td>
<td>+/-</td>
<td></td>
</tr>
<tr>
<td>d(CpA)$^d$</td>
<td>+/-</td>
<td></td>
<td>+/-</td>
<td></td>
</tr>
<tr>
<td>d(ApT)$^d$</td>
<td>+/-</td>
<td></td>
<td>+/-</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.2 continued.

<table>
<thead>
<tr>
<th></th>
<th>A(H2)</th>
<th>A(H8)</th>
<th>G(H8)</th>
<th>G(H1')</th>
</tr>
</thead>
<tbody>
<tr>
<td>G(2'–5')A</td>
<td>n</td>
<td>n</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>A(2'–5')G</td>
<td>n</td>
<td>n</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ApGpU</td>
<td>n</td>
<td>n</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ApGpCpU</td>
<td>n</td>
<td>n</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

---

a Symbols used in the tables are: +, positive polarization; -, negative polarization; w, weak polarization; n, no observed polarization.

b The d(CpG) and samples with nucleotides containing 5'-phosphates were observed at pH 5; all other samples were measured at pH 7.

c Polarization of the H8 proton in these dinucleotides is weaker than in other G-containing nucleotides (see text).

d The intensity and sign of the polarization are dependent on temperature, buffer concentration, and on the presence of ionized phosphate groups in the nucleotide (Scheek, et al., 1981).
the RNA dinucleotides and Cheng & Sarma, (1977) for the DNA dinucleotides. The magnitude of the sugar polarization in GpC or GpU does not change noticeably when G(H8) is deuterated or saturated with rf during the light pulse. The resonances from the sugar protons of UpG and the H5 from the U overlap at room temperature, but are well resolved at higher temperatures. Only the H1' resonance which is at slightly higher field shows polarization. The polarized resonance also has a larger $J_{1',2'}$ coupling constant, consistent with its assignment to the G residue (Ezra, et al., 1977). In d(pCpG) the lowfield H1' triplet is polarized, whereas in d(pGpA) the highfield H1' triplet is polarized. The assignment of these polarized resonances to the G(H1)''s is consistent with previous assignments for d(CpG) and d(GpA) (Cheng & Sarma, 1977). It is clear that CIDNP offers a very direct method for assigning the G(H1') NMR resonances.

The magnitude of the G(H1') polarization relative to the G(H8) polarization increases with decreasing temperature for GpA, GpU, CpG, dpGpA, ApGpU, and ApGpCpU. The ratio of the G(H1') polarization to the G(H8) polarization also increases in the following series: $0 = \text{mononucleosides} < \text{dinucleotides} < \text{ApGpU} < \text{ApGpCpU}$. For the deoxyribo-dinucleotides only the G(H1') to G(H8) polarization ratio for d(GpA) was measured, and the ratio for this dinucleotide is similar to the ratio for ApGpU. The ratio of the polarizations for G(2'-5')A and A(2'-5')G are similar to or greater than those of the corresponding 3'5' dinucleotides. The $T_1$ values of G(H8) and G(H1') in GpA at 10, 20, 30, and 40°C are 1.2, 2.1; 1.1, 1.8; 1.3,
It is assumed that the $T_1$ variation as a function of temperature is similar for the other nucleotides studied. The $T_1$ values of $G(H8)$ and $G(H1')$ in ApGpU are 1.1 s and 1.6 s at 20°C.

C. Tetranucleotide

Figure 4.8 shows a series of CIDNP spectra for the self-complementary oligonucleotide ApGpCpU at different temperatures. The CIDNP intensity of $G(H8)$ decreases dramatically below 20°C, and essentially no CIDNP is observed below 10°C. The melting temperature, $T_m$, of ApGpCpU under these conditions (1.3mM ApGpCpU, 5mM phosphate, 0.2M NaCl, 0.4mM riboflavin, pH 7) as determined optically is about 17-18°C and as determined by the average inflection of the proton chemical shift vs. temperature curves is $\sim 25°C$. Note that the $T_m$ value for this molecule is strongly dependent on buffer and nucleic acid concentration (compare, for example, Neilson, et al., 1979 and Bubienko, et al., 1981). The chemical shifts of this tetramer are unaffected by the presence of riboflavin at any temperature. The riboflavin chemical shifts and their temperature dependence were different in the presence of ApGpCpU, indicating some effect of the tetramer on the dye. The $T_1$ values of $G(H8)$ and $G(H1')$ are approximately 1.1 and 1.5 s at 30°C and 1.0 and 1.6 s at 15°C in our samples.

D. tRNA$^{phe}$

The secondary structure, as determined by X-ray crystallography, of yeast tRNA$^{phe}$ is shown in Figure 4.9. Two positively polarized peaks are observed in the aromatic region of the CIDNP
Figure 4.8. The 360 MHz $^1$H NMR spectra of ApGpCpU (1.3 mM) in 5 mM phosphate buffer, pH 7, 0.2 M NaCl, 0.4 mM riboflavin. The G(H8) protons were partially deuterated in this sample. The spectra are A) the light-minus-dark difference spectra taken at 5, 10, 15, 20, 25, and 30°C (4 accumulations each) and B) the dark spectrum (100 accumulations).
**LIGHT - DARK**

**A**

- 30°C
- 25°C
- 20°C
- 15°C
- 10°C
- 5°C

**Dark Spectrum, 25°C**

**B**

- U(H8)
- G(H8)
- A(H8)
- U(H1')
- U(H5)
- A(H1')
- C(H5)
- C(H1')
- A(H2)
- C(H6)

**PPM**

- 8.5
- 8.0
- 7.5
- 7.0
- 6.5
- 6.0
- 5.5
Figure 4.9. The secondary structure of yeast tRNA$^{\text{phe}}$. 
spectra of tRNA\textsuperscript{phe} below 60°C. The spectrum at 20°C is shown in Figure 4.10. A negatively polarized peak is also observed in the aromatic region at temperatures below 50°C. Addition of 0.15M phosphate or use of the dye 8BrF did not change the sign or the pattern of the polarization. The negatively polarized peak at 7.7 ppm disappeared when the intense negatively polarized resonance at 2.5 ppm was irradiated with rf during the light pulse. There was no change in the CIDNP pattern or intensity when any of the other polarized resonances were irradiated with rf during the light pulse. As the temperature was increased from 60° to 85°C, the number and intensity of polarized peaks increased dramatically (see Figure 4.11). The negatively polarized peak at 7.7 ppm was not observed at temperatures greater than or equal to 50°C. The chemical shifts and relative intensities of the polarized peaks as a function of temperature did not change substantially when the Y base was excised. However, the linewidth of the major polarized signal was somewhat greater for the tRNA\textsuperscript{phe}-Y between 40° and 60°C than the linewidth of the corresponding polarized signal for tRNA\textsuperscript{phe}.

Two negatively polarized resonances are present in the aliphatic region of the tRNA\textsuperscript{Phe} CIDNP spectrum shown in Figure 4.10. Both negative peaks are also observed upon addition of 0.15M phosphate, when 8BrF was used rather than the riboflavin, and in tRNA\textsuperscript{phe}-Y. It should be noted, however, that the relative intensities of the two peaks vary somewhat from sample to sample and as a function of temperature, phosphate concentration and the dye used. The polarized peak at 2.0 ppm is not observed above 50°C.
Figure 4.10. The 360 MHz $^1$H NMR spectra of yeast tRNA$^{\text{phe}}$ (1.5mM) in 5mM phosphate buffer, pH 7, 0.15M NaCl, 0.4mM 8BrF at 20°C. The R indicates resonances from the dye and the i indicates an impurity peak. The spectra are A) the light-minus-dark spectrum (1 accumulation) and B) the dark spectrum (500 accumulations).
Figure 4.11. The 360 MHz $^1$H NMR light-minus-dark spectra (1 accumulation each) of yeast tRNA$^{\text{phe}}$ (~1.5mM) in 5mM phosphate buffer, pH 7, 0.15M NaCl, 0.4mM 8BrF at 20, 50, 60, 70, and 80°C.
Above 60°C an additional negatively polarized resonance is observed (3.3 ppm at 60°C). The methyl and methylene resonances of tRNA^{phe} are assigned based on the chemical shifts, integrals, and splittings of the signals. The assignments agree with those of Davanloo, et al., (1979). Minor differences in the chemical shifts between our spectra and theirs are most likely due to differences in tRNA^{phe} concentration, buffer composition, and magnesium concentration. The Y base assignments are based on the absence of these resonances in the NMR spectrum of tRNA^{phe-Y} (see Figure 4.12c), which is otherwise very similar to the spectrum of tRNA^{Phe} at all temperatures (Davanloo, et al., 1979).

The polarization of aliphatic resonances in the tRNA^{Phe-Y} sample decreases at high temperature even though polarization of the aromatic resonances remains strong. There are several unassigned peaks between 2.9 and 3.3 ppm in the tRNA^{phe-Y} dark NMR spectrum at 85°C. These could be due to sample heterogeneity, chemical degradation of the molecule, or photochemical degradation of selected bases (m^2G26 and m^1G10). The latter explanation is very likely because this sample had been used repeatedly in CIDNP experiments. Similar peaks of much smaller intensity are seen in the spectrum of tRNA^{Phe} at 85°C (this sample had been subjected to far less laser irradiation). Degradation of other molecules (for example, ApGpCpU) has been observed after exposure to many light pulses. It has been reported that chemical degradation of tRNA^{Phe} occurs in the presence of magnesium at high temperatures. The samples used here contain ~5mM Mg^{2+}, so chemical degradation is also a possibility.
Figure 4.12. The 360 MHz $^1$H NMR spectra at 85°C. A) the light-minus-dark difference spectrum (1 accumulation) of yeast tRNA$^{\text{phe}}$ (~1.5mM) in 5mM phosphate buffer, pH 7, 0.15M NaCl, 0.4mM 8BrF. The s indicates spikes arising from the rf used to saturate the solvent resonance. These have random phase and thus are not present in spectra consisting of the sum of many accumulations. B) the dark spectrum (252 accumulations) of yeast tRNA$^{\text{Ph}}$e taken under the same conditions as spectrum A above. C) The dark spectrum of yeast tRNA$^{\text{phe}-Y}$ (252 accumulations) taken under the same conditions as spectrum A. The x indicates resonances that may be due to chemical degradation (see text).
Experiments were done from low to high temperature. If the temperature was subsequently decreased, the pattern of the polarization matched that taken previously but the intensity was greatly reduced. For this reason, samples that were used at temperatures greater than 40°C were not used in subsequent experiments at lower temperatures.

The NMR spectrum of tRNA^Phe is unaffected by the addition of dye, although the chemical shifts and linewidths of some of the flavin resonances are somewhat different in the presence of the tRNA^Phe.

5. DISCUSSION

A. Mechanism

As in previous studies of the mechanism of the reaction between the flavin triplet state and amino acids (Kaptein, 1980; McCord, et al., 1981), the first step is to determine if the radical formation occurs by electron transfer or hydrogen-atom abstraction. The reaction is clearly the latter for the pyrimidines, as no polarization is detected when N1 is substituted. It should be noted that the results presented here for Cyt and Ura differ from those obtained by Kaptein and co-workers. They report qualitatively that polarization is not observed for Cyt and Ura (Kaptein, et al., 1979), whereas Figure 4.5 shows a strong polarization for Cyt and Ura. Hydrogen-atom abstraction is also consistent with the observed polarization of pseudouridine, which retains the H1. When the free radical is formed, the electron is delocalized throughout the aromatic ring system. This delocalization can be described by an
electron density or spin density at each nucleus in the system. Adjacent nuclei in the system will have spin densities with opposite signs. For Thy, Cyt and Ura, the radical produced after hydrogen-atom abstraction can be considered as an allylic system with a large positive spin density predicted and observed by EPR at C5 of Cyt and at the 5-methyl protons of Thy and a negative spin density at H5 of Ura (Herak & Dulcic, 1972; Zehner, et al., 1976; Wethof & Van Rooten, 1976). This agrees with the positive polarization observed for Thy 5-methyl protons and the negative polarization observed for Ura and Cyt 5-protons.

The purines, adenine, guanine, and hypoxanthine show polarization regardless of the presence or absence of the ribose moiety. It is possible that either the imino or the amino protons are involved in a hydrogen-atom abstraction reaction in the purines. The polarization of m\textsubscript{2}G rules out the involvement of the amino protons, assuming the presence of the methyl groups does not alter the mechanism of the reaction. The m\textsubscript{1}G shows negative polarization of G(H8). This may be due to a change in the reaction mechanism, but it is most likely due to an increase in the g-factor of the radical (see discussion below). Thus it is reasonable to suggest that G resonances become polarized through an electron transfer mechanism. The electron transfer mechanism for A is supported by Kaptein's studies on the effect of phosphate concentration on the A polarization (Scheek, et al., 1981). In the work presented here there is no observable effect of phosphate concentration on the sign of the polarization of G(H8) in ApG or in G. The reaction for I is
most likely hydrogen-atom abstraction from N1, as m\(^1\)I does not show polarization with either 8BrF or riboflavin as the dye.

The g-factors of many of the radicals involved in these reactions have been reported: C\(^+\), 2.0055 (Dertinger & Hartig, 1972) or 2.0042 (Sevilla & Mohan, 1974); A\(^+\), 2.0040 (Dertinger & Hartig, 1972) or 2.0043 (Sevilla & Mohan, 1974); Thy\(^-\), 2.0053 (Herak & Dulcic, 1972) or 2.0044 (Zehner, et al., 1976), and the flavin anion radical, 2.0034 (Eriksson & Ehrenberg, 1964). The pyrimidine radicals in these studies are believed to be the result of hydrogen-atom abstraction from N1. However, definitive identification of the nature of the radical is rarely available. The radicals are produced by radiolysis or photolysis in the solid state, the spectra are often poorly resolved, and analyses often depend on calculated spin densities, which may be grossly in error (McCord, et al., 1981).

The CIDNP results presented here are not in agreement with the previously measured g-factors for purine radicals which were given above. This is based on the following analysis. Start with the assumption that the g-factor of 8BrF is higher than that of the riboflavin, due to increased spin-orbit coupling, and that the two flavins react with any given base by the same mechanism. The g-factors of the neutral A radical, the I radical and the m\(^1\)G radical must then be in-between the g-factors of the two flavin radical anions, as the polarization of these three molecules reverses sign upon switching dyes. Using the rule which predicts the sign of the polarization (Kaptein, 1971), one can conclude that the spin density
at C8 is positive in the A and G radical cations. This is consistent with most molecular orbital calculations for these radicals (Adams & Box, 1975). From this analysis there should be a positive spin density at C8 in the I radical cation also, which must then give a g-factor less than that of the 8BrF radical anion. It is likely that I reacts by hydrogen-atom abstraction of Hl, as m1 does not show polarization with either dye. Thus our experimental results are consistent with the following ordering of the g-factors:

\[ \text{A}^+ = \text{G}^+ < \text{riboflavin anion} < \text{m}^1\text{G neutral radical} = \text{A neutral radical} = \text{I neutral radical} < \text{8BrF anion radical}. \]

It is worth noting that the use of several different dye molecules, especially those with quite different g-factors, offers several advantages. As the g-factors of the target residues (nucleotide) are roughly constant, the sign of the spin polarization can be inverted by using a different dye as long as the dye changes the sign of the g-factor difference. Furthermore, it is desirable, for resolution purposes, to perform the CIDNP experiments at the highest possible magnetic field; however the intensity of the CIDNP polarization decreases with increasing field. This can be offset by using a dye whose g-factor is as close as possible to that of the target residue.

B. Adenine Polarization

The polarization of A is not observed in molecules which also contain G and is greatly attenuated when G-containing molecules are also present in solution. For example, A is strongly polarized in dinucleotides which do not contain G (see Table 4.2). This effect
is probably due to kinetic competition between G and A for the flavin triplet. When the flavin concentration is in excess of the combined concentrations of A and G both nucleosides do show polarization. This notion receives some support from EPR studies of mixed nucleotides in aqueous glasses, from which it has been determined that the electron donating ability of G is greater than that of A (Gregoli, et al., 1979). The A polarization arises from two different radicals with opposite signs for the polarization (Scheek, et al., 1981), thus, it is possible for the polarization of A to be zero because the polarization from each pathway exactly cancels. This effect for A itself is due to deprotonation of the A radical cation and is catalyzed by phosphate ions (Scheek, et al., 1981). In the studies presented here, no CIDNP is observed for A in GpA (2.5 mM, 0.4 mM riboflavin) where the phosphate concentration is varied from 5 to 750 mM. It is not very likely that in all these samples where G is present a fortuitous cancellation of A polarization results in a lack of observable polarization for A. Therefore, I favor the simple kinetic argument. This result has important consequences for the detection of CIDNP in more complex oligonucleotides. Polarization of A will only be observed in the absence of accessible G residues, or when the accessibilities of the G residues become sufficiently low to allow an A residue to compete successfully for the photo-excited dye.

C. Sugar Polarization

i. Mechanisms. Polarization of the G(H1') proton in G-containing dinucleotides and oligonucleotides can result from two
possible mechanisms: 1) a hyperfine interaction between the H1' nuclear spin and the unpaired electron in the radical intermediate or 2) cross relaxation between another spin polarized transition and H1' in the diamagnetic product. The spin must be from a proton because cross relaxation is only effective between nuclei with similar magnetic moments. The only proton that shows polarization is the HB. Because the polarization is positive for both H1' and H8, dipolar cross-relaxation is not likely, as it has been shown that dipolar cross-relaxation leads to polarization of the opposite sign for the interacting protons in small molecules (Kaptein, 1979). Consistent with this, no substantial change has been observed in the magnitude of the H1' polarization upon rf saturation of the G(H8) or substitution of the G(H8) with a deuteron in dinucleotides such as CpG or UpG. Note also that rf saturation of H8 during the light pulse or deuteration of H8 does not result in the appearance of sugar polarization in the mononucleosides 9MeGua, 5'-GMP, or G. Thus the negative polarization due to cross-relaxation is not canceling some other source of positive polarization for H1' in these molecules.

In the following paragraphs five possible mechanisms which could lead to spin density at H1' and might provide an explanation for the variation in the magnitude of the H1' polarization will be discussed.

(1) Interaction with N9.

Spin density in an aromatic ring can lead to spin density at a position by hyperconjugation, σ-bond polarization, and/or an
anisotropic dipolar interaction (Gordy, 1980). For aliphatic radicals in solution in which there is unit spin density at the α-carbon, the latter two mechanisms lead to little or no spin density at the β-hydrogen in comparison to the spin density due to hyperconjugation. Spin density due to hyperconjugation between the α-carbon orbital and the Cβ-H σ-bond has the angular dependence:

\[ A = (B_0 + B_2 \cos^2 \theta) \rho_\alpha, \]

where \( A \) is the hyperfine coupling constant, \( B_0 \) and \( B_2 \) are experimentally determined constants on the order of 0-5 and 40-45 gauss, respectively, (Fischer, 1973), \( \rho_\alpha \) is the spin density at the α position, and \( \theta \) is the dihedral angle between the α-orbital and the Cβ-H bond. In the present case, the dihedral angle, \( \theta \) (the angle between the H1'-Cl' bond and the N9-π orbital), is related to the glycosidic torsion angle, \( \chi \) (the angle between the O4'-Cl' bond and the N9-C4 bond; see Figure 4.3), by the following: \( \theta = \chi - 210^\circ \). The glycosidic torsion angle decreases when a nucleoside is incorporated into a dinucleotide (Lee, et al., 1976; Bangerter & Chan, 1969), increasing the value of \( \cos^2 \theta \) and the hyperconjugative spin density on H1'.

Contrary to this analysis, however, the 9-methyl protons of 9MeGua show barely detectable positive polarization. This is not due to an overall lack of reactivity for this derivative, as the H8 polarization is about twice that of the other mononucleosides (this is due, in part, to the substantially longer \( T_1 \) of 9MeGua(H8)). The average hyperfine coupling constant for a freely rotating methyl
Thus, one predicts that the hyperfine coupling constant and the polarization of the methyl protons would be comparable to, or greater than, that of a single proton at any fixed angle. Note that the ribose-containing monomers also do not exhibit HI' polarization, so the electronegative oxygen adjacent to Cl' does not appear to have an important effect on the polarization.

(2) Opposing Effects.

It is possible that there is one mechanism, such as σ-bond polarization, which leads to negative spin density on the HI' and another mechanism, such as hyperconjugation, which leads to positive spin density at the HI'. For mononucleosides, with more freely rotating HI' and in 9MeGua, the hyperconjugative spin density might cancel the σ-bond polarization spin density. For dinucleotides, however, where the angle θ might be very close to 90°, there would be little spin density via hyperconjugation, and the spin density due to σ-bond polarization would dominate. This mechanism is not likely as it requires two mechanisms of opposite sign to cancel in a variety of different molecules.

(3) Interaction with Spin Density at N3.

Baudet, et al., (1962) have calculated the spin density in the guanine cation radical and concluded that the spin density at the N9 position is zero, while that at the N3 position is 0.324, much larger than that at any other position in the radical. In order for this mechanism to be selective for the dinucleotides the N3 must be closer, on the average, to the HI' in dinucleotides than it is in G or than it is to the methyls in 9MeGua. Obviously the closest
possible approach of these protons to N3 in all the compounds is the same; however, the average distance may be quite different. For 9MeGua and similarly for G there is a possibility of free rotation. It has been suggested that the preferred conformation for G-containing mononucleosides and dinucleotides is syn (Tran-Dinh & Chachaty, 1977; Chachaty, et al., 1977; Neuman, et al., 1979) which places the H1' almost as far from N3 as possible. The likelihood of this mechanism being correct depends, then, on whether the H1' in the dinucleotides and higher oligonucleotides spends more time close to N3 than in a model mononucleoside.

(4) Dual or Sequential Radical Pairs.

It is conceivable that the radical intermediate for the mononucleoside is different from that in the dinucleotides, and that these radicals have very different spin densities at N9. For example, the radical cation might be stabilized in a stacked dinucleotide, whereas the mononucleoside could tend to deprotonate or react via hydrogen-atom abstraction.

(5) Sugar-centered Radical.

This possibility is highly unlikely for several reasons. (i) Only G sugar protons show polarization; (ii) G in deoxyribo-
dinucleotides also shows H1' polarization and it has no 2'-hydroxyl; (iii) polarization is never observed or expected for primary and secondary alcohols reacting with the triplet state of riboflavin, and ribose itself shows no polarization.

ii. The Verdict. Considering all of the above mechanisms the most likely is mechanism (3), which suggests that the spin density
at H1' in dinucleotides and higher oligonucleotides is a consequence of special geometrical constraints. Nucleic acid stacking interactions are thought to increase as the length of the oligonucleotide increases and also as the temperature is lowered. In a stacked structure, the anti conformation of the glycosidic angle, χ, is preferred, which brings H1' close to N3, whereas the mononucleosides appear to prefer a syn conformation (note: this discussion is only for G-containing mononucleosides, dinucleotides and oligonucleotides). The ratio of the enhancement of the G(H1') to that of the G(H8) increases as the length of the oligonucleotide increases and as the temperature is lowered (for example, see Figure 4.8). The difference in the temperature dependence of the T1's for G(H8) and G(H1') in ApG is probably not sufficient to lead to the change in the G(H8) to G(H1') polarization ratios as function of temperature. Neither is the difference in the T1's of 9MeGua, G, ApG, and ApGpU at 20°C sufficient to explain the variation in the amount of sugar polarization among these molecules. Obviously, a change in either the G(H8) or the G(H1') polarization intensity affects the ratio between the two. However, as there is no obvious change in the electronic structure between the G's in any of these molecules, I will assume that it is the H1' spin density that is changing and the G(H8) can act as an internal standard to correct for changes due to temperature, for example. It has been suggested from T1 measurements that deoxyribo-dinucleotides are more stacked than the corresponding ribo-dinucleotides (Ts'o, et al., 1975);
consistent with these expectations, d(GpA) has a higher G(H1') to G(H8) polarization ratio than does GpA.

The 2'5'-dinucleotides however, introduce a complication. G(2'-5')A, A(2'-5')G, and G(3'-5')A all have very similar G(H8) enhancement factors and a similar G(H1') to G(H8) polarization ratio, however, 2'5'-dinucleotides are thought to have somewhat greater base-base interactions than their 3'5' analogs. This idea is based on NMR chemical shift data and hypochromicity measurements (Kondo, et al., 1970). CD data demonstrate that G(2'-5')A is stacked to an appreciable extent at room temperature, but that the stacking geometry differs from that of the 3'5'-isomer (Warshaw & Cantor, 1970). In 2'5'-dinucleotides the glycosidic bond angle, χ, of the 5' residue is thought to be substantially different (~130°) from the 3'5' analog, based on calculations of the chemical shifts of A(2'-5')A, A(2'-5')C, and C(2'-5')A (Dhingra & Sarma, 1978). The 2' residue is thought to have a χ~240°. When χ = 240°, the H1' is at its closest approach to N3. According to mechanism (3), we would expect the G-polarization from G(2'-5')A to differ from that of A(2'-5')G. However, experimentally they are very similar. This could be an argument against mechanism (3). The results of Dhingra & Sarma are based on chemical shift calculations and are much less reliable than those based on direct coupling constant measurements, so it is quite possible that their analysis is in error or perhaps not applicable to all 2'5'-dinucleotides.

In order to provide further support for mechanism (3), CIDNP experiments have been carried out on a number of molecules in which
the glycosidic torsion angle, $\chi$, is constrained. The glycosidic torsion angle in the model compound V1 is fixed in the syn conformation ($\chi = 25^\circ$, as estimated from molecular models, $H_1'$ is near to $H_8$). $H_1'$ shows negative polarization in V1 and no polarization if $H_8$ is deuterated or saturated with rf during the light pulse. Thus, the negative polarization of $H_1'$ in this molecule is clearly the result of dipolar cross-relaxation from $H_8$. Because no polarization is observed when the source of the dipolar cross-relaxation is removed, this source of negative polarization must not be competing with a source of positive polarization (spin density at the $H_1'$ through interaction with the N9, N3, or $\sigma$-bond polarization) in this molecule. In compound S3, $\chi = 275-305^\circ$; this molecule retains considerable rotational flexibility about $\chi$. In this compound $H_1'$ shows weak negative polarization. In light of the strong $H_2'$ polarization, the $H_1'$ polarization is probably due to spin density at C2, transmitted to $H_1'$ through a $\sigma$-bond polarization mechanism. The strong polarization of $H_2'$ indicates that the radical of this molecule may well have a different spin density distribution from that of the G radical, and thus is not a good model for our purposes. Compounds S1 and S2 have $H_1'$ near to N3 ($\chi = 180-210^\circ$). These compounds show negligible polarization at pH 7, however, at higher pH they exhibit strong positive polarization of several protons, including positive polarization of $H_1'$. These are the only mononucleosides which show appreciable positive polarization of $H_1'$. Observation of positive polarization for the $H_5'$ and $H_5''$ in S1 and S2 is consistent with positive spin density at C8 and
suggests that the radicals of these molecules are similar to the G radical. To investigate the effect on the spin density of an amino group at the C8 position, the polarization in adenine and 8-aminomethyladenine has been studied (8-aminomethylguanosine was not readily available). The presence of the amino group at the H8 position does not change the sign of the spin density. Observation of H1' polarization in S1 and S2 supports mechanism (3) discussed above: the sugar polarization is due to the proximity of the H1' to N3.

D. Tetranucleotide

The reaction leading to the production of CIDNP requires accessibility of the base to the photo-excited flavin. Therefore, it is likely that CIDNP will only be observed in single-stranded oligonucleotides. The temperature dependence of the G(H8) enhancement factor of ApGpCpU (see Figure 4.8) as compared to that of G, GpA, or ApGpU provides a good example of this effect. The Tm of ApGpCpU under the conditions of the CIDNP experiment is about 20°C. A substantial drop in the CIDNP intensity is observed in this molecule as the temperature is decreased below the Tm; the polarization at 5°C is less than 15% of that at 20°C. This is in contrast to G, GpA, and ApGpU; in these compounds the CIDNP intensity of G(H8) at 5°C ranges from 55-85% of that at 20°C. This reduction for G, GpA, and ApGpU is probably the result of changes in T1's, diffusion rates, radical lifetimes, exchange rates, aggregation, etc. However, it is clear that the intensity decrease in the CIDNP of G(H8) in ApGpCpU as a function of temperature is much larger than
that of G, GpA, and ApGpU. This is probably due to a decrease in the accessibility of the bases in the double-stranded molecules. The temperature dependence of the ratio of the polarization of \textit{H}1' to that of \textit{H}8 is especially obvious for ApGpCpU as shown in Figure 4.8. Now that the mechanisms that are involved in the CIDNP of mononucleotides, dinucleotides, and several oligonucleotides have been explored, it is time to move on to bigger molecules such as tRNA.

E. \textit{tRNA^{Phe}}

Based on the previous discussion, I will assume that only the accessible G bases will show polarization, for example, in the single-stranded loop regions of the tRNA molecule. These regions of the tRNA are thought to be involved in critical contacts with other components in protein synthesis. The NMR spectrum of tRNA, especially the aromatic region, is very complex and poorly understood in contrast to the wealth of data available for the base-pairing imino protons (Hurd & Reid, 1979; Johnston & Redfield, 1981). Thus, the CIDNP method offers valuable information on this molecule.

Examination of a model based on the crystal structure of yeast \textit{tRNA^{Phe}} (see also Figure 4.9) suggests that three G's (G57, G20, and Gm34) are in loop regions and are not involved in tertiary inter-loop hydrogen bonding interactions (Holbrook, et al., 1978). G57 is in the interior of the TYC loop, whereas G20 and Gm34 are located on the outside of the molecule and are readily accessible to the
solvent. Chemical modification studies using N-cyclohexyl-N'-8(4-methyl-morpholinium)ethylcarbodiimide-p-toluenesulfonate at 37°C (Rhodes, 1975) and ketoxal at 25°C (Litt, 1971) have indicated that G20 and Gm34 are accessible (G18 is slightly accessible to the carbodiimide reagent). Chemical modification studies, using the reagent dimethyl sulfate at 37°C to attack at the G(N7) position, indicate that G18, G19, G30, G45, G71, and Gm34 are fully accessible, while G1 and G65 are partially accessible (Peattie & Gilbert, 1980). Measurements of the rate of tritium exchange into the HB positions of the purines at 37°C indicate that Gm34 and G18 are probably the most solvent accessible G's (the accessibility of G20 could not be determined in these experiments) (Gamble, et al., 1976). Chemical modification using the carbodiimide reagent has also been studied as a function of temperature up to 55°C (Rhodes, 1977). G4, G15, G18, G20, Gm34, G45, m7G46, and possibly m2G10, G22, and G65 were found to be accessible in the early stages of thermal unfolding. There was no evidence for modification of m2G26 at temperatures up to 55°C.

Because m2G26 is found to exhibit strong spin polarization, which does not seem to be consistent with the chemical modification studies, I will consider its chemistry in greater detail. m2G26 is thought to be hydrogen bonded to A44 through one or two hydrogen bonds with the two bases twisted in a propellor fashion relative to each other (Kim, et al., 1975). The dimethylamino group is oriented directly outward from the tRNA molecule into the solution. m2G26 has not been affected in the chemical modification studies, but the
appropriate atoms necessary for reaction may not be accessible to these reagents, and it is also possible that even a fully accessible \textit{m}_2\textit{G} would not react with these reagents. The kethoxal adduct involves dehydration with one of the amino hydrogens and may not be stable when these are not available. The carbodiimide reagent is thought to react at N1 and the dimethyl group on the \textit{m}_2\textit{G} could interfere sterically with this reaction.

By contrast, CIDNP is generated by reversible one electron transfer from G or its derivatives to the photo-excited flavin. This has the important advantage that the molecule being probed is not being changed chemically. \textit{m}_2\textit{G} was found to be nearly insoluble at 25°C; however, it does show a very weak negative polarization of the methyl protons. The solubility of this compound is very temperature dependent and above 40°C strong CIDNP was observed for H8 and for the methyl protons. At 40°C the ratio of the enhancements of the H8 to the methyl resonance is 1 to -2.5. Separately, \textit{m}_2\textit{G} and G have almost the same enhancement factor for the H8 proton when each is 0.3mM (0.25mM riboflavin, 66mM phosphate, pH 7, 40°C). In an equimolar mixture of these two compounds under the same conditions (except that the bases were each 1mM), the H8 again showed almost identical enhancements.

At 20°C and above it is evident from Figures 4.10a and 4.12a that the \textit{m}_2\textit{G}26 methyl protons show strong polarization. The CIDNP in the aromatic region indicates that at least two G's (presumably Gm34, G20, and/or \textit{m}_2\textit{G}26) have spin polarized H8 resonances and therefore react with photo-excited riboflavin at 20°C. The studies
of smaller oligonucleotides, discussed previously, show that A residues are not polarized in the presence of accessible G residues, thus it is not likely that the polarization in the aromatic region is due to A, in spite of the fact that several of the A residues are accessible. This is confirmed by the observation that the sign of the polarization in the aromatic region is unaffected by high phosphate concentrations, contrary to what has been shown for A polarization (Scheek, et al., 1981).

The negatively polarized resonance at 7.7 ppm is spin polarized through cross-relaxation from the $m^2G26$ methyl protons. The fact that the transferred polarization is of the same sign is expected given the long correlation time of the tRNA molecule (i.e. tRNA is a big molecule) (Kaptein & Edzes, 1979). This polarized aromatic resonance must be from a nonexchangeable proton in close proximity to the $m^2G26$ methyl protons. Examination of a model of the crystal structure (kindly provided by Dr. S.-H. Kim) suggests that the H6 of C27, the H8 of $m^2G10$, or the H2 of A44 are within 6 Å of the $m^2G26$ methyl groups. A nuclear Overhauser effect has been reported from the methyl group of $m^2G26$ to an aromatic proton at 8.4 ppm and was assigned to $m^2G10(H8)$ in a sample of yeast tRNA^{Phe} (0.1M NaCl, 10mM phosphate, pH 7, 95% H$_2$O, no added Mg$^{2+}$) (Sanchez, et al., 1980). Several other NOE's to the aromatic region were observed but were not assigned. Hare and Reid (personal communication) have observed NOE's at 32°C from the $m^2G26$ methyl group to three resonances due to nonexchangeable protons at 8.41, 7.94, and 7.71 ppm (10 mg yeast tRNA^{Phe} in 0.5ml of deuterated 10mM phosphate buffer, pH 7, Mg$^{2+}$
concentration ~5mM, DSS as an internal standard). The intensities of the NOE's increased from low to high field. Based on the work of Sanchez, et al. (1980), and on the distances of the protons from the m2G26 methyl groups as estimated from the model, I suggest that the 8.41 ppm resonance is m2G10(H8), the 7.94 ppm resonance is C27(H6), and the 7.71 ppm resonance is A44(H2). The strongest Overhauser enhanced resonance at 7.71 ppm corresponds nicely to the cross-polarized resonance we observe at 7.7 ppm in the CIDNP spectrum, suggesting that this is A44(H2). This cross-polarization is no longer observed at temperatures above 50°C, presumably due to local melting of the structure.

The origin of the polarization at 2 ppm (20°C) is not known. The fact that it is still present in the CIDNP spectrum of tRNAphe-Y rules out the possibility that it is from the Y base, in spite of the fact that it occurs at nearly the same chemical shift. The linewidth of this polarized peak is considerably narrower than the other peaks, so it is likely to be a minor impurity in the solvent.

The temperature dependence of the CIDNP spectrum of tRNAphe is shown in Figure 4.11. Several new polarized resonances are observed starting at about 60-65°C and a weak polarization of the m2G10 methyl protons is also observed beginning at this temperature (data not shown). It is thought that the structure of tRNAphe approximates the native, active form in the presence of high NaCl concentration whether Mg2+ is present or not (Rhodes, 1977). In the presence of Mg2+, the tertiary structure of tRNAphe is stable up to approximately 60°C, where it melts cooperatively. In the absence of
Mg$^{2+}$, the melt starts at lower temperatures in a sequence of several steps but the secondary structure stays relatively intact until above 40°C (Johnston & Redfield, 1981). Comparison of the dark spectra as a function of temperature with those in the literature (Davanloo, et al., 1979; Robillard, et al., 1977) suggests that the samples discussed here contain between 0 and 10mM Mg$^{2+}$. Upon melting of the tertiary structure, one would expect G15, G18, G19, G45, G46, and G57 to become more accessible. Most of these, as well as G4, m$^2$G10, G22, and G65 become accessible to chemical modification at higher temperatures (Rhodes, 1977). Consistent with this, I have observed several additional polarized aromatic resonances beginning at ~60°C (see Figure 4.11).

To date, there has been relatively little progress in analyzing the aromatic regions of tRNA NMR spectra (Schmidt & Edelheit, 1981; Johnston & Redfield, 1981), in contrast to the wealth of information available from the exchangeable base-pairing imino proton resonances (Hurd & Reid, 1979; Johnston & Redfield, 1981). It is noteworthy that CIDNP selectively detects the single-stranded regions. These are mechanistically of great importance and have thus far not been analyzed in detail by NMR. CIDNP can also monitor the accessibility of many of the modified bases, which may be important in determining protein-tRNA and ribosome-tRNA binding sites. This example demonstrates that higher oligonucleotides of nucleic acids are definitely amenable to study using CIDNP and suggests the possibility of much wider studies.
6. CONCLUSIONS

A major problem in NMR studies of large nucleic acid systems is the assignment of resonances to particular protons of the bases. This chapter demonstrates the use of CIDNP in identifying the G(H8) and G(H1') resonances in systems ranging from dinucleotides to tetranucleotides.

Because CIDNP requires interaction of the nucleic acid base with a photo-excited dye, this technique can also be used as a probe of accessibility. This is shown by the study of the single-strand to double-strand transition in ApGpCpU. The single-strand state is expected to be more accessible to the solvent, and this is confirmed by an increase in the observed polarization with increasing temperature. This property of the polarization could be applied to the study of nucleic acid molecules containing single-stranded regions including bulges, internal loops, hairpins and dangling ends. CIDNP would presumably be observed from these single-stranded regions even if the rest of the molecule is double-stranded. Indeed, this chapter demonstrated success in using CIDNP to investigate the single-stranded regions in a native folded molecule of tRNA\textsuperscript{Phe}. 
References


Hilbers, C. W., Garssen, G. J., Kaptein, R., Schoenmakers, J. G. G., 
Spectroscopy in Molecular Biology (Pullman, B., Ed.), D. Reidel 


Johnston, P. D., & Redfield, A. G. (1981) Biochemistry 20, 1147- 
1156.


Pulsos.: Alta Resoluço, pp. 385-407.

Commun., 1092-1094.


Kaptein, R., Stob, S., Scheek, R. M., Dijkstra, K., & Schleich, T. 

Kondo, N. S., Holmes, H. M., Stempel, L. M., & Ts'o, P. O. P. (1970) 
Biochemistry 9, 3479-3498.

Lee, C., Ezra, F. S., Kondo, N. S., Sarma, R. H., & Danyluk, S. S. 


Appendix A

Ethidium Ion Binding Studied by NMR

In this appendix, I will present preliminary work that uses NMR to study ethidium ion binding to the duplex, rCA₆G + dCT₆G. Ethidium is a frameshift mutagen (McCann et al., 1975) whose structure is shown in Figure A.1. My original interest in studying ethidium binding was to investigate the mechanism for frameshift mutations proposed by Streisinger et al. (1966). In this model it is proposed that a region of base repeats can slide forming unpaired (bulged) bases. This region might be stabilized by a frameshift mutagen, thus locking in the mistake created by the base pair slippage. The question then becomes, how does the frameshift mutagen stabilize this perturbation in the helix. Ethidium is known to intercalate into DNA and RNA (Le Pecq & Paoletti, 1967; Douthart et al., 1973). Thus it has been proposed that a possible mechanism for the mutagen to stabilize the perturbation would be to intercalate into the helix at the site of the perturbation. Lee & Tinoco (1978) have shown that ethidium can form a complex with the trimer, rCpUpG. In this complex the ethidium intercalates between the two C·G base pairs and leaves the U bulged into solution. Studies by Helfgott and Kallenbach (1979) have shown that there is increased binding to polymers which contain extra non-base-paired bases.

Before studying ethidium binding to a molecule containing an extra base, it was necessary to characterize the system by studying binding to a "perfect" helix (one with all base pairs formed).
Figure A.1  The chemical structure of the ethidium ion.
Ethidium Bromide

XBL 782-7422
1. MATERIALS AND METHODS

The rCA₆G was synthesized by enzymatic techniques (Martin et al., 1971) and the dCT₆G was synthesized using diester solution techniques (Khorana, 1968). The sample was prepared following the procedures described in Chapter II. The sample was 2mM per strand in 85% H₂O (15% D₂O). The sample was placed in a micro-NMR-tube (Wilmad #508-CP) with a volume of ~170μl (slightly more sample was used so that aliquots for optical samples could be removed during the experiment). A sample of ethidium was made and the concentration was determined optically, using ε₄₈₀ = 5.6 x 10³ M⁻¹ cm⁻¹. Appropriate amounts of ethidium were then pipetted into small centrifuge tubes and lyophilized. The tubes were tested for adsorption of ethidium by redissolving the ethidium in several of the tubes and measuring the optical absorbance; there was no significant loss of ethidium. The titration for the NMR experiments was then executed by removing the sample from the NMR tube using a teflon needle and depositing the sample in the centrifuge tube containing the ethidium. After allowing for the ethidium to dissolve, the sample was returned to the NMR tube. In this manner the ethidium could be added to the sample without substantially changing the concentration of the nucleic acid. Before each addition of ethidium, a small aliquot (2-3 μl) of the sample was removed for use in optical studies. The NMR spectra were obtained using the Redfield 214 pulse sequence as discussed in Chapter II. The results of the oligonucleotide-ethidium titration are shown in Figure A.2.
Figure A.2 360 MHz ¹H NMR spectra demonstrating the effects of titrating the duplex rCA₆G + dCT₆G with ethidium ion, at 10°C. The ratio of the concentrations of the two molecules are given on the right side of each spectrum. The oligonucleotide duplex concentration was kept constant at 2mM per single strand. Each spectrum is the sum of 400-500 accumulations.
Ethidium : Helix

- **e**: 1.5 : 1
- **d**: 1 : 1
- **c**: 0.75 : 1
- **b**: 0.5 : 1
- **a**: 0 : 1

PPM

16 14.5 14 13.5 13 12.5 12 11.5

7 3-8 2.8 1
After the experiment is completed the ethidium can be removed from the sample by extracting with n-BuOH. If the oligonucleotide duplex is stable above room temperature the sample must be heated during the extraction procedure in order to remove the ethidium. After the extraction, the samples were monitored optically (at 480nm) to check for remaining ethidium. Typically over 90% of the ethidium could be removed by this method.

The aliquots of sample removed during the NMR titration experiments were used to determine the extent of ethidium ion binding. Melting curves were obtained by monitoring the absorbance of these samples, at 260 nm and 286 nm (the isosbestic point), as a function of temperature. If the curves taken at 286 nm level off at low temperature it was assumed that the ethidium ion was fully bound (see Chapter IV of Nelson (1981) for further discussion). For the sample whose spectra are shown in Figure A.2, the ethidium ion was found to be fully bound at the concentrations given.

2. DISCUSSION

I have chosen to look at the lowfield imino resonances because there is only one imino resonance per base pair which simplifies the spectrum. Also, there are no resonances from the ethidium in this region of the spectrum.

The most obvious effect of the ethidium binding is seen at the resonance for base pair 1. As ethidium is added this resonance moves upfield by ~0.4 ppm. The double strand with and without ethidium bound must be in slow exchange, on the NMR time scale, as resonances for both species are observed. At one equivalent of
ethidium per double strand the two resonances are approximately the same intensity. This indicates that there is more than one binding site in this molecule. Unfortunately the resonances from the other base pairs have very similar chemical shifts, so it is difficult to detect changes in these resonances. However, if we assume there is a similar shift (~0.4 ppm upfield) upon ethidium binding at other sites then the effect could be at base pairs 7 or 3-6 and it would not be detected. This would place the resonances for the complex under resonances for the uncomplexed nucleotide. In the future these experiments could be attempted at higher field which may increase the resolution of these resonances or perhaps a molecule that does not contain overlapping resonances could be studied. Finally it should be noted that previous work has shown that there is a preference for binding ethidium to pyrimidine (3'-5') purine sites (Krugh & Reinhardt, 1975; Dahl, 1981). In the molecule studied here, the only pyrimidine-purine site is between base pair 1 and 2. The spectra shown in Figure A.2 are consistent with preferential binding to this site combined with binding at several other sites.
APPENDIX B

Distances between protons in B-form dCA₆G + dCT₆G

In this appendix, I present the information needed to calculate distances between atoms in the B-form helix dCA₆G + dCT₆G, shown in Figure B.1. The cartesian coordinates for this duplex are given in Table B.1. The B-form coordinates for dCA₆G + dCT₆G were calculated from data published by Arnott and Hukins (1972). Because this data was determined by X-ray diffraction, the proton coordinates are not available and had to be estimated. All hydrogen atoms are 1.09 Å away from the atoms to which they are covalently bound. The actual position of the hydrogen atom is determined for non-hydrogen-bonded atoms on six membered rings by extending a line which connects the atom to which the hydrogen atom is attached and the atom furthest away in the ring. In the purines the H₈ position is determined by extending the line connecting the C₈ and the midpoint of the C₄-C₅ bond. The hydrogen-bonding protons, imino protons and amino protons, are placed 1.09 Å from the atom to which they are covalently bound on a line connecting the atoms which are involved in the hydrogen bond. I have given both the cylindrical polar coordinates and the cartesian coordinates for all the base protons in Table B.2. These coordinates, along with those given by Arnott and Hukins (1972), can be used to generate the B-form coordinates for any sequence of bases.

Distances between imino protons and other base protons in the dCA₆G + dCT₆G duplex are given in Table B.3. Enlargements of three
Figure B.1  The duplex $\text{dCA}_6\text{G} + \text{dCT}_6\text{G}$ in B-form geometry. Coordinates are given in Table B.1. The sugar-phosphate backbone has been omitted for easier viewing of the base-pair stacking.
Table B.1  X, Y, Z Coordinates for B-form  G-T-T-T-T-T-T-Cd

<table>
<thead>
<tr>
<th>Atom</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>3.95</td>
<td>-2.41</td>
<td>24.08</td>
</tr>
<tr>
<td>C6</td>
<td>4.8</td>
<td>-1.36</td>
<td>24.18</td>
</tr>
<tr>
<td>C5</td>
<td>4.35</td>
<td>-0.08</td>
<td>24.13</td>
</tr>
<tr>
<td>N4</td>
<td>2.42</td>
<td>1.32</td>
<td>23.93</td>
</tr>
<tr>
<td>C4</td>
<td>2.94</td>
<td>0.1</td>
<td>23.98</td>
</tr>
<tr>
<td>N3</td>
<td>2.11</td>
<td>-0.94</td>
<td>23.88</td>
</tr>
<tr>
<td>O2</td>
<td>1.84</td>
<td>-3.94</td>
<td>23.84</td>
</tr>
<tr>
<td>C2</td>
<td>2.58</td>
<td>-2.21</td>
<td>23.93</td>
</tr>
<tr>
<td>H5</td>
<td>5.04</td>
<td>0.76</td>
<td>24.21</td>
</tr>
<tr>
<td>H6</td>
<td>5.87</td>
<td>-1.53</td>
<td>24.3</td>
</tr>
<tr>
<td>H4</td>
<td>1.36</td>
<td>1.46</td>
<td>23.77</td>
</tr>
<tr>
<td>N9</td>
<td>-4.62</td>
<td>0.37</td>
<td>23.24</td>
</tr>
<tr>
<td>C8</td>
<td>-4.49</td>
<td>1.74</td>
<td>23.16</td>
</tr>
<tr>
<td>N7</td>
<td>-3.26</td>
<td>2.17</td>
<td>23.24</td>
</tr>
<tr>
<td>C5</td>
<td>-2.97</td>
<td>1.01</td>
<td>23.58</td>
</tr>
<tr>
<td>O6</td>
<td>-0.22</td>
<td>1.7</td>
<td>23.53</td>
</tr>
<tr>
<td>C6</td>
<td>-1.11</td>
<td>0.84</td>
<td>23.51</td>
</tr>
<tr>
<td>N1</td>
<td>-0.76</td>
<td>-0.52</td>
<td>23.63</td>
</tr>
<tr>
<td>N2</td>
<td>-1.14</td>
<td>-2.79</td>
<td>23.76</td>
</tr>
<tr>
<td>C2</td>
<td>-1.66</td>
<td>-1.56</td>
<td>23.63</td>
</tr>
<tr>
<td>N3</td>
<td>-2.97</td>
<td>-1.41</td>
<td>23.5</td>
</tr>
<tr>
<td>C4</td>
<td>-3.33</td>
<td>-0.1</td>
<td>23.38</td>
</tr>
<tr>
<td>H8</td>
<td>-5.33</td>
<td>2.43</td>
<td>23.04</td>
</tr>
<tr>
<td>H1</td>
<td>0.31</td>
<td>-0.7</td>
<td>23.73</td>
</tr>
<tr>
<td>H2</td>
<td>-0.06</td>
<td>-2.94</td>
<td>23.79</td>
</tr>
</tbody>
</table>

* C COORDS

<table>
<thead>
<tr>
<th>Atom</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>N9</td>
<td>-4.62</td>
<td>0.37</td>
<td>23.24</td>
</tr>
<tr>
<td>C8</td>
<td>-4.49</td>
<td>1.74</td>
<td>23.16</td>
</tr>
<tr>
<td>N7</td>
<td>-3.26</td>
<td>2.17</td>
<td>23.24</td>
</tr>
<tr>
<td>C5</td>
<td>-2.97</td>
<td>1.01</td>
<td>23.58</td>
</tr>
<tr>
<td>O6</td>
<td>-0.22</td>
<td>1.7</td>
<td>23.53</td>
</tr>
<tr>
<td>C6</td>
<td>-1.11</td>
<td>0.84</td>
<td>23.51</td>
</tr>
<tr>
<td>N1</td>
<td>-0.76</td>
<td>-0.52</td>
<td>23.63</td>
</tr>
<tr>
<td>N2</td>
<td>-1.14</td>
<td>-2.79</td>
<td>23.76</td>
</tr>
<tr>
<td>C2</td>
<td>-1.66</td>
<td>-1.56</td>
<td>23.63</td>
</tr>
<tr>
<td>N3</td>
<td>-2.97</td>
<td>-1.41</td>
<td>23.5</td>
</tr>
<tr>
<td>C4</td>
<td>-3.33</td>
<td>-0.1</td>
<td>23.38</td>
</tr>
<tr>
<td>H8</td>
<td>-5.33</td>
<td>2.43</td>
<td>23.04</td>
</tr>
<tr>
<td>H1</td>
<td>0.31</td>
<td>-0.7</td>
<td>23.73</td>
</tr>
<tr>
<td>H2</td>
<td>-0.06</td>
<td>-2.94</td>
<td>23.79</td>
</tr>
</tbody>
</table>

* A COORDS

<table>
<thead>
<tr>
<th>Atom</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>N9</td>
<td>-1.07</td>
<td>-4.5</td>
<td>17.32</td>
</tr>
<tr>
<td>C8</td>
<td>0.25</td>
<td>-4.83</td>
<td>17.4</td>
</tr>
<tr>
<td>N7</td>
<td>1.05</td>
<td>-3.81</td>
<td>17.33</td>
</tr>
<tr>
<td>C5</td>
<td>0.19</td>
<td>-2.73</td>
<td>17.18</td>
</tr>
<tr>
<td>N6</td>
<td>1.64</td>
<td>-0.8</td>
<td>17.04</td>
</tr>
<tr>
<td>C6</td>
<td>0.42</td>
<td>-1.35</td>
<td>17.05</td>
</tr>
<tr>
<td>N1</td>
<td>-0.66</td>
<td>-0.55</td>
<td>16.93</td>
</tr>
<tr>
<td>C2</td>
<td>-1.87</td>
<td>-1.1</td>
<td>16.94</td>
</tr>
<tr>
<td>N3</td>
<td>-2.21</td>
<td>-2.37</td>
<td>17.06</td>
</tr>
<tr>
<td>C4</td>
<td>-1.11</td>
<td>-3.14</td>
<td>17.18</td>
</tr>
<tr>
<td>H8</td>
<td>0.63</td>
<td>5.84</td>
<td>17.52</td>
</tr>
<tr>
<td>H6</td>
<td>1.8</td>
<td>0.27</td>
<td>16.88</td>
</tr>
<tr>
<td>H2</td>
<td>-2.72</td>
<td>-0.43</td>
<td>16.84</td>
</tr>
</tbody>
</table>

* TA 2

<table>
<thead>
<tr>
<th>Atom</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>N9</td>
<td>1.78</td>
<td>-4.27</td>
<td>20.7</td>
</tr>
<tr>
<td>C8</td>
<td>3.04</td>
<td>-3.76</td>
<td>20.78</td>
</tr>
<tr>
<td>N7</td>
<td>3.09</td>
<td>-2.46</td>
<td>20.8</td>
</tr>
<tr>
<td>C5</td>
<td>1.76</td>
<td>-2.1</td>
<td>20.56</td>
</tr>
<tr>
<td>N6</td>
<td>1.8</td>
<td>0.32</td>
<td>20.42</td>
</tr>
<tr>
<td>C6</td>
<td>1.13</td>
<td>-0.84</td>
<td>20.43</td>
</tr>
<tr>
<td>N1</td>
<td>-0.21</td>
<td>-0.83</td>
<td>20.31</td>
</tr>
<tr>
<td>C2</td>
<td>-0.86</td>
<td>-1.99</td>
<td>20.32</td>
</tr>
<tr>
<td>N3</td>
<td>-0.39</td>
<td>-3.22</td>
<td>20.44</td>
</tr>
<tr>
<td>C4</td>
<td>0.94</td>
<td>-3.19</td>
<td>20.56</td>
</tr>
<tr>
<td>H8</td>
<td>3.94</td>
<td>-4.35</td>
<td>20.9</td>
</tr>
<tr>
<td>H6</td>
<td>1.3</td>
<td>1.28</td>
<td>20.26</td>
</tr>
<tr>
<td>H2</td>
<td>-1.94</td>
<td>-1.94</td>
<td>20.22</td>
</tr>
</tbody>
</table>

* A COORDS

<table>
<thead>
<tr>
<th>Atom</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>N9</td>
<td>-3.52</td>
<td>3.01</td>
<td>19.86</td>
</tr>
<tr>
<td>C8</td>
<td>-2.78</td>
<td>4.17</td>
<td>19.76</td>
</tr>
<tr>
<td>N7</td>
<td>-1.54</td>
<td>1.79</td>
<td>20.05</td>
</tr>
<tr>
<td>C5</td>
<td>-1.43</td>
<td>4.14</td>
<td>19.81</td>
</tr>
<tr>
<td>O4</td>
<td>0.5</td>
<td>2.77</td>
<td>20.01</td>
</tr>
<tr>
<td>C4</td>
<td>-0.72</td>
<td>2.89</td>
<td>19.96</td>
</tr>
<tr>
<td>N3</td>
<td>-1.54</td>
<td>1.79</td>
<td>20.05</td>
</tr>
<tr>
<td>O2</td>
<td>-3.56</td>
<td>0.74</td>
<td>20.1</td>
</tr>
<tr>
<td>C2</td>
<td>-2.92</td>
<td>1.78</td>
<td>20.01</td>
</tr>
<tr>
<td>H6</td>
<td>-3.28</td>
<td>5.13</td>
<td>19.64</td>
</tr>
<tr>
<td>H3</td>
<td>-1.04</td>
<td>0.82</td>
<td>20.17</td>
</tr>
</tbody>
</table>

* T COORDS

<table>
<thead>
<tr>
<th>Atom</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>N9</td>
<td>-3.52</td>
<td>3.01</td>
<td>19.94</td>
</tr>
<tr>
<td>C8</td>
<td>-2.64</td>
<td>4.06</td>
<td>14.02</td>
</tr>
<tr>
<td>N7</td>
<td>-1.39</td>
<td>3.7</td>
<td>13.95</td>
</tr>
<tr>
<td>C5</td>
<td>-1.45</td>
<td>2.32</td>
<td>13.8</td>
</tr>
<tr>
<td>N6</td>
<td>0.86</td>
<td>-1.62</td>
<td>13.66</td>
</tr>
<tr>
<td>C6</td>
<td>-0.45</td>
<td>1.33</td>
<td>13.67</td>
</tr>
<tr>
<td>N1</td>
<td>-0.86</td>
<td>-0.06</td>
<td>13.55</td>
</tr>
<tr>
<td>C2</td>
<td>-2.16</td>
<td>0.2</td>
<td>13.56</td>
</tr>
<tr>
<td>N3</td>
<td>-3.18</td>
<td>-0.62</td>
<td>13.68</td>
</tr>
<tr>
<td>C4</td>
<td>-2.74</td>
<td>-1.89</td>
<td>13.8</td>
</tr>
<tr>
<td>H8</td>
<td>-2.92</td>
<td>-5.09</td>
<td>14.14</td>
</tr>
<tr>
<td>H6</td>
<td>1.61</td>
<td>-0.84</td>
<td>13.5</td>
</tr>
<tr>
<td>H2</td>
<td>-2.45</td>
<td>1.25</td>
<td>13.46</td>
</tr>
</tbody>
</table>

* T COORDS

<table>
<thead>
<tr>
<th>Atom</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>1.78</td>
<td>4.27</td>
<td>13.1</td>
</tr>
<tr>
<td>C6</td>
<td>3.11</td>
<td>3.93</td>
<td>13.0</td>
</tr>
<tr>
<td>C5</td>
<td>4.92</td>
<td>2.21</td>
<td>12.94</td>
</tr>
<tr>
<td>C5</td>
<td>3.49</td>
<td>2.64</td>
<td>13.05</td>
</tr>
<tr>
<td>O4</td>
<td>2.79</td>
<td>0.38</td>
<td>13.25</td>
</tr>
<tr>
<td>C4</td>
<td>2.52</td>
<td>1.58</td>
<td>13.2</td>
</tr>
<tr>
<td>N3</td>
<td>1.22</td>
<td>2.02</td>
<td>13.29</td>
</tr>
<tr>
<td>O2</td>
<td>-0.39</td>
<td>3.62</td>
<td>13.34</td>
</tr>
<tr>
<td>C2</td>
<td>0.79</td>
<td>3.33</td>
<td>13.25</td>
</tr>
<tr>
<td>H6</td>
<td>3.86</td>
<td>4.71</td>
<td>12.88</td>
</tr>
<tr>
<td>H3</td>
<td>0.46</td>
<td>1.29</td>
<td>13.41</td>
</tr>
</tbody>
</table>
Table B.1 Continued.

<table>
<thead>
<tr>
<th>*</th>
<th>* TA 5</th>
<th>*</th>
<th>* TA 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>*</td>
<td>* A COORDS</td>
<td>*</td>
<td>* A COORDS</td>
</tr>
<tr>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>N9</td>
<td>-4.62</td>
<td>-0.37</td>
<td>10.56</td>
</tr>
<tr>
<td>C8</td>
<td>-4.52</td>
<td>-1.73</td>
<td>10.64</td>
</tr>
<tr>
<td>N7</td>
<td>-3.3</td>
<td>-2.17</td>
<td>10.57</td>
</tr>
<tr>
<td>C5</td>
<td>-2.54</td>
<td>-1.03</td>
<td>10.42</td>
</tr>
<tr>
<td>N6</td>
<td>-0.25</td>
<td>-1.81</td>
<td>10.28</td>
</tr>
<tr>
<td>C6</td>
<td>-1.15</td>
<td>-0.81</td>
<td>10.29</td>
</tr>
<tr>
<td>N1</td>
<td>-0.73</td>
<td>0.45</td>
<td>10.17</td>
</tr>
<tr>
<td>C2</td>
<td>-1.63</td>
<td>1.44</td>
<td>10.18</td>
</tr>
<tr>
<td>N3</td>
<td>-2.94</td>
<td>0.09</td>
<td>10.42</td>
</tr>
<tr>
<td>H8</td>
<td>-5.35</td>
<td>-2.41</td>
<td>10.76</td>
</tr>
<tr>
<td>H6</td>
<td>0.81</td>
<td>-1.63</td>
<td>10.12</td>
</tr>
<tr>
<td>H2</td>
<td>-1.25</td>
<td>2.45</td>
<td>10.08</td>
</tr>
<tr>
<td>*</td>
<td>* T COORDS</td>
<td>*</td>
<td>* T COORDS</td>
</tr>
<tr>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>N1</td>
<td>3.95</td>
<td>2.41</td>
<td>9.72</td>
</tr>
<tr>
<td>C6</td>
<td>4.82</td>
<td>1.36</td>
<td>9.62</td>
</tr>
<tr>
<td>C5ME</td>
<td>5.26</td>
<td>-1.1</td>
<td>9.56</td>
</tr>
<tr>
<td>C5</td>
<td>4.38</td>
<td>0.08</td>
<td>9.67</td>
</tr>
<tr>
<td>O4</td>
<td>2.46</td>
<td>-1.34</td>
<td>9.87</td>
</tr>
<tr>
<td>C4</td>
<td>2.97</td>
<td>-0.28</td>
<td>9.82</td>
</tr>
<tr>
<td>N3</td>
<td>2.18</td>
<td>0.91</td>
<td>9.91</td>
</tr>
<tr>
<td>O2</td>
<td>1.81</td>
<td>3.16</td>
<td>9.96</td>
</tr>
<tr>
<td>C2</td>
<td>2.59</td>
<td>2.23</td>
<td>9.87</td>
</tr>
<tr>
<td>H6</td>
<td>5.89</td>
<td>1.54</td>
<td>9.5</td>
</tr>
<tr>
<td>H3</td>
<td>1.11</td>
<td>0.74</td>
<td>10.03</td>
</tr>
<tr>
<td>*</td>
<td>* TA 7</td>
<td>*</td>
<td>* CG 8</td>
</tr>
<tr>
<td>*</td>
<td>* A COORDS</td>
<td>*</td>
<td>* C COORDS</td>
</tr>
<tr>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>N9</td>
<td>-1.78</td>
<td>4.27</td>
<td>3.8</td>
</tr>
<tr>
<td>C9</td>
<td>-3.04</td>
<td>3.76</td>
<td>3.88</td>
</tr>
<tr>
<td>N7</td>
<td>3.09</td>
<td>2.46</td>
<td>3.81</td>
</tr>
<tr>
<td>C5</td>
<td>-1.76</td>
<td>2.1</td>
<td>3.66</td>
</tr>
<tr>
<td>N6</td>
<td>-1.8</td>
<td>-0.32</td>
<td>3.52</td>
</tr>
<tr>
<td>C6</td>
<td>-1.13</td>
<td>0.84</td>
<td>3.53</td>
</tr>
<tr>
<td>N1</td>
<td>0.21</td>
<td>0.83</td>
<td>3.41</td>
</tr>
<tr>
<td>C2</td>
<td>0.86</td>
<td>1.99</td>
<td>3.42</td>
</tr>
<tr>
<td>N3</td>
<td>0.39</td>
<td>3.22</td>
<td>3.54</td>
</tr>
<tr>
<td>C4</td>
<td>-0.94</td>
<td>3.19</td>
<td>3.66</td>
</tr>
<tr>
<td>H8</td>
<td>-3.94</td>
<td>4.35</td>
<td>4.0</td>
</tr>
<tr>
<td>H6</td>
<td>-1.3</td>
<td>-1.28</td>
<td>3.36</td>
</tr>
<tr>
<td>H2</td>
<td>1.94</td>
<td>1.94</td>
<td>3.32</td>
</tr>
<tr>
<td>*</td>
<td>* T COORDS</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>N1</td>
<td>3.52</td>
<td>-3.01</td>
<td>2.96</td>
</tr>
<tr>
<td>C6</td>
<td>2.78</td>
<td>-4.17</td>
<td>2.86</td>
</tr>
<tr>
<td>C5ME</td>
<td>0.58</td>
<td>-5.37</td>
<td>2.8</td>
</tr>
<tr>
<td>C5</td>
<td>1.43</td>
<td>-4.14</td>
<td>2.91</td>
</tr>
<tr>
<td>O4</td>
<td>-0.5</td>
<td>-2.77</td>
<td>3.11</td>
</tr>
<tr>
<td>C4</td>
<td>0.72</td>
<td>-2.89</td>
<td>3.06</td>
</tr>
<tr>
<td>M3</td>
<td>1.54</td>
<td>-1.79</td>
<td>3.15</td>
</tr>
<tr>
<td>O2</td>
<td>3.56</td>
<td>-0.74</td>
<td>3.2</td>
</tr>
<tr>
<td>C2</td>
<td>2.92</td>
<td>-1.78</td>
<td>3.11</td>
</tr>
<tr>
<td>H6</td>
<td>3.28</td>
<td>-5.13</td>
<td>2.74</td>
</tr>
<tr>
<td>H3</td>
<td>1.04</td>
<td>-0.82</td>
<td>3.27</td>
</tr>
<tr>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>
TABLE B.2 The cylindrical polar and cartesian coordinates for the base protons. All coordinates refer to bases on one side of the dyad axis; negate Θ and z to generate the dyadically related base. Coordinates for successive nucleotides in a helix can be generated by adding $36^\circ$ to Θ and $3.38 \, \text{Å}$ to z.

<table>
<thead>
<tr>
<th>atom</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>r</th>
<th>Θ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H8</td>
<td>-0.63</td>
<td>5.84</td>
<td>0.62</td>
<td>5.87</td>
<td>96.2</td>
</tr>
<tr>
<td>N-H6</td>
<td>-1.8</td>
<td>-0.27</td>
<td>-0.02</td>
<td>1.82</td>
<td>188.53</td>
</tr>
<tr>
<td>H2</td>
<td>2.72</td>
<td>0.43</td>
<td>-0.06</td>
<td>2.75</td>
<td>8.98</td>
</tr>
<tr>
<td>Guanine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H8</td>
<td>-0.66</td>
<td>5.82</td>
<td>0.62</td>
<td>5.86</td>
<td>96.5</td>
</tr>
<tr>
<td>H1</td>
<td>0.57</td>
<td>-0.51</td>
<td>-0.07</td>
<td>0.76</td>
<td>-41.8</td>
</tr>
<tr>
<td>N-H2</td>
<td>2.81</td>
<td>-0.85</td>
<td>-0.13</td>
<td>2.94</td>
<td>-16.8</td>
</tr>
<tr>
<td>Cytosine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H5</td>
<td>-2.28</td>
<td>4.56</td>
<td>0.55</td>
<td>5.1</td>
<td>116.6</td>
</tr>
<tr>
<td>H6</td>
<td>-0.36</td>
<td>6.06</td>
<td>0.64</td>
<td>6.07</td>
<td>93.4</td>
</tr>
<tr>
<td>N-H4</td>
<td>-1.81</td>
<td>0.84</td>
<td>0.11</td>
<td>2.0</td>
<td>155.1</td>
</tr>
<tr>
<td>Thymine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H6</td>
<td>-0.36</td>
<td>6.08</td>
<td>0.64</td>
<td>6.09</td>
<td>93.4</td>
</tr>
<tr>
<td>H3</td>
<td>0.36</td>
<td>1.28</td>
<td>0.11</td>
<td>1.33</td>
<td>74.3</td>
</tr>
</tbody>
</table>
TABLE B.3 Calculated distances for B-form  
\[ \text{G-T-T-T-T-T-T-C \ d} \]

Intra-base-pair distances

\[
\begin{align*}
\text{T(H3)} & \rightarrow 2.39\text{Å} \rightarrow \text{A(N-H6)} \\
\text{T(H3)} & \rightarrow 2.91\text{Å} \rightarrow \text{A(H2)} \\
\text{G(H1)} & \rightarrow 2.4\text{Å} \rightarrow \text{C(N-H4)} \\
\text{G(H1)} & \rightarrow 2.63\text{Å} \rightarrow \text{G(N-H2)}
\end{align*}
\]

Inter-base-pair distances

\[
\begin{align*}
\text{G1(H1)} & \rightarrow 4.1\text{Å} \rightarrow \text{T2(H3)} \\
\text{T2(H3)} & \rightarrow 3.93\text{Å} \rightarrow \text{A3(H2)} \\
\text{G1(H1)} & \rightarrow 4.12\text{Å} \rightarrow \text{A2(N-H6)} \\
\text{G1(H1)} & \rightarrow 4.35\text{Å} \rightarrow \text{A2(H2)} \\
\text{T2(H3)} & \rightarrow 4.37\text{Å} \rightarrow \text{G1(N-H4)} \\
\text{T2(H3)} & \rightarrow 4.38\text{Å} \rightarrow \text{A3(N-H6)} \\
\text{T2(H3)} & \rightarrow 5.31\text{Å} \rightarrow \text{G1(N-H2)} \\
\text{Tn(H3)} & \rightarrow 3.48\text{Å} \rightarrow \text{Tn-1(H3)*} \\
\text{Tn(H3)} & \rightarrow 3.48\text{Å} \rightarrow \text{Tn+1(H3)} \\
\text{Tn(H3)} & \rightarrow 3.85\text{Å} \rightarrow \text{A}_{n-1}(N-H6) \\
\text{Tn(H3)} & \rightarrow 4.96\text{Å} \rightarrow \text{A}_{n-1}(H2) \\
\text{Tn(H3)} & \rightarrow 3.93\text{Å} \rightarrow \text{A}_{n+1}(H2) \\
\text{Tn(H3)} & \rightarrow 4.38\text{Å} \rightarrow \text{A}_{n+1}(N-H6) \\
\text{T7(H3)} & \rightarrow 3.39\text{Å} \rightarrow \text{G8(H1)} \\
\text{T7(H3)} & \rightarrow 3.83\text{Å} \rightarrow \text{G8(N-H2)} \\
\text{T7(H3)} & \rightarrow 3.85\text{Å} \rightarrow \text{A6(N-H6)} \\
\text{T7(H3)} & \rightarrow 4.42\text{Å} \rightarrow \text{C8(N-H4)} \\
\text{T7(H3)} & \rightarrow 4.97\text{Å} \rightarrow \text{A6(H2)} \\
\text{G8(H1)} & \rightarrow 3.98\text{Å} \rightarrow \text{A7(N-H6)} \\
\text{G8(H1)} & \rightarrow 4.4\text{Å} \rightarrow \text{A7(H2)} \\
\end{align*}
\]

* where n = 3, 4, 5 and 6
base pair stacks are shown in Figures B.2-B.4. These three figures illustrate all the possible stacks in the duplex. The distances for which NOE has been observed are shown in these figures with dashed lines. As discussed in Chapter II, the intensity of the NOE is inversely related to the sixth power of the distance between two atoms; therefore this appendix will serve as a guide in determining the relative magnitudes of NOE expected for the duplex in a B-form geometry. It should be remembered that there are other factors which contribute to the magnitude of the NOE. One important competing effect is chemical exchange. Therefore, the observed magnitude of the NOE to imino and amino protons may be less than expected because these protons exchange with H₂O. The chemical exchange rates of the imino protons is discussed in Chapter III.
Figure B.2  The three base pair stack dC-A-A + dT-T-G. The dashed lines represent distances for which NOE has been observed. The actual distances are a) G1(imino) → T2(imino), 4.1Å; b) Tn(imino) → Tn+1(imino), 3.48Å; c) G1(imino) → A2(H2), 4.35Å; d) Tn(imino) → An+1(H2), 3.93Å.
Figure B.3  The three base pair stack dA-A-A + dT-T-T. The dashed lines represent distances for which NOE has been observed. The distances are b) $T_n(\text{imino}) \rightarrow T_{n+1}(\text{imino})$, 3.48Å; d) $T_n(\text{imino}) \rightarrow A_{n+1}(\text{H2})$, 3.93Å.
Figure B.4  The three base pair stack dA-A-G + dC-T-T. The dashed lines represent distances for which NOE has been observed. The distances are b) $T_n(\text{imino}) \rightarrow T_{n+1}(\text{imino})$, 3.48Å; d) $T_n(\text{imino}) \rightarrow A_{n+1}(\text{H2})$, 3.93Å; e) $T_7(\text{imino}) \rightarrow G_8(\text{imino})$, 3.39Å; f) $G_8(\text{imino}) \rightarrow A_7(\text{H2})$, 4.4Å.
APPENDIX C

Theoretical Calculations of Chemical Shifts

In the first part of this appendix, I will present the results of theoretical calculations of shielding parameters. These values are used in Chapter II to predict the effect of having the extra cytosine stacked in the helix. In the second part of the appendix, the results of an empirical determination of the high temperature chemical shifts of rCA₆G and rCA₃CA₃G will be presented.

1. THEORETICAL SHIELDING PARAMETERS

The calculated values for the shielding parameters of various protons are presented in Table C.1 for dCA₆G + dCT₆G and in Table C.2 for dCA₃CA₃G + dCT₆G. These values were determined by a method similar to that of Arter and Schmidt (1976). However the values given in Tables C.1 and C.2 also include the effects of local magnetic anisotropy and static electric polarization. The effect of the sugar-phosphate backbone is not included in these calculations. The values given in these tables were determined by David Keller in our laboratory. I will not go into the details of how these values were actually determined; the values will only be used to show trends in the shielding rather than to calculate exact chemical shifts.

A view of dA-A + dT-T is shown in Figure C.1 looking down the helix axis to show the effect of the base overlap on the shielding of various base protons. Figures C.2 and C.3 are also looking down the helix axis at the two stacks that would be present if the extra

<table>
<thead>
<tr>
<th>Proton</th>
<th>Nearest Neighbor Effect</th>
<th>Cross Strand Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>induced near stack ring current</td>
<td>induced next near stack ring current</td>
</tr>
<tr>
<td></td>
<td>magnetic anisotropy</td>
<td>static electric polarization</td>
</tr>
</tbody>
</table>

|       | 1 2 3 4 5 6 7 8 |

| C1(H6) | TG  | AC  | 0.021 | 0.008 | 0.005 | 0.008 |
|        | AG  | TC  | 0.013 | 0.005 | 0.005 | 0.000 |
|        | AA  | TT  | 0.049 | 0.015 | -0.024 | -0.256 | -0.110 | 1.524 |
|        | TT  | AA  | 1.125 | 0.265 | 0.224 |
|        | A4(H2) or A5(H2) | AA  | 0.049 | 0.154 | -0.024 | -0.256 | -0.110 | 1.456 |
|        | TT  | AA  | 1.125 | 0.265 | 0.224 |
|        | A6(H2) | AA  | 0.049 | 0.154 | -0.024 | -0.256 | -0.110 | 1.368 |
|        | TT  | AA  | 1.125 | 0.265 | 0.136 |
Table C.2  Theoretical Calculations of the Shielding Parameters for Protons in

<table>
<thead>
<tr>
<th>proton</th>
<th>nearest neighbor effects</th>
<th>next nearest neighbor effects</th>
<th>cross strand effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>near stack</td>
<td>induced ring current</td>
<td>magnetic anisotropy</td>
</tr>
<tr>
<td>Cl(H6)</td>
<td>TG</td>
<td>-0.008</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td>AC</td>
<td>-0.018</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>C8(H6)</td>
<td>-0.008</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td>AC</td>
<td>0.082</td>
<td>-0.008</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>0.082</td>
<td>-0.008</td>
</tr>
<tr>
<td></td>
<td>Bulge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3(H2)</td>
<td>TT</td>
<td>1.125</td>
<td>0.265</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>0.049</td>
<td>0.154</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>0.049</td>
<td>0.154</td>
</tr>
</tbody>
</table>

Continued
Table C.2  Continued.

<table>
<thead>
<tr>
<th>Proton</th>
<th>Near Stack</th>
<th>Nearest Neighbor Effects</th>
<th>Next Nearest</th>
<th>Cross Strand Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Induced Ring Current</td>
<td>Magnetic Anisotropy</td>
<td>Static Electric Polarization</td>
</tr>
<tr>
<td>A4(H2)</td>
<td>AA TT</td>
<td>0.049</td>
<td>0.154</td>
<td>-0.034</td>
</tr>
<tr>
<td></td>
<td>T CA</td>
<td>0.122</td>
<td>0.217</td>
<td>-0.033</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A5(H2)</td>
<td>GA T</td>
<td>-0.007</td>
<td>-0.040</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>TT AA</td>
<td>1.125</td>
<td>0.265</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A6(H2)</td>
<td>AA TT</td>
<td>0.049</td>
<td>0.154</td>
<td>-0.034</td>
</tr>
<tr>
<td></td>
<td>TT AA</td>
<td>1.125</td>
<td>0.265</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure C.1  A view of $\frac{dA-A}{T-T}$ looking down the helix axis.
The intersection of the dyad axis and the helix axis is
indicated with an x.
Figure C.2  A view of $\frac{dC}{dT}$ looking down the helix axis.
The intersection of the dyad axis and the helix axis is indicated with an x.
Figure C.3  A view of $d_{\alpha-C}$ looking down the helix axis. The intersection of the dyad axis and the helix axis is indicated with an x.
cytosine were stacked in the helix (note: there is no G on the cross strand in these figures). These figures can be used along with Figure 1 in Giessner-Prettre and Pullman (1976), which shows shielding values for the four bases, to better visualize the effects of various stacks on the shielding values of the base protons.

2. EMPIRICALLY DETERMINED HIGH TEMPERATURE CHEMICAL SHIFTS

The values in Tables C.3 and C.4 were determined using the method of Hader et al., (1982). This method involves calculating chemical shifts for oligonucleotides starting with the monomer chemical shift and adding in empirically determined values for the shielding parameters of nearest and next-nearest base stacks. The empirical shielding values were determined by comparing experimental chemical shifts in dinucleoside monophosphates and trinucleoside diphosphates to the chemical shifts found for mononucleotides. These values are determined for ribo-oligonucleotides and I have made no correction for this when comparing them to the experimental data for the deoxyribo-oligonucleotides. Also the dCA₆G and the dCA₃CA₃G may still be partially stacked at 60°C; values used in the calculations are determined from dinucleotides and trinucleotides at 70°C, where there may be significantly less stacking. Both of these facts may account for some of the differences between the experimental and calculated chemical shifts.

The values in Tables C.3 and C.4 were calculated in an attempt to assign the A(H2) resonances at high temperature. However, the difference between the calculated and the experimental chemical shifts is not negligible as is shown by the values of δₘₐₓ - δₑₓₚₜ. 

<table>
<thead>
<tr>
<th>Base(proton)</th>
<th>Nearest Neighbor Effects</th>
<th>Next Nearest Neighbor Effects</th>
<th>Total Calculated Shift</th>
<th>Experimental Chemical Shift</th>
<th>δ$<em>{\text{calc}}$ - δ$</em>{\text{expt}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chemical Shift in Monomer 3'</td>
<td>5'</td>
<td>3' 5'</td>
<td>3' 5'</td>
<td></td>
</tr>
<tr>
<td>A2(H2)</td>
<td>8.260</td>
<td>-0.148</td>
<td>0.000</td>
<td>-0.077</td>
<td>8.035</td>
</tr>
<tr>
<td>A3(H2)</td>
<td>8.260</td>
<td>-0.148</td>
<td>-0.077</td>
<td>-0.015</td>
<td>7.943</td>
</tr>
<tr>
<td>A4(H2)</td>
<td>8.260</td>
<td>-0.148</td>
<td>-0.077</td>
<td>-0.089</td>
<td>7.869</td>
</tr>
<tr>
<td>A5(H2)</td>
<td>8.260</td>
<td>-0.148</td>
<td>-0.077</td>
<td>-0.089</td>
<td>7.869</td>
</tr>
<tr>
<td>A6(H2)</td>
<td>8.260</td>
<td>-0.148</td>
<td>-0.077</td>
<td>-0.029</td>
<td>7.917</td>
</tr>
<tr>
<td>A7(H2)</td>
<td>8.260</td>
<td>-0.074</td>
<td>-0.077</td>
<td>-0.089</td>
<td>8.020</td>
</tr>
</tbody>
</table>

* Calculated chemical shifts are determined for rCA$_6$G at 70°C.

† Experimental chemical shifts are for dCA$_6$G at 60°C.

<table>
<thead>
<tr>
<th>Base(proton)</th>
<th>Chemical Shift in Monomer</th>
<th>Nearest Neighbor Effects (3')</th>
<th>Nearest Neighbor Effects (5')</th>
<th>Next Nearest Neighbor Effects (3')</th>
<th>Next Nearest Neighbor Effects (5')</th>
<th>Total Calculated Shift</th>
<th>Experimental Chemical Shift</th>
<th>(\delta_{\text{calc}} - \delta_{\text{expt}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>A2(H2)</td>
<td>8.260</td>
<td>-0.148</td>
<td>0.000</td>
<td>-0.077</td>
<td>-----</td>
<td>8.035</td>
<td>7.969</td>
<td>0.066</td>
</tr>
<tr>
<td>A3(H2)</td>
<td>8.260</td>
<td>-0.148</td>
<td>-0.077</td>
<td>-0.006</td>
<td>-0.015</td>
<td>8.014</td>
<td>7.848</td>
<td>0.166</td>
</tr>
<tr>
<td>A4(H2)</td>
<td>8.260</td>
<td>-0.050</td>
<td>-0.077</td>
<td>-0.077</td>
<td>-0.089</td>
<td>7.967</td>
<td>7.759</td>
<td>0.208</td>
</tr>
<tr>
<td>A5(H2)</td>
<td>8.260</td>
<td>0.000</td>
<td>-0.148</td>
<td>-0.077</td>
<td>-0.089</td>
<td>7.946</td>
<td>7.739</td>
<td>0.207</td>
</tr>
<tr>
<td>A6(H2)</td>
<td>8.260</td>
<td>-0.148</td>
<td>-0.077</td>
<td>-0.029</td>
<td>-0.015</td>
<td>7.991</td>
<td>7.825</td>
<td>0.166</td>
</tr>
<tr>
<td>A7(H2)</td>
<td>8.260</td>
<td>-0.074</td>
<td>-0.077</td>
<td>-----</td>
<td>-0.089</td>
<td>8.020</td>
<td>7.877</td>
<td>0.143</td>
</tr>
</tbody>
</table>

* Calculated chemical shifts are determined for rCA₃CA₃G at 70°C.

† Experimental chemical shifts are for dCA₃CA₃G at 60°C.
The $A(H_2)$ resonances were therefore assigned by matching the experimental shifts to the calculated shifts so that they both follow the same trends. For example, if $A_2(H_2)$ has the largest calculated chemical shift it will be assigned to the largest experimental chemical shift. The results of assigning the $A(H_2)$ resonances in this manner are discussed in Chapter II and are shown in Figures 2.11 and 2.12.
References


This report was done with support from the Department of Energy. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the Department of Energy.

Reference to a company or product name does not imply approval or recommendation of the product by the University of California or the U.S. Department of Energy to the exclusion of others that may be suitable.