REFERENCES
Alix, A. J. P., Bernard, L., & Manfait, M. (1981) *Inorg. Chim. Acta*, 147-152.
Benevides, J. M., & Thomas, G. J., Jr. (1983) *Nucleic Acids Res. 11*, 5747-5761.
Brown, E. B., & Peticolas, W. L. (1975) *Biopolymers 14*, 1259-1271.
Dickerson, R. E., & Drew, H. R. (1981) *J. Mol. Biol. 149*, 761-786.
Erfurth, S. C., Kiser, E. R., Kiser, J., & Peticolas, w. L. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 938-941.
Erfurth, S. C., Bond, P. J., & Peticolas, W. L. (1975) *Biopolymers 14*, 1245-1257.
Germann, M. W., Schoenwaelder, K. H., & van der Sande, J. H. (1985) *Biochemistry 24*, 5698-5702.
Goodwin, D. C., & Brahms, J. (1978) *Nucleic Acids Res. 5*, 835-850.
Nishimura, Y., Tsuboi, M., Nakano, S., Hiquchi, T., Sato, T., Shido, S., Uesugi, S., Ohtsuka, E., & Ikehara, M. (1983) *Nucleic Acids Res. 11*, 1579-1588.
Thomas, G. A., & Peticolas, W. L. (1983) *J. Am. Chem. Soc. 105*, 986-992.
Thomas, G. A., & Peticolas, W. L. (1984) *Biochemistry 23*, 3202-3207.
Wartell, R. M., & Harrell, J. T. (1986) *Biochemistry 25*, 2664-2671.

Wobble dG·dT Pairing in Right-Handed DNA: Solution Conformation of the d(C-G-T-G-A-A-T-T-C-G-C-G) Duplex Deduced from Distance Geometry Analysis of Nuclear Overhauser Effect Spectra†

Dennis Hare

Infinity Systems, 14810 216th Avenue NE, Woodinville, Washington 98072

Lawrence Shapiro and Dinshaw J. Patel*

Department of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, New York, New York 10032

Received May 29, 1986; Revised Manuscript Received July 25, 1986

ABSTRACT: We report below on features of the three-dimensional structure of the d(C-G-T-G-A-A-T-T-C-G-C-G) self-complementary duplex (designated 12-mer GT) containing symmetrical G·T mismatches in the interior of the helix. The majority of the base and sugar protons in the 12-mer GT duplex were assigned by two-dimensional nuclear Overhauser effect (NOESY) spectra in H_2O and D_2O solution. A set of 92 short (<4.5-Å) proton–proton distances defined by lower and upper bounds for one symmetrical half of the 12-mer GT duplex were estimated from NOESY data sets recorded as a function of mixing time. These experimental distances combined with nucleotide bond length parameters were embedded into Cartesian space; several trial structures were refined to minimize bond geometry and van der Waals and chirality error. Confidence in this approach is based on the similarity of the refined structures for the solution conformation of the 12-mer GT duplex. The G and T bases pair through two imino–carbonyl hydrogen bonds, and stacking is maintained between the G·T wobble pair and adjacent Watson–Crick G·C pairs. The experimental distance information is restricted to base and sugar protons, and hence structural features such as base pair overlap, glycosidic torsion angles, and sugar pucker are well-defined by this combination of NMR and distance geometry methods. By contrast, we are unable to define the torsion angles about the bonds C3′–O3′–P–O5′–C5′–C4′ in the backbone of the nucleic acid.

The ready availability of synthetic oligonucleotides of defined sequence has permitted a systematic attempt to probe the conformation of DNA in solution by NMR techniques [see reviews by Patel et al. (1982a), Kearns (1983), and Wemmer and Reid (1985)] and in the solid state by X-ray methods [see reviews by Dickerson et al. (1982), Rich et al. (1984), and Kennard (1985)]. These approaches have been extended to studies of errors in DNA such as base pair mismatches in an attempt to delineate the conformational basis for their recognition and repair (Patel et al., 1982b).

Chart I

C1 — G2 — T3 — G4 — A5 — A6 — T7 — T8 — C9 — G10 — C11 — G12

G12 — C11 — G10 — C9 — T8 — T7 — A6 — A5 — G4 — T3 — G2 — C1

The G·U mismatch in RNAs and the G·T mismatch in DNAs drew initial attention on the basis of the proposal of wobble base pair formation by Crick (1976). Support for this model came from identification of G·U pairing in transfer RNA by crystallographic (Quigley & Rich, 1976; Jack et al., 1976) and NMR solution (Schimmel & Redfield, (1980) investigations. An NMR chemical shift study has been reported for poly(dG-dT) in solution (Early et al., 1978).

† The research was supported by National Institutes of Health Grants GM 34504-02 to D.J.P. and GM 35620-02 to D. H. The NMR spectrometers were purchased from funds donated by the Robert Woods Johnson, Jr., Trust and the Matheson Trust toward setting up the NMR Center in the Basic Medical Sciences at Columbia University.

0006-2960/86/0425-7445$01.50/0

The structural analysis could be undertaken at higher resolution for oligonucleotides of defined sequence, and hence efforts were made with deoxy oligomers up to the dodecanucleotide level. Nuclear Overhauser experiments on the d(C-G-T-G-A-A-T-T-C-G-C-G) duplex (henceforth called 12-mer GT, Chart I) demonstrated wobble base pair formation for the symmetrically related G·T mismatches with intact base pairing on either side of the mismatch site (Patel et al., 1982c). These studies were followed by hydrogen exchange measurements on the resolved imino protons at and adjacent to the dG·dT mismatch site in the same duplex (Pardi et al., 1983b; Patel et al., 1984).

The thermodynamics of helices containing G·T mismatches have been probed from calorimetric (Patel et al., 1982c) and optical (Tibanyenda et al., 1984; Aboul-ela et al., 1985) measurements in an effort to account for the decrease in transition midpoint on replacing a dG·dC Watson–Crick base pair by a dG·dT wobble base pair. The dG·dT mismatch in the interior of helices has also been investigated by theoretical computations with wobble pairing preferred over alternate pairing modes (Chuprina & Poltev, 1983; Keepers et al., 1984).

Structural studies on the dG·dT mismatch at atomic resolution have been deduced from single-crystal X-ray investigations of right-handed (Brown et al., 1985; Kneale et al., 1985; Kennard, 1985) and left-handed (Ho et al., 1985; Brown et al., 1986; Kennard, 1985) oligonucleotide duplexes. These studies showed only a small distortion of the DNA helix resulting from incorporation of the G·T mismatch. Further, the higher resolution crystal structures demonstrate the role of organized water in both grooves in stabilizing the G·T pair (Ho et al., 1985; Kneale et al., 1985).

The majority of structural information present in NMR has been used in a qualitative manner since the approximate interproton distance data from nuclear Overhauser effect (NOE) experiments are not easy to interpret in terms of three-dimensional structure. This situation has changed dramatically with the introduction and application of distance geometry approaches by Crippen and co-workers to biological problems (Crippen, 1977; Havel et al., 1979). The structural information available from proton NMR consists of approximate interproton distances for proton pairs up to about 5 Å; above this distance the cross relaxation is usually too small to detect. Because of the approximate and incomplete nature of the distance data, one cannot expect distance geometry to yield a unique structure from the input data. Rather, the distance geometry algorithm is used to generate a "family" of structures that are all consistent with the experimental measurements. The degree of similarity between these structures is then used to evaluate the extent of structural determination implicit in the data. This approach offers the ability to study molecules in solution, an advantage exclusive to NMR methods. Thus, distance geometry analysis of NMR data has been successfully applied to determine the solution structures of several proteins (Braun et al., 1983; Williamson et al., 1985) and a DNA hairpin (Hare and Reid, unpublished results). Although distance geometry calculations require considerable computer resources, we have found that it takes less than a month from data collection and assignment to final structures when a microVAX II is used for refinement.

This paper reports on a detailed two-dimensional NMR study of the 12-mer GT duplex in H_2O and D_2O solution. The exchangeable and nonexchangeable base and sugar protons have been assigned from an analysis of the correlated spectroscopy (COSY) and nuclear Overhauser effect spectroscopy (NOESY) data sets of this 12-mer GT duplex. We have recorded the buildup of the NOE cross-peaks in a set of NOESY spectra of the 12-mer GT duplex as a function of mixing time and estimated the proton–proton distances defined by lower and upper bounds. A distance geometry analysis of these experimental distance constraints has been undertaken on the symmetrical hexanucleotide segment of the self-complementary 12-mer duplex. The resulting solution conformation of the symmetrical hexanucleotide segment provides details of the conformation at and adjacent to the dG·dT mismatch site.

Experimental Procedures

Sample Preparation. The d(C-G-T-G-A-A-T-T-C-G-C-G) 12-mer GT duplex was synthesized by solution-phase phosphotriester chemistry as reported previously (Patel et al., 1982c). The NMR sample was made up by dissolving 175 A_{260} units of 12-mer GT in 0.4 mL of 0.1 M NaCl, 10 mM phosphate, and 1 mM ethylenediaminetetraacetic acid (EDTA) buffer for recording NMR spectra in D_2O and H_2O solution.

NMR Data Collection and Processing. Two-dimensional proton NMR spectra were collected on a Bruker AM 500 spectrometer. The data sets were transferred to magnetic tape and processed on the Columbia VAX 11-780 by using Fourier transform (FT) two-dimensional NMR processing software (Hare, unpublished results).

Two-dimensional magnitude NOESY spectra (mixing time, 120 ms) of the 12-mer GT duplex in H_2O solution at 5 °C were recorded by using a 90° excitation pulse, and 1–1 pulses for mixing and detection. The carrier frequency was 4000 Hz downfield from the H_2O resonance, and the 1–1 delay was optimized for water suppression. The time domain data set consisted of 2K complex data points in the t_2 dimension and 512 data points in the t_1 dimension. Sixty-four scans were collected for each t_1 increment. The data were apodized in t_2 and t_1 with an unshifted sine bell function 1024 points in length and Fourier transformed in both dimensions.

The two-dimensional magnitude COSY spectrum of the 12-mer GT duplex in D_2O solution was recorded at 25 °C with a repetition delay of 2.0 s. We collected 512 t_1 increments over a spectral width of 5000 Hz using 1024 complex points in the t_2 dimension. The data were apodized in t_2 and t_1 with an unshifted sine bell function 600 points in length and Fourier transformed in both dimensions.

Two-dimensional phase-sensitive NOESY spectra of the 12-mer GT duplex in D_2O solution were recorded at 25 °C with a repetition delay of 2.0 s for mixing time values of 50, 75, 100, 150, 250, and 300 ms. The carrier was placed on the HDO resonance, and 1024 complex data points were acquired in the t_2 dimension with a spectral width of 5000 Hz. For each NOESY experiment, real and imaginary data were collected for 512 t_1 values by using the method of States et al. (1982). The data in the t_2 dimension were apodized by using an exponential window of 1 Hz line broadening and then Fourier transformed and phase corrected. Since Bruker spectrometers acquire real and imaginary data at different points in time, it was necessary to correct the data prior to Fourier transformation by using the method of Redfield and Kunz (1975). The data in the t_1 dimension are apodized by using a 90° phase-shifted sine bell. This window function was used to avoid truncation effects in t_1 without distorting the relative intensities of resonances having different line widths. After apodization, t_1 data were Fourier transformed and a small phase correction was applied.

Distance Constraints. The volume integrals for each resolved cross-peaks corresponding to each detectable proton–

proton interaction were measured from the NOESY contour plots as a function of mixing time. The simple relationship between cross relaxation rate (σ) and interproton separation (r)

$$\sigma_{ij}/\sigma_{kl} = (r_{kl})^6/(r_{ij})^6 \quad (1)$$

was used to estimate the interproton distances, where r_{kl} corresponds to the fixed distance yardstick and r_{ij} corresponds to the experimentally determined distance. The ratio σ_{ij}/σ_{kl} could be estimated by measuring the volume integrals of cross-peaks corresponding to interactions i–j and k–l and together with the known fixed distance r_{kl} yields an estimate of the unknown distance r_{ij}.

The fixed distances used to calibrate volume integrals in the NOESY data sets are shorter than the distances for which they serve as calibrants. This is essential to the accurate setting of upper and lower bounds for experimental distances through minimizing the effects of spin diffusion. At low mixing times, NOEs from the shorter fixed distances which build up faster are much more intense than those representing the distances to be measured according to their calibration, and hence at low mixing times these distances seem artificially long, serving to provide an upper bound. At high mixing times, due to the nonlinear nature of the NOE buildup curve, the measured distances are artificially short and serve to provide a lower bound distance.

The dependence of cross relaxation rates on rotational motion and spin–spin relaxation required the use of various distance yardsticks in order to estimate interproton distances. We have found that cross relaxation between geminal protons is too fast to allow proper calibration of these buildup rates except at very short mixing times. The thymidine H6–CH_3 distance of 3.00 Å was used as a yardstick for all NOEs involving CH_3 protons, the sugar H2′–H2″ distance of 1.85 Å was used as a yardstick for all NOEs involving sugar H2′ and H2″ protons, and the cytidine H5–H6 distance of 2.45 Å was used as a yardstick for all remaining NOEs between nonexchangeable protons. The thymidine imino to adenosine H2 distance of 2.81 Å was used as a yardstick for NOEs between imino protons and for NOEs between imino and adenosine H2 protons.

The interproton distances were measured on the basis of eq 1 for the NOESY spectra recorded as a function of mixing time. These estimated distances were found to be dependent on mixing time within the initial buildup region so that it has not been possible to estimate an exact distance. Rather, each interproton distance was constrained by upper and lower bounds which were set to d + 0.1 Å and d – 0.1 Å, where d is the distance range estimated from short mixing time NOESY data. Somewhat larger bounds of d + 0.3 Å and d – 0.3 Å were set for weak cross-peaks which were observable in NOESY spectra only at longer mixing times.

Distance Geometry Algorithm. The distance geometry algorithm accepts input in the form of upper and lower distance bounds between pairs of atoms. Initially, we generated the distances implicit in the primary chemical structure of the molecule: (1) The distance between each pair of bonded atoms was fixed at an appropriate value derived from small-molecule crystallographic data. (2) Bond angle information was converted to distances between each pair of substituents bonded to a given atom. (3) Each atom having four distinguishable substituents behaves as a chiral center. The chirality of each asymmetric center was encoded in the form of a signed volume (Crippen, 1977). (4) Distances between atoms in rigid groups (DNA bases) were fixed at appropriate values. (5) The lower distance bound for each unbonded atom pair was set to the sum of their van der Waals radii. Next, the experimental distance constraints were added to the existing bounds. In addition to distances measured from observed NOEs, exact distance values for hydrogen bonding interactions were also included. At this point all known distance information had been entered, and all undefined upper distance bounds were set to an arbitrarily large value (300 Å). The bounds were next "smoothed" by using two different algorithms. The first algorithm attempted to deduce improved bounds between atoms forming sets of rigid triangles sharing common sides. The triangle inequality and inverse triangle inequality (Havel et al. 1977) were then used to complete the bound smoothing.

In order to generate a random set of starting structures, several random "trial" distance matrices were generated by selecting random distances between upper and lower distance bounds for each interatomic distance. Each of the trial distance matrices were then embedded in three-dimensional space as described by Havel et al. (1979), yielding the corresponding trial structures. The trial structures were locally distorted due to the fact that the randomly chosen trial distances define structures that cannot exist in three-dimensional space. These structures were each refined by using nonlinear optimization techniques to minimize the disagreement between the structure and the distance bounds. Nonlinear optimization algorithms are iterative, so the refinements require a significant amount of computer time. These algorithms are also prone to converging to false minima. In our experience, each trial structure would encounter several false minima en route to a "real" minimum having near-zero disagreement with the bounds.

We found several methods useful for escaping from false minima. Our refinement algorithm weights each type of distance constraint separately, and it was often possible to escape from a false minimum by simply changing the value of one or more weights. A false minimum is commonly encountered when the movement of a group of atoms is blocked by van der Waals bounds. If the van der Waals weight is temporarily set to zero, the group is free to reorient and the refinement will proceed. One can think of the weighting scheme as a means of changing the shape of the hypersurface being searched; a "true" minimum will still be a true minimum regardless of weights whereas false minima change with weights. Another method we found very useful in refinements involved adding random vectors to all or selected atomic positions. Instead of changing the shape of the search hypersurface, this operation simply changed location on the hypersurface. This technique also allows one to explore the conformational freedom of fully refined structures by successive randomization and refinement. Two different algorithms were used for refinement. The first is a simple conjugate-gradient algorithm that moves each Cartesian coordinate for each atom. This algorithm is very fast but tends to be of little use for initial refinement. The second algorithm is more intelligent, but much slower than the first, and cannot generally refine a whole molecule simultaneously. This algorithm can move groups of atoms as units with translational and rotational freedom and is well suited for dealing with the distorted trial structures. After the rough refinement is done, the conjugate-gradient algorithm is used for final refinement.

RESULTS

The numbering system for the self-complementary 12-mer GT duplex is outlined in Chart I and applies equally well for spectral assignments and distance geometry analysis.

Exchangeable Proton Assignments. We have previously reported on the temperature dependence of the exchangeable imino protons in the 12-mer GT duplex (Patel et al., 1982c).

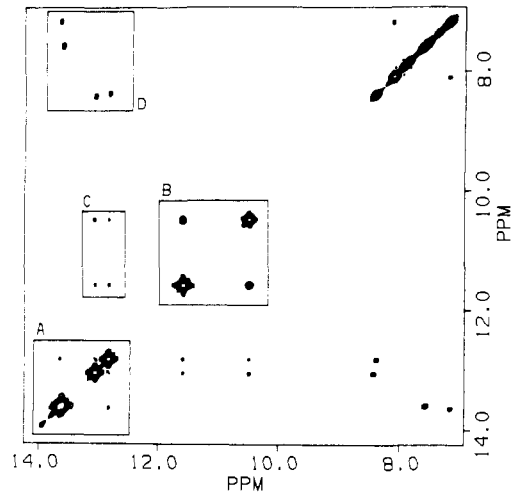

FIGURE 1: Contour plot of the magnitude NOESY spectrum (mixing time, 120 ms) of the 12-mer GT duplex in 0.1 M NaCl and 10 mM phosphate, in H_2O at 5 °C.

Table I: Proton Chemical Shifts of Nonterminal Imino, Cytidine Amino, and Adenosine H2 Protons in the 12-mer GT Duplex in 0.1 M NaCl and 10 mM Phosphate, in H_2O at 5 °C

	chemical shift (ppm)			
	H1	H3	H4[a]	H2
G2·C11	13.05		8.42	
T3·G10	10.50	11.59		
G4·C9	12.82		8.38	
A5·T8		13.64		7.16
A6·C7		13.59		7.56

[a] Chemical shift is for hydrogen-bonded cytidine H4 amino proton.

The spectral data can be conveniently collected in two dimensions, and the magnitude NOESY spectrum (120-ms mixing time) of the 12-mer GT duplex in 0.1 M NaCl in H_2O at 5 °C is plotted in Figure 1 for the symmetrical range extending from 7.0 to 14.0 ppm.

The thymidine imino protons at 13.64 and 13.59 ppm exhibit intrabase pair NOEs to adenosine H2 protons at 7.16 and 7.56 ppm, respectively (box D, Figure 1). In addition, the 13.64 ppm thymidine imino proton exhibits an NOE to an adjacent guanosine imino proton at 12.82 ppm (box A, Figure 1), which permits assignment of the 13.59 ppm thymidine imino proton to the A6·T7 base pair, the 13.64 thymidine imino proton to the A5·T8 base pair, and the 12.82 ppm guanosine imino proton to the G4·C9 base pair in the 12-mer GT duplex.

The guanosine imino protons at 12.82 and 13.05 ppm exhibit intra base pair NOEs to the hydrogen-bonded cytidine amino protons at 8.38 and 8.42 ppm, respectively (box D, Figure 1). The 12.82 ppm imino and 8.35 ppm amino protons are assigned to the G4·C9 base pair while the 13.05 ppm imino and 8.42 ppm amino protons correspond to the G2·C11 base pair.

The imino protons from the T3·G10 interaction resonate at 10.50 and 11.59 ppm and exhibit a strong NOE between themselves (box B, Figure 1), and each in turn exhibits a weak NOE to the 12.82 ppm guanosine imino proton of flanking G4·C9 and the 13.05 ppm guanosine imino proton of flanking G2·C11 base pair (box C, Figure 1). These intra and inter base pair NOEs are best visualized by taking slices through the 11.59 thymidine imino proton (Figure 2B) and 10.50 ppm guanosine imino proton (Figure 2C) of the T3·G10 interaction in the 12-mer GT duplex NOESY spectrum at 5 °C.

The chemical shifts of the nonterminal base pair imino, cytidine amino, and adenosine H2 protons in the 12-mer GT duplex at 5 °C are listed in Table I.

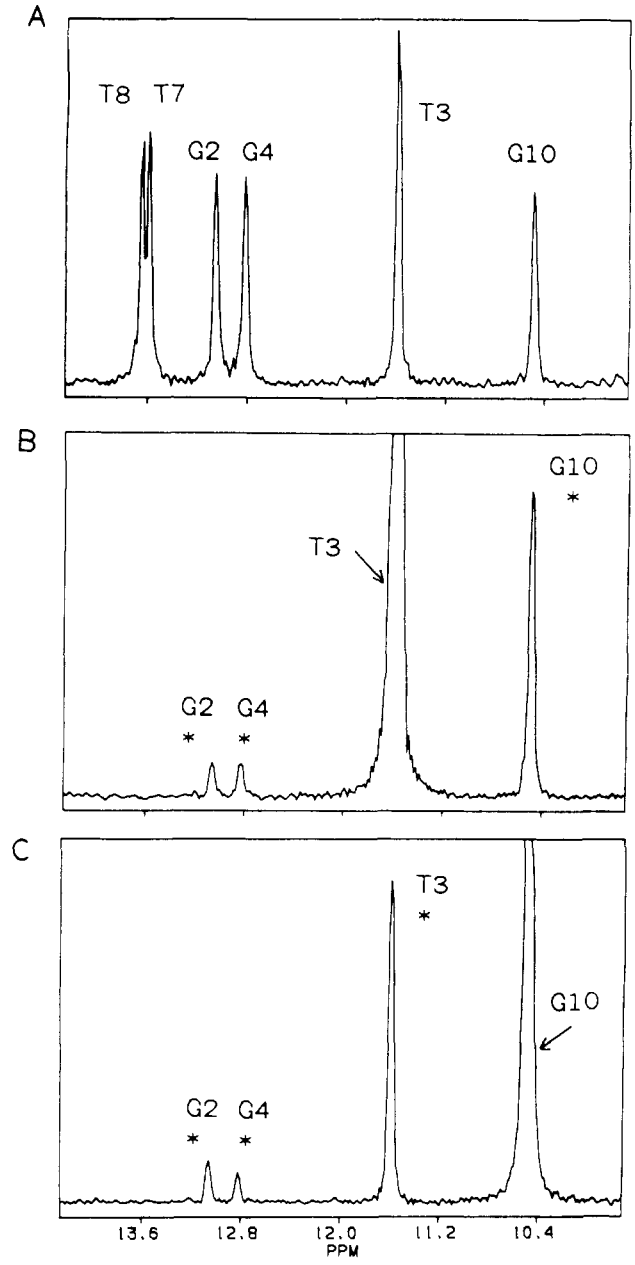

FIGURE 2: One-dimensional slices (6.5–14.0 ppm) establishing distance connectivities between Watson–Crick and wobble imino protons in the 12-mer GT duplex NOESY spectrum (120-ms mixing time) in 0.1 M NaCl and 10 mM phosphate, in H_2O at 5 °C. The control spectrum is shown in (A) while (B) and (C) represent one-dimensional slices through the wobble thymidine imino (11.59 ppm) and wobble guanosine imino (10.50 ppm) protons.

Nonexchangeable Proton Assignments. The earlier one-dimensional NOE studies on the 12-mer GT duplex are able to classify the nonexchangeable base proton resonances by type but could not unambiguously assign them to given positions in the sequence. Further, no assignments could be made for any of the sugar protons in the 12-mer GT duplex (Patel et al., 1982c). Resolved cross-peaks are observed in the two-dimensional spectra of oligonucleotide duplexes (Pardi et al., 1983a; Feigon et al., 1983), and procedures have been worked out for the assignment of base and sugar protons in right-handed helices (Hare et al., 1983; Scheek et al., 1984; Weiss et al., 1984). We have undertaken through-bond-correlated (COSY) and through-space (NOESY) two-dimensional NMR experiments on the 12-mer GT duplex in 0.1 M NaCl in D_2O solution at 25 °C. A contour plot of the phase-sensitive NOESY spectrum (250-ms mixing time) of the 12-mer GT

Table II: Nonexchangeable Proton Chemical Shifts of the 12-mer GT Duplex in 0.1 M NaCl and 10 mM Phosphate, in D_2O at 25 °C

	chemical shift (ppm)							
	H8	H2	H6	H5/CH_3	H1′	H2′	H2″	H3′
C1			7.65	5.94	5.80	2.00	2.42	4.70
G2	7.93				6.11	2.61	2.81	4.98
T3			7.21	1.73	5.44	1.97	2.15	4.84
G4	7.89				5.55	2.73	2.79	5.02
A5	8.12	7.26			5.97	2.70	2.92	5.07
A6	8.12	7.63			6.16	2.58	2.92	5.01
T7			7.11	1.29	5.90	1.98	2.56	4.83
T8			7.40	1.57	6.10	2.20	2.59	4.92
C9			7.54	5.72	5.80	2.19	2.49	4.88
G10	7.86				5.93	2.55	2.73	4.98
C11			7.29	5.44	5.74	1.80	2.30	
G12	7.96				6.18	2.64	2.40	4.68

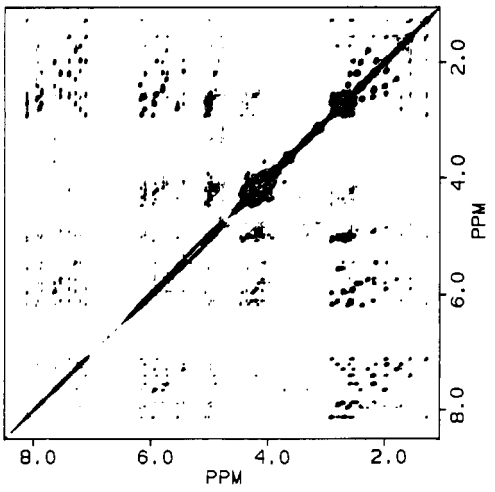

FIGURE 3: Contour plot of the phase-sensitive NOESY spectrum (mixing time, 250 ms) of the 12-mer GT duplex in 0.1 M NaCl and 10 mM phosphate, in D_2O at 25 °C.

duplex is shown in Figure 3. The one-dimensional spectrum is along the diagonal with off-diagonal cross-peaks representing through-space interactions between protons separated by <4.5 Å.

We have followed standard procedures for assignment of the base and sugar H1′, H2′, H2″, and H3′ protons (Hare et al., 1983; Scheek et al., 1984; Weiss, et al., 1984), and these are briefly outlined below for the 12-mer GT duplex. Expanded NOESY contour plots establishing distance connectivities from the base protons (7.1–8.2 ppm) to the sugar H1′ and cytidine H5 protons (5.4–6.2 ppm) and to the thymidine CH_3 protons (1.2–1.9 ppm) are shown in panels A and B, respectively, of Figure 4. The strong cross-peaks linking the H6 and H5 protons of the three cytidines (separation, 2.45 Å) in the 12-mer GT duplex are designated by asterisks in Figure 4A as are the strong cross-peaks linking the H6 and CH_3 protons of the three thymidines (separation, 3.00 Å) in Figure 4B.

We have demonstrated that the purine H8 or pyrimidine H6 proton exhibits an NOE to its own and 5′-flanking sugar H1′ protons for right-handed DNA (Hare et al., 1983; Scheek et al., 1984; Weiss et al., 1984) so that it is possible to walk down the 12-mer GT helix from base C1 to G12 by stepping through the sugar H1′ protons. This tracing is outlined in Figure 4A and is straightforward except for the sugar H1′ protons of C1 and C9, which exhibit the same chemical shift. The base and sugar H1′ chemical shifts in the 12-mer GT duplex at 25 °C are listed in Table II.

We observe NOEs between purine H8 and pyrimidine H5 CH_3 protons on adjacent bases in purine–(3′–5′)–pyrimidine segments for the G2–T3 step (peak H, Figure 4B), the A6–T7

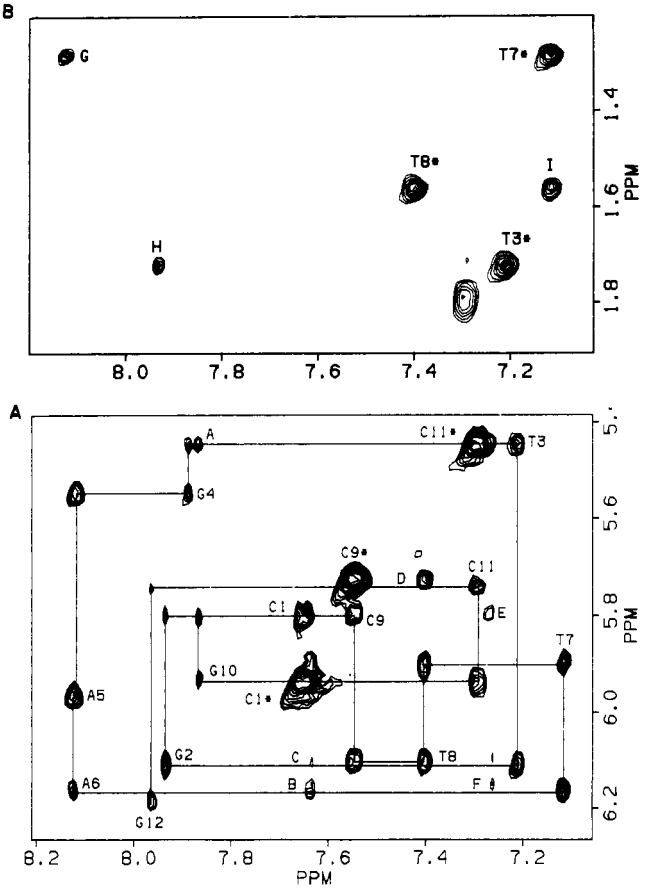

FIGURE 4: Expanded contour plot of the phase-sensitive NOESY spectrum (mixing time, 250 ms) of the 12-mer GT duplex in 0.1 M NaCl and 10 mM phosphate, in D_2O at 25 °C, establishing distance connectivities (A) between the base protons (7.0–8.2 ppm) and the sugar H1′ and cytidine H5 protons (5.0–6.4 ppm) and (B) between the base protons (7.0–8.2 ppm) and the thymidine CH_3 protons (1.2–1.9 ppm). The lines in (A) follow connectivities between adjacent base protons through their intervening sugar H1′ protons. The cytidine H5–H6 cross-peaks are designated by asterisks in (A) and the thymidine H6–CH_3 cross-peaks are designated by asterisks in (B). The cross-peaks A–I are discussed in the text.

step (peak G, Figure 4B), and the G10–C11 step (peak A, Figure 4A) in the 12-mer GT duplex. In addition, NOEs are observed between pyrimidine H6 and pyrimidine H5 CH_3 protons on adjacent bases in pyrimidine–(3′–5′)–pyrimidine segments for the T7–T8 step (peak I, Figure 4B) and the T8–C9 step (peak D, Figure 4A) in the 12-mer GT duplex. These intrastrand base–base NOEs exhibit directionality and are characteristic of right-handed helices.

We also detect NOEs between the 7.26 ppm H2 proton of A5 and the H1′ proton of A6 on the same strand (peak F, Figure 4A) and the H1′ proton of C9 on the partner strand

Table III: Intranucleotide Nonexchangeable Proton Distance Constraints (in Angstroms) in the 12-mer GT Duplex at 25 °C

	H1′–H8/H6	H2′–H8/H6	H2″–H8/H6	H3′–H8/H6	H1′–H2′	H1′–H2″	H1′–H3′	H1′–H2
C1	3.0–3.3	2.0–2.3	2.4–3.3		2.7–3.3	2.05–2.4		
G2	3.8–4.3	2.1–2.55	2.4–2.95	3.15–3.75		2.05–2.35	3.3–4.1	
T3	2.95–3.7	1.9–2.2	2.65–3.35		2.8–3.4	2.0–2.3		
G4	3.7–4.2	2.1–2.6	2.3–3.0	3.9–4.4	2.35–2.95	2.0–2.45	3.3–4.0	
A5	3.0–4.0			3.2–3.7		2.05–2.4	3.25–3.8	
A6	3.5–4.0	2.3–2.6		3.45–3.9		2.05–2.3	3.45–3.85	3.7–4.3
T7	3.55–4.05	2.05–2.35			2.3–2.9			
T8	3.4–3.9	1.9–2.15	2.15–2.95	3.55–4.05	2.35–2.95			
C9	3.1–3.7		2.25–3.15	3.5–4.4	2.4–3.0	2.0–2.4		
G10	3.8–4.3	2.2–2.45	2.4–2.95	3.6–4.1		1.95–2.3		
C11	3.25–3.65	1.95–2.35	2.5–3.3		2.65–3.25	1.95–2.3		
G12	3.55–4.05	2.05–2.45	2.4–3.1			2.0–2.3		

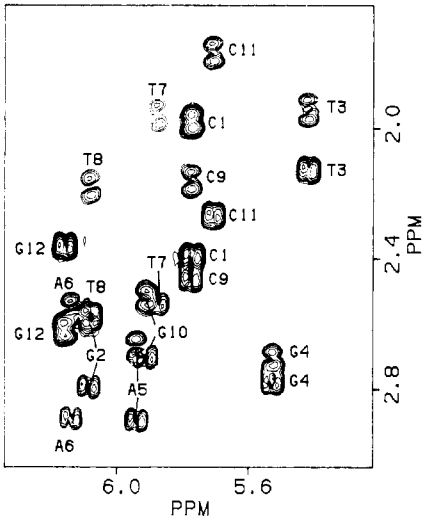

FIGURE 5: Expanded contour plot of the magnitude COSY spectrum of the 12-mer GT duplex in 0.1 M NaCl and 10 mM phosphate, in D_2O at 25 °C, establishing coupling connectivities between the sugar H1′ protons (5.2–6.2 ppm) and the sugar H2′ and H2″ protons (1.6–3.0 ppm).

(peak E, Figure 4A) and NOEs between the 7.63 ppm H2 proton of A6 and the H1′ proton of A6 on the same strand (peak B, Figure 4A) and the H1′ proton of T8 on the partner strand (peak C, Figure 4A). The directionality of the NOEs between protons across the strands is also characteristic of right-handed duplexes.

The expanded magnitude COSY spectrum of the 12-mer GT duplex in 0.1 M NaCl at 25 °C establishing coupling connectivities between the sugar H1′ protons (5.2–6.3 ppm) and the sugar H2′ and H2″ protons (1.7–3.0 ppm) is plotted in Figure 5. Each sugar H1′ proton exhibits cross-peaks to its own sugar H2′ and H2″ protons and since the former assignments have been made from analysis of the NOESY data it is straightforward to assign the latter sugar protons. The cross-peak patterns for the upfield sugar H2′ proton and downfield sugar H2″ protons are quite different and are only reversed at the 3′-terminal G12 residue (Figure 5). We can assign the sugar H3′ protons since we detect cross-peaks between these protons and the assigned sugar H2′ protons in the 12-mer GT duplex. The sugar H2′, H2″, and H3′ proton chemical shifts in the 12-mer GT duplex at 25 °C are listed in Table II, and these assignments have been checked through analysis of the remaining regions of the NOESY spectrum of the 12-mer GT duplex in Figure 3.

Distance Constraints—Nonexchangeable Protons. The proton NOE buildups for the 12-mer GT duplex in 0.1 M NaCl, in D_2O at 25 °C, can be estimated from the NOESY data sets recorded at mixing times of 50, 75, 100, 150, 250, and 300 ms. We have translated these NOE parameters into approximate distances defined by lower and upper bounds by using the method outlines under Experimental Procedures.

Table IV: Nonexchangeable Proton Distance Constraints (in Angstroms) between Adjacent Nucleotides in the 12-mer GT Duplex at 25 °C

	Same Strand				
	H1′–H8/H6	H2′–H8/H6	H2″–H8/H6	H8/H6–H5/CH3	H2′–H5
C1–G2	3.7–4.3		2.35–3.0		
G2–T3	3.0–3.75	2.55–3.4	2.25–2.55	3.5–4.15	
T3–G4	3.7–4.3				
G4–A5	3.3–3.65		2.5–3.2		
A5–A6	3.2–4.3				
T7–T8	3.3–3.9	2.4–2.95	2.1–2.9	3.15–3.6	
T8–C9	3.3–4.2		2.35–2.95	3.3–3.95	2.8–3.4
C9–G10	3.65–4.15	2.9–3.75	2.5–2.95		
G10–C11	3.45–3.95		2.2–2.7	3.85–4.3	
C11–G12	3.9–4.6				

Partner Strands	H2–H1′
A6–T8	3.9–4.3
A5–C9	3.8–4.4

Table V: Exchangeable Proton Distance Constraints (in Angstroms) between Adjacent Nucleotides in the 12-mer GT Duplex at 5 °C

	same strand, H1/H3–H1/H3	partner strands, H1/H3–H1/H3
G2–T3	3.15–3.75	
T3–G4	3.45–4.05	
G2–G10		3.3–3.9
T3–G10		1.95–2.5
G4–T8		3.15–3.75
G4–G10		3.15–3.75

The experimental distance data for intranucleotide NOEs between nonexchangeable protons in the 12-mer GT duplex are listed in Table III. We have been unable to estimate distances when NOE cross-peaks overlap, and these represent the various blank spots in Table III.

The corresponding distance data for NOEs between nonexchangeable protons on adjacent nucleotides in the 12-mer GT duplex are listed in Table IV. The majority of these distances are between protons on the same strand, but two cross-strand distances are also listed in Table IV.

Distance Constraints—Exchangeable Protons. We have recorded the NOEs among exchangeable protons and between exchangeable and nonexchangeable protons in the 12-mer GT duplex in 0.1 M NaCl in H_2O at a single mixing time (120 ms) and at a low temperature (5 °C). The estimated distances were given a lower bound d − 0.3 Å and an upper bound d + 0.3 Å and are listed in Table V. We have assumed that these distance constraints for the 12-mer GT at 5 °C remain unchanged at 25 °C, the temperature at which we collected the nonexchangeable distance constraints listed in Tables III and IV.

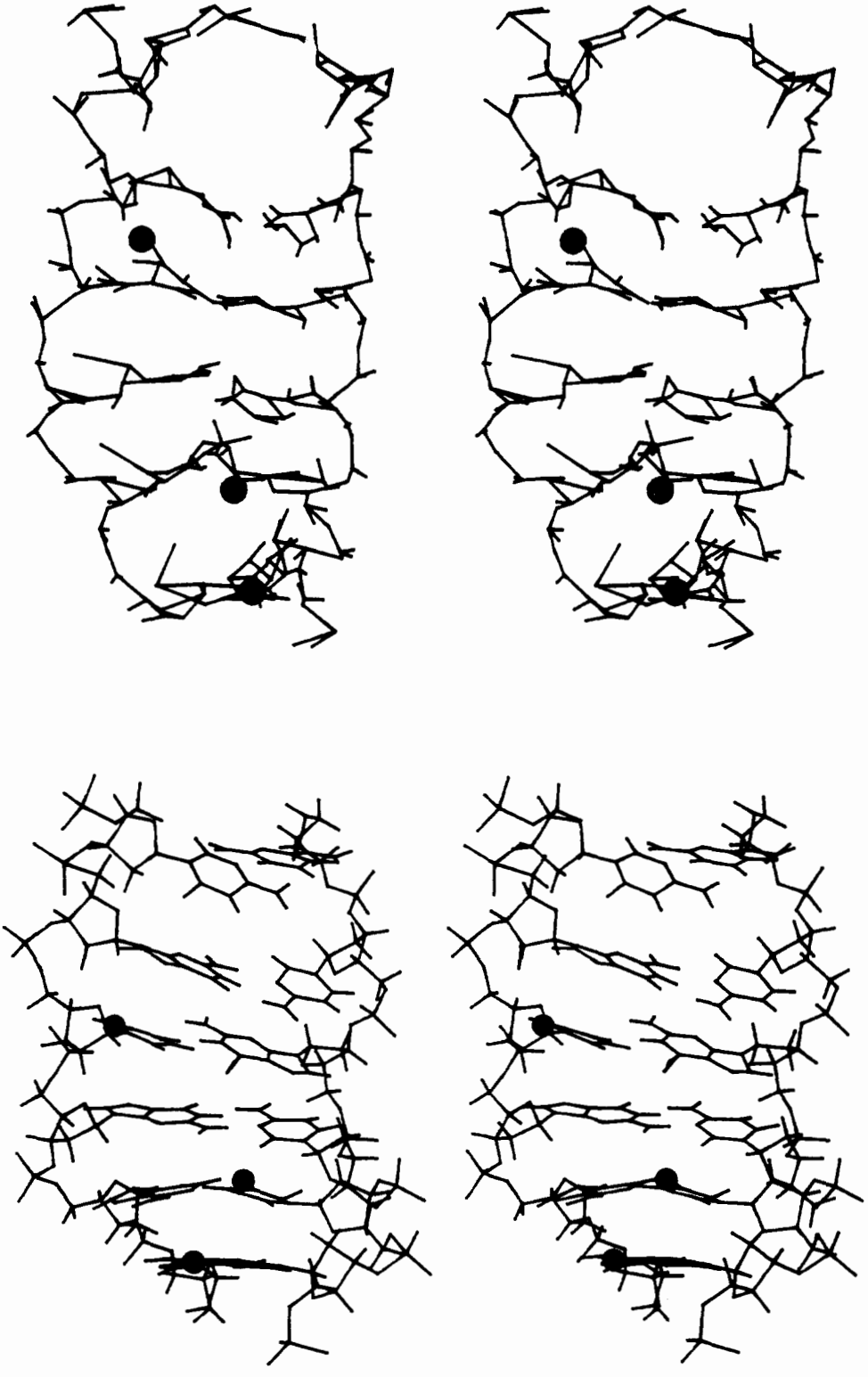

FIGURE 6: Stereopairs (top) of an initial embedded structure and (bottom) of the refined structure of one symmetrical half of the 12-mer GT duplex. These views correspond to the refinement designated C. The thymidine CH_3 groups are included as darkened balls for recognition purposes. The chain runs C1-G2-T3-G4-A5-A6 from top to bottom of left strand, and the chain runs T7-T8-C9-G10-C11-G12 from bottom to top of right strand in stereopair.

Distance Geometry Refinement. A description of the distance geometry algorithm and its usage is outlined under Experimental Procedures. The refinements have been undertaken on a set of 92 short (<4.5 Å) proton–proton distances defined by lower and upper bounds (Tables III–V) for one symmetrical half of the 12-mer GT duplex. A typical embedded starting structure (designated C) based on the experimental NOE based distance constraints is plotted in stereo in Figure 6 (top) with the thymidine CH_3 residues shown as darkened balls. The resulting structure following distance geometry refinement (designated C) is plotted in stereo in Figure 6 (bottom). Two views of the pairing of T3 and G10 at the T3·G10 mismatch site in the refined 12-mer GT duplex are plotted in stereo in the top and bottom panels of Figure 7.

The same procedure was repeated with several embedded starting structures involving random selection of each experimental distance from within the lower and upper bounds and

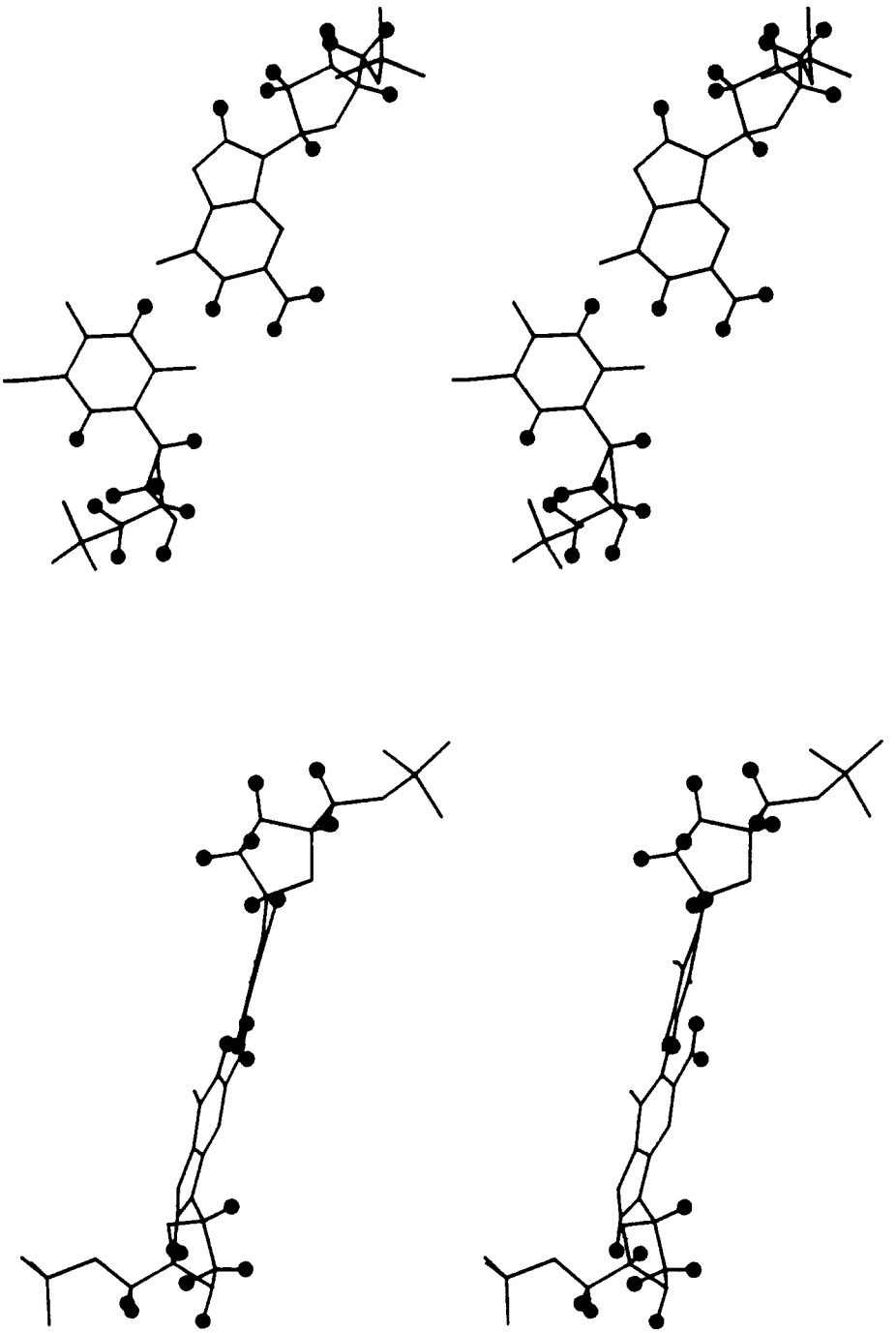

FIGURE 7: Stereoviews of the T3·G10 mismatch pair in the 12-mer GT refinement designated C. (Top) This view demonstrates wobble base pair formation of T3 and G10 at the mismatch site. (Bottom) This view demonstrates propeller twisting at T3·G10 pair.

yielded several refined structures. All refined structures had total sum-of-square bond length disagreements of less than 0.9 Å^2, and the largest single bond length disagreements were less than 0.1 Å. The chirality of each asymmetric center was also correct.

Comparison of Refined Structures. Three refinements of one symmetrical half of the 12-mer GT duplex designated A, C, and E are plotted as stereo pairs in Figure 8. These three refinements yield strikingly similar solution conformations, providing confidence in the reproducibility of the results independent of the starting embedded structure.

We have compared the local and global structures of these three refinements by analyzing their conformations in pairs. The average differences in Å for atoms separated by a given average distance ($d \pm 0.5$ Å) between pairs of refined structures of the 12-mer GT duplex are compared in Figure 9A. We find that since the NOE approach provides a short distance ruler, the average difference is small for short interatom separations but continues to increase for longer interatom separations (Figure 9A).

We have also superpositioned pairs of refined structures to the greatest extent possible and calculated the average coordinate difference for each atom at a given residue along the oligomer sequence. These results are plotted in Figure 9B for pairs A and E, pairs C and E, and pairs A and C in one symmetrical half of the 12-mer GT duplex. The average coordinate difference is centered about a value of approximately 1 Å and fluctuates by approximately ±0.5 Å at residues along the oligomer chain (Figure 9B).

Torsion Angles. We can compute the sugar–phosphate backbone torsion angles and the glycosidic torsion angles at all residues for the refined structure of the 12-mer GT duplex.

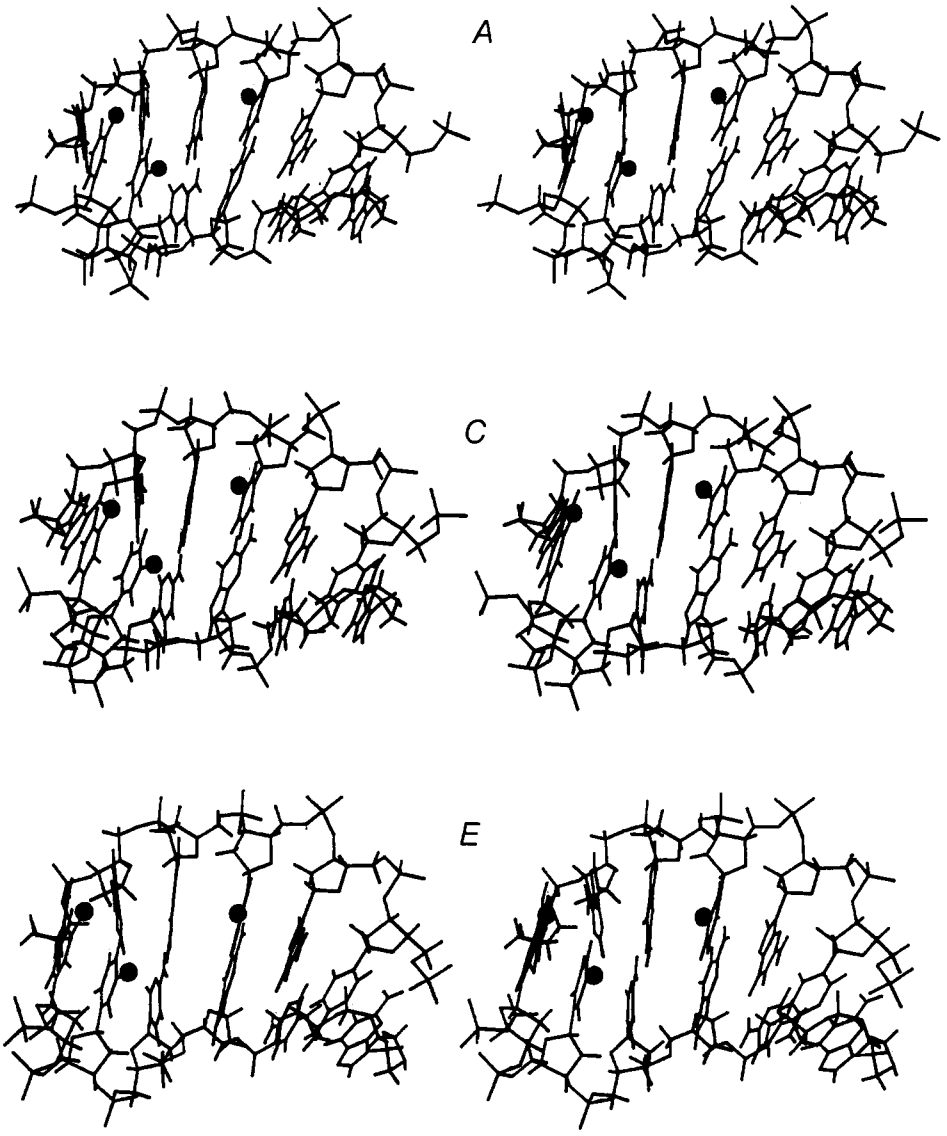

FIGURE 8: Stereopairs of refined structures designated A, C, and E of one symmetrical half of the 12-mer GT duplex.

We limit our analysis to the glycosidic torsion angle and the pucker of the sugar ring since they are defined by the experimental proton NOE constraints.

We have plotted the lower and upper ranges for the glycosidic torsion angle χ and the sugar pucker pseudorotation value P at individual residue positions in the 12-mer GT refinements A, C, and E (Figure 10). The variations in glycosidic torsion angles at a given position between the refined structures cover a smaller range than the pseudorotation for sugar pucker in the 12-mer GT duplex. The glycosidic torsion angles are centered about 260° while the pseudorotation values are centered about 160° (Figure 10).

DISCUSSION

Spectral Assignments. The earlier proton NMR study of the 12-mer GT duplex yielded imino proton assignments but was unable to assign the nonexchangeable base and sugar protons (Patel et al., 1982). We have applied COSY and NOESY procedures to successfully assign the base and sugar H1′, H2′, H2″, and H3′ protons in the 12-mer GT duplex, and these chemical shifts at 25 °C are listed in Table II.

Distance Constraints. We found it imperative to measure the NOESY spectra as a function of mixing time to estimate the nonexchangeable proton–proton distance constraints defined by lower and upper bounds listed in Tables III and IV. Since the cross-peak intensity is spread out in two dimensions, volume integration is necessary to extract accurate cross-peak intensities from the data.

The distance estimates between protons on adjacent bases constitute an important set of constraints. These include constraints defined by the base H8/H6–(3′–5′)–base H5/CH_3 step on the same strand as observed for the G2–T3, T7–T8, T8–C9, and G10–C11 steps (Table IV) and intrastrand and interstrand constraints between imino protons on adjacent G2·C11, T3·G10, G4·C9, and A5·T8 base pairs (Table V).

We have established Watson–Crick pairing for the dA·dT and dG·dC base pairs by observing NOEs between the thymidine imino and adenosine H2 protons for the former pair and NOEs between the guanosine imino and hydrogen-bonded cytidine amino protons for the latter pair. The two strands can then be related to each other by standard Watson–Crick pairing distance constraints. The strong NOE between the thymidine imino and guanosine imino protons at the dG·dT wobble site also linked the two strands at this mismatch site. Finally, the cross-strand NOEs between adjacent base pairs as observed between the adenosine H2 and sugar H1′ protons (Table IV) and among the imino protons (Table V) put additional constraints on the linkage of the two strands to form the duplex.

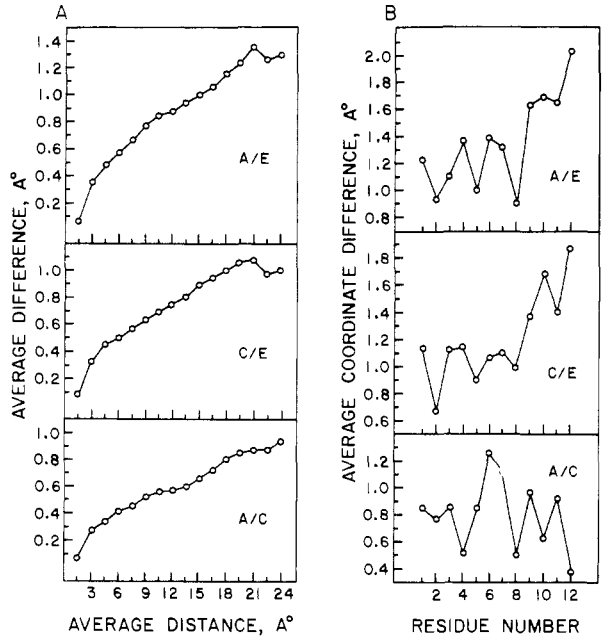

FIGURE 9: (A) Comparison of the average difference in angstroms for atoms separated by an average distance $d \pm 0.5$ Å between pairs of refined structures of the 12-mer GT duplex. The comparison was undertaken for pairs A and E, pairs C and E, and pairs A and C. (B) Comparison of the average coordinate difference in angstroms for all atoms of a given residue between pairs of refined structures of the 12-mer GT duplex.

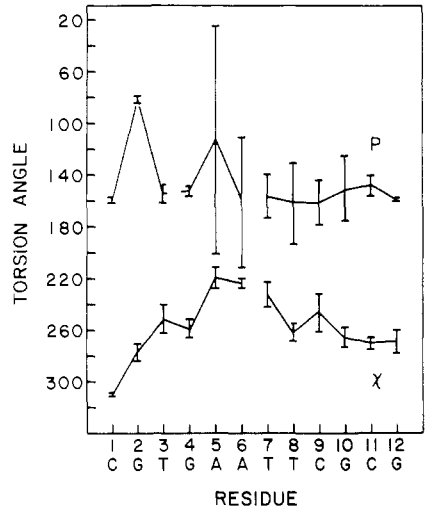

FIGURE 10: Plot of the lower and upper ranges for the glycosidic torsion angle angle χ and the sugar pucker pseudorotation value P at individual residue positions in the 12-mer GT refinements A, C, and E.

We note that the base (purine H8 or pyrimidine H6) protons exhibit stronger NOEs to their own sugar H2′ compared to H2″ proton (Table III) while they exhibit the stronger NOE to the 5′-flanking sugar H2″ compared to H2′ proton (Table IV). The relative order of these NOEs has been used in distinguishing between general classes of A-DNA and B-DNA conformations (Haasnoot et al., 1984).

Refined Structures. Since visual examination of structures is quite subjective, we have found it useful to compare the torsional angles for each structure instead. We observe that the torsion angles within the base and deoxyribose ring were quite consistent between structures, while the torsion angles of the phosphate backbone showed considerable variation (see also Hare and Reid, unpublished results). This is to be expected since the only protons in this segment are the H5′ and H5″ methylene protons, which are poorly resolved and could not be used to generate distance constraints in this study. Therefore, the C3′–O3′–P–O5′–C5′–C4′ segment was merely constrained to bridge the gap between deoxyribose rings without violating any bond geometry or van der Waals constraints.

Although distance geometry produced several structures that are consistent with experimental data, it is important to qualify the meaning of these results. Since the distance information obtained by using NMR is approximate and incomplete, one should realize that the structures themselves reflect the nature of the data used to generate them. All of the final structures exhibit very similar local structure; for example, one will observe very similar relative orientations of adjacent base pairs between structures (Figure 8). However, considerable differences can be observed between the structures if one compares regions further apart in space (Figure 9A). This is not surprising since the maximum range of the NOE is about 5 Å, and uncertainties over long distances tend to propagate.

Pairing at G·T Mismatch Site. We observe the same pairing of T3 and G10 at the mismatch site for all refined structures of the 12-mer GT duplex. This is outlined for refined conformation designated C in Figure 7 (top) where the H3 imino proton of T3 pairs to the carbonyl-6 of G10 and the H1 imino proton of G10 pairs to carbonyl-2 of T3 at the mismatch site. This is exactly the pairing that was inferred for the G·T mismatch from solution one-dimensional NOE NMR studies of tRNA (Schimmel & Redfield, 1980) and DNA oligomer duplexes (Patel et al., 1982c), from theoretical computations of mismatches (Chuprina & Poltev, 1983; Keepers et al., 1984), and from more recent single-crystal X-ray studies of DNA oligomers (Brown et al., 1985, 1986; Kneale et al., 1985; Ho et al., 1985). For refined conformation C of the 12-mer GT duplex, the T3 imino to G10 imino intra base pair separation is 2.0 Å, while the T3 imino proton is 3.9 and 3.8 Å from the imino protons of G2·C11 and G4·C9 and the G10 imino proton is 3.8 and 3.4 Å from the imino protons of G2·C11 and G4·C9 flanking base pairs. These separations account for the observed NOEs between the imino protons of the mismatch pair and flanking Watson–Crick base pairs (Figure 2).

Structure Adjacent to Mismatch Site. The imino proton data demonstrate formation of G2·C11 and G4·C9 base pairs adjacent to the T3·G10 mismatch site in the 12-mer GT duplex at low temperature (Figure 2). We note that the bases at the T3·G10 mismatch site stack with adjacent Watson–Crick G2·C11 and G4·C9 base pairs in the refined structures of the 12-mer GT duplex (Figure 6, bottom). We are unable to define the phosphodiester backbone but note that the X-ray studies of G·C vs. G·T pairing have shown minimal perturbation in backbone torsion angles on replacing a Watson–Crick base pair by a wobble base pair (Kneale et al., 1985; Ho et al., 1985).

Glycosidic Torsion Angles. The glycosidic torsion angles for B-DNA and A-DNA are 257° and 201°, respectively. We observe that the glycosidic torsion angles for the 12-mer GT duplex are centered about 260° (Figure 10), indicative of B-DNA-like glycosidic angles at and adjacent to the mismatch site.

Sugar Pucker. The sugar pucker pseudorotation angles are 148° and 7° for B-DNA and A-DNA, respectively. The sugar pucker pseudorotation angles are centered about 150°, indicative of a B-DNA-type pucker at and adjacent to the mismatch site.

Refinement Methods. The distance information derived from NOE data sets has been primarily translated into solution

conformations of proteins and nucleic acids by using distance geometry (DG) (Braun et al., 1983; Arseniev et al., 1984; Williamson et al., 1985; Hare and Reid, unpublished results) and restrained molecular dynamics (Zuiderweg et al., 1985; Clore et al., 1985; Nilsson et al., 1986) computations. One of the main advantages of distance geometry as opposed to restrained molecular dynamics is that the former method incorporates into its error function constraints obtained entirely through experiment. Restrained molecular dynamics relies on a theoretical energy potential superimposed on a penalty function generated by input measured distances. Although this technique may lead to a higher level of structure determination due to the larger number of error parameters to be minimized, the validity of this determination is subject to the accuracy of the theoretical energy potential.

Current Limitations and Future Potential of NMR-DG Methods. The NMR-DG method applied in this paper depends critically on the realistic measurement of distances defined by upper and lower bounds. There are various pitfalls associated with these measurements (Bax & Lerner, 1986), and these are considered below. We have symmetrized our NOESY plots but are aware that volume integral measurements give distances that differ by up to 0.2 Å when measured on either side of the diagonal prior to symmetrization. This may be due to experimental or instrumentation deficiencies and may require future clarification. We have also assumed a single correlation time to describe the tumbling of the dodecanucleotide in solution. This is an approximation, but not a serious one, based on earlier results where Keepers and James (1984) demonstrated that octanucleotide duplexes are adequately represented by an effective overall correlation time. A more serious problem could surface if the dodecanucleotide or segments of it were in fast exchange between two or more conformations. Then, the NOE-based distances would be weighted toward the form with the shorter interproton separations due to the inverse sixth power dependence of the NOE buildup rate. We have observed two instances involving slow exchange between hairpin monomer–bulge duplex dimer structures and fast interconversions between stacked in and looped out orientations of extrahelical pyrimidines in the interior of helices. In each case, temperature-dependent chemical shift changes pointed toward this equilibrium, and NMR-DG refinements were not attempted. By contrast, such interconversions were not observed for the 12-mer GT duplex in this paper and the adenosine 13-mer duplex in the following paper (Hare et al., 1986). In the event that NMR-DG was carried out on a distance data set collected from a system in rapid exchange, it is extremely unlikely that these distances would be self-consistent. The distance geometry algorithm will identify sets of distances that are not self-consistent and would alert the user to some problem in the distance data set.

We wish to stress that structures determined by using distance geometry do not provide unique solutions to the conformation of oligonucleotide duplexes in solution. Rather, families of structures are deduced and their common properties evaluated. Due to the incomplete and approximate nature of the distance data sets currently available refinements may show considerable variation. The accuracy of distance geometry refinements could be substantially improved by refinement techniques that iteratively minimize the difference between experimental NOESY spectra and NOESY spectra simulated from the structure being refined. A back-calculation refinement technique of this nature would be analogous to crystallographic reciprocal space refinement methods because the structures would be refined against actual data rather than our interpretation of that data into distance bounds.

We are aware that the helices deduced from these NMR-DG studies are underwound compared to those observed by single-crystal X-ray methods. We are uncertain at this time whether this reflects the limitations and accuracy associated with the incomplete distance set currently available or reflects differences between solution and crystal structures. Additional studies employing interative comparisons between observed and computed NOESY data sets may clarify these observations. Although the technique is obviously capable of determining overall structural features, i.e., helix handedness, base pairing, and conformations of extrahelical bases, we have refrained from tabulating helix parameters due to the degree of determination of our current refinements.

Registry No. G, 73-40-5; T, 65-71-4; 12-mer GT, 80416-07-5.

REFERENCES

Aboul-ela, F. Koh, D., Tinoco, I., Jr., & Martin, F. H. (1985) *Nucleic Acids Res. 13*, 4811–4824.

Arseniev, S. A., Kondakov, V. I., Maiorov, V. N., & Bystrov, V. F. (1984) *FEBS Lett. 165*, 57–62.

Bax, A., & Lerner, L. (1986) *Science (Washington, D.C.) 232*, 960–967.

Braun, W., Wider, G., Lee, K. H., & Wuthrich, K. (1983) *J. Mol. Biol. 169*, 921–948.

Brown, T., Kennard, O., Kneale, G., & Rabinovich, D. (1985) *Nature (London) 315*, 604–606.

Brown, T., Kneale, G., Hunter, W. N., & Kennard, O. (1986) *Nucleic Acids Res. 14*, 1801–1809.

Chuprina, V. P., & Poltev, V. I. (1983) *Nucleic Acids Res. 11*, 5205–5222.

Clore, G. M., Gronenborn, A. M., Brunger, A. T., & Karplus, M. (1985) *J. Mol. Biol. 186*, 435–455.

Crick, F. H. (1976) *J. Mol. Biol. 19*, 548–555.

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

Dickerson, R. E., Drew, H. R., Connor, B. N., Wing, R. M., Fraitini, A. V., & Kopka, M. L. (1982) *Science (Washington, D.C.) 216*, 475–485.

Early, T. A., Olmsted, J., Kearns, D. R., & Lezius, A. G. (1978) *Nucleic Acids Res. 5*, 1955–1970.

Feigon, J., Leupin, W., Denny, W. A., & Kearns, D. R. (1983) *Biochemistry 22*, 5943–5951.

Haasnoot, C. A., Westerink, H. P., van der Marel, G. A., & van Boom, J. H. (1984) *J. Biomol. Struct. Dyn. 2*, 345–360.

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

Hare, D., Shapiro, L., & Patel, D. J. (1986) *Biochemistry* (following paper in this issue).

Havel, T. F., Crippen, G. M., & Kuntz, I. D. (1979) *Biopolymers 18*, 73–81.

Ho, P. S., Fredrick, C. A., Quigley, G. J., van der Marel, G. A., van Boom, J. H., Wang, A. H., & Rich, A. (1985) *EMBO J. 4*, 3617–3623.

Jack, A., Ladner, J. E., & Klug, A. (1976) *J. Mol. Biol. 108*, 619–649.

Kearns, D. R. (1984) *CRC Crit. Rev. Biochem. 15*, 237–290.

Keepers, J. W., & James, T. L. (1984) *J. Magn. Reson. 57*, 404–426.

Keepers, J., Schmidt, P., James, T., & Kollman, P. A. (1984) *Biopolymers 23*, 2901–2929.

Kennard, O. (1986) *J. Biomol. Struct. Dyn. 3*, 205–226.

Kneale, G., Brown, T., Kennard, O., & Rabinovich, D. (1985) *J. Mol. Biol. 186*, 805–814.

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

Pardi, A., Walker, R., Rapoport, H., Wider, G., & Wuthrich, K. (1983a) *J. Am. Chem. Soc. 105*, 1652-1653.
Pardi, A., Morden, K. M., Patel, D. J., & Tinoco, I., Jr. (1983b) *Biochemistry 22*, 1107-1113.
Patel, D. J., Pardi, A., & Itakura, K. (1982a) *Science (Washington, D.C.) 216*, 581-590.
Patel, D. J., Kozlowski, S. A., Rice, J. A., Marky, L. A., Breslauer, K. J., Broka, C., & Itakura, K. (1982b) in *Topics in Nucleic Acid Structure* (Neidle, S., Ed.) Vol. 2, pp 81-136, MacMillan, London.
Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Dallas, J., Itakura, K., & Breslauer, K. J. (1982c) *Biochemistry 21*, 437-444.
Patel, D. J., Kozlowski, S. A., Ikuta, S., & Itakura, K. (1984) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 43*, 2663-2670.
Quigley, G. J., & Rich, A. (1976) *Science (Washington, D.C.) 194*, 796-806.
Rich, A., Nordheim, A., & Wang, A. H. (1984) *Annu. Rev. Biochem. 53*, 791-846.
Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H., & Kaptein, R. (1984) *Biochemistry 23*, 1371-1376.
Schimmel, P. R., & Redfield, A. G. (1980) *Annu. Rev. Biophys. Bioeng. 9*, 181-221.
Tibanyenda, N., de Bruin, S. H., Haasnoot, C. A., van der Marel, G., van Boom, J. H., & Hilbers, C. W. (1984) *Eur. J. Biochem. 139*, 19-27.
Weiss, M. A., Patel, D. J., Sauer, R. T., & Karplus, M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 130-134.
Wemmer, D. E., & Reid, B. R. (1985) *Annu. Rev. Phys. Chem. 36*, 105-137.
Williamson, M. P., Havel, T. F., & Wuthrich, K. (1985) *J. Mol. Biol. 182*, 295-315.
Zuiderweg, E. R., Scheek, R. M., Boelens, R., van Gunsteren, W., & Kaptein, R. (1985) *Biochimie 67*, 707-715.

Extrahelical Adenosine Stacks into Right-Handed DNA: Solution Conformation of the d(C-G-C-A-G-A-G-C-T-C-G-C-G) Duplex Deduced from Distance Geometry Analysis of Nuclear Overhauser Effect Spectra†

Dennis Hare
Infinity Systems, 14810 216th Avenue NE, Woodinville, Washington 98072
Lawrence Shapiro and Dinshaw J. Patel*
Department of Biochemistry and Molecular Biophysics, College of Physicians and Surgeons, Columbia University, New York, New York 10032

Received May 29, 1986; Revised Manuscript Received July 25, 1986

ABSTRACT: This paper reports on features of the three-dimensional structure of the d(C-G-C-A-G-A-G-C-T-C-G-C-G) self-complementary duplex (designated adenosine 13-mer), which contains symmetrical extrahelical adenosines in the interior of the helix. The majority of the protons have been assigned from two-dimensional nuclear Overhauser effect (NOESY) spectra of the adenosine 13-mer in H_2O and D_2O solution. The measurement of NOESY cross-peak volume integrals as a function of mixing time has yielded a set of 96 short (<4.5-Å) proton–proton distances defined by lower and upper bounds, which have served as input parameters for a distance geometry analysis of one symmetric half of the adenosine 13-mer duplex. We demonstrate that the extrahelical adenosine stacks into the duplex for all refined structures without disruption of base pairing on either side of the modification site. The distance geometry refinement yields two classes of conformations consistent with distance measurements but which differ in orientation of the stacked extrahelical adenosine at the modification site.

The conformation of DNA duplexes containing helical interruptions resulting from the presence of an unpaired base on one strand is of considerable interest in mechanisms of frame shift mutagenesis (Streisinger et al., 1966). The extrahelical base has the potential of either stacking into the duplex or looping out into solution so that addition or deletion mutations can result due to partial misalignment of strands during replication. The repair mechanisms are also governed by the conformation and stability of the extrahelical base at the helix interruption site in the interior of the duplex. Recent results on the role of aberrant nucleic acid conformations in frame shift mutagenesis have stressed the formation of extrahelical bases and mismatched base pairs in the stem of hairpin loops generated at DNA sequences containing quasipalindromic sites (Ripley, 1982; de Boer & Ripley, 1984).

Some structural features at extrahelical sites have been determined on the basis of analysis of one-dimensional NMR chemical shift and distance dependent nuclear Overhauser effect (NOE) parameters. Thus, it has been shown that the extrahelical adenosine stacks into the self-complementary d(C-G-C-A-G-A-A-T-T-C-G-C-G) duplex (Patel et al., 1982). Hydrogen exchange measurements at the individual base pair level demonstrate that stacking of the extrahelical adenosine results in faster hydrogen exchange rates and higher activation barriers at the flanking, as well as more distant, base pairs (Pardi et al., 1982). The thermodynamics and kinetics of the

† The research was supported by National Institutes of Health Grants GM 34504-02 to D.J.P. and GM 35620-02 to D.H. The NMR spectrometers were purchased from funds donated by the Robert Woods Johnson, Jr., Trust and the Matheson Trust toward setting up the NMR Center in the Basic Medical Sciences at Columbia University.

0006-2960/86/0425-7456$01.50/0