Solution Structure of GCCAAT Recognition Motif by 2D NMR, Spectral Simulation, Molecular Modeling, and Distance Geometry Calculations

R. Nibedita, R. Ajay Kumar, A. Majumdar, R. V. Hosur,* and Girjesh Govil

Chemical Physics Group, Tata Institute of Fundamental Research, Bombay 400005, India

K. Majumder and V. S. Chauhan

International Centre for Genetic Engineering and Biotechnology, New Delhi 110067, India

Received December 23, 1992; Revised Manuscript Received May 5, 1993

ABSTRACT: Solution conformation of a self-complementary 14-mer DNA duplex (d-GGATTGGCCAATCC) containing the GCCAAT recognition motif of several transcription factors has been investigated by NMR spectroscopy. Complete resonance assignment of all the protons (except H5′,H5″ protons) has been obtained following standard procedures based on two-dimensional NMR techniques. Three-bond coupling constants have been determined by spectral simulation procedures. New strategies have been described and employed for quantifying NOE intensities from the structural point of view. Approximate ranges of γ torsion angles have been obtained from a selective NOESY experiment, by estimating the $J(4'–5')$, $J(4'–5'')$, or their sum in the H1′–H4′ cross peaks of the spectrum. Likewise, ranges of ϵ torsion angles have been obtained by monitoring the H3′ multiplicities in the H8/H6–H3′ cross peaks in selective NOESY spectra. With the help of such a total of 73 coupling constraints, 79 NOE intensity constraints, and 108 H-bond constraints, model building has been carried out to obtain a structure which satisfies the constraints. Starting from such a structure, an expanded distance constraint set has been created which has been used for the distance geometry calculations using the program TANDY. In the best structure thus derived, interesting irregularities similar to a B_I–B_{II} transition have been observed in the center. The molecule exhibits a bend. The overall base stacking is different from that in either B- or A-DNA models. The base pairs are tilted with respect to the local helix axes. The observed structural features are likely to have important implications for the recognition mechanism of the GCCAAT motif.

In many eukaryotic genes, a hexanucleotide sequence, GCCAAT (often referred to as the CCAAT box), is present in the promoter sequence about 60–80 base pairs upstream of the transcription initiation site. This plays an important role in genetic expression. Mutation and methylation interference studies have indicated that the GCCAAT sequence is recognized by a group of proteins collectively known as CCAAT-binding transcription factors (CTFs) (Jones et al., 1985). These proteins act as transcription initiation factors for many vertebrate globin genes. From HeLa cell extracts, three proteins, namely, CP1, CP2, and NFI, have been isolated which bind to the GCCAAT sequence, making specific contacts (Chodosh et al., 1988). A recent study has shown that a single CCAAT binding protein can act as an RNA polymerase II promoter element as well as an initiation factor for DNA replication (Santoro et al., 1988). It has been suggested that all CCAAT binding proteins may have originated from the same gene but developed differently because of splicing of mRNA. It is interesting that the same GCCAAT sequence acts as a control sequence for transcription in some species (e.g., human) but as a DNA replication initiation control sequence for some other species (e.g., adenovirus) (Jones et al., 1987).

In view of the interesting biochemical observations on the GCCAAT sequence, we have studied the structure of this motif by incorporating it into a self-complementary oligonucleotide d-GGATTGGCCAATCC (hereafter referred to as the 14-mer), by two-dimensional (2D) NMR and computer simulation approaches. In addition to the six unit long stretch ATTGGC (complementary to GCCAAT), the sequence contains a CC stretch at the 3′-end and a corresponding GG stretch at the 5′-end. This has been done to eliminate the influence of end effects on the structure of the CCAAT box in the middle.

There are two basic steps in the determination of solution structures by NMR: first, interpretation of NMR spectra to extract structural parameters such as *J*-coupling constants and internuclear distances; second, model building so as to satisfy the above constraints. In this context, both the accuracy and the number of the constraints play crucial roles in determining the precision of the structure. The number of experimentally derivable constraints for DNA is rarely adequate to define a structure uniquely. The cross peaks appear in very narrow regions, and the overlaps further reduce the number of interpretable peaks. However, H-bonded base pairs impose additional constraints in a duplex DNA and significantly reduce the number of possible structures. It is difficult to simultaneously satisfy the NMR-derived constraints, which are usually within the same strand of the DNA duplex, and the condition of duplex formation. Invariably, the NOE[1] and *J*-derived constraints are limited in number, and several structures satisfy the experimental constraints; only a few of them form duplexes. With these considerations, we have developed software packages for (i) simulating 2D

[1] Abbreviations: COSY, correlated spectroscopy; NOE, nuclear Overhauser effect; NOESY, NOE correlated spectroscopy; TANDY, torsion angle approach to nucleic acid distance geometry; CORMA, complete relaxation matrix analysis; COMATOSE, complete matrix analysis torsion optimized structure; MARDIGRAS, matrix analysis of relaxation for discerning geometry of an aqueous structure; IRMA, iterative relaxation matrix approach; SICOS, simulation of correlated spectra; BURP, band-selective uniform response and phase; DG, distance geometry; RMSD, root mean square deviation.

0006-2960/93/0432-9053$04.00/0

NMR spectra so that the accuracy of the parameters derived can be substantially enhanced (Majumdar & Hosur, 1992) and (ii) building DNA models with any kind of user-defined constraints. The structure of the 14-mer oligonucleotide has been refined by a combined use of 2D NMR, molecular modeling, spectral simulations, and distance geometry calculations.

MATERIALS AND METHODS

(*a*) *Synthesis and Purification.* The 14-mer (d-GGATTGGCCAATCC) was synthesized using an automated DNA synthesizer (Applied Biosystems Model 381A) using the standard β-cyanoethyl phosphoramidite chemistry (Lewin, 1987). The 5′-dimethoxytrityl (DMTr) group was not removed at the end of the synthesis. After usual ammonia treatment, the oligomer was purified twice by reversed-phase HPLC, using a volatile buffer system (Majumder et al., 1986). The final purified product was dried in vacuo, solubilized in water, and passed through a Dowex 50X8 (sodium form) column in order to get the sodium salt of the oligomer. The purity of this compound was reconfirmed by both RP-HPLC and ^{32}P-labeling at the 5′-end, followed by polyacrylamide gel electrophoresis and autoradiography.

(*b*) *NMR Experiments.* NMR spectra were recorded on a Bruker AM 500 spectrometer operating at 500 MHz for ^{1}H. The sample was prepared by dissolving 14 mg of the oligonucleotide in 0.4 mL of 20 mM phosphate buffer containing 0.2 M NaCl and 1 mM ethylenediaminetetraacetate (EDTA). The DNA concentration on the duplex basis is approximately 4 mM. The pH of the solution was 7.0. For experiments in ^2H_2O, the solution was lyophilized several times from ^2H_2O and finally made up with 99.95% ^2H_2O. For the experiments in H_2O, the lyophilized sample was dissolved in a mixture of 80% H_2O and 20% ^2H_2O. In such experiments, the strong H_2O signal was suppressed by the jump-and-return sequence (Plateau & Gueron, 1982). All experiments have been carried out at 25–30 °C. The chemical shifts have been expressed with respect to the internal standard TSP.

Data for two-dimensional phase-sensitive COSY (90°–t_1–90°–t_2) spectra have been collected with 2048 t_2 points for each of the 512 t_1 experiments. The data were collected in quadrature along t_1 using the TPPI (time-proportional phase incrementation) procedure (Redfield & Kunz, 1975; Marion & Wuthrich, 1983). The data were zero filled to 1024 along t_1, multiplied by sine-bell window functions phase shifted by $\pi/16$, and Fourier transformed along the two axes in a sequential manner. Two ω_1-scaled phase-sensitive COSY-45 [Hosur et al., 1985a; see review by Hosur (1990)] spectra of the sample in ^2H_2O have been recorded with the shift scaling factor $\alpha = 0.5$ and 0.8, respectively. The data consisted of 2048 points along t_2 and 450 points along t_1. This was zero filled to 4096 along t_2 and 2048 along t_1, prior to window multiplication and 2D Fourier transformation. Sine-bell windows phase shifted by $\pi/32$ along t_2 and $\pi/16$ along t_1 were employed for this purpose. Final digital resolution in the $\alpha = 0.5$ spectrum was 1.83 Hz/point along both ω_2 and ω_1 axes. A relayed COSY spectrum (Eich et al., 1982) of the 14-mer has also been recorded with the same parameters. To enhance the ω_1 resolution in COSY spectra, a semiselective COSY experiment was performed with the first 90° pulse being replaced by a band-selective BURP pulse (Geen & Freeman, 1991). The BURP pulse was positioned in the middle of the H2′,H2″ region and a bandwidth of 900 Hz was selected using a pulse of 10-ms duration. A total of 192 t_1 experiments were performed with 2048 t_2 points each, and the data were processed as in other cases. The final digital resolution in the processed spectrum was 1.95 and 0.88 Hz/point along the ω_2 and ω_1 axes, respectively.

Phase-sensitive NOESY spectra (90°–t_1–90°–τ_m–90°–t_2) have been recorded with several mixing times (τ_m) of 40, 50, 80, 120, 150, 270, and 300 ms of the sample in ^2H_2O solution. The data consisted of 2048 points along t_2 and between 400 and 512 points along t_1 for different spectra. The data at 40, 80, 120, and 150 ms have been collected in one go with identical experimental conditions for the purpose of NOE buildups and structure analysis. These data have been processed on a X-32 data station or the Silicon Graphics IRIS 4D/70G work station, using UXNMR and MADNMR (P. Darba, unpublished) softwares, respectively. A 200-ms NOESY spectrum of the sample was recorded with 2048 data points along t_2 and 512 points along t_1. The data were zero filled to 1024 along t_1, multiplied by sine-bell window functions phase shifted by $\pi/2$, and Fourier transformed along both the axes. The final digital resolution in the processed spectrum was 1.97 and 3.94 Hz/point along the ω_2 and ω_1 axes, respectively. This NOESY spectrum was used for the NOESY simulation procedure. A 150-ms NOESY-BURP (Mukhopadhyay et al., 1992) was performed with the pulse sequence 90–t_1–90–τ_m–[BURP]–acquire (t_2) to select a vertical section centered around the H3′, H4′, H5′,H5″ region of the full NOESY spectrum. The H1′–H4′ and H8/H6–H3′ cross peaks in this region of the spectrum were used to get rough estimates of H4′ and H3′ multiplet structures, respectively.

Spin–lattice relaxation time (T_1) measurements have been carried out on the oligonucleotide using the inversion recovery 180°–τ–90°–acquire pulse sequence for estimating the correlation times in the molecule. τ values of 0.04, 0.06, 0.08, 0.2, 0.6, 0.8, 2.0, 4.0, 6.0, 10.0, and 14.0 s were used for monitoring the recovery of the magnetization. Lower bounds of T_2 values have been obtained from line-width measurements. The correlation times were then calculated using the formula (Woessner et al., 1962)

$$\tau_c \simeq (3T_2/T_1)^{-1/2} 2\omega_0^{-1} \quad (1)$$

where ω_0 is the spectrometer frequency of measurement. The correlation times τ_c thus obtained have been used for the simulation of NOESY spectra.

(*c*) *Simulation of Cross-Peak Patterns in J-Correlated Spectra.* A software package, SICOS, has been developed for the simulation of *J*-correlated spectra to obtain the *J*-coupling constant information (Majumdar, 1990; Majumdar & Hosur, 1992). Figure 1 shows an outline of such an algorithm. Starting from an input of the pulse sequence of the experiment and approximate values of chemical shifts and coupling constants relevant for a particular cross peak, a table of frequencies, amplitudes, and phases in the cross peak is generated by explicit density matrix calculations. The above output is converted into a 2D time domain data set $S(t_1, t_2)$ into which are incorporated the line-width parameters. This discrete data set is then Fourier transformed in a fashion identical to that used in processing experimental data, and a digitized frequency domain cross peak is produced. This is compared with the experimental peak at the contour plot level. The coupling parameters are varied until a superimposable fit is obtained. Our experience with the above algorithm is that if the components in the cross peak are well resolved, the simulations can sense a variation of coupling constants to the extent of ±0.2 Hz.

(*d*) *Quantification of Experimental NOEs.* Integration of the NOE cross peaks is a crucial step in structure determination by NMR spectroscopy (Wüthrich, 1986). The peak intensities carry information on the interproton distances, and any error

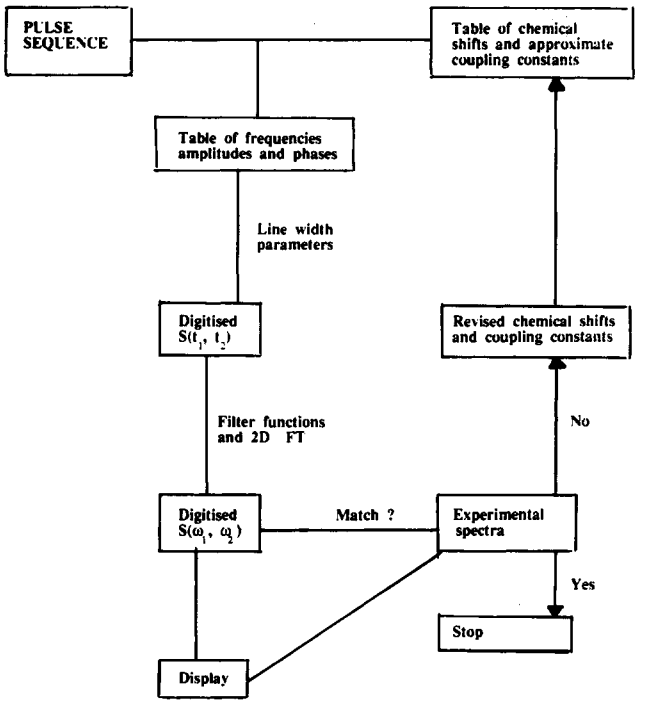

FIGURE 1: Flow chart for simulating cross peaks in *J*-correlated spectra using the program SICOS.

in peak quantification percolates into distance information and finally in the structure. Because the experimental spectra are digital in nature, the most common approach to peak quantification has been to sum up the values of the points spanned by any particular peak. With this approximation, the higher the digital resolution in the spectra, the better the estimate of the peak intensity. In addition, the noise level in the spectrum plays a crucial role in the accuracy of estimated peak intensities. Since the noise level may be different in different portions of the spectrum, each peak should be treated differently for estimating their volumes. In our approach a separate threshold level is selected for each peak, and the volume above this threshold is determined by summing up the magnitudes of the points in excess of the threshold value. These are stored along with the respective thresholds for all the usable cross peaks in the spectrum. Further, the peak intensities are calibrated with respect to a chosen cross peak for internal consistency and ease of comparisons.

(*e*) *Calculation of NOEs and Their Scaling To Match Experimental Intensities.* The process of calculation of NOEs involves the solution of a set of coupled differential equations (Solomon, 1955; Macura & Ernst, 1980)

$$\partial \mathbf{M}/\partial t = -\mathbf{RM} \qquad (2)$$

where **M** is the magnetization vector describing the deviation from thermal equilibrium of the various spins and **R** is the relaxation matrix, whose elements represent dipole–dipole cross relaxations between pairs of spins. Under conditions of isotropic tumbling and a single correlation time for the motions, the elements of **R** are explicitly given as

$$R_{kl} = K\tau_c/r_{kl} \qquad (3)$$

where K is a constant, τ_c is the correlation time, and r_{kl} is the distance between spins k and l.

Two approaches have been used previously for solution of eq 2 and consequently for simulation of NOESY spectra: (i) numerical integration of the system of equations [program BKCALC of Dennis Hare, unpublished, but used by Banks et al. (1989), Reid et al. (1989), and Nerdal et al. (1989)]; (ii) explicit relaxation matrix calculations (Keepers & James, 1984; Borgias & James, 1988, 1990). This approach has been used in the programs CORMA, COMATOSE, IRMA, and MARDIGRAS [see Borgias et al. (1990) for review]. These programs differ in the intricate details of dealing with experimental spectra in making actual comparison and in the iterative procedures of simulations. The simulation algorithm SIMNOE developed in our laboratory (Majumdar & Hosur, 1992; Nibedita et al., 1992a) adopts the relaxation matrix approach with explicit diagonalization of **R**. In this case, the solution of eq 2 is written as

$$\mathbf{M}(\tau_m) = \exp(-\mathbf{R}\tau_m)\mathbf{M}(0) \qquad (4)$$

The NOE intensities are proportional to $\mathbf{M}(\tau_m)$, which in turn is dependent on all the cross-relaxation rates in the spin network (Macura & Ernst, 1980). If Ω is the diagonal matrix of eigenvalues of **R** and χ is the matrix of eigenvectors of **R**, then the matrix of NOE cross-peak intensities, **C**, is calculated as

$$\mathbf{C} = \chi \exp(-\Omega\tau_m)\chi^{-1} \qquad (5)$$

Figure 2 shows the flow chart of the simulation algorithm. Starting from an initial structure of the DNA, a matrix of interproton distances is generated. From these, the relaxation matrix, whose elements are the cross-relaxation and autorelaxation rates, is generated on the assumption that the molecular motions in solution satisfy the condition $\omega_0\tau_c \gg 1$, which is usually true at 500 MHz for duplex oligonucleotides longer than 5–6 base pairs. From the various cross-relaxation rates, the expected NOE intensities (C_{ij}) are calculated.

The NOE intensities thus calculated are to be compared with experimental intensities. However, these two are not directly comparable since the latter are "digital" in nature while the former are "analog" numbers. To circumvent this problem, the calculated intensities are scaled in one or the other ways by different authors (Baleja et al., 1990a,b; Gochin et al., 1990; Zhou et al., 1987). These procedures, however, do not attempt to convert the calculated intensities into digital intensities. We have shown recently that, for a reliable comparison of the two intensities, digitization of the calculated peaks is the best approach (Nibedita et al., 1992a). The simplest method to achieve this is to generate time domain data and Fourier transform it to the frequency domain as described in the case of *J*-correlated spectra. However, for an iterative structure optimization, this becomes very slow and computationally intensive. As a remedy, we have developed a new scaling procedure (Nibedita et al., 1992a) which eliminates the need for repetitive generation of time domain data, Fourier transformation to generate frequency domain data, etc., and speeds up the exercise significantly. This procedure is also shown schematically in Figure 2. The underlying idea is that, in a cross peak with in-phase multiplet structure and finite line widths, the digital intensity calculated as a sum of points scales by the same factor as the analog intensity of the peak. The digital intensities are calculated above the same threshold values as has been done for the experimental peaks. Comparison of the digitial intensities so obtained with experimental intensities, after suitable normalization, permits optimization of the structure of the DNA segment. The DNA structure is altered and the process of comparison of experimental and calculated intensities repeated until a best fit is obtained between the two intensities. In order to facilitate this procedure, a separate program called MODEST (R. Ajay Kumar, unpublished) is used which changes the torsion angles α, β, γ, ϵ, ζ, and χ and the sugar pucker P in real time and provides an analysis of peak intensities by a color-coded display. Hard contacts are also checked

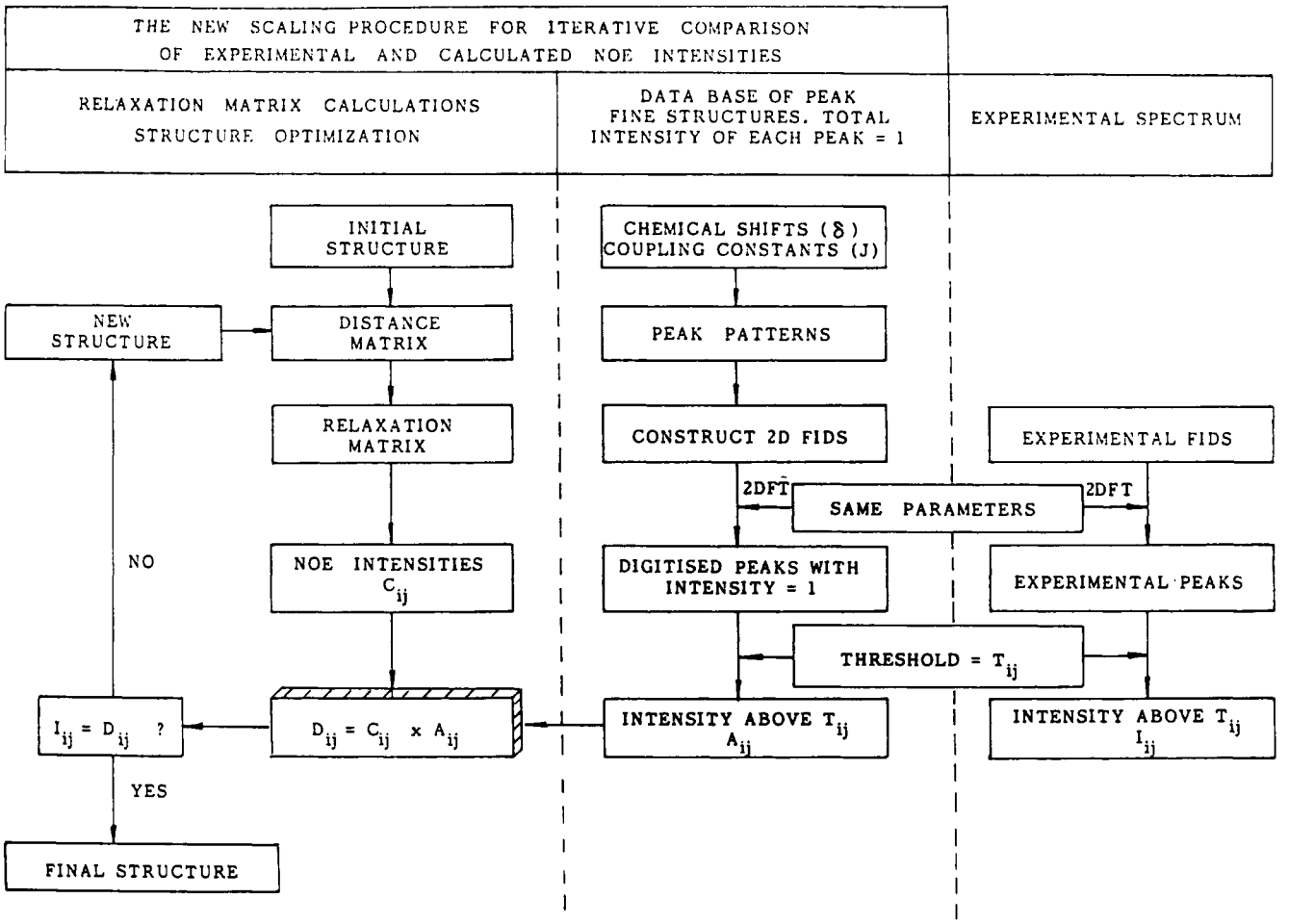

FIGURE 2: Flow chart for the NOESY simulation algorithm SIMNOE.

during the process so that unacceptable structures are eliminated straightway. The fit between experimental and simulated peak intensities is checked after each iteration. In this context we have recently investigated the relation between the goodness of fit and the goodness of the derived structure (Nibedita et al., 1992b). While the fit cannot be demanded to be better than the uncertainties in the experimental intensities, the higher the accuracy of the fit and the larger the number of peaks between adjacent nucleotide units, the more unique and precise is the structure characterization. Certain torsion angles are better determined than others in this exercise.

(f) Distance Geometry Calculations in Torsion Angle Space. The NOE simulation and structure optimization process provides a structure consistent with the peak intensities. However, because of the limited number of internucleotide NOE peaks generally available, the conformational sampling is usually not exhaustive, and a greater search will be needed to find a family of structures consistent with the NMR data. This can be achieved by distance geometry calculations. Several authors have used distance geometry algorithms based on the metric matrix approach, but there has been some amount of confusion and uncertainty in the reliability of the results obtained by these calculations (Metzler et al., 1989; Kim et al., 1992). We have used the distance geometry programs TANDY (Ajay Kumar et al., 1991) and TANDY-2S (R. Ajay Kumar, unpublished) developed in our laboratory, the former being applicable to single-stranded DNA and the latter to duplex DNA. Both work in the torsion angle space and minimize a variable target function to a level of satisfaction defined by the user. Information about particular torsion angles is imposed as a constraint so that the search operation can be reduced and sampling can be better. The distance constraint set required for the program is obtained from the NOE fit structure, and certain bounds are specified on these as tolerable variations in the process of target function minimization. TANDY-2S utilizes an angle geometry algorithm (Ajay Kumar, 1992) for optimization of H-bond geometry and base-pair formation. The distance geometry generated structures are again used to calculate NOE spectra so that better fits with experimental spectra can be obtained, and this process of NOE simulation and DG calculation is carried out in succession repeatedly until a good convergence and satisfactory fit with experimental data are obtained.

RESULTS

(a) Resonance Assignment. Figure 3a shows the one-dimensional NMR spectrum of the 14-mer molecule. Sequence-specific assignment of the nonexchangeable proton resonances has been obtained using well-established procedures based on 2D COSY and NOESY spectra (Hare et al., 1983; Scheek et al., 1983; Hosur et al., 1985b). Figure 4 shows an indicative sequential $(H2'')_{i-1} \rightarrow (H8/H6)_i \rightarrow (H2'')_i \rightarrow (H8/H6)_{i+1}$ connectivity pattern in a NOESY (τ_m = 300 ms) spectrum of the oligomer. Similar connectivities have been observed in other portions of the spectrum, leading to overall self-consistent assignment of all nonexchangeable protons except the H5′,H5″ protons of a few nucleotide units. Stereospecific assignment of the H2′/H2″ protons in the sugar rings has been obtained from a 50-ms NOESY spectrum by monitoring the relative intensities of H1′–H2′ and H1′–H2″ cross peaks (Hosur et al., 1986).

Assignment of the exchangeable imino protons (Figure 3b) has been obtained on the basis of chemical shifts, relative

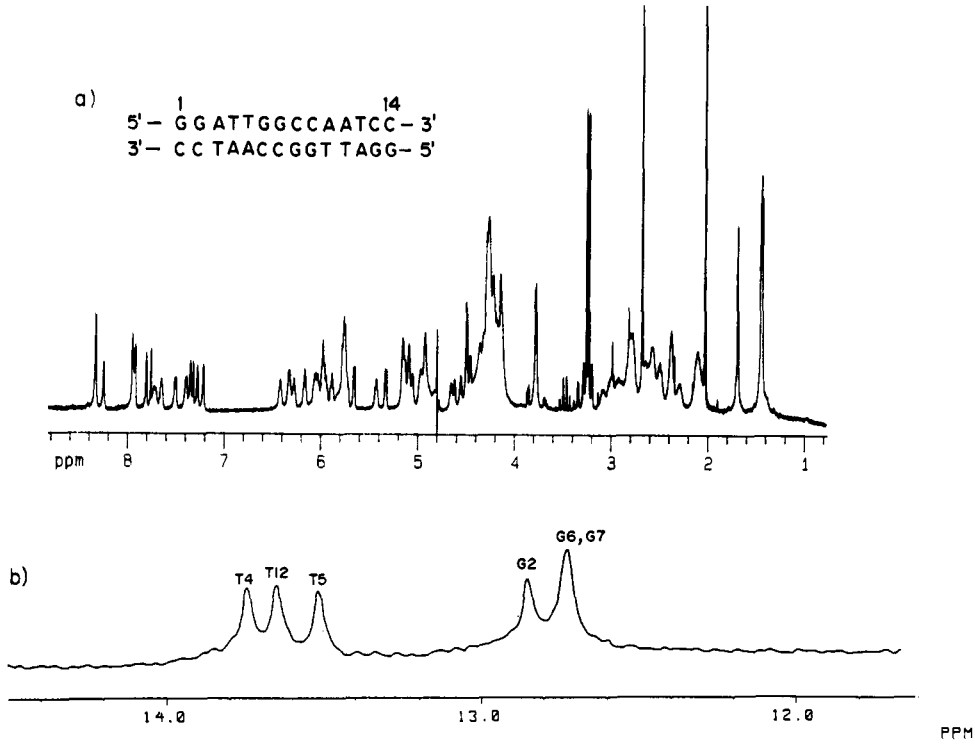

FIGURE 3: 1D NMR spectrum of d-GGATTGGCCAATCC (14-mer) in (a) 99.95% D_2O solution and (b) 80:20 mixture of H_2O and D_2O.

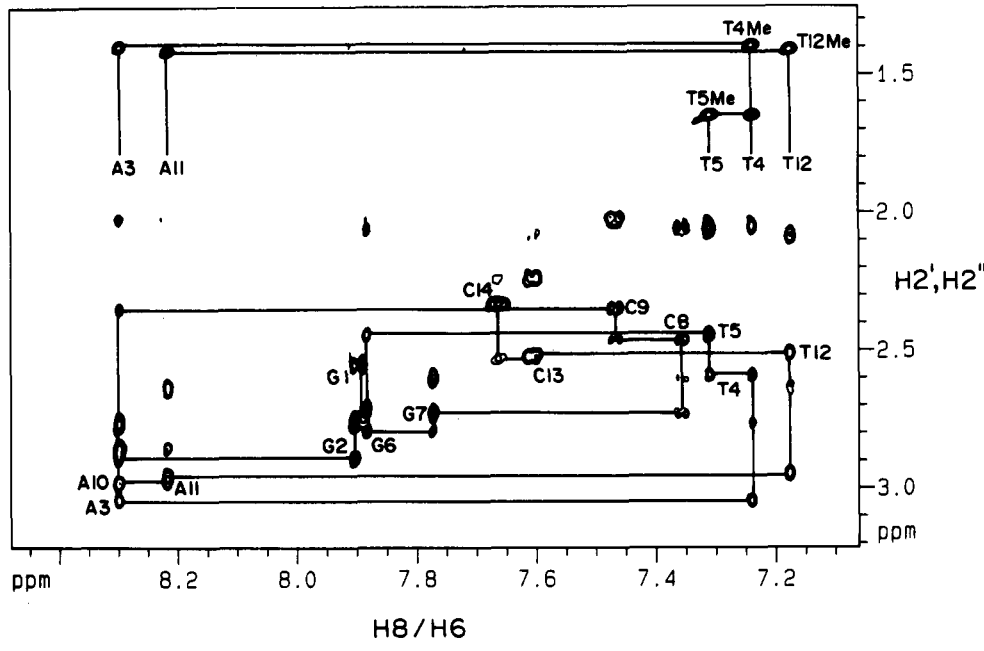

FIGURE 4: H6/H8–H2′,H2″ region of a 300-ms NOESY spectrum of d-GGATTGGCCAATCC showing sequential resonance assignments via $H2''_{i-1}$–$(H6/H8)_i$ connectivity.

intensities, and line-width considerations. The general guidelines in this context are as follows: (i) The exterior imino protons exchange more rapidly with water and hence have larger line widths. (ii) The T-NH protons in the Watson–Crick base-paired duplexes resonate at lower fields compared to the G-NH protons. (iii) Purines produce larger upfield shifts on the adjacent nucleotide units than pyrimidines. (iv) In self-complementary sequences, a 2-fold symmetry in the structure results in equivalences of base pairs at equal distances from the ends. Figure 3b shows two groups of resonances which can be identified as belonging to A-T and G-C pairs. The number and intensities of the resonances in the two groups permit easy identification of the A-T and G-C groups. Within each group, specific assignments have been based on intensities, nearest-neighbor considerations, and comparison with assignments in similar stretches (-ATTA- and -ATC-) in other oligonucleotides (Patel et al., 1983; Chazin et al., 1986). The chemical shifts of the assigned protons are listed in Table I.

(*b*) *Coupling Constants in the Deoxyribose Rings.* Several methods have been reported to obtain values of the coupling constant from experimental *J*-correlated spectra [see Hosur et al. (1988a) and van de Ven and Hilbers (1988) for reviews]. In earlier work, the relative intensities of various cross peaks in COSY spectra were used (Hosur et al., 1985b, 1986; Ravikumar et al., 1985). Rinkel and Altona (1987) obtained rough estimates of coupling constants by estimating the sums of coupling constants from partially resolved cross-peak fine structures in the absolute value COSY spectra. Subsequently,

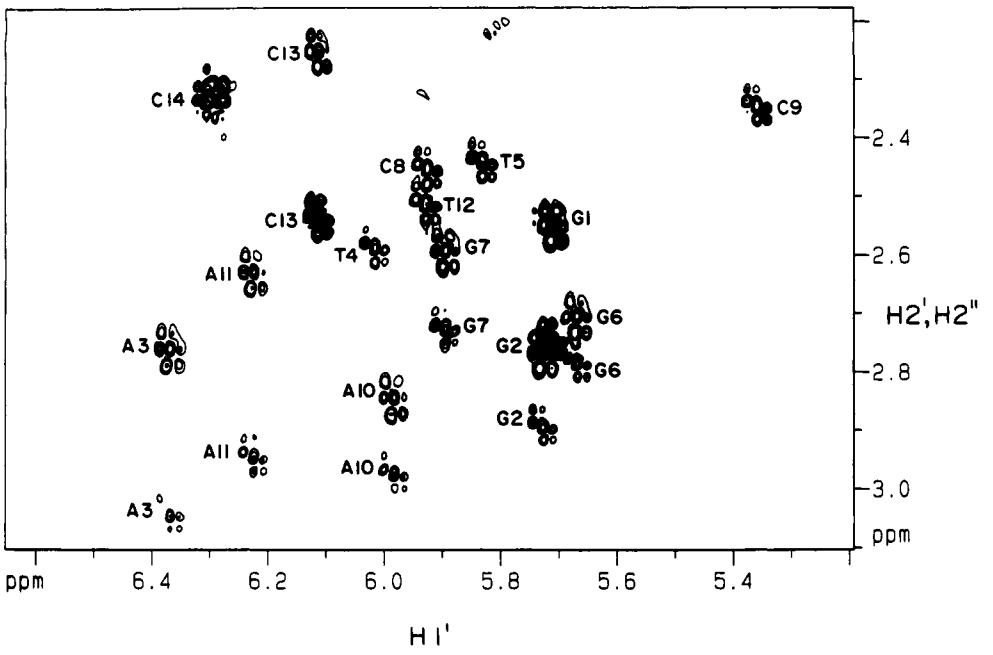

FIGURE 5: H1′–H2′,H2″ cross-peak region of a BURP-COSY-45 spectrum of the 14-mer.

Table I: Chemical Shifts (ppm)

residue	H1′	H2′	H2″	H3′	H4′	H2	H5	H6/H8	TMe	NH
G1	5.71	2.55	2.75	4.89	4.26			7.89		
G2	5.73	2.76	2.88	5.10	4.46			7.91		12.85
A3	6.37	2.76	3.04	5.12	4.58	7.92		8.30		
T4	6.02	2.05	2.59	4.92	4.29			7.23	1.40	13.74
T5	5.84	2.06	2.44	4.94	4.19			7.31	1.65	13.52
G6	5.67	2.71	2.78	5.06	4.43			7.87		12.73
G7	5.90	2.59	2.72	5.02	4.45			7.77		12.73
C8	5.93	2.05	2.46	4.84	4.21		5.28	7.36		
C9	5.37	2.02	2.35	4.86	4.10		5.61	7.47		
A10	5.98	2.85	2.97	5.12	4.46	7.27		8.30		
A11	6.23	2.63	2.94	5.06	4.52	7.72		8.22		
T12	5.93	2.08	2.50	4.89	4.23			7.18	1.42	13.65
C13	6.12	2.25	2.53	4.89	4.23		5.70	7.61		
C14	6.30	2.32	2.32	4.62	4.10		5.74	7.68		

procedures based on addition and subtraction of cross peaks and diagonal peaks in double-quantum-filtered COSY spectra were employed (Zhou et al., 1987). These procedures yield qualitative estimates of coupling constants, and the structural information derived therefrom is also qualitative. Efforts to improve the quality of *J*-correlated spectra, by resolving the components in the cross peaks, led to the development of ω_1-scaling pulse sequences (Hosur et al., 1985a,c, 1988b, 1990) and use of selective excitations [see Ernst et al. (1987) and Kessler et al. (1991) for reviews and Geen and Freeman (1991)]. These have been successfully used in several systems for measurement of coupling constants (Chary et al., 1988, 1989; Sheth et al., 1989; Bax & Lerner, 1988). In parallel, procedures have been developed for simulating cross-peak fine structures in *J*-correlated spectra (Widmer & Wüthrich, 1986; Hosur et al., 1989; Majumdar & Hosur, 1992). These developments provide significant improvement in the accuracy of coupling constant values derived from the spectra (Chazin et al., 1986; Gochin et al., 1990; Celda et al., 1989; Sheth et al., 1989).

We have simulated different kinds of *J*-correlated spectra of the 14-mer. Figure 5 shows as an example the H1′–H2′/H2″ cross-peak region of a phase-sensitive BURP-COSY-45 spectrum. We believe that this is a better spectrum to simulate than the double-quantum-filtered COSY spectrum (Chazin et al., 1986; Gochin et al., 1990) since it has a smaller number of components and the resolution is better.

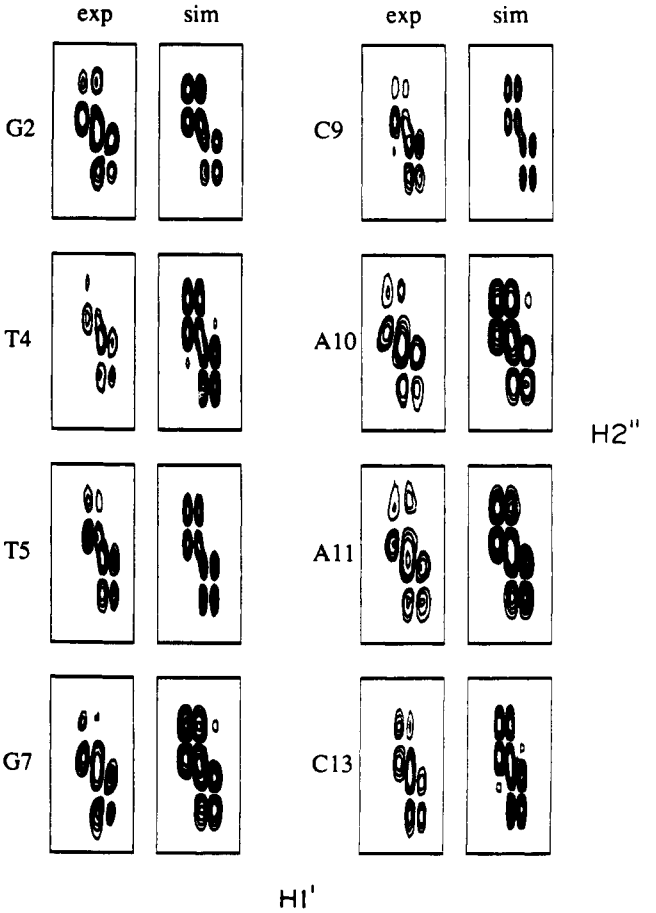

FIGURE 6: Best fit simulations of some H1′–H2″ cross peaks of the BURP-COSY-45 spectrum. sim = simulated peaks and exp = corresponding experimental peaks. A uniform line width of 2.7 Hz along both axes has been used in all the simulations.

(*1*) *H1′–H2″ Peaks.* These peaks possess the $J(1'2')$ and $J(1'2'')$ information along the ω_2 axis and $J(1'2'')$, $J(2'2'')$, and $J(2''3')$ information along the ω_1 axis. In most cases the $J(2''3')$ coupling constant is very small and does not show explicit splitting in the spectra. Figure 6 shows some best fit simulations of H1′–H2″ peaks along with the experimental peaks on an enlarged scale. In most cases the fits were very

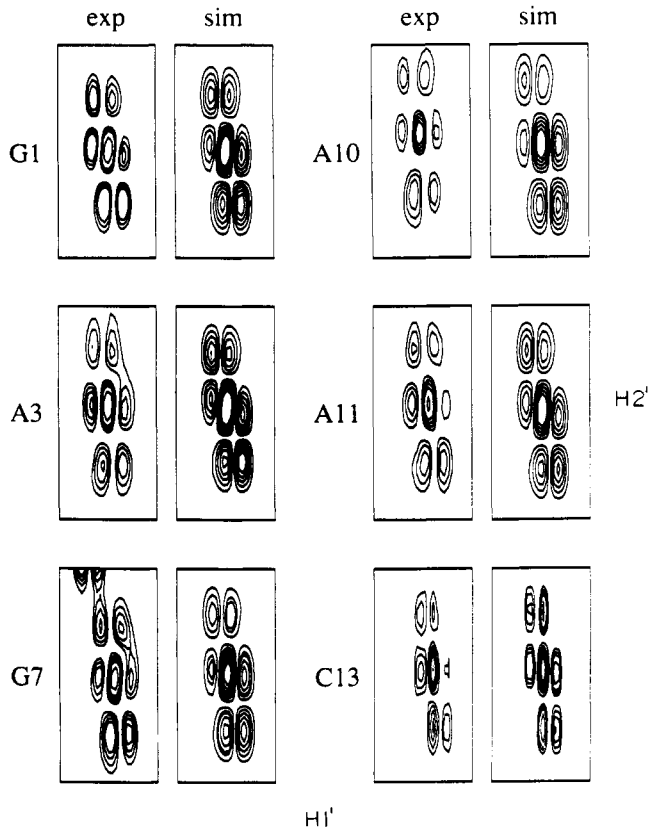

FIGURE 7: Best fit simulations of some H1′–H2′ cross peaks of the BURP-COSY-45 spectrum. sim = simulated peaks and exp = corresponding experimental peaks. A uniform line width of 2.7 Hz has been used in all the cases.

good. The search directions for optimizing the coupling constants for obtaining best fits were facilitated by a data base of simulations (Majumdar & Hosur, 1992) for different sugar geometries, for different line widths, and for different digital resolutions. By and large the individual peaks were simulated separately. In a few cases which showed extensive overlap, the overlapping peaks were simulated together.

(*2*) *H1′–H2′ Peaks.* These peaks possess the $J(1'2')$ and $J(1'2'')$ information along the ω_2 axis and $J(1'2')$, $J(2'2'')$ and $J(2'3')$ information along the ω_1 axis. The simulation of many of these peaks was facilitated by the prior simulation of the H1′–H2″ peaks of the corresponding nucleotide units; since $J(1'2')$, $J(1'2'')$, and $J(2'2'')$ were already available, $J(2'3')$ alone was varied to obtain the best fits. Simultaneous fitting of both H1′–H2′ and H1′–H2″ cross-peak patterns eliminated some of the ambiguities and resulted in a unique set of coupling constants. Figure 7 shows a few best fit simulations of these peaks along with their experimental counterparts.

The coupling constants derived from the above simulations are listed in Table II. From the sensitivity of the fits to changes in coupling constant values, we could ascertain the accuracy of the above values to be ±0.2 Hz. This is a significant improvement over previous reports based on measurement of *J*-values from peak-to-peak separations, where the accuracy is at best limited to ±0.5 Hz.

(*c*) *Coupling Constants in the DNA Backbone.* There are three coupling constants in the DNA backbone which are significant from the structural point of view, namely, H4′–H5′,H5″, H3′–P, and P–H5′,H5″. Strategies have been proposed to measure the latter heteronuclear ^1H–^{32}P coupling constants (Sklenar & Bax, 1987; Hosur et al., 1990). We have recently proposed a strategy based on the use of NOESY spectra for the measurement of H4′–H5′,H5″ coupling

Table II: Coupling Constants (Hz) Obtained from Simulation of Cross Peaks[a]

residue	1′–2′	1′–2″	2′–2″	2′–3′	2″–3′	Σ4′	Σ3′
G1	9.0	6.0	−13.5	6.0	1.0	13	13
G2	9.7	7.3	−15.0	6.0	2.3	12	14
A3	9.5	7.3	−14.0	6.7	3.5	12	*b*
T4	9.0	7.4	−15.0	6.0	4.0	14	14
T5	9.0	7.0	−15.0	6.0	3.0	12	18
G6	9.0	6.0	−13.0	6.5	4.0	14	14
G7	9.0	6.8	−14.0	6.8	3.0	11	16
C8	8.5	6.9	−16.0	6.0	4.0	*b*	*c*
C9	9.0	6.3	−15.3	6.0	2.5	*c*	*c*
A10	9.5	7.5	−14.5	7.5	3.0	10	*b*
A11	8.5	7.2	−15.0	6.5	4.0	13	15
T12	9.0	6.4	−16.0	6.0	4.5	*b*	12
C13	7.0	6.5	−14.5	7.0	3.5	16	16
C14	*b*	*b*	*b*	*b*	*b*	12	15

[a] The data in the last two columns have been obtained from measurement of total widths of the cross peaks in the BURP-NOESY spectrum of Figure 8. [b] These coupling constants could not be evaluated due to severe overlap of peaks. [c] These coupling constants could not be evaluated due to very weak cross peaks.

constants (Mukhopadhyay et al., 1992). The basic idea is that the NOESY spectrum also has coupling constant information in its cross peaks and there are a much larger number of peaks in the spectrum, implying that the multiplet structures of particular protons are available at many places in the spectrum in combination with different kinds of protons. Consequently, multiplet structures of some protons which are not easily accessible in *J*-correlated spectra are determinable from the NOESY spectrum. The multiplets of interest here are those of the H4′ and H3′ protons. Among the available cross peaks which carry this information we focus on the H1′–H4′ and H8/H6–H3′ cross peaks in the spectrum. These peaks are far removed from the diagonal and thus are devoid of complications due to tails. These peaks also possess simple fine structures. To resolve the multiplet components of these peaks, high-resolution data were collected by using a semiselective BURP observe pulse in the NOESY pulse sequence. These regions of an experimental spectrum are shown in Figure 8. Some of these peaks have been simulated to get a feel for their structure, and subsequently, upper bounds of the relevant coupling constants have been obtained by monitoring the total widths of the peaks and using prior knowledge of the other coupling constants. The derived coupling constant data are also listed in Table II, along with those obtained from *J*-correlated spectra. From these data, it can be concluded that the H3′–P, H4′–H5′, and H4′–H5″ coupling constants are less than 5 Hz in all the nucleotide units.

(*d*) *Correlation Time Measurements.* The correlation time (τ_c) constitutes an essential input in the calculation of NOE intensities from known structures. We have obtained these inputs for as many protons as possible by measuring the T_1 and T_2 relaxation times. Figure 9 shows illustrative inversion–recovery profiles, and the T_1's have been estimated to lie between 0.7 and 1.4 s in most of the base protons. From the estimates of T_2's obtained from line-width measurements, the correlation times have been estimated to be around 2–3 ns for most of the base protons. We could not obtain similar information on the sugar protons due to overlap problems in these resonances. Thus we have used a single correlation time of 2 ns for all the protons in our NOE calculations; some justification for this comes from the recent work of Reid et al. (1989).

(*e*) *Molecular Modeling and DG Structures.* Calculation of the structures consistent with the above coupling and NOE intensity data by DG methods requires proper translation of these data into distance constraints. Initial rate distances

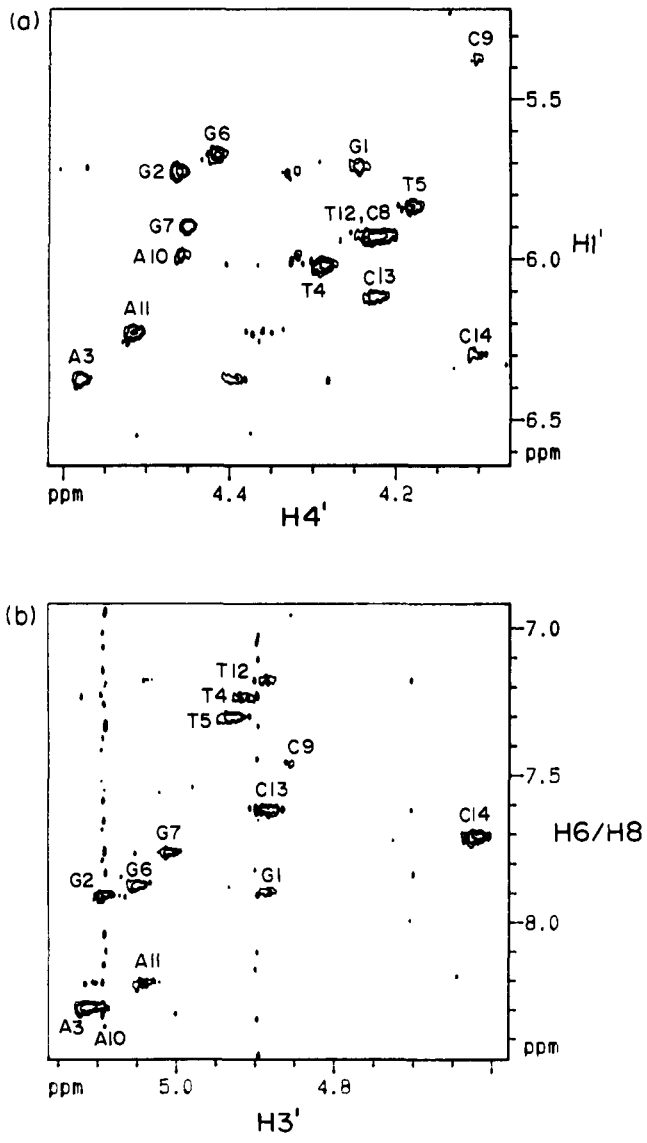

FIGURE 8: H8/H6–H3′ and H1′–H4′ cross-peak regions of a BURP-NOESY spectrum of the 14-mer. These have been used for estimating H4′ and H3′ multiplicities.

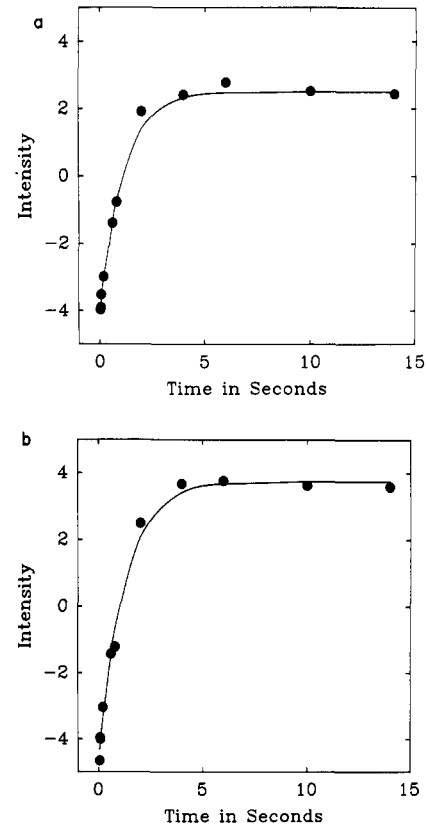

FIGURE 9: Inversion recovery profile for determination of T_1 of base protons using the equation $m_z = m_0[1 - b\{\exp(-t/T_1)\}]$: (a) A3H8 (parameters for the fit are $m_0 = 2.50$, $b = 2.65$, and $T_1 = 1.11$); (b) T12H6 (parameters for the fit are $m_0 = 3.75$, $b = 2.22$, and $T_1 = 1.23$).

Table III: Summary of Initial Constraints Used for Model Building and DG Calculations

constraint	no.	allowed deviation
coupling constants[a]	73	±0.1 Hz[b]
NOE intensities[c]	79	±20%
H-bonds	108	standard geometry[d]
steric contacts	all	yes/no

[a] The coupling constants include three-bond H–H coupling constants in the sugar ring, H4′–H5′,H5″ coupling constants in the backbone, and the H3′–P coupling constants. [b] For the H3′–P and H4′–H5′,H5″ coupling constants, the allowed deviation is 1 Hz. [c] The NOEs considered include intranucleotide and internucleotide H8/H6–H1′, H2′, H2″, H3′, and H1′–H4′ correlations. [d] Standard H-bond geometry taken from Saenger (1984).

have been used by some authors to derive the "initial" constraint set, and the structures have been refined using NOE intensities. Our approach deviates slightly from this. We have built a model interactively (using the program MODEST; R. Ajay Kumar, unpublished), continuously monitoring the calculated NOE intensities, H-bond parameters, coupling constants, and steric constraints, to derive the distance constraint set; a summary of the input constraints used in the above exercise is given in Table III. For each constraint certain bounds are defined as tolerance limits, and this depends on the confidence level of the determined parameter, namely, the NOE intensity or the coupling constant. The maximum deviation allowed in this range is 20% of the value of the parameter. Once all the specified constraints are satisfied in the interactive model building exercise, a new constraint set is derived from the structure at that instant, and this set includes all the initial constraints and also several other distance constraints which could not be determined from the data earlier. DG calculations are then carried out with such a constraint set. The constraint set is refined whenever the NOE fit is seen to be better than before, and the process is repeated to obtain better structures. Figure 10 shows a few converged DG generated structures superimposed. Figure 11 shows the best structure. The quality of the structures can be described in terms of an RMSD with respect to NOE intensities, which we have defined as

$$\mathrm{RMSD}_{\mathrm{NOE}} = [\sum_i (1 - E(i)/S(i))^2/N]^{1/2} \qquad (6)$$

where the summation runs over all the peaks (N), and $E(i)$ and $S(i)$ are the experimental and calculated intensities, respectively. For the structures shown in Figure 10 the RMSDs were in the range of 0.3–0.5, considering all the peaks belonging to the central 12 nucleotide units of the molecule, and for the best structure it was 0.27. The peak-to-peak fit of intensities for this structure is also shown in Figure 11. The terminal units were excluded since the terminal fraying and other motions were not included in our calculations.

DISCUSSION

We noted that the several structures of the 14-mer molecule, derived from the NMR data described above, differ from

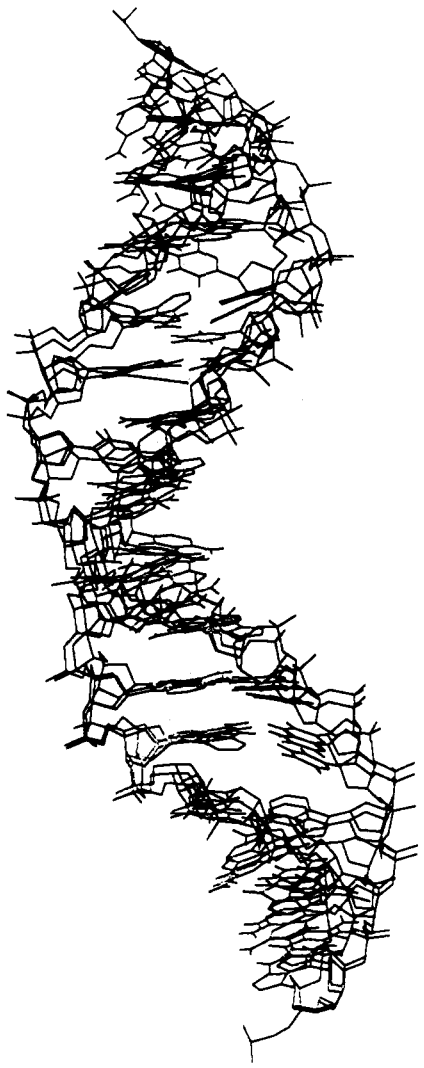

FIGURE 10: Superposition of four DG structures consistent with the NOE data and other constraints. The $RMSD_{NOE}$ values are in the range 0.3–0.5.

Table IV: Torsion Angles (deg) in the Final Structure

residue	α	β	γ	P	χ	ϵ	ζ
G1		208.5	30.8	121.0	227.3	170.4	275.2
G2	315.5	204.2	32.7	152.6	251.0	166.4	261.1
A3	313.4	219.7	25.5	180.2	249.1	158.6	256.7
T4	308.5	190.7	74.6	139.9	224.3	161.4	261.7
T5	271.7	208.7	71.9	119.1	233.3	160.2	277.2
G6	314.0	215.1	36.9	156.3	248.4	293.5	173.9
G7	234.8	140.4	73.6	168.1	240.6	185.1	262.7
C8	291.8	216.1	55.0	129.4	225.1	195.8	242.3
C9	278.5	164.0	97.0	126.5	225.6	163.7	246.4
A10	334.8	193.6	39.4	164.0	249.2	165.8	244.4
A11	342.0	181.7	28.3	161.1	234.7	177.2	218.5
T12	326.5	160.7	70.7	177.4	231.4	183.6	268.1
C13	316.8	186.6	31.7	122.5	240.8	169.7	274.4
C14	301.7	210.4	36.0	148.8	260.2		

each other only marginally in their overall features. In the distance geometry calculations, it was observed that irrespective of the starting points, and the randomization process inherent in the TANDY program (Ajay Kumar et al., 1991), the structures tended to converge to similar structures. Table IV lists the various torsion angles in the best structure mentioned above. These torsion angles have also been plotted against the base sequence in Figure 12 for better appreciation of the structural variations along the sequence of the molecule. In the following, we analyze the structural features of the molecule so as to put the results in the proper perspective, taking into consideration the strengths and weaknesses of the calculation procedures.

(*a*) *Conformations of the Deoxyribose Rings.* The conformations of the five-membered deoxyribose rings can be independently determined from two different inputs: (i) by interpretation of the five vicinal H–H coupling constants in the rings which depend on the pseudorotation angle P and (ii) from the NOE-derived intrasugar and internucleotide H–H distances. While both the NOE intensities and the coupling constants will be time averages, the NOEs will be either a $\langle r^{-3} \rangle$ or a $\langle r^{-6} \rangle$ average depending upon the relative rates of internal and overall tumbling motions [Krishnan et al., 1991; see review by James (1991)]. Because of such a strong distance dependence, the major contributor to the NOE will generally be a single conformation. Therefore, to demand consistency of results on sugar geometries from the two approaches, we first obtained the ranges of sugar geometries from the *J* values and then imposed these as constraints in the interactive model building to fit NOE intensity data and derive distance constraint sets for the DG calculations. Hence, the *P* values listed in Table IV are in a way an NOE-based selection of the possible sugar geometries in the molecule. It is seen that, in all cases (G2–C13), the sugar ring adopts a conformation between C1′-exo and C2′-endo geometries.

(*b*) *Glycosidic Torsion Angle* (χ). This is one of the best determined torsion angles from the NOE data. All the intranucleotide H8/H6–sugar NOEs and also the internucleotide NOEs involving base protons help in characterizing this torsion angle. In the 14-mer molecule we observe that the glycosidic torsion angle shows a systematic oscillation along the sequence of the molecule. All the values are in the anti domain.

(*c*) *Backbone Torsion Angles.* As in the case of sugar geometries, some of the backbone torsion angles, specifically γ and ϵ, are determinable from both coupling and NOE data. It is in fact necessary to use both these inputs here so as to be able to provide better and reliable estimates of the other torsion angles, which are derivable exclusively from NOE data which often may not be adequate to fix all the backbone torsion angles uniquely. Thus we have used estimates or upper limits of H3′–P and H4′–H5′,H5″ coupling constants to fix domains of ϵ and γ torsion angles in the interactive model building exercise; ϵ was fixed to the range 140–220° or 230–310° and γ was fixed to the range 20–100°.

In the final best NOE fit structure most of the backbone torsion angles are seen to be in the domain expected of right-handed structures, with small variations along the sequence, except at the G6–G7 step which shows some significant changes (see Figure 12). The ϵ angle adopts a rare, but energetically allowed value around 290° (Govil & Hosur, 1982; Gupta et al., 1980; Hosur et al., 1990), and likewise, the torsion angle ζ is also in the rare trans domain. The α and β values at this place are rather low, and the torsion angle γ is on the higher side. These variations in torsion angles indicate a structural transition, resembling a B_I–B_{II} transition, at the beginning of the GCCAAT recognition motif in the molecule. Recent molecular dynamics studies have investigated the profile of B_I–B_{II} transitions in great detail (Ravishanker et al., 1989).

(*d*) *Base Stacking and Helical Parameters.* The structural models displayed in Figures 10 and 11 indicate that the base pairs are inclined to the helical axis, as is normally seen in A-DNA. But the sugars and many of the torsion angles belong to the B-DNA family. The structure has a wide major groove and a very narrow minor groove. Determination and analysis of the base-pair characteristics and the helical parameters of the best structure shown in Figure 11 have been performed

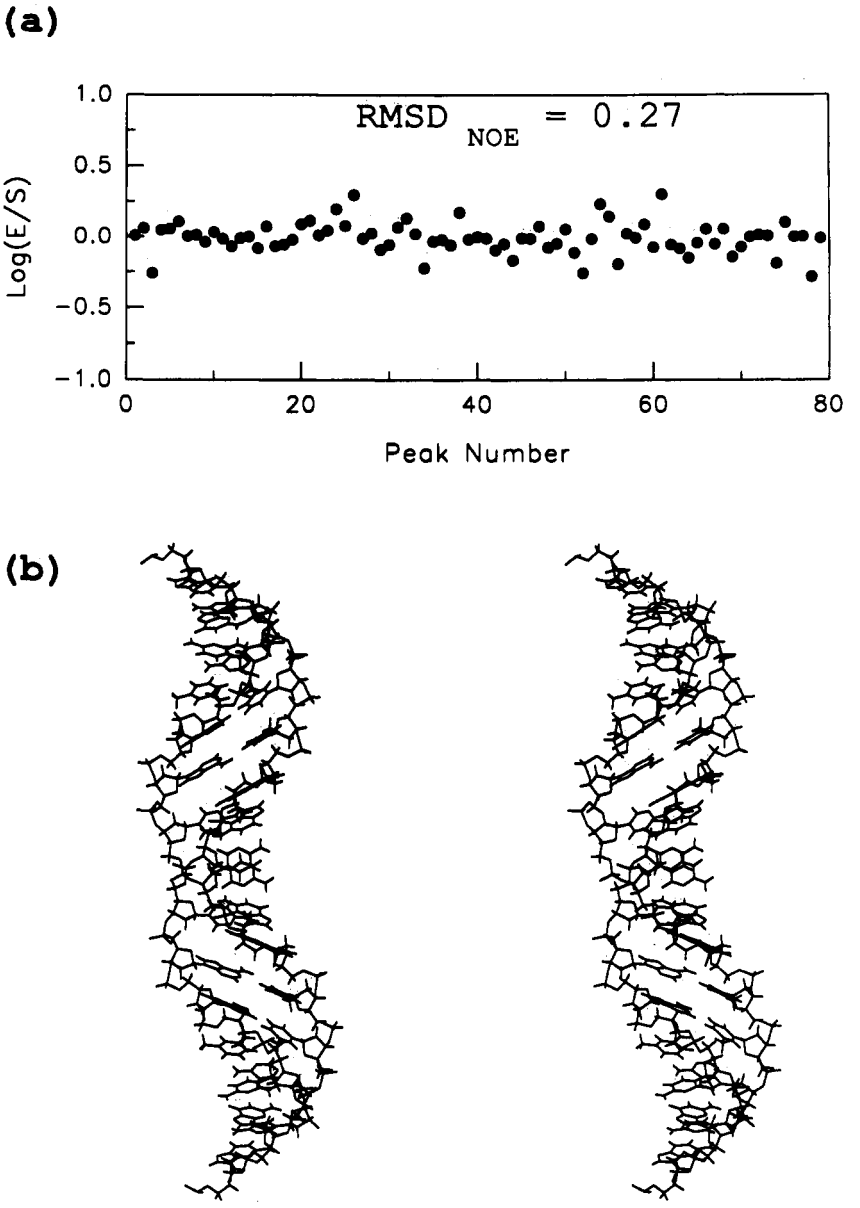

FIGURE 11: Stereoview of the best structure derived from the NMR data. The top panel shows the peak-to-peak fit of the calculated and experimental NOE intensities for this structure. The base pairs are tilted as in A-DNA models, and a bend is seen at the center of the molecule. The structure has a wide major groove and a narrow minor groove.

using the program NUPARM (kindly supplied by Dr. M. Bansal) (Bhattacharyya & Bansal, 1989), and the details are shown in Figure 13. The helical axis is not seen to be well-defined, and the molecule bends in the middle. A local helical axis can be defined, and this changes direction continuously through the length of the molecule. In all the parameters, the duplex is symmetric with respect to the center, but in each half there is a systematic variation. The symmetry with respect to the center must be expected in view of the self-complementarity of the sequence, but the variation in each half is an interesting feature of the structure.

In the above discussion we have not explicitly considered local motions and their influence on the NOE intensities. Variabilities in correlation times, anisotropic motions, variabilities in leakage rates, etc. (Withka et al., 1992; Baleja et al., 1990a,b) have not been explicitly included in the calculations, and expectedly, the NOE fit could not be obtained to better than an $RMSD_{NOE}$ of 0.27. However, as has been recently shown (Nibedita et al., 1992b), the NOE fit is a sensitive function of the structure, and a small change in the structure (less than 10° variation in torsion angles) can change the $RMSD_{NOE}$ dramatically. Thus, it is expected that time averaging over slightly different conformations can significantly reduce the $RMSD_{NOE}$ values. Explicit inclusion of anisotropic motions (Withka et al., 1992) is also expected to improve the NOE fit. While these factors help in improving the fit between calculated and experimental NOE intensities, the gross features of the structure as have been derived in the present analysis are not likely to change significantly.

CONCLUSIONS

We have determined in this paper the solution structure of a 14-mer DNA segment containing the recognition motif GCCAAT of several transcription factors in various bacterial systems. Several important observations have been made from the present structure determination. First, the NOE data can indeed define a narrow family of structures for a DNA duplex. Repeated distance geometry (DG) calculations on the duplex produced very similar structures irrespective of the starting structure, although the NOE fits showed some variations. We do not claim that the conformational search is exhaustive, but we do feel from the progress of DG

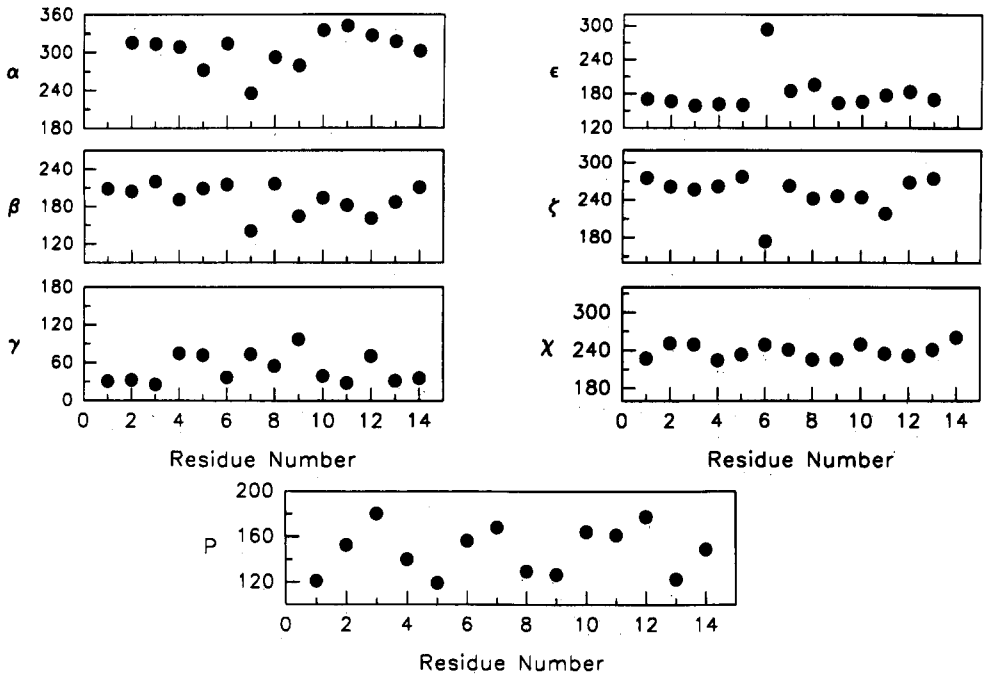

FIGURE 12: Plot of the various torsional angles vs sequence in the final structure of Figure 11. Large deviations can be seen in the ϵ, ζ, α, β, and γ torsion angles in the central portion of the molecule.

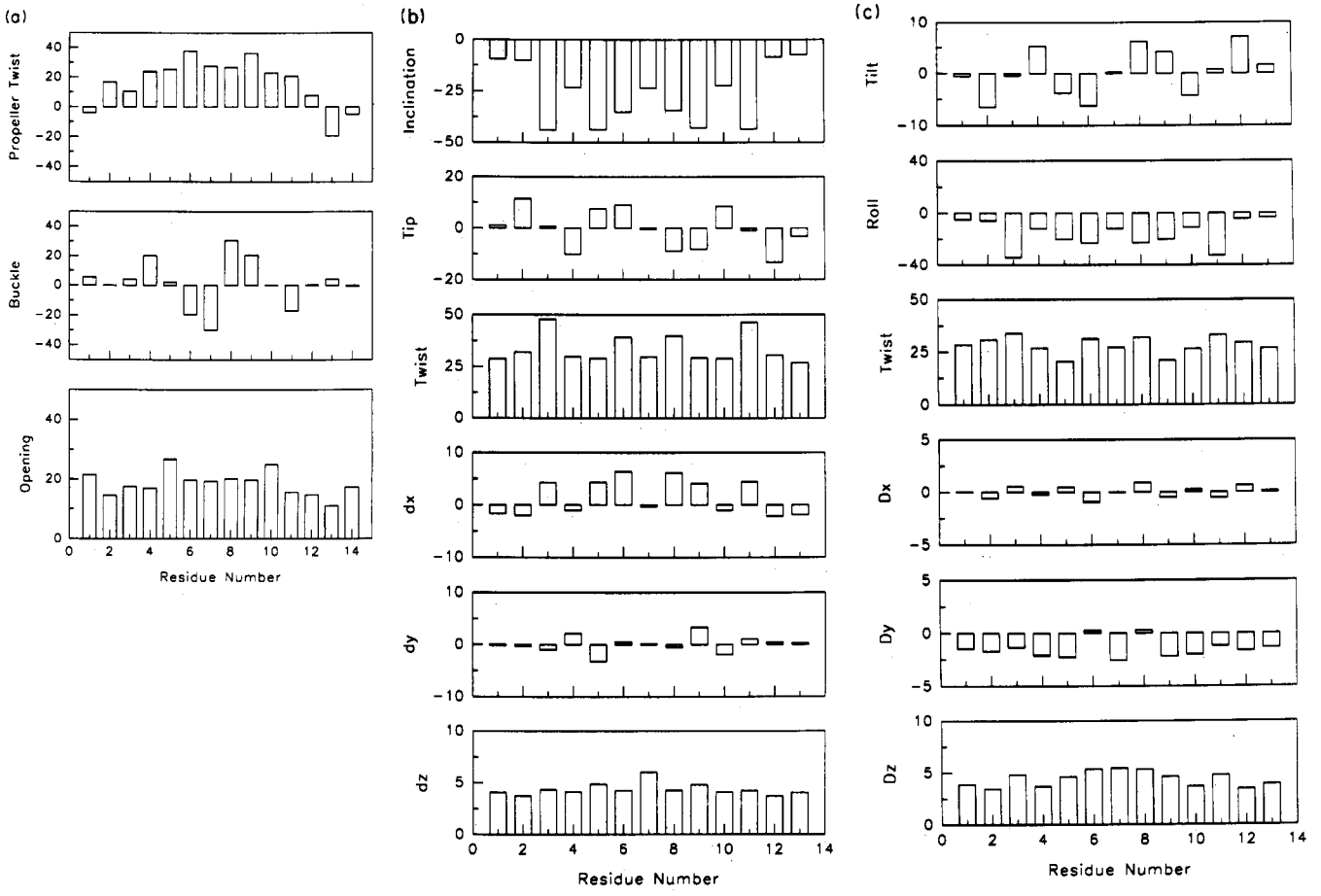

FIGURE 13: Helicoidal parameters vs sequence for the final structure shown in Figure 11. (a) Intra-base-pair, (b) local, and (c) local step helicoidal parameters of the final structure were calculated using the program NUPARM [see Bhattacharyya and Bansal (1989) for the definition of the various parameters]. The structure shows a 2-fold symmetry in all the parameters, but significant variations are seen along the sequence of the molecule.

calculations that the structural family is quite narrow. The best structure having the lowest $RMSD_{NOE}$ shows interesting variations along the sequence, with a significant B_I–B_{II}-type deviation at the central portion marking the beginning of the GCCAAT recognition motif in the molecule. The structure has a wide major groove and a narrow minor groove. The base pairs are oriented as in a A-DNA, but the sugar geometries and other torsion angles are similar to those in a B-DNA model. The molecule exhibits a nonuniform helical axis with a bend in the center. These features could have important implications for the specific recognition of the GCCAAT motif by the relevant proteins.

ACKNOWLEDGMENT

The facilities provided by the 500-MHz FTNMR National Facility supported by the Department of Science and Technology are gratefully acknowledged.

SUPPLEMENTARY MATERIAL AVAILABLE

Coordinates of the best fit NOE final structure (11 pages). Ordering information is given on any current masthead page.

REFERENCES

Ajay Kumar, R. (1992) *J. Biomol. NMR 2*, 519–526.
Ajay Kumar, R., Hosur, R. V., & Govil, G. (1991) *J. Biomol. NMR 1*, 363–378.
Baleja, J. D., Pons, R. T., & Sykes, B. D. (1990a) *Biochemistry 29*, 4828–4839.
Baleja, J. D., Germann, M. W., van de Sande, J. H., & Sykes, B. D. (1990b) *J. Mol. Biol. 215*, 411–428.
Banks, K. M., Hare, D. R., & Reid, B. R. (1989) *Biochemistry 28*, 6996–7010.
Bax, A., & Lerner, L. (1988) *J. Magn. Reson. 79*, 429–438.
Bhattacharyya, D., & Bansal, M. (1989) *J. Biomol. Struct. Dyn. 6*, 635–653.
Borgias, B. A., & James, T. L. (1988) *J. Magn. Reson. 79*, 493–512.
Borgias, B. A., & James, T. L. (1990) *J. Magn. Reson. 87*, 475–487.
Borgias, B. A., Gochin, M., Kerwood, D. J., & James, T. L. (1990) *Prog. Nucl. Magn. Reson. Spectrosc. 22*, 83–100.
Celda, B., Widmer, H., Leupin, W., Chazin, W. J., Denny, W. A., & Würthrich, K. (1989) *Biochemistry 28*, 1462–1471.
Chary, K. V. R., Hosur, R. V., Govil, G., Chen, C. Q., & Miles, H. T. (1988) *Biochemistry 27*, 3858–3867.
Chary, K. V. R., Modi, S., Hosur, R. V., Govil, G., Chen, C. Q., & Miles, H. T. (1989) *Biochemistry 28*, 5240–5249.
Chazin, W. J., Wüthrich, K., Rance, M., Hyberts, S., Denny, W. A., & Leupin, W. (1986) *J. Mol. Biol. 190*, 439–453.
Chodosh, L. A., Baldwin, A. S., Carthew, R. W., & Sharp, P. A. (1988) *Nature 334*, 218–224.
Eich, G., Bodenhausen, G., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 3731–3732.
Ernst, R. R., Bodenhausen, G., & Wokaun, A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford.
Geen, H., & Freeman, R. (1991) *J. Magn. Reson. 93*, 93–141.
Gochin, M., Zon, G., & James, T. L. (1990) *Biochemistry 29*, 11161–11171.
Govil, G., & Hosur, R. V. (1982) *Conformation of Biological Molecules: New Results from NMR*, Springer-Verlag, Heidelberg.
Gupta, G., Bansal, M., & Sasiekharan, V. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 6486–6490.
Hare, D. R., Wemmer, D. E., Chou, S. H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319–336.
Hosur, R. V. (1990) *Prog. NMR Spectrosc. 22*, 1–53.
Hosur, R. V., Chary, K. V. R., & Ravikumar, M. (1985a) *Chem. Phys. Lett. 116*, 105–108.
Hosur, R. V., Ravikumar, M., Roy, K. B., Zu-Kun, T., Miles, H. T., & Govil, G. (1985b) in *Magnetic Resonance in Biology and Medicine* (Govil, G., Khetrapal, C. L., & Saran, A., Eds.) pp 243–260, Tata McGraw-Hill, New Delhi.
Hosur, R. V., Chary, K. V. R., Kumar, A., & Govil, G. (1985c) *J. Magn. Reson. 65*, 375–381.
Hosur, R. V., Ravikumar, M., Chary, K. V. R., Sheth, A., Govil, G., Zu-Kun, T., & Miles, H. T. (1986) *FEBS Lett. 205*, 71–76.
Hosur, R. V., Govil, G., & Miles, H. T. (1988a) *Magn. Reson. Chem. 26*, 927–944.
Hosur, R. V., Sheth, A., & Majumdar, A. (1988b) *J. Magn. Reson. 76*, 218–223.
Hosur, R. V., Majumdar, A., & Patel, D. J. (1989) *J. Am. Chem. Soc. 111*, 5482–5483.
Hosur, R. V., Chary, K. V. R., Saran, A., Govil, G., & Miles, H. T. (1990) *Biopolymers 29*, 953–959.
James, T. L. (1991) *Curr. Opin. Struct. Biol. 1*, 1042–1053.
Jones, K. A., Yamamoto, K. R., & Tjian, R. (1985) *Cell 42*, 559–572.
Jones, K. A., Kadonaga, J. T., Rosenfeld, P. J., Kelly, T. J., & Tjian, R. (1987) *Cell 48*, 79–89.
Keepers, J. W., & James, T. L. (1984) *J. Magn. Reson. 57*, 404–426.
Kessler, H., Mronga, S., & Gemmecker, G. (1991) *Magn. Reson. Chem. 29*, 527–557.
Kim, S. G., Lin, L. J., & Reid, B. R. (1992) *Biochemistry 31*, 3564–3574.
Krishnan, V. V., Shekar, J. C., & Anil Kumar (1991) *J. Am. Chem. Soc. 113*, 7542–7550.
Lewin, B. (1987) *Genes III*, p 211, John Wiley and Sons, New York.
Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.
Majumdar, A. (1990) Ph.D. Thesis, University of Bombay, Bombay, India.
Majumdar, A., & Hosur, R. V. (1992) *Prog. NMR Spectrosc. 24*, 109–158.
Majumder, K., Latha, P. K., & Brahmachari, S. K. (1986) *J. Chromatogr. 355*, 328–334.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Metzler, W. J., Hare, D. R., & Pardi, A. (1989) *Biochemistry 28*, 7045–7052.
Mukhopadhyay, N., Majumdar, A., & Hosur, R. V. (1992) *Spectrochim. Acta 48A*, 1731–1737.
Nerdal, W., Hare, D. R., & Reid, B. R. (1989) *Biochemistry 28*, 10008–10021.
Nibedita, R., Ajay Kumar, R., Majumdar, A., & Hosur, R. V. (1992a) *J. Biomol. NMR 2*, 467–476.
Nibedita, R., Ajay Kumar, R., Majumdar, A., & Hosur, R. V. (1992b) *J. Biomol. NMR 2*, 477–484.
Patel, D. J., Kozlowski, S. A., & Bhatt, R. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3908–3912.
Plateau, P., & Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.
Ravikumar, M., Hosur, R. V., Roy, K. B., Miles, H. T., & Govil, G. (1985) *Biochemistry 24*, 7703–7711.
Ravishanker, G., Swaminathan, S., Beveridge, D. L., Lavery, R., & Sklenar, V. (1989) *J. Biomol. Struct. Dyn. 6*, 689–699.
Redfield, A. G., & Kunz, S. D. (1975) *J. Magn. Reson. 19*, 250–254.
Reid, B. R., Banks, K., Flynn, P., & Nerdal, W. (1989) *Biochemistry 28*, 10001–10007.
Rinkel, L. J., & Altona, C. (1987) *J. Biomol. Struct. Dyn. 4*, 621–649.
Saenger, W. (1984) *Principles of Nucleic Acid Structures*, Springer-Verlag, New York.
Santoro, C., Mermod, N., Andrews, P. C., & Tjian, R. (1988) *Nature 334*, 218–224.
Scheek, R. M., Russo, N., Boelens, R., Kaptein, R., & van Boom, J. H. (1983) *J. Am. Chem. Soc. 105*, 2914–2916.
Sheth, A., Hosur, R. V., Govil, G., Hosur, M. V., Kannan, K. K., Zu-Kun, T., & Miles, H. T. (1989) *Biochemistry 28*, 7275–7282.
Sklenar, V., & Bax, A. (1987) *J. Am. Chem. Soc. 109*, 7525–7526.
Solomon, I. (1955) *Phys. Rev. 99*, 559–565.
van de Ven, F. J. M., & Hilbers, C. W. (1988) *Eur. J. Biochem. 178*, 1–38.
Widmer, H., & Wüthrich, K. (1986) *J. Magn. Reson. 70*, 270–279.
Withka, J. M., Swaminathan, S., Srinivasan, J., Beveridge, D. L., & Bolton, P. H. (1992) *Science 255*, 597–599.
Woessner, D. E. (1962) *J. Chem. Phys. 36*, 1–4.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York.
Zhou, N., Bianucci, A. M., Pattabiraman, N., & James, T. L. (1987) *Biochemistry 26*, 7905–7913.