Volume 16 Number 2 1988 Nucleic Acids Research

Sequence specific molecular recognition by a monocationic lexitropsin of the decadeoxyribonucleotide d-[CATGGCCATG]$_2$: structural and dynamic aspects deduced from high field ^{1}H-NMR studies

Moses Lee[1], John A.Hartley[1], Richard T.Pon, Krzysztof Krowicki[1] and J.William Lown[1]

Regional DNA Synthesis Laboratory, University of Calgary, Calgary, Alberta T2N 4N1 and [1]Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada

Received October 30, 1987; Accepted November 20, 1987

ABSTRACT

All ^{1}H-NMR resonances of d-[CATGGCCATG]$_2$ and the 1:1 complex of lexitropsin **1** and the DNA were assigned by the NOE difference, COSY and NOESY methods. Addition of **1** causes the base and imino protons for the sequence 5'-CCAT to undergo the most marked drug-induced chemical shift changes, thereby indicating that **1** is located in this base pair sequence. NOEs confirmed the location and orientation of the drug in the 1:1 complex, with the amino terminus oriented to C(6). The van der Waals interaction between H12a,b of **1** and AH2(8) may be responsible for reading of the 3' A.T base pair in the 5'-CCAT sequence. Exchange NMR effects allow an estimate of $\simeq 62\ s^{-1}$ for the intramolecular "slide-swing" exchange of the lexitropsin between two equivalent binding sites with $\Delta G^{\ddagger} = 58 \pm 5$ kJ mol^{-1} at 301°K.

INTRODUCTION

The sequence specific molecular recognition of DNA by proteins is central to the regulation of many biological processes (1-5). Consequently there is considerable interest in studying the structure and dynamics of complexes of polypeptides with nucleic acids in solution. Netropsin, a natural oligopeptide which is a fermentation product of streptomyces netropsis (6), has proven useful as a model for such sequence specificity. It exhibits antibiotic, antiviral and antitumor activity (7-9) by binding to $(A.T)_4$ sequences in the minor groove of double-stranded B-DNA (9-11) and it interferes with the replication and transcription processes in the cell.

The single crystal X-ray analysis of the 1:1 complex of netropsin to a dodecadeoxyribonucleotide showed that it interacts with four adjacent A.T base pairs by displacing the tightly bound water molecules lining the minor groove (12). The complex formation reaction is an exothermic process (13,14). In the analysis of the structural requirements for the binding of netropsin to DNA, it was predicted by us, and independently by Kopka et al. (12), that replacement of one or more pyrrole rings with hydrogen bonding acceptor heterocycles, such as imidazole moieties, should alter the recognition

properties from A.T to G.C. This prediction is confirmed (15), in which the imidazole containing analogues of netropsin were shown to exhibit a decrease in specificity of A.T sites and concomitant increase in the acceptance of G.C base pairs. This implies the formation of new hydrogen bonds between the 2-amino group of guanine and the N_3 imidazole nitrogen. The firm and site-specific binding between these information reading oligopeptides, or lexitropsins, and DNA is a net result of hydrogen-bonding, electrostatic attraction and van der Waals interactions (11,12).

Recently, we have developed ^{1}H-NMR methodology for studying the structure and dynamics of the binding of oligopeptides to DNA (16). The reported study showed that a prototype lexitropsin containing an imidazole at the carboxyl terminus of netropsin binds specifically to the AATT sequence of the decamer d-[CGCAATTGCG]$_2$, which is in accord with the footprinting results (15).

The deepest negative potential well in the minor groove occurs at A.T. rich runs (14). Therefore doubly positively charged ligands such as netropsin will have a bias for such A.T rich runs regardless of the sequence specificity imposed by hydrogen bonding. As a further logical step in altering the A.T to G.C specificity, one of the positive charges is removed from the lexitropsins. Accordingly the guanidiniumacetyl group at the N-terminal of the netropsin analogues was replaced by a formyl group to give the monocationic analog **1** (17,18) which is the lexitropsin used for this present study. DNase I footprinting analysis of **1** on a 139 b.p. Hind III/Nci I restriction fragment from pBR322 DNA showed that this compound is highly specific in binding to the sequence 5'-CCGT (17). The terminal A.T binding is presumably determined by the van der Waals contacts of the methylenes at the C-terminal of the lexitropsin (17). These methylenes would enter into steric contact with the guanine amino group of the G.C base pair, thus preventing the binding of the C-terminal to a 3' G.C site; to read 5'-CCGC or 5'-CCGG.

In attempting to uncover the structural elements important to the molecular recognition process between the lexitropsin **1** and DNA we have undertaken in the present study to examine the structure and dynamics of the decadeoxyribonucleotide d-[CATGGCCATG]$_2$ and the 1:1 complex of **1** and this decadeoxyribonucleotide using ^{1}H-NMR techniques.

MATERIALS AND METHODS

Synthesis and Purification of d-CATGGCCATG]$_2$

This decadeoxynucleotide (30 μmol) was synthesized using our previously described large scale synthesis cycle on an Applied Biosystems 380A triple

column DNA synthesizer (16,19). After completion of the deprotection by heating (50°, 16 h) in 15 M NH_3OH, 1988 OD units of crude oligonucleotide mixture was recovered. Purification was then performed on a NACS-20 column (16 mm × 100 mm) containing 30 g (dry wt.) of resin. The column was eluted with a gradient formed between 0.25 M NaCl, 12 mM NaOH (175 mL) and 0.45 M NaCl, 12 mM NaOH (175 mL) at a flow rate of 1 mL/min. The desired oligonucleotide eluted with 0.37 M NaCl. The collected product was neutralized with NH_4Cl solution, concentrated and dialysed in SpectraPore 1000 dialysis tubing to yield 1090 OD units of oligonucleotide. This material was desalted a second time by elution on a Sephadex G-25 column and 25.6 mg (8 μmol) of pure oligonucleotide was obtained.

Sample Preparation and NMR Spectroscopy

The NMR sample of the decadeoxyribonucleotide was prepared by dissolving 12.2 mg of the oligomer in a 99.8% D_2O solution (Aldrich) containing 25 mM potassium phosphate buffer (pH 7.0) and 0.15 mM EDTA. The solution was lyophilized twice with 99.8% D_2O and once with 99.996% D_2O (Aldrich), and finally made up to 0.4 mL with 99.996% D_2O. For exchangeable proton NMR studies, the sample was dissolved in a 9:1 (v/v) H_2O:D_2O solution.

For the 1:1 complex of lexitropsin **1** and the oligomer, the sample was prepared by dissolving the decadeoxyribonucleotide (7.7 mg) in a 99.8% D_2O solution (0.4 mL) containing 25 mM potassium phosphate (pH 7.0) and 0.15 mM EDTA and then lyophilizing to dryness. The lyophilization step was repeated twice more (99.8% and 99.996% D_2O, respectively) and the sample redissolved in 0.4 mL of 99.996% atom purity D_2O (aldrich). For detecting the exchangeable ^{1}H-NMR signals, the sample was made up in a 9:1 (v/v) H_2O:D_2O solution. A fresh stock solution of lexitropsin **1** in the same phosphate buffer solution was prepared before the titration experiment.

NMR Spectroscopy

NMR spectra were obtained at 21°C on Bruker WH-400 and WM-360 cryospectrometers, both of which were interfaced with Aspect 2000 data systems. Two-dimensional spectra [COSY (20)] were recorded with magnitude detection. In the NOESY experiment, the mixing time (0.3 s) was randomly varied by 18% to eliminate spin-coupling effects (21). The 2D-spectra were symmetrized but not apodized. To detect the imino and amido proton NMR signals, the binomial 1-3-3-1 pulse sequence (22) was employed to suppress the intense HDO peak.

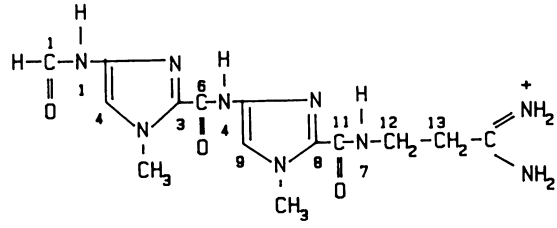
Computer Modeling

The computer generated diagram of the 1:1 complex of **1** and d-[CATGGCCATG]$_2$ was obtained on a Zenith Z-158 microcomputer (Zenith Data

A)

5′-CATGGCCATG- 3′
I II IIIIVV V IVIIII I
3′-GTACCGGTAC- 5′

B)



Scheme I. A) Structure of the self-complementary decamer d-[CATGGCCATG]$_2$.
B) Structure and numbering system of the lexitropsin **1**.

Systems, MI) using the molecular graphics program PC Model (Academic Press Inc., Toronto) and Arnott's coordinates for B-DNA (23).

RESULTS AND DISCUSSION

The decadeoxyribonucleotide is numbered d($C_1A_2T_3G_4G_5C_6C_7A_8T_9G_{10}$) (Scheme IA) while the five imino protons are designated by Roman numerals. The numbering system for the protons and nitrogen in **1** is adapated from that used in the single crystal X-ray studies on netropsin (24) and is shown in Scheme IB.

^{1}H-NMR Chemical Shift Assignments for d-[CATGGCCATG]$_2$ at 21°C in Aqueous Solution

The recently developed techniques using combinations of one-dimensional [NOE differences (25)] and two-dimensional (COSY and NOESY) techniques have proven invaluable in the ^{1}H-NMR analysis of oligodeoxyribonucleotides. By employing these and related techniques (26-28) complete assignment of all exchangeable and non-exchangeable protons in the sequence was possible, as summarized in Table I. Individual intra-residue connectivities deduced from the COSY and NOESY data may be obtained from the authors upon request. It was deduced that, as expected, the decamer exists in solution in a B-type conformation (29-31).

Titration of DNA with Lexitropsin **1**

The results of adding increasing amounts of **1** to the decamer d-[CATGGCCATG]$_2$ are shown in Figure 1 along with the reference spectra of the

Table I. ^{1}H Chemical Shifts (ppm) of d-[CATGGCCATG]$_2$ in Aqueous Solution at 294°K

Base	N1H	N3H	H8	H2	H6	H5	H1'	H2'1	H2'2	H3'	H4'	H5'	T-CH_3
C1					7.73	5.96	5.68	2.67	2.54	4.96	4.39	3.72	
A2			8.45	7.85			6.28	2.92	2.74	5.03	4.40	4.14	
T3		13.51			7.13		5.70	2.32	1.88	4.84	4.41	4.11	1.44
G4	12.78		7.82				5.67	2.70	2.62	4.97	4.34	4.13	
G5	12.90		7.71				5.84	2.67	2.54	4.92	4.39	4.16	
C6					7.35	5.23	5.87	2.43	2.05	4.76	4.37	4.16	
C7					7.49	5.59	5.45	2.42	2.07	4.78	4.13	4.04	
A8			8.36	7.74			6.25	2.91	2.68	5.03	4.41	4.08	
T9		13.61			7.13		5.78	2.32	1.88	4.83	4.41	4.11	1.54
G10			7.88				6.13	2.68	2.32	4.64	4.13	4.08	

free decamer (Figure 1D) and free lexitropsin **1** (Figure 1E). With increasing proportions of **1**, the ^{1}H-NMR resonances undergo chemical shift changes and the lines are broadened. The changes in chemical shifts arise from the changes in chemical environment of the two components upon formation of the lexitropsin:DNA complex, while the line broadening is a result of binding of a small ligand to a macromolecule with consequent increase in the rotational correlation time (32). After the addition of 1.0 mol. equiv. (**1**:DNA duplex), see Figure 2A, it is evident that there is loss of degeneracy (i.e. doubling of resonances) at the sites CH6(6), CH6(7), AH8(8) and TH6(9) for the DNA component and all the aromatic (H4 and H9) and the prochiral methylene (H12 and H13) protons of the lexitropsin component in the 1:1 complex. The spectra shown in Figure 1 were recorded at 21°C, and the presence of exchanging signals indicate that the exchange process of the lexitropsin on the DNA is slow on the NMR time scale at ambient temperatures. The exchange mechanism will be discussed later. The doubling of base resonances for $C_6C_7A_8T_9$ is the first indication that the drug is located on this sequence.

Assignment of Non-Exchangeable Proton Resonances in the 1:1 **1**:DNA Complex: Defining the Location and Orientation of the Lexitropsin and Conformation of the DNA

The assignment of the ^{1}H-NMR resonances of the lexitropsin protons and the base protons of the DNA were made using COSY (Figure 2) and 1D-NOE difference NMR spectroscopy (Figure 3). The results from the COSY and NOESY

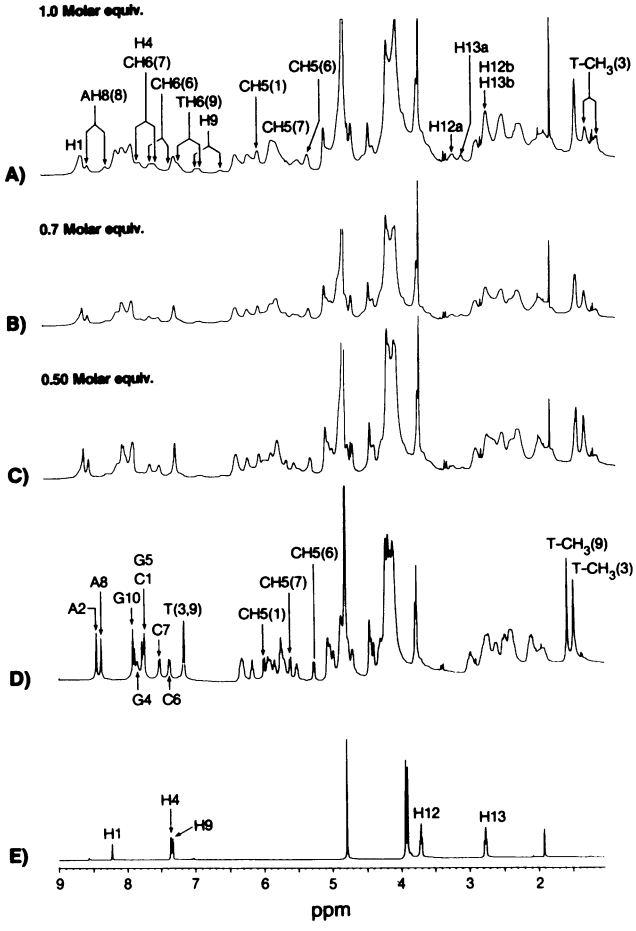


Figure 1. Titration of d-[CATGGCCATG]$_2$ with lexitropsin **1**. Selected proton NMR assignments are indicated in the spectra. Parts A, B and C correspond to the addition of 1.0, 0.7 and 0.5 molar equivalent of **1**, respectively. Parts D and E represent the ^{1}H-NMR spectra of the free DNA and free lexitropsin **1**, respectively.

experiments are limited since very few cross peaks were observed. This is presumably due to the long effective correlation time and effective cross-relaxation between many protons, a phenomenon known as spin diffusion (33). From the partial COSY of the 1:1 complex (Figure 2), it is evident that only one of the three cytosine CH5-CH6 cross peaks is observed. It has been reported that the intensity of the CH5-CH6 correlation is highly sensitive to the local mobility of the cytosine residue (34). Therefore binding of ligands

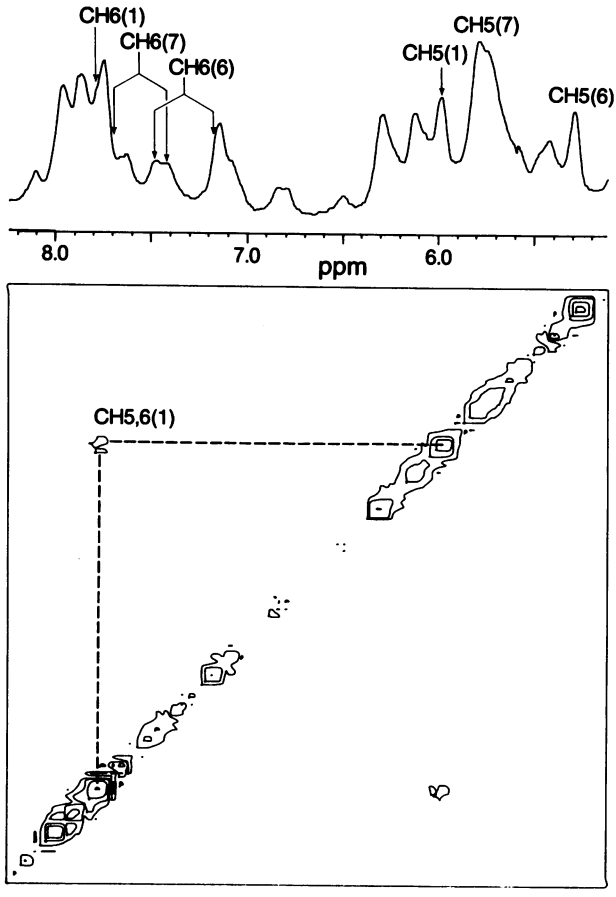


Figure 2. Partial COSY spectrum of the 1:1 complex of **1**:decadeoxyribonucleotide in D_2O at 21°C, from 8.25 to 5.11 ppm. Experimental conditions: initial t_1 = 3 μs, Δt_1 = 0.29 ms, sweep width = 3424 Hz, and 32 FIDs were collected for each of the 256 experiments.

on or near the cytosine residue in question will reduce the mobility and it increases the correlation time of the residue. As a result, the intensity of the CH5-CH6 cross peak in the COSY spectrum is reduced. Since the CH5-CH6 cross peaks for C(6) and C(7) are totally diminished, and only that for C(1) is observable, the results strongly suggest that the lexitropsin **1** is located in the $C_6C_7A_8T_9$ sequence of the DNA as indicated earlier.

The results of the NOE difference experiments are depicted in Figure 3 together with the reference spectrum. Saturation of the low field signal at 8.48 ppm (Figure 3B) gave NOE signals at 6.27 ppm (-3.9% for H1'), 5.04 (-3.2%

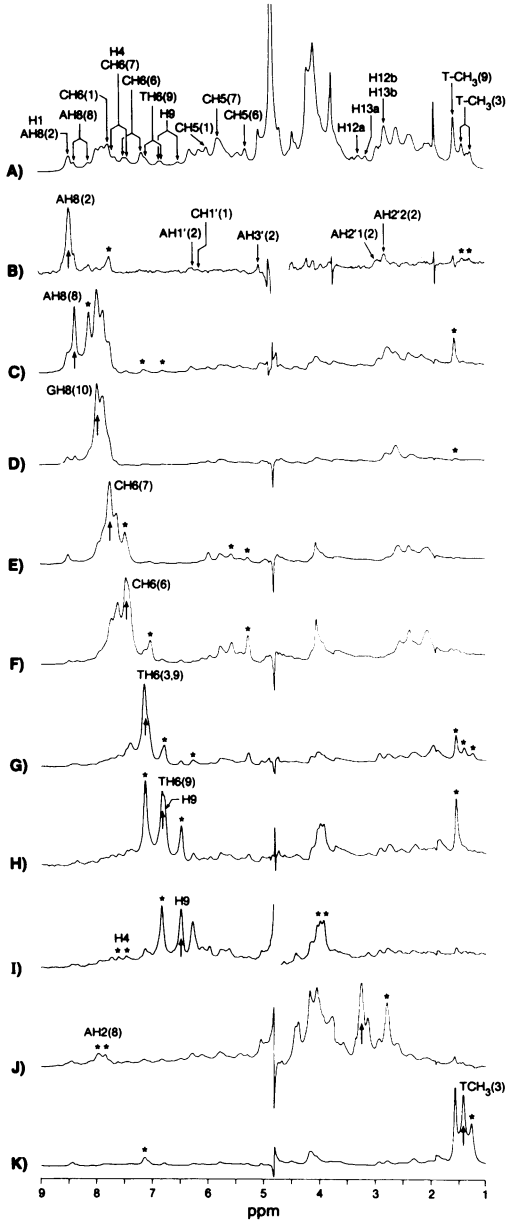


Figure 3. One-dimensional and 1D-NOE difference spectra of the 1:1 complex at 21°C. Experimental cnditions: irradiation time = 0.6 s, spectral width = 3012 Hz and 1600 FIDs were collected for each experiment. The arrows indicate the peaks that were saturated, and important NOEs and exchange signals are denoted by asterisks.

for H3'), 2.93 (-12.9% for H2'1) and 2.79 (-16.1% for H2'2). In addition, NOE peaks are seen at 1.43 and 1.30 ppm for the exchanging signals of T-CH_3(3), and at 6.13 ppm (-1.5%) and 7.71 ppm which is assigned to CH6(1) from the COSY spectrum. Thus the base proton at 8.48 ppm can be assigned to AH8(8). The observed pattern of NOE intensities for H2'>>H1'>H3' when AH8(2) is irradiated indicates that the DNA in the 1:1 complex with **1** adopts a conformation that belongs to the B-family (29). Furthermore, the presence of NOE between AH8(2) and CH1'(1) at 6.15 ppm confirms the B conformation of the DNA (30,31).

It may be noted that the T-CH_3(3) signal is split into two resonances due to the complexation of **1**. Saturation of the AH8 signal at 8.37 ppm gave a strong exchange signal at 8.10 ppm (Figure 3C) in addition NOEs to the T-CH_3(9) and TH6(9) signals at 1.54 and 7.13, respectively. The NOE seen between T-CH_3(9) and GH8(10) (Figure 3D) confirms the assignment of the signal at 7.95 ppm to the latter proton.

Saturation of the peak at 7.59 ppm (Figure 3E) gave an exchange peak at 7.45, and NOEs to a CH5 signal at 5.77 as well as CH5(6) at 5.27 ppm which is assigned by analysis of the titration ^{1}H-NMR spectra given in Figure 1A-C. Thus the exchanging peaks at 7.45 and 7.59 ppm are ascribed to CH6(7). In addition, a NOE peak is observed at 4.05 ppm for a 5-methyl group on the convex face of the drug. Saturation of the signal at 4.05 ppm shows NOE at 7.62 and 7.46 ppm (data not shown); this led to the assignment of the latter signals to the exchange signals for H4 of the lexitropsin. Subsequent irradiation of the signal at 7.49 ppm gave rise to an exchange signal at 7.03 ppm in addition to the CH5(6) peak (Figure 3F); therefore, the former two signals are assigned to the exchanging peaks CH6(6). Saturation of the signal at 7.13 ppm gave NOEs to the T-CH_3 signals (Figure 3G) and thus it is assigned to the TH6 proton of T(3) and T(9). In Figure 3G, there is an exchange TH6 signal observed at 6.82 ppm. Accordingly, saturation of the TH6 signal at 6.82 ppm (Figure 3H) gave an exchange signal to the TH6 peak at 7.13 ppm, and NOE to T-CH_3(9). Thus, this permitted the assignment of exchanging signals for TH6(9). Since the TH6(9) position at 6.82 ppm is near another peak, NOEs with respect to this peak are also seen; an exchange signal at 6.48 ppm, and NOEs at 4.05 and 3.99 ppm. These chemical exchange and dipolar interactions are confirmed by the irradiation of the peak at 6.48 ppm (Figure 3I). The exchanging signals at 6.48 and 6.78 ppm are ascribed to H9 of **1**, and the signals at 4.05 and 3.99 ppm must belong to the 5- and 10-methyl signals of the lexitropsin.

A weak NOE is observed between the exchanging signals of H9 and H4 of the

lexitropsin (Figure 3I), indicating that **1** must be propeller twisted between the two imidazole rings in the 1:1 complex. The propeller twist observed between the two heterocycles could provide the structural flexibility for the lexitropsin in binding compactly in the minor groove of the DNA double helix. It may be noted that a NOE is seen for the two imidazole-methyl groups in the free lexitropsin **1** (data not shown), indicating that the two heterocyclic rings are also propeller twisted in the free form of the drug.

As can be seen in Figure 1A-C, two peaks appear between 3-3.5 ppm as the drug concentration is increased. Accordingly saturation of the lower field signal at 3.26 ppm (Figure 3J) gave an exchange signal at 2.80 ppm, and NOE to the AH2(8) exchange signals at 7.95 and 7.82 ppm. These results indicate that the prochiral methylene protons (H12a,b) are in close proximity to the AH2 proton of the dA(8).dT(3) base pair. This result is consistent with the previous suggestion by us (15,16) that it is this van der Waals interaction that prevents the binding of the carboxyl-terminus of **1** to G.C sites thus forcing the carboxyl terminus to read A.T sites.

The NOEs observed between H12a,b and AH2(8) provide additional support for locating the drug in the minor groove of the decamer along the $C_6C_7A_8T_9$ sequence. These results provide evidence for the orientation of the drug in the 1:1 complex such that the amino to carboxyl termini of **1** are bound to the 5'-CCAT-3' sequence reading left to right. Additional evidence will be presented later to support the latter contention. Saturation of the methyl signal of T(3) (Figure 3K), gave expected NOEs for AH8(2), and the sugar protons of the two nucleotide units. In addition, a weak NOE is observed at 7.81 ppm which is assigned to GH8(4).

The location of **1** on d-$[CATGGCCATG]_2$ was confirmed by chemical shift measurements. The lexitropsin-induced changes in the chemical shifts of individual protons of the DNA duplex versus the sequence are plotted in Figure 4. The graphs demonstrate that the nucleotides C(6) to T(9) are clearly influenced by the binding of **1**. The change in chemical shifts ($\Delta\delta$) for imino proton I for the dC(1).dG(10) base pair are not shown in Figure 4B because the ^{1}H-NMR signal for this proton of the free DNA cannot be detected even at 2°C, presumably due to fraying of the G.C termini.

From the non-selective inversion-recovery experiments of the complex (data not shown), the signals with long spin-lattice relaxation times are observed for AH2 of the DNA and the aromatic (H4 and H9) and formyl (H1) protons of **1**. These results are in accord with the NOE studies.

The NOEs observed for the non-exchangeable protons permitted the location

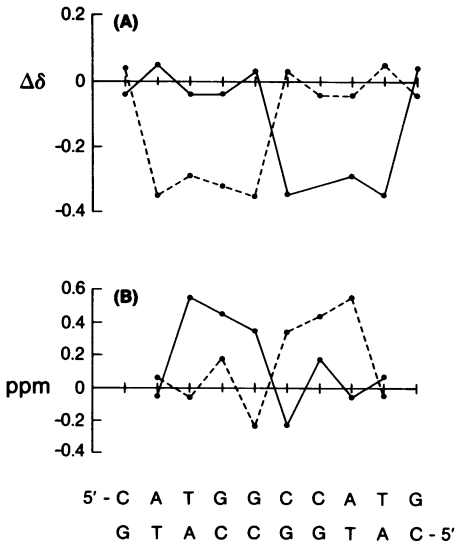


Figure 4. Graphical presentation of the drug-induced chemical shift changes for selected protons in the DNA sequence. Positive values represent that the resonances in the 1:1 complex are at lower field than that in the free DNA. Both strands of the DNA are shown at the bottom, and the drug-induced chemical shift changes ($\Delta\delta$) for each of the two strands are drawn.

of the binding site of **1**, and the NMR assignments for selected protons in the 1:1 lexitropsin:DNA complex, as summarized in Table II.

Assignment of the Exchangeable Protons: Confirmation of the Binding Site and Orientation of the 1:1 Complex

Since both lexitropsin **1** and the DNA contain exchangeable protons, NMR experiments were performed in 9:1 H_2O:D_2O solution to detect the signals of drug-amide and DNA-imino protons.

The ^{1}H-NMR spectrum of the exchangeable protons in the 1:1 complex at 21°C reveals nine imino proton and three amide proton signals as shown in Figure 5A. NOE measurements of the imino protons permitted assignment of all these resonances (28). Saturation of the furthest downfield signal at 14.07 ppm (Figure 5B) gave a strong exchange signal at 13.46 ppm and weak NOEs to 13.51 for an A.T base pair, 12.92 for a G.C pair, and 7.95 ppm for AH2(8). Thus the signals at 14.07 and 13.46 ppm are ascribed to imino proton III. Subsequent irradiation of the signal at 13.68 ppm gave a strong exchange signal at 13.57 ppm, and weak NOEs to the signals for imino proton III (Figure 5C). Therefore the former set of exchange resonances can be ascribed to imino

Table II. ^{1}H-NMR Chemical Shift Assignments of Individual Protons of **1** and d-[CATGGCCATG]$_2$ in the 1:1 Complex at 21°C in D_2O Solution

DNA Resonances			
Base	**DNA**	**1:1 Complex**	**Chemical Shift Change**
CH6(1)	7.73	7.71	-0.02
AH8(2)	8.45	8.49	0.04
TH6(3)	7.17	7.13	-0.04
GH8(4)	7.85	7.81	-0.04
GH8(5)	7.74	7.77	0.03
CH6(6)	7.38	7.03, 7.49	0.11, -0.35
CH6(7)	7.52	-	-
AH8(8)	8.39	8.10, 8.37	-0.02, -0.29
TH6(9)	7.17	6.82, 7.13	-0.04, -0.35
GH8(10)	7.91	7.95	0.04
Imino Protons			
C(1)-G(10) I	-	13.20	-
A(2)-T(9) II	13.61	13.57, 13.68	-0.04, 0.07
T(3)-A(8) III	13.51	13.46, 14.07	-0.05, 0.56
G(4)-C(7) IV	12.78	12.97, 13.23	0.19, 0.45
G(5)-C(6) V	12.90	12.67, 13.25	-0.23, 0.35
Lexitropsin Protons		Free Drug	1:1 Complex
H1		8.22	8.47
H4		7.36	7.46, 7.62
H9		7.33	6.48, 6.78
H12		3.72	2.80, 3.26
H13		1.79	2.80, 3.14
N-CH_3(5)		3.94	4.05
N-CH_3(10)		3.91	3.99

proton II. The assignment of imino protons II and III is in accord with the general observation that the chemical shift of the imino proton of thymine in an A.T base pair resonates at lower field than that of a G.C base pair (27). In Figure 5D, the broad imino proton signal at 13.23 (resolution enhanced spectrum reveals three peaks at 13.20, 13.23 and 13.25 ppm) gave exchange peaks at 12.97 and 12.67 ppm. Subsequent irradiation of the latter two imino

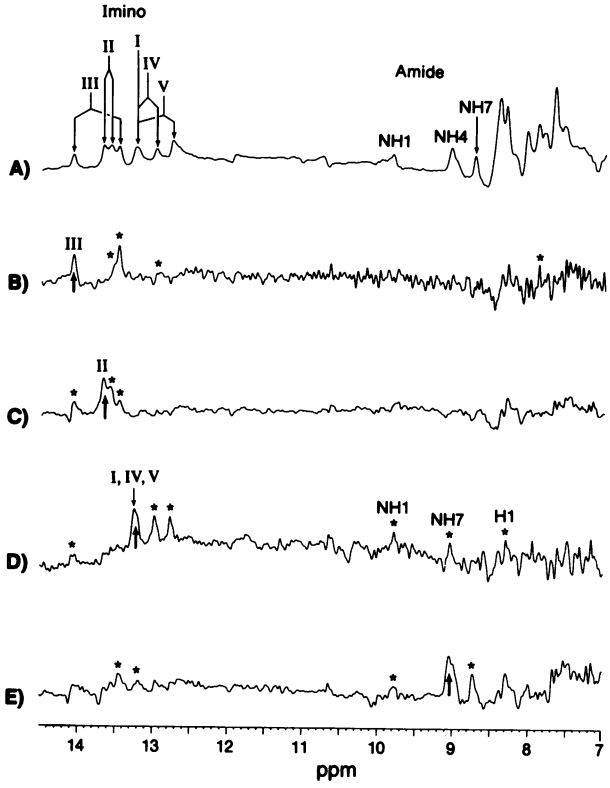


Figure 5. One-dimensional imino proton NMR and 1D-NOE difference spectra of the 1:1 complex of 1 and the decadeoxyribonucleotide, d-$[CATGGCCATG]_2$, at 21°C. Experimental conditions: sweep width = 8064 Hz, irradiation time = 0.2 s, and 1200 FIDs were collected for each spectrum. The arrows indicate the peaks saturated, and the asterisks represent the important NOEs observed.

proton resonances (data not shown) permitted assignment of the exchange signals at 12.97-13.23 and 12.67-13.25 to imino protons IV and V, at 21°C, while only four imino proton signals can be detected even at 2°C for free d-$[CATGGCCATG]_2$. This result suggests that there is kinetic stabilization of the DNA double helix not only at the binding site, $C_6C_7A_8T_9$, but also at the adjacent base pairs. This observation is in agreement with the thermodynamic studies of the 1:1 complex of netropsin to poly(dA).poly(dT) in which the change in energy of complexation is -10.7 Kcal/mol (13). The stabilization of the DNA double helix and increase in the lifetimes for exchange of the imino protons in the DNA to the medium upon binding of netropsin have also been studied using ^{1}H-NMR (35). To provide supporting

evidence for stabilization of the DNA double helix, thermal denaturation studies by determination of UV hyperchromicity at 260 nm of the 1:1 **1**:DNA complex and free DNA were performed. The **1**:decadeoxyribonucleotide complex exhibits a sharper increase in absorbance with increasing temperature with pronounced T_M of 56°C than that of the free decadeoxyribonucleotide whose T_M is 46°C. This suggests that the duplex DNA is considerably stabilized by minor groove binding of **1**. The binding constant of **1** to calf thymus DNA determined by the ethidium displacement assay is 4.91×10^5 M^{-1}. Under these conditions the binding constant for netropsin is 1.87×10^6 M^{-1} (36,37).

Also shown in Figure 5D are NOEs between the G.C imino protons and the drug amide protons at 9.82 and 9.04 ppm, and the formyl proton, H1, in **1** at 8.47 ppm. Irradiation of the amide proton at 9.04 ppm (Figure 5E) gave NOE peaks for the other two amide protons at 9.82 and 8.74 ppm; thus the resonance at 9.04 ppm can be ascribed to NH4 of **1**. Saturation of the amide proton peak at 8.74 ppm only gave NOEs to imino proton III and amide NH4 (data not shown). Since NH1 should have a large NOE with the formyl proton at 8.47 ppm (38), the amide proton signals at 9.82 and 8.74 ppm can be assigned to NH1 and NH7, respectively. The NOEs observed between NH1, NH4 and imino protons IV and V are in agreement with the proposed model for the 1:1 complex, in which the N to C termini of **1** are bound to the 5' and 3' termini of the 5'-CCAT-3' sequence, respectively, as depicted in Figure 6b. For comparison, the model of a 1:1 complex of netropsin and $(A.T)_4$ is given in Figure 6a. Since there are four sets of exchanging imino proton signals, this result provides additional evidence for locating the lexitropsin in the CCAT sequence. Thus, due to the symmetry of the DNA, binding of **1** to 5'-GGCC and 5'-GCCA sequences would result in two and three sets of exchanging imino proton signals, respectively.

Dynamics of the Lexitropsin:DNA Complex

In Figure 1A, one can detect exchange signals for CH6(6), CH6(7), AH8(8), TH6(9), and $T\text{-}CH_3(3)$ for the DNA, as well as H4, H9, H12, and H13 for the lexitropsin. It is evident from the variable temperature ^{1}H-NMR experiments (data not shown) that as the temperature increasees the exchange rate accelerates and as a result the exchanging peaks coalesce and the signals sharpen. Coalescence for the $T\text{-}CH_3(3)$ signals are observed at 301°K. From the standard formula for two-site exchange with equal population described by Sutherland (39), the rate of exchange of **1** between the two equivalent 5'-CCAT sites on the decadeoxyribonucleotide is estimated to be about 62 s^{-1} at 301°K. A $\Delta G^{\ddagger}$ of 58 ± 5 kJ mol^{-1} for the exchange process is calculated for

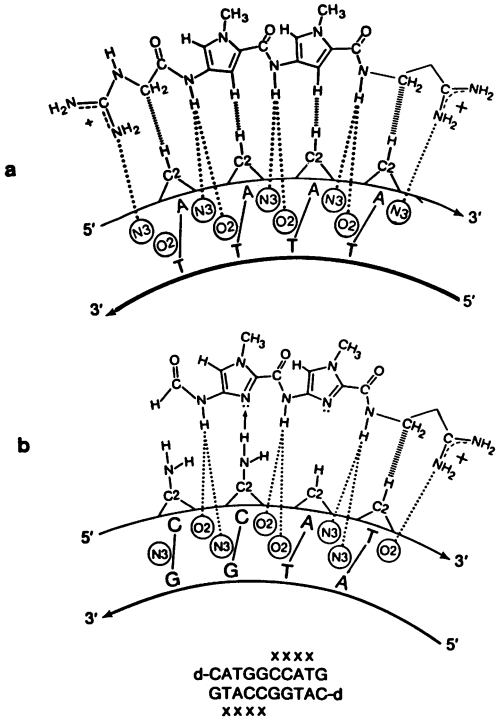


Figure 6. Depiction of molecular recognition components of (a) netropsin bound to 5'-AATT-3' deduced from x-ray diffraction analysis and (b) alignment of lexitropsin **1** on the 5'-CCAT sequence of the decadeoxyribonucleotide. Dotted lines represent hydrogen bonding between the drug amide and amidinium protons and cytosine-O_2, thymine-O_2 and adenine-N_3. The thick arrow indicates the new hydrogen bond between the 2-amino moiety of guanine and the lone pair of electrons on imidazole-N_3 (N2). The NOE observed between AH2(8) and H12a,b of **1** is denoted by the dashed line.

the T-CH_3(3) signal at 301°K (40). It has been suggested that the rate determining step for the two-site exchange process is the departure of the lexitropsin transiently from its binding site and it is then reoriented to become associated with the opposite strand (38). These authors have also estimated that in the flip-flop exchange process of distamycin with the d-$[CGCGAATTCGCG]_2$, the drug flips on the DNA at half the frequency at which it dissociates and estimated $k_{flip} \approx 2\ s^{-1}$ and $k_{ex} \approx 4\ s^{-1}$ at 27°C. Recently we have reported that the estimated exchange rate of an imidazole containing analog of netropsin, bound to a 5'-AATT-3' sequence in a flip-flop process is about 44 s^{-1} at 294°K (16).

The exchange mechanism for the 1:1 complex of lexitropsin **1** and

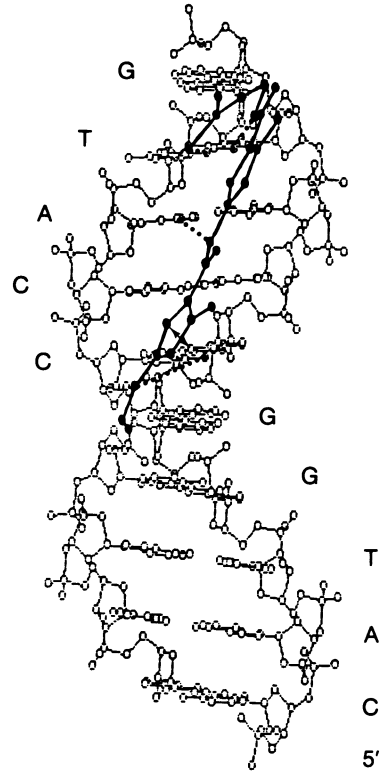


Figure 7. Computer generated depiction of the 1:1 complex. The dotted lines indictate hydrogen bonds between the lexitropsin and DNA, and the double headed arrow represents the hydrogen bond between G(4)-2-NH_2 and imidazole N_3(N2).

d-[CATGGCCATG]$_2$ is more complicated than the flip-flop mechanism at centrally located sites (16), because the two equivalent binding sites of the opposite strands are located towards the two ends of the duplex DNA, as shown in Figure 8. The doubling of the aromatic protons of **1** in the complex suggests that during the exchange process the lexitropsin experiences different chemical environments. The loss of degeneracy of some DNA signals can only arise as a result of drug binding wherein the two strands have separate identities. The two-step "slide-swing" mechanism as shown in Figure 8 is the simplest model to fit the data. The sliding step would change the chemical environment of **1**, and this could explain the loss of degeneracy of the lexitropsin protons. The

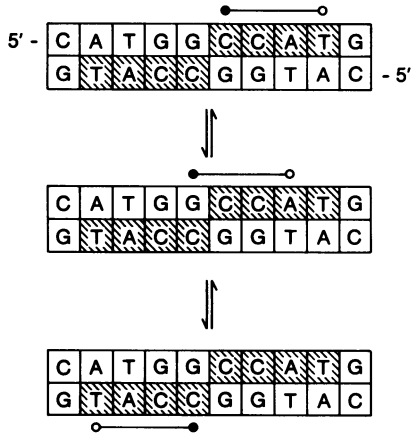


Figure 8. Possible "slide-swing" mechanism of exchange in the 1:1 complex. The disymmetric ligand is represented by two linked circles.

slide-swing mechanism may occur through an intramolecular process, or it may involve transient dissociation of the complex, and perhaps involving exchange of drug molecules with different DNA duplexes. Even though a definite assignment of a mechanism cannot yet be made, there are indications that the exchange mechanism is largely an intramolecular process. These are the kinetic stabilization of the DNA base pairs and slow exchange of the imino protons in the complex, which indicates that the drug considerably reduces the solvent accessibility of the imino protons of the DNA. Therefore, the exchange mechanism is most likely dominated by an intramolecular process on this type of self-complementary receptor.

Molecular Modeling

The computer generated depiction of the 1:1 complex of **1** and d-[CATGGCCATG]$_2$ is shown in Figure 7 with the lexitropsin residing snugly in the minor groove at the 5'-CCAT sequence. The peptide amide protons and the imidazole-N_3 groups are directed toward the concave face and to the base pair edges in the minor groove of DNA. The hydrogen bonding interactions between the drug amide protons and either adenine-N_3 or thymine-O_2 or cytosine-O_2 are illustrated by the dotted lines, while the interaction between imidazole-N_3 (N2) and the 2-amino moiety of G(5) is emphasized by the double headed arrow. Generation of the 1:1 complex is constrained by NOE measurements to establish the contacts between the lexitropsin and DNA.

CONCLUSIONS

Based on the drug-induced chemical shifts of selected protons across the DNA sequence, and the appearance of specific two-site exchange signals in the ^{1}H-NMR spectrum of the 1:1 complex, the lexitropsin molecule is located in the 5'-CCAT sequence of the DNA. This conclusion as well as the orientation placing the amino terminus of the drug at 5'-cytosine in the 1:1 complex are confirmed by NOE studies of both the exchangeable and non-exchangeable protons. NOE difference measurements confirmed that the DNA in the complex exists in the right-handed B-form, which is similar to that of the free decamer. The lexitropsin is subject to propeller twisting about the C(6)-N(4) bond in both the free and bound forms. These results are in good agreement with our previous ^{1}H-NMR studies of a prototype imidazole containing analogue of netropsin and d-[CGCAATTGCG]$_2$ (16).

Further NOE measurements permitted the generation of a diagram of the complex, in which the drug resides on the floor of the minor groove and forms hydrogen bonds to the adenine-N_3, thymine-O_2, cytosine-O_2 with the amide protons, and the new hydrogen bond between the imidazole-N_3 (N2) and G(5)-2-NH_2 thus displacing the shell of hydration on the DNA. It is the hydrogen bonding interactions between N2 of **1** to G(5)-NH_2 and replacement of the positively charged guanidiniumacetyl group with the formyl moiety that are responsible for G.C reading at the N-terminus. Close van der Waals contacts between H12a,b and the adenine-H2 of the dA(8).dT(3) base pair may be responsible for the 5'-CCAT (or 5'-CCGT) reading of lexitropsin **1**. This information has led to the design of new types of lexitropsins that have the methylene groups at the carboxyl terminus successively removed to permit reading of 5'-N_n(C.G)-3' sequences. Such experiments are in progress.

On the basis of the increase in lifetimes of the imino proton NMR signals and the T_M studies of the complex it is evident that the DNA double helix is stabilized by binding of the lexitropsin. Exchange ^{1}H-NMR effects allowed an estimate of a rate of about 62 s^{-1} and $\Delta G^{\ddagger}$ of 58 ± 5 kJ mol^{-1} at 301°K for the two equivalent sites in the dynamic equilibrium binding of lexitropsin **1** to d-[CATGGCCATG]$_2$. The experimental data are in accord with the proposed intramolecular slide-swing mechanism for the exchange process.

These studies show that ^{1}H-NMR is a powerful technique for investigating the structural and dynamic aspects of the factors contributing to the molecular recognition and binding of oligopeptides to DNA. The lexitropsins of future interest include ones that can read longer predetermined base pair sequences as well as rationally designed sequence specific agents that can

serve as vectors for the delivery of DNA effectors, such as anticancer agents. The results of studies in these directions will be reported in due course.

ACKNOWLEDGEMENT

This investigation was supported by grants (to JWL) from the National Cancer Institute of Canada, and the Biotechnology Strategic Grants Programme of the Natural Sciences and Engineering Research Council of Canada. JAH thanks the Alberta Heritage Foundation for Medical Research for a fellowship.

REFERENCES

1. Caruthers, M.H. (1980) Acc. Chem. Res. **13**, 155-160.
2. Frederick, C.A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobson, L., Wang, B.C., Greene, P., Boyer, H.W., and Rosenberg, J.W. (1984) Nature **309**, 327-331.
3. Gurskii, G.V., Tumanyan, V.G., Zavedatelev, A.S., Zhuze, A.L., Grokhovsky, S.L., and Gottikh, B.P. (1977) in Nucleic Acid-Protein Recognition (Vogel, H.J. ed.) pp. 189, Academic Press, New York.
4. Takeda, Y., Ohlendorf, D.H., Anderson, W.F., and Matthews, B.W. (1983) Science **221**, 1020-1026.
5. Kim, S.H., Sussman, J.L., and Church, G.M. (1974) Structure and Conformation of Nucleic Acids and Protein-Nucleic Acid Interactions (Sundaralingam, M., and Rao, S.T. eds.) pp. 571-575, University Park Press, Baltimore, MD.
6. Finlay, A.C., Hochstein, F.A. Sobin, B.A., and Murphy, F.X. (1951) J. Am. Chem. Soc. **73**, 341-343.
7. Hahn, F.E. (1975) in Antibiotics III. Mechanism of Action of Antimicrobial and Antitumor Agents (Corcoran, J.W., and Hahn, F.E. eds.) pp. 79-100.
8. Krey, A. (1980) Prog. Mol. Subcell. Biol. **7**, 43-87.
9. Zimmer, C., and Wähnert, U. (1986) Prog. Biophys. Molec. Biol. **47**, 31-112.
10. Wartell, R.M., Larson, J.E., and Wells, R.D. (1974) J. Biol. Chem. **249**, 6719-6731.
11. Zimmer, C. (1975) Prog. Nucleic Acid Res. Mol. Biol. **15**, 285-318.
12. Kopka, M.L., Yoon, C., Goodsell, D., Pjurra, P., and Dickerson, R.E. (1985) Proc. Natl. Acad. Sci. USA **82**, 1376-1380.
13. Marky, L.A., Snyder, J.G., Remeta, D.P., and Breslauer, K.J. (1983) in Biomolecular Stereodynamics (Sarma, R.H. ed.) Vol. 1, pp. 487, Academic Press, New York.
14. Zakrewska, K., Lavery, R., and Pullman, B. (1983) Nucl. Acids Res. **11**, 8825-8839.
15. Lown, J.W., Krowicki, K., Bhat, U.G., Skorobogaty, A., Ward, B., and Dabrowiak, J.C. (1986) Biochemistry **25**, 7408-7416.
16. Lee, M., Chang, D.-K., Hartley, J.A., Pon, R., Krowicki, K., and Lown, J.W. (1987) Biochemistry in press.
17. Kissinger, K., Krowicki, K., Lown, J.W., and Dabrowiak, J.C. (1987) Biochemistry **26**, 5590-5595.
18. Krowicki, K., and Lown, J.W. (1987) J. Org. Chem. **52**, 3493-3501.
19. Debart, F., Rayner, B., Imbach, J.-L., Lee, M., Chang, D.K., Pon, R.T., and Lown, J.W. (1987) J. Biomol. Struct. Dyn. **5**, 47-65.

20. Aue, W.P., Bartholdt, E., and Ernst, R.R. (1976) J. Chem. Phys. **71**, 2229-2246.
21. Macura, S., and Ernst, R.R. (1980) Molec. Phys. **41**, 95-117.
22. Hore, P.J. (1983) J. Magn. Res. **55**, 283-300.
23. Arnott, S., Campbell-Smith, P.J., and Chandrasekasan, R. (1976) in Handbook of Biochemistry and Molecular Biology, (Fasmman, O.D. ed.) pp. 411, CRC Press, Cleveland, Ohio.
24. Berman, H.M., Neidle, S., Zimmer, C., and Thrum, N. (1979) Biochim. Biophys. Acta **561**, 124-131.
25. Noggle, J.H., and Schirmer, R.D. (1971) The Nuclear Overhauser Effect: Chemical Applications, Academic Press, New York.
26. Kearns, D.R. (1984) C.R.C. Critical Rev. in Biochem. **15**, 237-290.
27. Debart, F., Rayner, B., Imbach, J.-L., Chang, D.K., and Lown, J.W. (1986) J. Biomol. Struct. Dyn. **4**, 343-363.
28. Roy, S., and Redfield, A.G. (1981) Nucl. Acids Res. **9**, 7073-7083.
29. Gronenborn, A.M., and Clore, G.M. (1985) Prog. Nucl. Magn. Reson. Spectrosc. **17**, 1-32.
30. Hare, D.R., Wemmer, D.E., Chou, S.H., Drobny, G., and Reid, B.R. (1983) J. Mol. Biol. **171**, 319-336.
31. Patel, D.J., and Shapiro, L. (1986) J. Biol. Chem. **261**, 1230-1240.
32. James, T.L. (1975) Nuclear Magnetic Resonance in Biochemistry pp. 220, Academic Press, New York.
33. Kalk, A., and Berendsen, H.C.J. (1976) J. Magn. Res. **24**, 343-366.
34. Borah, B., Roy, S., Zon, G., and Cohen, J.S. (1985) Biochem. Biophys. Res. Commun. **133**, 380-388.
35. Pardi, A., Morden, K.M., Patel, D.J., and Tinoco, I. Jr. (1983) Biochemistry **22**, 1107-1113.
36. Morgan, A.R., Lee, J.S., Pulleyblank, D.E., Murray, N.L., and Evans, D.H. (1979) Nucl. Acids Res. **7**, 547-569.
37. Le Pecq, J.-B., and Paoletti, C. (1967) J. Mol. Biol. **27**, 87-106.
38. Klevit, R.E., Wemmer, D.E., and Reid, B.R. (1986) Biochemistry **25**, 3296-3303.
39. Sutherland, I.O. (1971) Ann. Rep. NMR Spectroscopy **4**, 71-225.
40. Günther, H. (1980) NMR Spectroscopy pp. 234-280, Wiley, New York.