Solution Structure of the T_nA_n DNA Duplex GCCGTTAACGCG Containing the *Hpa*I Restriction Site[†,‡]

Seong-Gi Kim[§] and Brian R. Reid*[,§,∥]

Departments of Chemistry and Biochemistry, University of Washington, Seattle, Washington 98195

Received June 26, 1992; Revised Manuscript Received September 28, 1992

ABSTRACT: The solution structure of the self-complementary DNA duplex $[d(GCCGTTAACGGC)]_2$, which contains the *Hpa*I restriction site GTTAAC, has been elucidated by two-dimensional NMR, distance geometry (DG), and NOE back-calculation methods. Initial distance constraints were determined by polynomial fitting the two-spin initial NOE rates; backbone constraints from NOE and *J*-coupling observations (Kim et al., 1992) were included. RMSDs between initial-distance-refined structures derived from random-embedded DG, A-DNA, and B-DNA starting structures were all in the range 0.5–1.0 Å, indicating good convergence properties of the algorithm, regardless of the starting structure. A semiautomatic back-calculation refinement procedure was developed and used to generate more refined structures for which the BKCALC-simulated NOE volumes matched the experimental data. The six final structures refined from various starting structures exhibit very good agreement with the experimental data (R values = 0.18) and converge well to within 0.8-Å RMSD differences for the central 8 base pairs. The torsion and pseudorotation phase angles were found to be well determined by the data, and the local helical parameters for each base step converged quite well. The final structures show that the central T6–A7 step is somewhat underwound (twist angle ca. 29°), with a large negative cup and a normal (wide) minor groove width, while the T5–T6 and A7–A8 steps have a partially narrowed minor groove.

Determining the detailed structural variations in DNA is the most important step in elucidating the effects of sequence on local duplex structure and understanding the molecular basis of sequence-specific recognition of DNA by proteins. Single-crystal X-ray studies of short DNA duplexes show extensive local deviations from ideal B-DNA structure (Drew et al., 1981; Dickerson & Drew, 1981; Calladine, 1982; Dickerson, 1983, 1990; Prive et al., 1991; Yanagi et al., 1991). However, since DNA recognition takes place in solution, the structure determined in the crystalline state may not represent the biological relevant structure; solid-state DNA structures are known to be affected by molecular packing forces (DiGabriele et al., 1989; Shakked et al., 1989) and hydration level (Franklin & Gosling, 1953). High-resolution nuclear magnetic resonance (NMR) spectroscopy is probably the only method that can provide structural information at the atomic level directly in solution. Thus, the determination of the structure of short DNA duplexes in solution by NMR has become the goal of several laboratories [for reviews, see Wüthrich (1986), Reid (1987), Hosur et al. (1988), van de Ven and Hilbers (1988), and Clore and Gronenborn (1989)]. The most common method used for determining solution structure relies heavily on the dipole–dipole coupling (nuclear Overhauser effect) interaction between protons separated by less than 5 Å. The nuclear Overhauser effect (NOE) is proportional to r^{-6}, and distance information can be obtained by analysis of the crosspeak intensities in two-dimensional NOE experiments.

A major concern in elucidating the solution structures of DNA is the precision of determining local structural detail, i.e., torsion angles and helical parameters. Clore, Gronenborn, and co-workers reported that variations in DNA helical parameters are strongly dependent on the sequence (Nilsson et al., 1986; Happ et al., 1988; Clore et al., 1988). Their approach involved energy minimization and molecular dynamics constrained by two-spin approximated NOE distance estimates, together with α, β, γ, ϵ, and ζ backbone torsion angle constraints derived from X-ray studies of B-form or A-form DNA. Back-calculation of their proposed structures to see if they reproduced the experimental NOESY spectra was not performed in these studies. In contrast, Pardi and co-workers (Pardi et al., 1988; Metzler et al., 1990) reported that the local conformation is not particularly dependent on the sequence and varies widely within a set of structures generated from the same distance constraints. They used distance geometry, molecular dynamics, and back-calculation methods, but without any backbone constraints. The major differences between the two studies are (i) the presence of backbone angle constraints (even if assumed from solid-state data) and (ii) the use of back-calculation methods in structure refinement. Back-calculation per se does not improve convergence; rather, it takes into account spin diffusion and produces a model structure that better fits the experimental data. The use of backbone constraints (ad hoc or otherwise) may be the major contributor to structural convergence and the precision of the resulting local geometry. Recently we have developed a reliable method (Kim et al., 1992) whereby the backbone angles β, γ, and ϵ (and indirectly α and ζ) can be constrained experimentally via a combination of measurable *J* couplings, line widths, and lower bound distance constraints in NOESY and COSY spectra; this method relies only on proton NMR data without invoking any structural assumptions from crystal data or energetic considerations and leads to much-improved structural convergence in a set of refined final structures that originate from different starting conformations.

[†] Support from National Institutes of Health Grant GM32681 is gratefully acknowledged.

[‡] The coordinates of a representative structure will be deposited in the Brookhaven data bank.

* Author to whom correspondence should be addressed at the Department of Chemistry.

[§] Department of Chemistry.

[∥] Department of Biochemistry.

0006-2960/92/0431-12103$03.00/0

In this paper we present the family of closely related solution structures for a 12-base DNA duplex containing the *Hpa*I restriction site, d[GCCGTT↓AACGGC]$_2$, produced by distance geometry (DG) calculations that incorporate these additional experimentally determined backbone constraints. In these structure calculations we refined the initial starting structures by a semiautomatic back-calculation/simulated annealing method until their simulated crosspeak volumes matched the experin ntal crosspeak volumes. The iterative refinement procedure was performed on a variety of starting structures, including random-embedded DG, A-form, and B-form structures. The final structures converge well regardless of the starting structure. The present paper presents the torsion angles and helical parameters of the final refined structures.

MATERIALS AND METHODS

Sample Preparation. The DNA dodecamer containing the *Hpa*I restriction sequence [d(GCCGTT↓AACGGC)]$_2$ was synthesized and purified as described previously (Kintanar et al., 1987). The DNA sample (28 mg) was dissolved in 0.4 mL of buffer containing 0.7 mM EDTA, 30 mM sodium phosphate, and 75 mM NaCl. The sample was repeatedly lyophilized to dryness and finally dissolved in 0.4 mL of 99.996% D_2O to produce a final DNA concentration of ca. 9.6 mM.

NMR Spectroscopy. NMR experiments were performed at 500 MHz either on a Bruker AM-500 spectrometer or on a home-built NMR spectrometer (Gladden and Drobny, unpublished results). Four NOESY spectra with mixing times of 40, 80, 120, and 200 ms were collected using the phase-sensitive method (States et al., 1982) within a single 4-day time period, without removing the sample from the spectrometer. In each NOESY spectrum, the mixing time was randomly varied over ±10% of the designated mixing time to eliminate zero-quantum coherence. For all NOESY experiments, 1024 complex points in t_2 and 400 points in t_1 were collected with a relaxation delay (RD) of 2.0 s and a spectral width (SW) of 4386 Hz. The Exclusive COSY (E. COSY) spectrum (Griesinger et al., 1986, 1987) was recorded in the phase-sensitive mode with time-proportional phase incrementation (Drobny et al., 1979; Marion & Wüthrich, 1983). To increase sensitivity and maximize resolution in t_2, the E. COSY spectrum was acquired with 2048 complex points in t_2 and with 48 scans per t_1 experiment. TOCSY spectra (Braunschweiler & Ernst, 1983) were collected in the phase-sensitive mode with time-proportional phase incrementation and a 150-ms isotropic mixing time. An MLEV-16 pulse sequence (Levitt et al., 1982) was used for long-range coherence transfer by isotropic mixing within the scalar-coupled spin systems i.e., furanose rings. These spectra were acquired with 1024 complex points in t_2 and 64 scans per t_1 experiment. The NMR data were transferred to an IRIS 4D computer and processed using the FTNMR and FELIX programs (Hare Research, Inc., Woodinville, WA). NOESY data sets were zero-filled to 2048 points in each dimension, apodized using 3 Hz of Gaussian broadening in t_2 and a sine-squared 90° phase-shifted function in t_1, and Fourier-transformed. E. COSY data were processed using 6 Hz of exponential narrowing and 6 Hz of Gaussian broadening in the t_2 domain and a skewed-sinebell window function in t_1. Spectral resolution was enhanced by zero-filling to 4096 points in t_2 (1.07 Hz per point) and 2048 points in the t_1 domain. TOCSY data were processed using 90°-shifted sine-squared window functions in t_2 and t_1. In all spectra, t_1 ridges were attenuated by multiplying the first t_1 slice by 0.5 (Otting et al., 1985).

The resolved crosspeak volumes corresponding to each detectable proton–proton dipole–dipole interaction were obtained from the NOESY spectra (vide infra). Assuming a two-spin approximation, i.e., no spin diffusion, the distance (r_{ij}) can be estimated relative to a fixed distance (r_{ref}) from the relative cross relaxation rates Rc_{ij} and Rc_{ref} from the relationship $r_{ij} = r_{ref}(Rc_{ref}/Rc_{ij})^{1/6}$. A cytosine H5–H6 distance of 2.46 Å was used as the reference (r_{ref}).

From the E. COSY spectrum, $^3J_{H1'-H2'}$ and $^3J_{H1'-H2''}$ were determined from H1′–H2″ and H1′–H2′ crosspeaks, respectively (Griesinger et al., 1987; Bax & Lerner, 1988). Although $J_{H2''-H3'}$ and $J_{H3'-H4'}$ cannot be accurately determined due to passive couplings, including phosphorus coupling, the intensities of the COSY crosspeaks are directly related to the magnitudes of their J couplings, provided T_2 relaxation is similar for all spins. The relative intensities of H2″–H3′ and H3′–H4′ crosspeaks (approximate J coupling constants) were determined, and the sums of J coupling constants (ΣJ) for each proton were also determined.

Distance Geometry Calculations. The concepts and details of the distance geometry algorithm DSPACE have been described previously (Hare & Reid, 1986; Hare et al., 1986a,b; Nerdal et al., 1988, 1989; Banks et al., 1989). All distance geometry calculations were performed on an IRIS 4D workstation. To initiate the distance geometry calculation, upper and lower bounds were assembled for all distances implicit in the primary structure and stored in a bounds matrix, along with the experimentally determined distance bounds; these latter were determined by polynomial fitting (vide infra). The entire bounds matrix containing all of the available distance information was subjected to several smoothing procedures including the triangle and inverse triangle inequalities (Crippen, 1981; Havel & Wüthrich, 1985). A series of distance matrices were then generated by choosing random distances between the lower and upper bounds for each each matrix element, to produce a trial matrix, which was converted to a metric matrix and embedded in 3-space (Crippen, 1981; Havel et al., 1983). The initially embedded structures were refined by a combination of conjugate gradient refinement and a simulated annealing algorithm (Metropolis et al., 1953; Kirkpatrick et al., 1983), which uses the deviation from the distance bounds as its primary penalty parameter (Nerdal et al., 1988). The penalty (target) function is a sum of distance-based pseudoenergy violations consisting of covalent bond, experimental distance, and nonbonded contact errors. The covalent penalty includes bond length, bond angle, chirality, and planarity errors; the nonbonded penalty is a van der Waals contact error. The conjugate gradient refinement is simply a penalty minimizer and may not overcome local minima; the simulated annealing steps move all atoms of the system by Newton's equations of motion for a suitable time at a temperature that is directly proportional to the kinetic energy (penalty) of the system (Metropolis et al., 1953; Kirkpatrick et al., 1983; Nerdal et al., 1988). Thus, simulated annealing involves heating the system followed by slow cooling to overcome local minima and access the global minimum region of the target function (penalty). To test the convergence properties of the method, canonical B-form and A-form DNA structures (Arnott & Hukins, 1972a,b) were also randomized prior to refinement against the distance file via a combination of simulated annealing and conjugate gradient.

Back-Calculation Refinement. Back-calculations of NOESY spectra from various trial structures were performed with

Table I: Chemical Shift of Protons of *Hpa*I Dodecamer

residue	H6/H8	H5/H2/M5	H1′	H2′	H2″	H3′	H4′
G1	7.91		5.92	2.59	2.73	4.84	4.26
C2	7.46	5.31	6.06	2.13	2.48	4.85	4.23
C3	7.43	5.57	5.56	2.10	2.41	4.84	4.11
G4	7.91		6.00	2.65	2.80	4.98	4.39
T5	7.21	1.44	5.95	2.01	2.50	4.82	4.18
T6	7.31	1.66	5.64	1.99	2.36	4.85	4.07
A7	8.24	6.88	5.86	2.72	2.87	5.03	4.38
A8	8.07	7.54	6.02	2.53	2.76	4.98	4.39
C9	7.08	5.09	5.48	1.71	2.18	4.73	4.05
G10	7.75		5.55	2.60	2.68	4.93	4.27
G11	7.67		5.93	2.48	2.69	4.94	4.35
C12	7.26	5.10	6.09	2.21	2.21	4.48	3.98

the program BKCALC (Hare Research), previously discussed in detail by Banks et al. (1989). The simulated NOESY spectra were calculated from the proton *xyz* coordinates of any given molecular structure by numerical integration of all magnetization transfer pathways between neighboring protons. This algorithm takes into account all spin diffusion pathways and generates NOESY crosspeak volumes at any desired mixing time. NOESY spectra were then generated from the calculated crosspeak volumes using the appropriate chemical shifts and line widths, assuming a Gaussian line shape for each crosspeak.

The NOESY spectra of a particular structure were calculated at various mixing times, and the crosspeak volumes were then compared to the experimental NOESY crosspeak volumes at the corresponding mixing times. The resolved base–sugar and base–base proton crosspeaks were subjected to an automated refinement program in which the distance bounds are automatically adjusted by quantitative comparison between experimental and simulated volume integrals (vide infra). The individual component volumes of overlapped crosspeaks cannot be accurately determined and, for such multiple peaks, visual comparisons/manual distance adjustments were performed. The new distance bounds file was then used to create a new structure by simulated annealing and conjugate gradient refinement against the new distance bounds. These back-calculated refinement cycles were performed until the simulated NOESY crosspeak volumes matched the experimental data, as judged by the *R* factor (vide infra).

Calculation of Helical Parameters. All helical parameters for each resulting structure were calculated by the program NEWHELIX90 (kindly provided by R. E. Dickerson), derived originally from HELIX (J. M. Rosenberg), and BROLL, CYLIN, and DTORAN (R. E. Dickerson). The designation of helical parameters follows the 1989 Cambridge convention [*EMBO J.* (1989) *8*, 1–4]. Some of the helical parameters are dependent on an overall helical axis, and the best fit helix axis was calculated using all vectors between C1′ atoms on successive bases; all 12 base pairs were used to define the helix axis.

RESULTS AND DISCUSSION

Resonance Assignment. The resonance assignments of nonexchangeable protons in the *Hpa*I restriction sequence [d(GCCGTTAACGGC)]$_2$ were made (Table I) using standard assignment strategies developed previously (Hare et al., 1983; Scheek et al., 1983). Contour plots of the base proton to H1′ region and the base proton to H2′/H2″ region NOESY spectra are shown in Figures 1 and 2, respectively. Note the relatively weak NOEs between A7 H8 and T6 H1′ and between G10 H8 and C9 H1′. Additionally, the contour plot of the

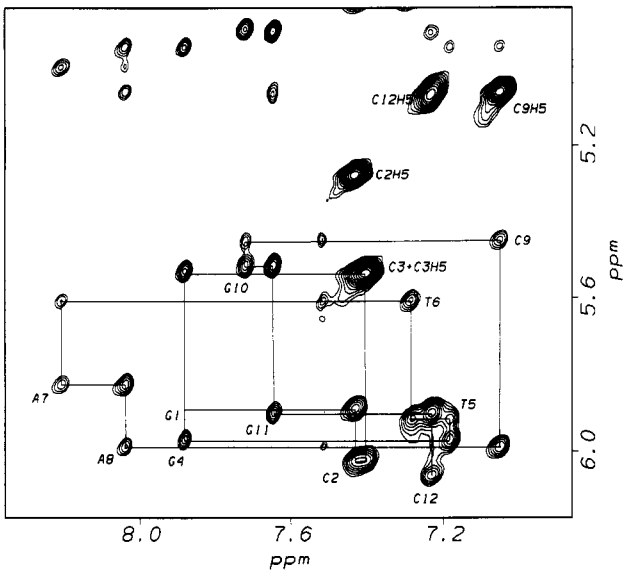

FIGURE 1: Contour plot of the H6/H8–H1′/H3′ NOESY region at 200-ms mixing time for the *Hpa*I dodecamer [d(GCCGTTAACG-GC)]$_2$. The intraresidue H6/H8-H1′ crosspeaks are labeled with the residue number. The sequential connectivity, (*n*)H6/H8–(*n*)-H1′ → (*n*+1)H6/H8–(*n*)H1′ → (*n*+1)H6/H8–(*n*+1)H1′, is connected by a solid line. The NOEs from A8 H2 (7.54 ppm) to A8 H1′, T6 H1′, and C9 H1′ are clearly detectable, with the latter two obviously stronger than the former.

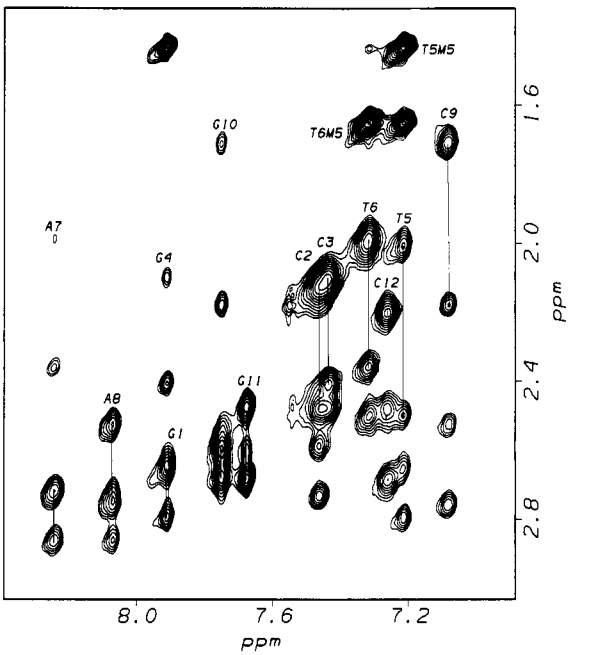

FIGURE 2: Contour plot of the H6/H8–H2′/H2″ region in the NOESY spectrum at 200-ms mixing time for the *Hpa*I dodecamer. The residues are labeled at the position of the H6/H8 chemical shift. For each residue, a vertical bar connects the (*n*)H6/H8–(*n*)H2′ crosspeak (upfield) to the (*n*)H6/H8–(*n*)H2″ crosspeak (downfield).

H1′–H2′/H2″ region in the E. COSY spectrum is shown in Figure 3. Apart from some overlaps due to the similar chemical shifts of C2 H6 and C3 H6, the spectra are reasonably well-resolved. These assignments were all cross-checked for consistency in NOESY, E. COSY, and TOCSY spectra.

Initial Distance Determination. In a two-spin system, the NOE crosspeak volume (intensity) V_{ij} at mixing time τ_m is

$$V_{ij}(\tau_m) = 0.5V_{jj}(0)(1 - e^{-Rc_{ij}\tau_m})e^{-R_L\tau_m}$$

where V_{ij} is the crosspeak intensity correlating proton j (ω_1 domain) and proton i (ω_2 domain), Rc_{ij} is the cross relaxation

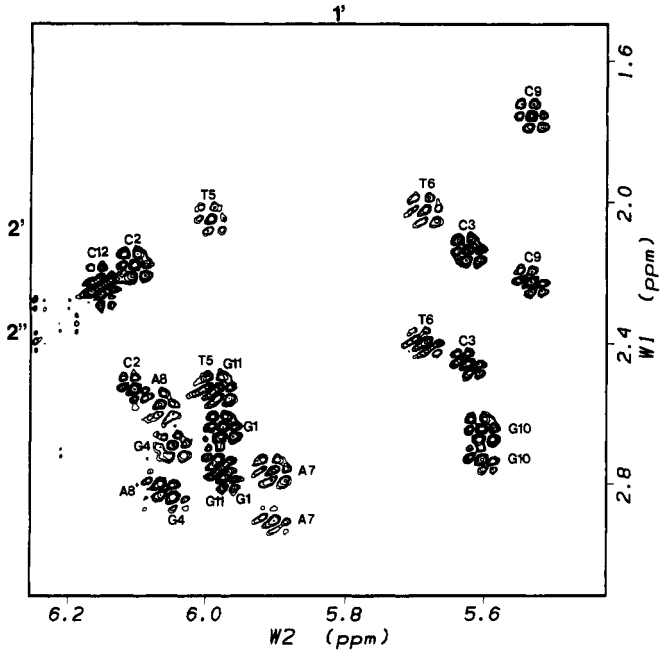

FIGURE 3: Contour plot of the H1′–H2′/H2″ region in the E. COSY spectrum of the *Hpa*I dodecamer. The H2′ resonance is upfield of the H2″ resonance.

rate, and R_L is the leakage rate constant (Ernst et al., 1987). In a multiple-spin system, spin diffusion effects contribute to the ij crosspeak intensity due to common neighbors $j \leftrightarrow k \leftrightarrow i$. The exponential term can be expanded in a Taylor series; the leakage rate and the spin diffusion effect contribute in the second- and higher-order terms as

$$V_{ij}(\tau_m) = (\mathrm{Rc}_{ij}\tau_m - 0.5\sum \mathrm{Rc}_{ik}\mathrm{Rc}_{kj}\tau_m^2 + \ldots) \times (1 - R_L\tau_m + \ldots)V_{jj}(0)/2$$

The contribution to the crosspeak volume buildup rate from terms higher than second-order can be ignored over typical experimental mixing times, and the buildup rate of crosspeak volumes can be estimated fairly accurately by a second-order polynomial function (Fejzo et al., 1989; Hyberts & Wagner, 1989). The second-order term is composed of (i) the multiple-step magnetization transfer (spin diffusion) via one or more common neighbors and (ii) the magnetization leakage to the surroundings, even for the isolated spin pair. The first-order term is not affected by spin diffusion or leakage and is thus used for determining Rc_{ij}.

To obtain the relative Rc_{ij} rates, eight crosspeak volumes (the above- and below-diagonal $j \rightarrow i$ and $i \rightarrow j$ peaks at four mixing times) were used for the second-order polynomial fitting. The slope (first-order term) was determined and then converted to a distance by calibration with respect to the reference H5–H6 distance and corresponding cross-relaxation rate (see Materials and Methods). The average of the C2, C9, and C12 H5–H6 crosspeak buildup rates was used as the reference Rc. Base–sugar and base–base proton crosspeaks were subjected to the second-order polynomial analysis to determine distances, which were used as input for DSPACE. Typical initial-estimate distance uncertainties used to define the range between upper and lower bounds were 0.2–0.4 Å for base proton–sugar proton crosspeaks and 0.6–0.8 Å for base proton–base proton crosspeaks. Even though the second-order polynomial approach partially accounts for spin diffusion, the base proton–(*n*)H2″/H3′ initial distance bounds estimates were withheld from the distance file in calculating initial DG structures due to strong spin diffusion pathways, e.g., H6/H8 ↔ (*n*)H2′ ↔ (*n*)H2″ and H6/H8 ↔ (*n*)H2′ ↔ (*n*)H3′; with rather wide distance bounds, they were used for later back-calculation refinement of structures. For crosspeaks involving a *rotating* methyl group, crosspeak intensities are dependent not only on the distance but also on the actual geometry, i.e., polar angles in the molecular frame of the internuclear vector (Tropp, 1980; Tropp & Redfield, 1981). Thus, thymine M5 pseudoatom to proton distance bounds were not used in the initial DG studies, and wide distance bounds were used during later back-calculation refinement of structures. Thus, in the calculation of initial DG structures only the initial distance bounds between base protons and base/sugar protons, i.e., H6/H8/H5–(*n*)H1′/H2′, H6/H8/H5–(*n*–1)H1′/H2′/H2″/H6/H8, and H2–(*n*+1)H1′/(*m*+1)H1′, as well as the backbone constraints, were used. The interproton distances, both initial and after refinement, together with their bounds ranges are listed in the supplementary material.

For volume comparisons between experimental NOE and simulated NOE crosspeaks, the experimental crosspeak volumes at all mixing times were fitted to a second-order polynomial to average the volume of the above- and below-diagonal crosspeaks, V_{ij} and V_{ji}, at each mixing time, and to smooth the volume differences between different mixing times. The fitted volumes at each mixing time were then calculated and used as experimental volumes for comparison with the simulated volumes of the current structure (see later).

The potential sources of error in interproton distance estimates from NOESY crosspeaks are (a) variations in correlation times due to differential motion, (b) incorrect estimations of the initial cross relaxation rate due to significant spin diffusion contamination, (c) differences in initial magnetization due to different proton spin–lattice relaxation times (T_1) in short relaxation delay experiments, and (d) rapid relaxation, i.e., line broadening, of NOESY crosspeaks due to paramagnetic impurities. First, although Clore and Gronenborn (1984a,b) claimed that the correlation time (motion) of sugar protons is 3 times shorter (faster) than that of base protons based on NOESY crosspeak buildup rates at 100–800-ms mixing times, their conclusion is based on attributing a spin diffusion effect to local motion. More recently, Reid et al. (1989) demonstrated (from 15-ms NOESY spectra) that sugar protons and base protons actually have the same correlation time. Second, to avoid spin diffusion contributions, NOESY spectra with very short mixing time, i.e., ca. 15 ms, should be used to estimate the initial Rc (Reid et al., 1989). In practice the ability to sample the true initial cross-relaxation rates is limited by the signal-to-noise ratio of the small crosspeaks (especially for long interproton distances) at very short mixing times. In the present study 40 ms was used as the shortest mixing time in NOESY spectra. To partially eliminate the spin diffusion contribution to the initial Rc in such data, fitting to a second-order polynomial function was used and, furthermore, H6/H8–H2″/H3′ distance constraints were not used in calculating the initial structures due to the expectation of large spin diffusion contributions to these pathways. In addition, back-calculation simulation of the NOESY spectrum including all multiple-spin pathways was performed to correct errors in the initial distance estimates (vide infra). Third, while the duty cycle in the NOESY spectra (2.3 s) is not much longer than the longest proton T_1, if the various proton T_1's are similar, then all protons will have the same partially relaxed initial magnetization as the reference protons. However, if they are not, then protons with a T_1 longer than the reference will have less initial magnetization [$V_{jj}(0) < V_{ref}(0)$] and will produce less intense crosspeaks

(see equation above), resulting in a distance overestimation. Experimentally, we determined that the relative intensity [$V_{jj}(0)$] ratios of 1-D spectra accumulated with 15- vs 2.3-s relaxation delays are not significantly different, i.e., except for the adenine H2 protons, which were reduced by ca. 23% in the 2.3-s RD spectra, the T_1's are quite similar. Because of the longer relaxation time of H2 protons, the H2–H1′ crosspeak intensity (but not the H1′–H2 intensity) was corrected for partial relaxation by a factor of 1.3 before distance estimation. Fourth, paramagnetic impurities can broaden site-specific proton line widths due to very short spin–spin relaxation times (T_2) and lead to reduced crosspeak intensities (Sletten et al., 1991; Lane et al., 1991). In the present [d(GCCGTTAACGGC)]$_2$ study, the G10 H8 resonance was initially rather broad in solutions containing only 0.1 mM EDTA, suggesting a local paramagnetic ion site, but was restored to the same line width as the other H8 resonances after the DNA sample was titrated to 0.7 mM EDTA before the NOESY data were acquired.

Sugar Conformations. Many of the intrasugar NOE crosspeaks in [d(GCCGTTAACGGC)]$_2$ are severely overlapped, making it difficult to integrate them and to measure a sufficient number of diagnostic distances to define the sugar conformation. We therefore used a combination of J coupling constants and H1′–H4′ and H2″–H4′ distances (when available) to independently determine the individual sugar conformations. The intrasugar proton–proton distances corresponding to that particular sugar conformation were then used in the distance file to constrain the sugar pseudorotational phase angle P, assuming a single sugar conformation (Kim et al., 1992; Hosur et al., 1986; Chary & Modi, 1988; van de Ven & Hilbers, 1988). The J coupling measurements from the E. COSY spectrum (Figure 3) indicate that in every residue $J_{H1'-H2'}$ is greater than $J_{H1'-H2''}$; this limits P to 90–190°. $J_{H3'-H4'}$ and $J_{H2''-H3'}$ are even more diagnostic concerning the sugar conformation. While these J couplings cannot be accurately determined due to passive couplings, including phosphorus coupling, the intensities of the COSY crosspeaks are directly related to the magnitude of the active J couplings. In the [d(GCCGTTAACGGC)]$_2$ COSY spectrum, C9 and C12 have strong H3′–H4′ crosspeaks and weak H2″–H3′ crosspeaks; C2, C3, T5, and T6 and the purines G1 and G11 have respectable H3′–H4′ and very weak or nonexistent H2″–H3′ crosspeaks, while G4, A7, A8, and G10 have weak H3′–H4′ and nonexistent H2″–H3′ crosspeaks [see Figure 2 of Kim et al. (1992)]. From this J coupling analysis, the P values of C9 and C12 are in the range 90–126°; for C2, C3, T5, T6, G1, and G11, P is 120–150°; and for G4, A7, A8, and G10, P is 144–162°.

The TOCSY experiment may provide further information on coupling "bottlenecks" in the J-coupled coherence transfer pathway among deoxyribose protons. While coherence transfer oscillates harmonically between two coupled spins at the frequency of the coupling constant J, coherence transfer between H1′ and H4′ is near maximal at a 150-ms isotropic mixing time for sugar conformations in the S-type (C1′-exo to C2′-endo) conformational range (Flynn et al., 1988; Remerowski et al., 1989; Cavanagh et al., 1990). The 150-ms mixing time TOCSY spectrum of [d(GCCGTTAACGGC)]$_2$ has weak H1′–H4′ peaks for all pyrimidines and nonexistent H1′–H4′ crosspeaks for all purines; this suggests that purine $J_{H3'-H4'}$ coupling constants (and possibly $J_{H2''-H3'}$) are the bottleneck for coherent magnetization transfer between H1′ and H4′ and is consistent with the E. COSY observation of weak H3′–H4′ coupling in G4, A7, A8, and G10.

The sugar conformation can also be determined independently from NOESY spectra (Chary & Modi, 1988; van de Ven & Hilbers, 1988). The H1′–H4′ and H2″–H4′ distances are the most diagnostic indicators of the pseudorotation phase angle. For $90° \leq P \leq 144°$, the H1′–H4′ (≤3.0 Å) crosspeak intensity should be greater than the H1′–H2′ intensity (ca. 3.0 Å) in the initial rate regime of the NOE buildup (van de Ven & Hilbers, 1988; Kim et al., 1991). The experimental H1′–H4′ distances for the purines range from 3.0 to 3.3 Å, while for pyrimidines the range is 2.7–2.9 Å. Further, NOESY spectra at 40- and 80-ms mixing times were used to visually compare the H2′–H4′ and H2″–H4′ crosspeak intensities. When the H2″–H4′ crosspeaks (≤3.85 Å) are more intense than the H2′–H4′ crosspeaks (3.75–3.90 Å throughout the 0–200° range of pseudorotational angle), the P value is less than 126°. From the combined J coupling and NOE analyses of each individual residue, the pseudorotational phase angles for C9 and C12 are in the range 100–126°; for C2, C3, T5, and T6, P is 120–135°; for G1 and G11, P is 130–150°; and for G4, A7, A8, and G10, P is 144–162°. The interproton H1′–H2′, H1′–H2″, H1′–H3′, H1′–H4′, H2′–H3′, H2′–H4′, H2″–H3′, H2″–H4′, and H3′–H4′ distance bounds corresponding to these P values were used as input distance bounds for the program DSPACE.

The approach outlined above relies on the assumption of a single fixed sugar conformation. Altona and co-workers (Rinkel & Altona, 1987) and James and co-workers (Schmitz et al., 1990) have reported that a two-site jump model involving conformational jumps between N (C3′-endo; P = ca. 18°) and S (C2′-endo; P = ca. 162°) sugar conformers is necessary to rationalize DNA sugar conformations based on experimental J couplings or COSY multiplet patterns. These early studies were performed on small-sized DNA duplexes (≤4 base pairs) or on single strands where the natural line width is narrow enough to determine accurate J coupling constants; however, one cannot assume a rigid single structure in such small oligomers. In moderate size DNA duplexes, i.e., 12 base pairs, the J coupling constants cannot be easily determined due to broad line widths (line width at half-height ≥ J coupling constant), passive coupling contributions, and overlapping peaks. Thus, COSY multiplet simulations were used to determine the N–S conformational ratio, but the simulations depend not only on the J coupling constants but also on the natural line width, window function, truncation, etc. (Celda et al., 1989; Gochin et al., 1990). Recently James and co-workers (Gochin et al., 1990; Gochin & James, 1991) showed that, in contrast to their previous interpretation of COSY crosspeaks (Schmitz et al., 1990), a single fixed sugar conformation is perfectly consistent with the COSY multiplet patterns in octameric DNA duplexes. From the E. COSY spectrum of [d(GCCGTTAACGGC)]$_2$, relatively precise $J_{H1'-H2'}$ and $J_{H1'-H2''}$ couplings and qualitative $J_{H2''-H3'}$ and $J_{H3'-H4'}$ couplings were determined, and these values can be explained quite adequately by a single sugar conformation. Furthermore, rapid conversion between N–S sugar conformations, if it occurred, would lead to averaging of both J coupling constants and NOESY crosspeak intensities. Interestingly, the distance between H1′ and H4′ is 3.5 Å for both N and S conformers; regardless of the ratio between N and S conformations, the H1′–H4′ distance should be ca. 3.5 Å. However, in [d(GCCGTTAACGGC)]$_2$, the H1′–H4′ distances are 2.7–2.9 and 3.0–3.3 Å for pyrimidine and purine nucleosides, respectively. This alone argues against a two-site jump model for internal sugars since these crosspeaks are readily measurable at 40 ms and there are no efficient indirect

H1′–H4′ spin diffusion pathways. The terminal residues G1 and C12 appear to exhibit conformational flexibility (Orbans & Altona, 1986; Celda et al., 1989). This fraying effect will introduce errors in estimating distances from the time-averaged NOE crosspeaks of terminal residues and possibly next-to-terminal residues (C2 and G11); hence, the structures derived for terminal residues may not be accurate.

Backbone Angle Constraints. The backbone linker C3′–O–P–O–C5′–C4′–C3′ was constrained by sums of J couplings, line widths, and NOE proton–proton distances in E. COSY and NOESY spectra (Kim & Reid, 1991; Kim et al., 1992). The γ angle was limited by the observed lower distance bounds of 3.3 and 3.8 Å from H2′ and from H6/H8 to any unassigned H5′/H5″, respectively. Furthermore, the γ angle was also independently constrained to 20–100° by the observed $\sum J_{H4'}$. The ϵ angle was constrained to lie in the range 130–200° or 280–350° from the observed $\sum J_{H3'}$ in E. COSY spectra (Kim et al., 1992). The β angle was limited to the range 180 ± 70° by line-width analysis of H5′/H5″ resonances in NOESY ω_1 slices. The α and ζ angles were indirectly constrained by H1′–(n+1)H5′/H5″ NOE lower distance bounds. None of these backbone constraints, which are derived from couplings or NOEs, are precise single values but consist only of conservative, partially restricted, allowed zones based on experimental data. However, the combination of these ranges of allowed values work synergistically to produce fairly highly constrained overall allowed values. The detailed analyses of backbone angle constraints are discussed elsewhere (Kim et al., 1992).

Initially Generated Structures. The distance bounds mentioned above were used to generate a bounds matrix file in DSPACE, along with implicit distances such as bond lengths. At this stage distances suspected of contamination by spin diffusion such as (n)H6/H8–(n)H2″ and (n)H6/H8–(n)H3′ were withheld from the distance file. In addition, interstrand NH···N and NH···O distances were entered into the distance matrix as standard hydrogen bond lengths without quantitative NMR experimental measurement. The hydrogen bond bounds ranges used were H···O 1.85–2.15 Å and H···N 1.75–2.00 Å. The bounds matrix was then smoothed, converted to a metric matrix, and embedded in 3-D space. The random-embedded DG structure was refined by conjugate gradient and simulated annealing against the distance file. To measure the similarity between structures, coordinate root-mean-square deviations (RMSD) per atom between pairwise combination of structures were determined. The coordinate RMSDs between DG and ideal B-form structures and between DG and ideal A-form structures are 4.5 and 4.2 Å for all 12 base pairs and 2.7 and 3.0 Å for the middle 8 base pairs, respectively. These values indicate that the initial DG structure is quite different from both canonical B-DNA and A-DNA structures.

To test the convergence properties of the simulated annealing and conjugate gradient method, classical B- and A-form DNAs (Arnott & Hukins, 1972a,b) were used instead of the DG coordinates as starting structures. The coordinate RMSD per atom between these two model starting structures is 6.5 Å. The canonical B- or A-form DNA structure was randomized (by 5 Å) and refined against the experimental initial distance bounds matrix by simulated annealing and conjugate gradient methods. Three superimposed refined structures originating from (i) a random-embedded distance geometry (DG) structure, (ii) the canonical A-form structure (designated RA), and (iii) the B-form DNA (designated RB) are shown in Figure 4. The coordinate RMSDs between structures are given in Table II. The coordinate RMSDs among the DG, RA, and RB structures are 0.5–1.0 Å for all 12 base pairs and 0.3–0.4 Å for the middle 8 base pair segment. These values indicate that the simulated annealing procedure has sampled conformational space well and converges to essentially the same structures regardless of the initial starting structures. Although Clore, Gronenborn, and co-workers (Nilsson et al., 1986; Clore et al., 1988; Happ et al., 1988) and Sykes and co-workers (Baleja et al., 1990) reported that restrained molecular dynamics methods produce well-converged structures (≤0.8 Å RMSD) from different starting structures, i.e., ideal B-DNA and A-DNA, these final structures must reflect, at least partially, some contribution from as yet inadequately parametrized nonbonded potentials; such contributions are absent from purely distance-refined structures which satisfy only chemical bonding and NMR-derived distances (whether right or wrong).

The initial distance bounds for the refinement process were determined from the second-order polynomial fitting procedure, which partially takes into account spin diffusion contributions. However, at this stage the crosspeaks known to be strongly contaminated by spin diffusion, i.e., H6/H8–(n)H2″ and H6/H8–(n)H3′, were not present in the distance file, although the somewhat less contaminated H6/H8–(n–1)H2′ distance estimates were used, albeit as lower bounds with conservatively large bounds ranges. This strategy, in itself, reduces the number of spin diffusion distance errors incorporated into the structure but does not reproduce all spin diffusion effects or distances. To further correct the remaining spin diffusion distance errors, and to match the experimental spectra, back-calculation refinement of the structure was next performed.

Iterative Back-Calculation Refinements. The back-calculation of NOESY crosspeak intensities from the interim molecular structures was performed using the program BKCALC, which carries out numerical stepwise integration (Marion et al., 1987) of all cross-relaxation paths to all proximal protons over time. This numerical integration incorporates all spin diffusion pathways and generates the theoretical NOE crosspeak intensities corresponding to that structure. To compare the simulated NOE intensities V_{ij}^c with the experimentally observed NOE intensities V_{ij}^e, a single scaling factor S was determined by comparing the H5–H6 reference crosspeak intensities at 40-ms mixing time, i.e., $S = \sum V_{ref}^e / \sum V_{ref}^c$. Then, all simulated NOE intensities were multiplied by this scaling factor.

The NOE residual factor R was then used to quantitatively determine the fit of the simulated NOEs to the experimental NOE intensities as

$$R = \frac{1}{N}\sum \frac{|V_{ij}^e - V_{ij}^c|}{V_{ij}^e}$$

where N is the number of observables. The R values were determined at every mixing time and were segregated into three distance classes, namely short (≤3.0 Å), medium (3.0 < d ≤ 3.6 Å), and long (>3.6 Å) distances. Weighting factors based on the standard deviations in the NOE intensities are not included. For R value calculations, only the resolved base–sugar and base–base proton crosspeak intensities were used since the intrasugar proton distances (conformations) are independently fixed. Thus, while the R value appears to be better (lower) when all distances, including intrasugar interproton distances, are used, this is a meaningless artificial improvement given the locked sugar conformations. A total of 80 interproton distances, corresponding to 160 crosspeaks,

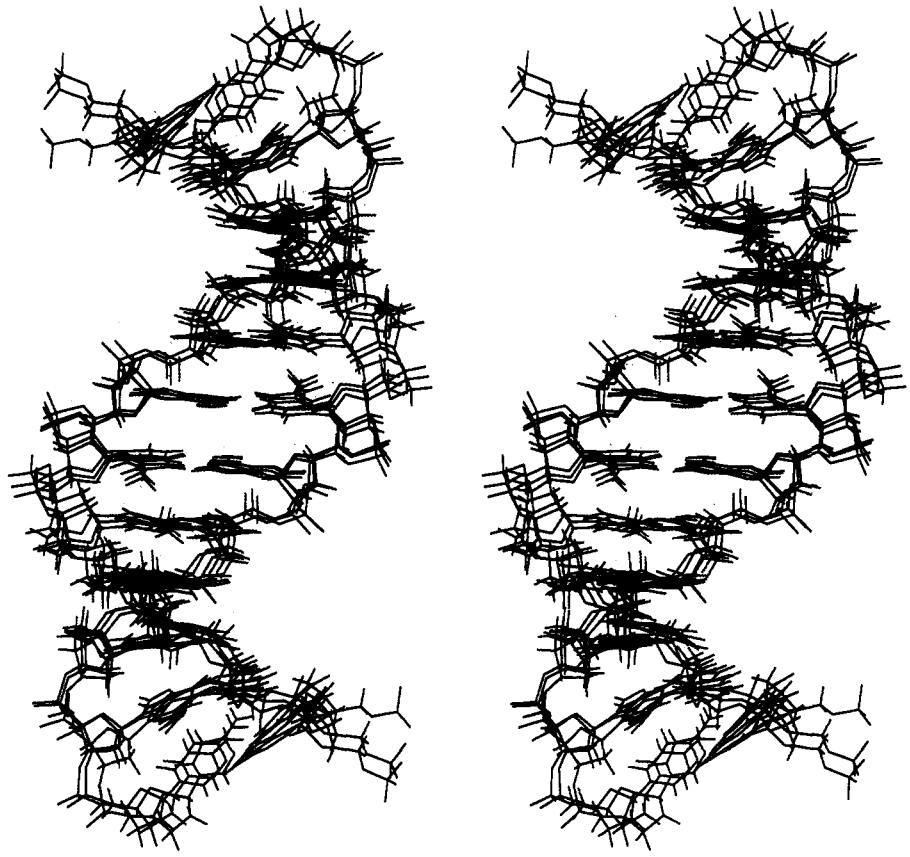

FIGURE 4: Stereoview (cross-eyed) of three superimposed structures of the *HpaI* dodecamer generated from (a) a random-embedded distance geometry DNA, (b) ideal B-DNA, and (c) ideal A-DNA by DPSACE refinement against the initial distance estimates. These are initial structures that have not yet been subjected to back-calculation refinement.

Table II: Coordinate RMSDs per Atom (Angstroms) of $[d(GCCGTTAACGGC)]_2$ Structures[a]

	B	A	DG	RB	RA	FDG_1	FDG_2	FDG_3	FRB	FRA_1	FRA_2
B		5.93	4.50	4.04	4.00	3.43	3.46	2.88	3.68	3.74	3.10
A	4.02		4.16	4.51	4.39	5.78	5.25	5.91	5.26	5.28	6.06
DG	2.70	3.00		1.02	0.74	2.21	1.81	2.67	1.61	1.62	2.65
RB	2.62	3.03	0.37		0.50	1.77	1.15	2.10	1.17	1.36	2.09
RA	2.48	3.07	0.30	0.30		1.84	1.28	2.15	1.20	1.37	2.15
FDG_1	2.18	3.67	1.02	0.99	0.93		**0.86**	**0.84**	**0.83**	**0.81**	**0.74**
FDG_2	2.29	3.50	0.84	0.71	0.70	**0.46**		**1.06**	**0.48**	**0.78**	**1.05**
FDG_3	1.92	3.85	1.32	1.23	1.13	**0.54**	**0.63**		**1.22**	**1.35**	**0.49**
FRB	2.29	3.44	0.77	0.68	0.64	**0.48**	**0.16**	**0.65**		**0.56**	**1.15**
FRA_1	2.39	3.54	0.80	0.83	0.80	**0.42**	**0.45**	**0.76**	**0.42**		**1.28**
FRA_2	2.03	3.95	1.34	1.21	1.16	**0.55**	**0.62**	**0.38**	**0.68**	**0.76**	

[a] RMSD values for all 12 base pairs are above the diagonal and those for the middle 8 base pairs are below the diagonal. B and A are ideal B-form and A-form DNA structures; DG is the random-embedded distance geometry structure; RB and RA are derived from B-DNA and A-DNA by refinement against the initial distance bounds; FDG_1, FDG_2, and FDG_3 are derived from DG starting structures by back-calculation refinement against the experimental NOESY crosspeak volumes; and FRB, FRA_1, and FRA_2 are derived from the initial-distance RB and RA structures by back-calculation refinement against the experimental NOESY crosspeak volumes.

Table III: Average *R* Value[a]

distance, Å	B	A	DG	RB	RA	FDG_1	FDG_2	FDG_3	FRB	FRA_1	FRA_2
$d \leq 3.0$ ($n = 24$)	0.64	0.79	0.19	0.21	0.21	0.13	0.14	0.15	0.15	0.15	0.15
$3.0 < d \leq 3.6$ ($n = 24$)	0.57	1.67	0.28	0.28	0.28	0.15	0.15	0.15	0.15	0.15	0.15
$d > 3.6$ ($n = 32$)	0.66	1.73	0.34	0.33	0.37	0.23	0.23	0.23	0.23	0.23	0.23
av ($n = 80$)	0.63	1.43	0.28	0.28	0.29	0.18	0.18	0.18	0.18	0.18	0.18

[a] See Table II, footnote *a*, for the definitions of B, A, DG, RB, RA, FDG_1, FDG_2, FDG_3, FRB, FRA_1, and FRA_2. *R* values are the averages from 40-, 80-, 120-, and 200-ms mixing time data.

were subjected to *R*-factor analyses. Tables III and IV show the average *R* values (short, medium, and long distances) from four mixing times for the DG, RA, and RB structures refined only against the initial distance bounds. The average *R* values of the RB and RA structures against the experimental NOE intensities are rapidly reduced from 0.63 (B-DNA) to 0.28 and from 1.43 (A-DNA) to 0.29, respectively. By distance categories, the average *R* values over four mixing times for the DG, RB, and RA structures are 0.19–0.21 ($n = 24$) for short distances (≤3.0 Å), 0.28 ($n = 24$) for medium distances ($3.0 < d \leq 3.6$ Å), and 0.33–0.37 ($n = 32$) for long distances (>3.6 Å). The experimental NOESY spectra at 120-ms

Table IV: Average R Value[a]

mixing time, ms	B	A	DG	RB	RA	FDG_1	FDG_2	FDG_3	FRB	FRA_1	FRA_2
40	0.76	2.31	0.34	0.34	0.36	0.21	0.21	0.21	0.20	0.21	0.21
80	0.67	1.50	0.26	0.26	0.28	0.16	0.16	0.16	0.16	0.16	0.16
120	0.59	1.12	0.24	0.24	0.25	0.16	0.16	0.16	0.16	0.17	0.16
200	0.49	0.77	0.25	0.27	0.27	0.19	0.19	0.20	0.20	0.19	0.19
av	0.63	1.43	0.28	0.28	0.29	0.18	0.18	0.18	0.18	0.18	0.18

[a] See Table II, footnote *a*, for definitions of B, A, DG, RB, RA, FDG_1, FDG_2, FDG_3, FRB, FRA_1, and FRA_2. A total of 80 distances, which represent 160 crosspeaks, are used for the R analyses.

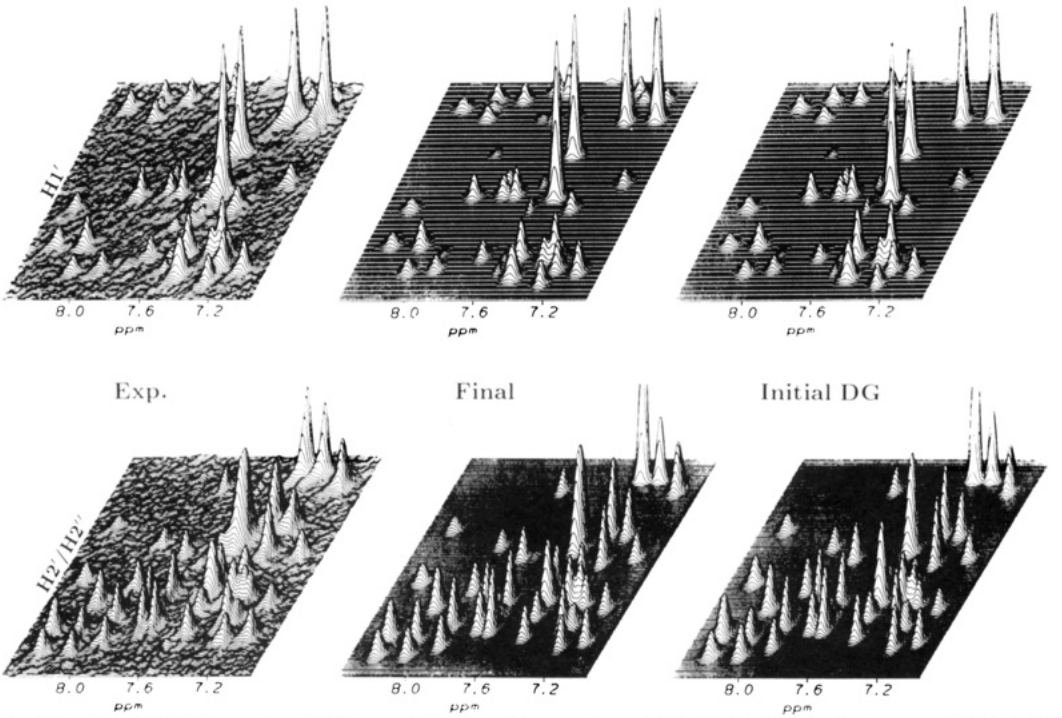

FIGURE 5: Stack plots of the NOESY spectra at 120-ms mixing time for experimental (left), final refined structure FDG_1 (middle), and initial DG structure (right). The H6/H8–H1′/H3′ (top) and H6/H8–H2′/H2″ (bottom) regions are shown. The corresponding NOESY spectra at 40-ms mixing time are shown in the supplementary material.

mixing time for the H6/H8–H1′/H3′ and H6/H8–H2′/H2″ regions are shown in Figure 5, together with simulated NOESY spectra of the initial DG structure (right) and the back-calculation-refined FDG_1 structure (center). The NOESY spectra at 40-ms mixing time are presented in the supplementary material.

For comparison, the NOESY spectra of the A-DNA, the back-calculation-refined A-DNA (FRA_1), B-DNA, and the back-calculation-refined B-DNA (FRB) are also presented in the supplementary material. The crosspeak volumes in the simulated spectra of the initial DG structure match the experimental data surprisingly well, except for some overlapped or weak (long distance) crosspeaks. This indicates that the polynomial modification of the two-spin approximation is a fairly effective approach to removing some of the spin diffusion effects and estimating the real distances. It is worth noting that the R factor volume comparison, which is independent of line shape or line width, is a more rigorous criterion than visual comparison with simulated spectra containing imperfect line shapes/line widths; i.e., the experimental crosspeaks are complex multiplets of varying line shapes, but the simulated crosspeaks are given Gaussian line shapes. Interestingly, the rather poor R value of 2.31 for an ideal A-DNA structure at 40 ms improves to an apparently much more respectable R value of 0.77 at 200 ms, for *exactly the same structure*, demonstrating that bad structures expose their defects rather poorly at longer mixing times compared to shorter mixing times. This is due to extensive spin diffusion and partial equilibration of the multiple-spin system at long mixing times, where R factors become poor indicators of similarity in structure. Beyond 200 ms, almost any structure generates a reasonably acceptable simulation of the almost steady-state distribution of magnetization.

The iterative back-calculation refinement was carried out in two ways: (a) by visual inspection of the simulated crosspeaks and manual adjustment of the distances as used previously (Nerdal et al., 1989; Banks et al., 1989) and (b) by an automated back-calculation refinement program that adjusts distances automatically only on the basis of the *volume* differences between experimental and simulated peaks (Kim & Reid, 1992). The automated refinement algorithm compares experimental and simulated volumes and changes the corresponding upper and lower distance bounds according to the relationship

$$d_{ij}(\text{new}) = d_{ij}(\text{old})(V^s_{ij}/V^c_{ij})^{1/6}$$

New distance bounds were created by a combination of quantitative automatic refinement of resolved crosspeaks and visual comparison of the overlapped crosspeaks (including intrasugar proton crosspeaks); the new values were then used as input distance bounds to generate the next structure in the

DSPACE iteration cycle. The simulated annealing algorithm is utilized to let the molecule sample the new distance constraints, and then the conjugate gradient algorithm is used to minimize the deviation from the new distance bounds. These back-calculation refinement procedures were performed until the simulated NOESY matched the experimental data, i.e., there was no further improvement in the R factor. During the iterative back-calculation refinement, short distances (≤3.0 Å) are adjusted first and the structure is refined until the R values in the short distance range do not further improve. Next medium and then long distances are adjusted in the same manner, *without changing the short distances*. Since NOE intensities corresponding to longer distances can be matched either by altering the spin diffusion pathway or by changing the two-spin *ij* distance, it is essential to match the accurate short distances first and lock them during subsequent distance adjustments. After the close distances in the spin diffusion networks are fixed, the NOE intensity of the longer distances can then be accurately refined without altering the short components of the spin diffusion path.

Five final refined structures were generated from the initial distance DG structure by back-calculation refinement of the distance file. Three of these (designated FDG_1, FDG_2, and FDG_3) were chosen on the basis of low penalty value (including backbone angle constraints) for further investigation. The modified distance file generated during the back-calculation refinement of the DG structure was then also used to further refine the initial distance refined RB and RA structures using simulated annealing and conjugate gradient refinement. This, in turn generated a final structure FRB (with classical B-DNA as its initial structure ancestor) and two additional final structures (FRA_1 and FRA_2) with A-DNA as their initial structure ancestors; these structures were also subjected to helical parameter analysis as described above. The internal 10 base pairs of the three final structures FDG_1, FRA_1, and FRB—each with different initial ancestors—are shown superimposed in Figure 6; coordinate RMSDs per atom between pairwise structures are shown in Table II. All *xyz* coordinates of the final structures are available upon request and are being deposited in the Brookhaven data bank. RMSDs among the final refined structures are less than 0.8 Å for the middle 8 base pairs; the final structures have converged very well. The R values were significantly lowered, from 0.28–0.29 to 0.18, during the back-calculation refinement (Tables III and IV), and the simulated spectra of the final structures match the experimental data better than do those of the initial structure (center spectra in Figure 5 and supplementary material). The R values of the six final refined structures are quite consistent for all mixing times and for all distance ranges (0.14–0.15 for short, 0.15 for medium, 0.23 for long distances); the relatively high R value for long distances at 40 ms (0.29–0.31) is simply a reflection of the poor experimental signal-to-noise ratio at such short mixing times.

The iterative back-calculation refinement procedure is an essential component in all NMR-based structural studies of biopolymers. Its main function is to correct inaccurate two-spin initial distance estimates. It is obvious that erroneous initial distance estimates, whether used in distance geometry or in restrained molecular dynamics, will produce incorrect structures. The automated refinement procedure used in this study (Kim & Reid, 1992) rapidly and efficiently reduces the differences between experimental and simulated NOESY intensities and yields quantitative R values directly without having to multiply resimulate the NOESY spectra. Minimization of the difference between the experimental and simulated volumes, i.e., reducing the R value, is the major target of this approach, which is similar to that described by others (Yip & Case, 1989; Baleja et al., 1990b; Bonvin et al., 1991). The method not only speeds up structural studies but also minimizes the errors induced by visual/manual methods which depend on the line widths/line shapes used in the simulation. However, this automated method requires accurate crosspeak volume integration, and thus overlapped peaks cannot be analyzed automatically (overlapped peaks are still amenable to visual comparison and adjustment). Deconvolution of overlapped peaks may be necessary to derive more accurate volumes. Intensity errors due to poor signal-to-noise ratios and/or differences between above- and below-diagonal crosspeaks were not treated separately in this analysis (the second-order polynomial was used to average and smooth crosspeak volumes over several mixing times). Thus, automatic back-calculation refinement might be further improved in the future by including all NOE crosspeaks, including deconvoluted overlapped peaks.

NOE-Refined Structures. The torsion angles and helical parameters of all six final refined structures were determined with the program NEWHELIX90. The mean values and the RMS deviations for the torsion and pseudorotation phase angles of the refined structures are listed in Table V, and the local helical parameters are plotted in Figure 7 for the central 8 base pairs. The global helical parameters inclination, tip, and displacement are presented in the supplementary material. While tilt, twist, buckle, slide, and cup show relatively little variation among the six structures, propeller twist appears to be less well-defined by the available NMR observables. The somewhat lower twist for C3 (Figures 6 and 7) may simply reflect the loss of tight distance constraints due to overlapped crosspeaks involving C3 H6; unfortunately, local underdetermination caused by redundant chemical shifts is frequently unavoidable in the NMR approach to structure determination.

The pseudorotational phase angles were all found to lie in the range 103–157° (C1′-exo to C2′-endo) with the pyrimidines tightly clustered in the 103–134° C1′-exo region and the purines clustered in the 143–157° region. These constraints are mainly derived from a combination of COSY J couplings, NOESY H1′–H4′ distances and H2″–H4′ NOESY crosspeaks. The refined distances and P values are well satisfied in the final structures. The glycosidic torsion angle χ for each residue converges well to a range of less than 10°, except for the terminal G1, and all backbone angles are quite well-determined in each residue by the combination of conservative ranges of backbone distance constraints in the absence of exact constraints.

Previous studies of DNA structure from this laboratory (Nerdal et al., 1989; Banks et al., 1989) and from Pardi and co-workers (Pardi et al., 1988; Metzler et al., 1990) used only NOE-derived distances within and between nucleosides, *without* any backbone angle constraints. The resulting structures (even after NOESY back-calculation refinement) exhibit a not surprising wide range of underdetermined backbone conformations and also quite large variations in base pair helical parameters (Metzler et al., 1990). Furthermore, RMSD differences between back-calculation-refined structures starting from A-DNA, B-DNA, and DG-DNA range from 1.5 to 3.4 Å in two 10-base pair duplexes (Metzler et al., 1990). Clore, Gronenborn, and co-workers (Nilsson et al., 1986; Happ et al., 1988; Clore et al., 1988) used distance-constrained energy minimization methods in which they incorporated backbone angle constraints taken from crystallographic studies rather than from solution data;

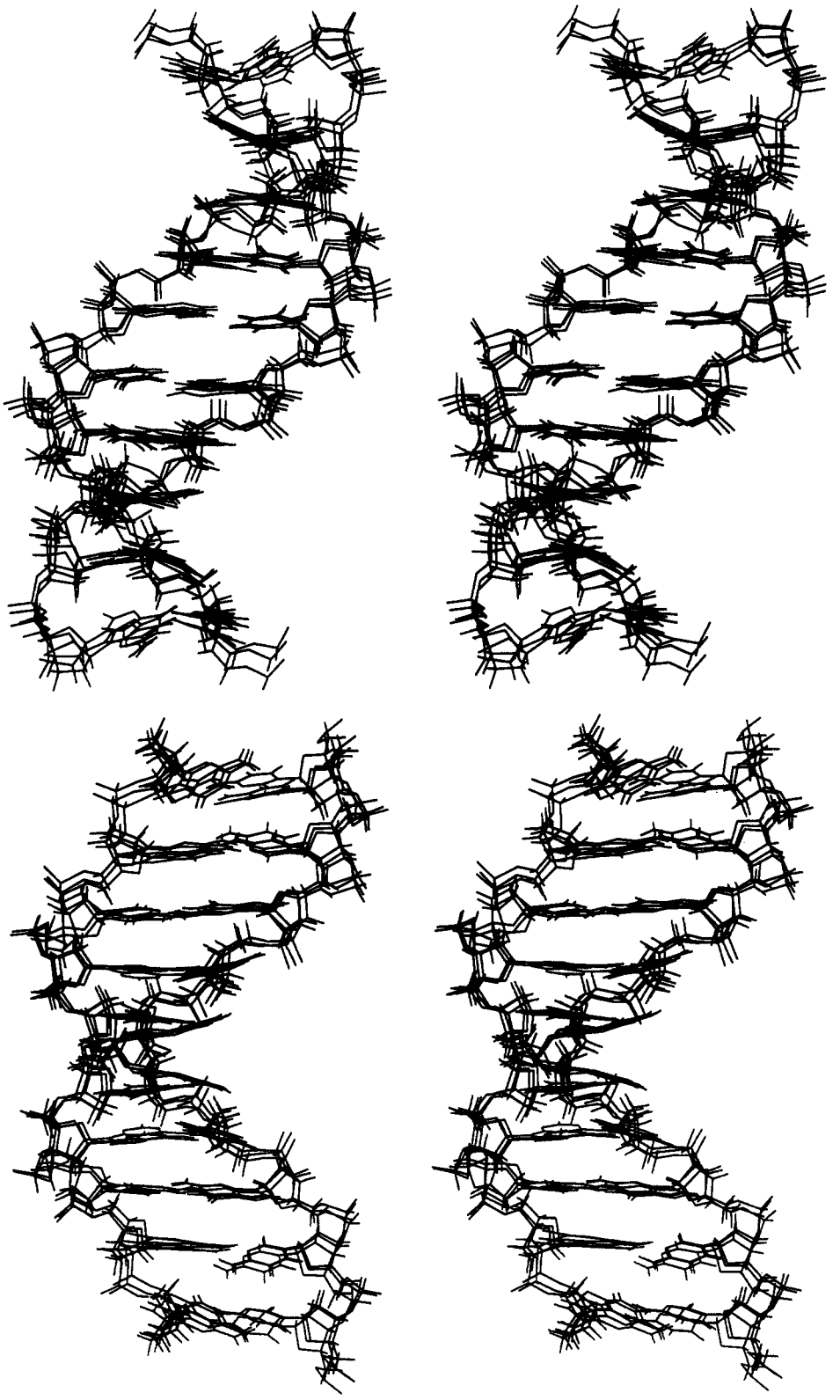

FIGURE 6: Stereoview (cross-eyed) of the superimposed **final** structures (FDG_1, FRA_1, and FRB) of the *Hpa*I dodecamer refined from structures shown in Figure 4 using back-calculation/DSPACE. (B, Bottom) has been rotated by 90° from (A, top). Only the middle 10 base pairs are shown.

the resulting helical parameters converged to quite narrow ranges (that were similar to the X-ray structure), indicating that determination of the backbone linker conformation, by one means or another, is essential. To find the relative position and orientation of two adjacent nucleoside residues, three translations and three rotations must be known. However, the number of useful and reliable internucleotide NMR observations in a B-DNA dinucleotide step is relatively small, i.e., H6/H8–(n–1)H1′/H2′/H2″ and H6/H8–(n–1)H6/H8 proton crosspeaks [if the 3′ residue is cytosine, additional H5–(n–1)H1′/H2′/H2″/H6/H8 distances are often also available]. Frequently only a subset of the above is measurable due to crosspeak overlaps. Without backbone angle constraints, the limited number of interresidue NOE observables available is incapable of uniquely defining the five additional backbone torsion angle variables of the internucleoside linker,

Table V: Pseudorotation and Torsion Angles of [d(GCCGTTAACGGC)]$_2$ Structures[a]

residue	P	χ	ε	ζ	α	β	γ	δ
G1	152 (2)	267 (4)	185 (3)	256 (2)			353 (131)	136 (1)
C2	131 (1)	238 (2)	170 (6)	272 (3)	243 (4)	200 (5)	63 (1)	120 (2)
C3	134 (2)	241 (3)	165 (4)	288 (5)	274 (3)	196 (5)	55 (2)	116 (1)
G4	144 (1)	235 (4)	145 (6)	299 (4)	271 (10)	185 (6)	67 (2)	113 (0)
T5	126 (2)	228 (1)	145 (3)	303 (3)	235 (6)	215 (6)	81 (2)	105 (2)
T6	131 (1)	240 (3)	156 (10)	299 (9)	246 (4)	214 (2)	69 (3)	107 (2)
A7	146 (1)	246 (3)	187 (3)	253 (2)	265 (12)	192 (7)	71 (4)	124 (1)
A8	152 (3)	253 (2)	182 (20)	246 (18)	284 (4)	179 (3)	52 (3)	133 (5)
C9	130 (2)	230 (6)	175 (3)	289 (1)	294 (21)	178 (28)	62 (8)	123 (1)
G10	157 (2)	265 (6)	172 (4)	260 (3)	259 (4)	202 (3)	37 (3)	139 (2)
G11	143 (6)	260 (6)	186 (1)	257 (2)	260 (6)	216 (6)	47 (2)	131 (3)
C12	103 (1)	195 (4)			271 (6)	183 (6)	51 (4)	103 (1)

[a] The results are means of six structures followed by their RMS deviations in parentheses. Final back-calculation-refined structures derived from distance geometry starting structure (FDG_1, FDG_2 and FDG_3), from the B-form starting structure (FRB), and from the A-form starting structure (FRA_1 and FRA_2) were used for these analyses.

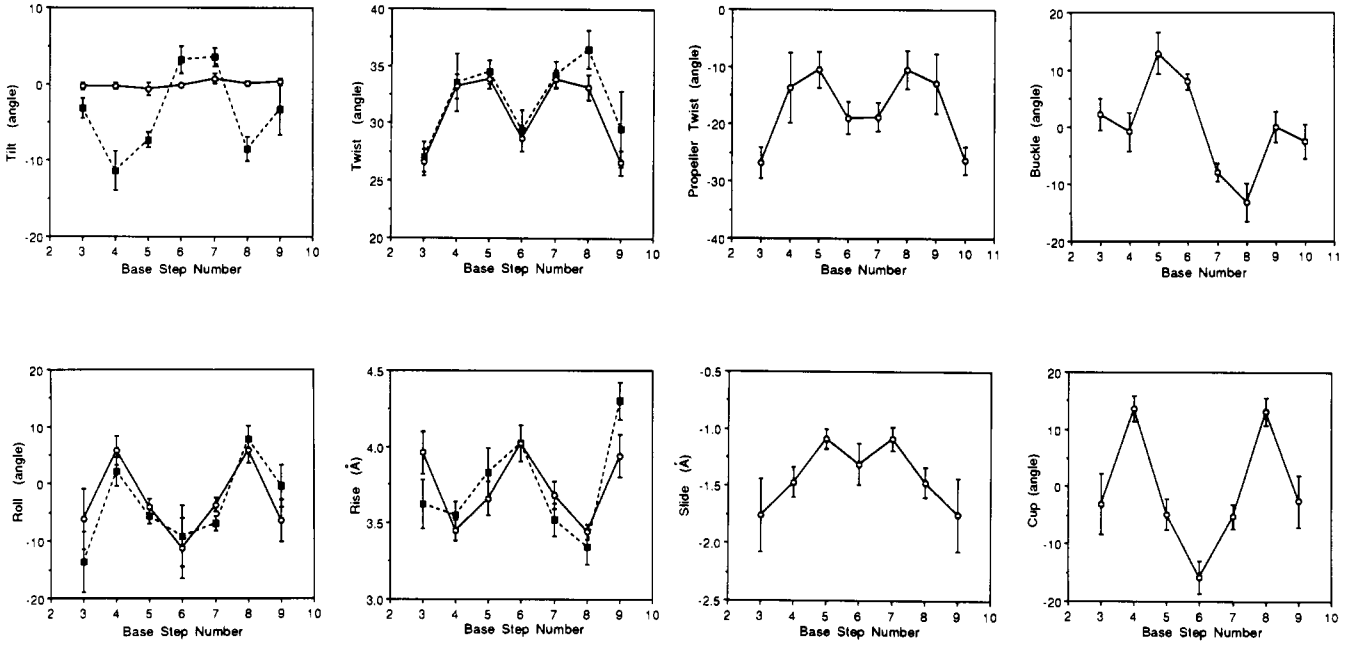

FIGURE 7: Helical parameters at each step for the six final back-calculation-refined structures of the *HpaI* dodecamer. The bar indicates RMS deviations for the helical parameters. Solid squares represent the average values of the parameter for a single strand, while open circles are base pair parameters.

resulting in large variations in both backbone conformation and helical parameters. Thus, it is essential to experimentally determine, or at least partially constrain, as many torsion angles as possible, including the backbone linker, as accurately as possible. In the absence of any backbone constraints, the large variations in helical parameters reported in earlier distance geometry studies (Pardi et al., 1988; Nerdal et al., 1989; Banks et al., 1989; Metzler et al., 1990) are simply to be expected.

The helical parameters of the six refined structures of the *HpaI* dodecamer (derived separately from DG, B-DNA, and A-DNA starting structures) exhibit fairly narrow ranges; RMS deviations for tip, inclination, roll, tilt, cup, buckle, and twist are less than 4° and for propeller twist less than 6°. Rise deviations are less than 0.2 Å, slide deviations are less than 0.3 Å, and those for displacement are less than 0.6 Å. Although the global helical parameters also show quite narrow variability (supplementary material), it is not clear at this stage how well these helical parameters are really determined by NMR observables. Since the NOE input data provide distance information only in the 2–5-Å range, these NMR observations are more directly related to local conformation than to global conformation. Furthermore, the values for the global helical parameters are dependent on, and sensitive to, the helix axis and therefore depend on how this axis is defined, *especially* for structures that do not have a perfectly straight helix axis. It is obvious that global helical parameters related to the helix axis (tip, inclination, and displacement) cannot be determined as well as local helical parameters related to only a single base step (roll, tilt, twist, slide, and rise). The local helical parameters of the final refined structures that are fairly directly related to the interresidue NMR observables are well-converged. Related single strand/duplex helical parameters (Figure 7) are equally well-determined and have similar strand/duplex patterns, with the exception of tilt (which averages the two base planes of a base pair in a duplex). The three local rotations, twist, roll, and tilt, and one local translation, slide, appear to be quite well-determined, probably reflecting the level of constraint on the torsion angles, including backbone torsion angles. The helical parameters relating structure within a single base pair, i.e., propeller twist and buckle, are not easily related to NOE observables and are generally determined mostly by the combination of the structures of the two strands with the requirement for twofold symmetry; the only related experimental constraints are the cross-strand H2–H1′ and H2–imino NOEs from adenine residues. For example, the propeller twist of the G4–C21 base pair in FRA_1 is −22°, while in FRA_2 it is −5°, despite

there being no detectable differences in NMR-observable proton–proton distances. Thus, propeller twist is a rather poorly determined parameter, especially for GC pairs. The relative precision of the tilt, roll, twist, and slide data is more reliable than that for displacement, inclination, and propeller twist. These observations are consistent with the conclusions of Baleja et al. (1990) on a 10-base pair duplex using distance-restrained molecular dynamics combined with crystallographically derived backbone angle constraints and also with a computer simulation study by Pardi et al. (1988) using distance geometry methods on a known structure without any backbone angle constraints. Baleja et al. (1990) reported that displacement and propeller twist showed greater variation than twist, roll, slide, tilt, and rise; Pardi et al. (1988) reported that displacement and propeller twist are much less well-determined than tilt and twist. From the present work, the helical parameters of the NMR-determined *Hpa*I solution structure do not follow Calladine's rules (1982), neither do they conform to the high-twist profile (HTP) and low-twist profile (LTP) rules of Yanagi et al. (1991).

The distance between the adenine H2 and the cross strand (m+1) H1′ (the H1′ of the 3′ neighbor of the complementary T residue) is a relatively direct measure of the minor groove width for B-like DNA in solution (Kintanar et al., 1987; Katahira et al., 1990; Chuprina et al., 1991a). The (n)H2–(m+1)H1′ distance is also an important helix-twisting determinant in distance geometry refinement. Unfortunately in the present sequence, the A7 H2–A19 H1′ NOE lies across the symmetry axis and is therefore exactly superimposed on the A7 H2–A7 H1′ crosspeak. However, the fact that this very weak double peak was almost totally accounted for by the A7 H2–A7 H1′ intensity (always 4.5 ± 0.1 Å for all reasonable χ values in the anti range) allowed us to deduce that the A7 H2–A19 H1′ intensity is extraordinarily weak and therefore justified the use of a bounds range of 4.5–5.5 Å in the distance file for the A7 H2–A19 H1′ cross-strand distance. The A8 H2–T18 H1′ crosspeak is well-resolved, and its intensity corresponds to 4.0–4.3 Å. This indicates that the central 6–7 T_pA step has a wide (normal) minor groove width typical of B-DNA, while the 7–8 A_pA step has a partially narrowed minor groove [for a discussion of H2 NOEs and groove width see Chuprina et al. (1991a)]. Note that in classical B-DNA the (n)H2–(m+1) H1′ distance is 5.1 Å. The resulting cross-strand A7–G22 and G10–A19 phosphate–phosphate distances in the six final refined structures are 4.4 ± 0.7 Å (excluding the P–P van der Waals 5.8-Å distance), while the A8 P–C21 P and C9 P–A20 P cross-strand phosphate distances are 6.0 ± 0.6 Å. These results agree well with the relationship between (n)H2–(m+1)H1′ and P–P distances in X-ray crystallographic structures (Chuprina et al., 1991b). Furthermore, rather weak A7 H8–T6 H1′/H2″ crosspeaks are observed at the T_pA step (see Figures 1 and 2); this may be closely related to the observed underwinding and large negative cup at this step. We have observed a similar result at the T_pA step in a DNA dodecamer containing a TTTAAA sequence (unpublished results). From polyacrylamide gel electrophoretic studies, $(NT_4A_4N)_n$ (Hagerman, 1986) and $(NT_4A_5)_n$ (Haran & Crothers, 1989) oligomers are not retarded and are assumed not to be bent, while $(N_2T_3A_3N_2)_n$ oligomers are weakly but significantly retarded (Hagerman, 1988). However $(NT_2A_5N_2)_n$ and $(NT_3A_5N)_n$, and obviously $(NT_5A_2N_2)_n$ and $(NT_5A_3N)_n$, are strongly retarded (Haran & Crothers, 1989), while $(NT_5A_4)_n$ (the phase-shifted complement of NT_4A_5) is obviously not; to our knowledge palindromic $(N_3T_2A_2N_3)_n$ duplexes have not been studied electrophoretically. Theoretical studies using the nucleic acid potentials of Zhurkin et al. (1980) led Chuprina and co-workers to predict that T_pA steps in T_nA_n sequences should produce bends, provided n were large enough to produce B′-DNA structure at the outer ends of such a tract (Chuprina & Abagyan, 1988; Chuprina et al., 1991a).

In the present T_2A_2 case, n = 2 may be too short to fully effect the B to B′ structural transition. In this context, the observed (n)H2–(m+1)H1′ distance of 4.2 Å at the 5–6 and 7–8 steps is intermediate between the ca. 3.8 Å found for narrow groove B′-DNA and the >4.6 Å characteristic of normal groove B-DNA (Chuprina et al., 1991b). Thus, although we observe an increase in this minor groove parameter from 4.2 to ≥4.5 Å at the T_pA step (a widening of the groove from ca. 4.4 to ca. 6.0 Å), the flanking residues have attained only partial B′ character, and we might therefore expect a more limited kinking/bending of the duplex at this step. In Figure 6B, one can discern tht the wide minor groove at the ends of the duplex narrows toward the middle of the molecule and then abruptly widens at the central step, producing a modest bend between the top half and the bottom half of the duplex. This mild bend appears to be produced by a combination of negative cup and negative roll, together with decreased twist and increased rise at the T_pA step (Figure 7).

Summary. In the present study we have shown that distance geometry combined with additional backbone constraints derived from J coupling generates very similar final structures regardless of the starting structure. The structures refined against initial distance constraints were subjected to iterative NOE back-calculation refinement—a procedure that has now been automated for resolved crosspeaks. The automated back-calculation refinement compares the experimental and simulated crosspeak volumes and adjusts the distance bounds according to the volume differences. This method speeds up the iterative refinement process and provides quantitative R values without simulating any intermediate NOESY spectra. The procedure was performed iteratively until the simulated NOE volumes from the final structure matched the experimental data, i.e., the R factors were minimized. The six final refined structures derived from various starting structures have R values of 0.18 for the resolved base–sugar and base–base proton crosspeaks (n = 80). Importantly, the R values are consistent for all mixing times and for all distance ranges. The final structures converge well with RMSD differences of 0.8 Å for the central 8 base pairs. The pseudorotational phase angles of the individual sugars were all well-determined; all pyrimidine nucleosides exhibited C1′-exo sugar conformations (P = 103–134°), while the purine nucleosides have higher P values (143–157°) close to the C2′-endo conformation. The glycosidic χ angles all converged to within 10° variability, and the individual backbone angles converged to deviations of 30° or less. On the basis of only one experimental observable, i.e., $J_{H3'-P}$, the phosphodiester backbone, in principle, could have either a B_I (ϵ = ca. 180°, ζ = ca. 270°) or a B_{II} (ϵ = ca. 270°, ζ = ca. 180°) conformation—yet all backbone conformations were found to be in the B_I conformation (presumably as a result of other indirect constraints on ϵ, i.e., bond lengths and angles superimposed on proton–proton distances). The local helical parameters tilt, roll, twist, and slide appear to be rather well-determined, while propeller twist is poorly determined. The final structures show that the central T6–A7 step has a wide (normal) minor groove width, is underwound, and has a large negative cup, while the A7–A8 step has a partially narrowed minor groove. The outer transitions between B-type structure in the flanking regions

and partial B′ structure in the central tetramer, as well as abrupt widening of the minor groove at the central T_pA step, produce internal kinks within the duplex that, although quite mild, may nevertheless be the basis of recognition of this duplex by the Hpa nuclease.

ADDED IN PROOF

In T_2A_2 and T_3A_3 DNA sequences, the first A residue routinely exhibits significantly broadened NMR lines which may reflect motion at this residue (Kennedy & Reid, manuscript in preparation) and could further complicate the analysis of gel mobility in T_nA_n sequences.

ACKNOWLEDGMENT

We thank Susan Ribeiro and Julie Miller for synthesizing the DNA dodecamer and Dennis Hare (and Hare Research) for providing FTNMR, FELIX, and DSPACE programs. We also thank Peter Flynn and Einar Sletten for stimulating discussions and Mary Coventry for typing the manuscript.

SUPPLEMENTARY MATERIAL AVAILABLE

Five stack plots of the simulated NOESY spectra for the final back-calculated structures of the *HpaI* dodecamer and one plot of the tip, inclination, and displacement parameters at each step as well as the initial and final distance bounds constraints file (12 pages). Ordering information is given on any current masthead page.

REFERENCES

Arnott, S., & Hukins, D. W. L. (1972a) *Biochem. Biophys. Res. Commun. 47*, 1504–1509.
Arnott, S., & Hukins, D. W. L. (1972b) *J. Mol. Biol. 81*, 93–105.
Baleja, J. D., Pon, R. T., & Sykes, B. D. (1990a) *Biochemistry 29*, 4828–4839.
Baleja, J. D., Moult, J., & Sykes, B. D. (1990b) *J. Magn. Reson. 87*, 375–384.
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.
Bonvin, A. M. J. J., Boelens, R., & Kaptein, R. J. (1991) *J. Biomol. NMR 1*, 305–309.
Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.
Calladine, C. R. (1982) *J. Mol. Biol. 161*, 343–352.
Cavanagh, J., Chazin, W. J., & Rance, M. (1990) *J. Magn. Reson. 87*, 110–131.
Celda, B., Widmer, H., Leupin, W., Chazin, W. J., Denny, W. A., & Wüthrich, K. (1989) *Biochemistry 28*, 1462–1471.
Chary, K. V. R., & Modi, S. (1988) *FEBS Lett. 233*, 319–324.
Chuprina, V. P., & Abagyan, R. A. (1988) *J. Biomol. Struct. Dyn. 6*, 121–138.
Chuprina, V. P., Fedoroff, O. Y., & Reid, B. R. (1991a) *Biochemistry 30*, 561–568.
Chuprina, V. P., Lipanov, A. A., Fedoroff, O. Y., Kim, S.-G., Kintanar, A., & Reid, B. R. (1991b) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9087–9091.
Clore, G. M., & Gronenborn, A. M. (1984a) *FEBS Lett. 172*, 219–225.
Clore, G. M., & Gronenborn, A. M. (1984b) *FEBS Lett. 175*, 117–123.
Clore, G. M., & Gronenborn, A. M. (1989) *Crit. Rev. Biochem. Mol. Biol. 24*, 479–564.
Clore, G. M., Oshkinat, H., McLaughlin, L. W., Benseler, F., Happ, C. S., Happ, E., & Gronenborn, A. M. (1988) *Biochemistry 27*, 4185–4197.
Crippen, G. M. (1981) *Distance Geometry and Conformational Calculations*, Research Studies/Wiley, Chichester, U.K.
Dickerson, R. E. (1983) *J. Mol. Biol. 166*, 419–441.
Dickerson, R. E. (1990) In *Structure and Methods* (Sarma, R. H., & Sarma, M. H., Eds.) Vol. 3, pp 1–38, Adenine Press, Schenectady, NY.
Dickerson, R. E., & Drew, H. R. (1981) *J. Mol. Biol. 149*, 761–786.
DiGabriele, A. D., Sanderson, M. R., & Steitz, T. A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1816–1820.
Drew, H. R., Wing, R. M., Takano, T., Broka, C., Tanaka, S., Itakura, K., & Dickerson, R. E. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2179–2183.
Drobny, G., Pines, A., Sinton, S., Weitkamp, D. P., & Wemmer, D. E. (1979) *Faraday Symp. Chem. Soc. 13*, 49–55.
Ernst, R. R., Bodenhausen, G., & Wokaun, A. (1987) *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Oxford University Press, Oxford, U.K.
Fejzo, J., Zolnai, Z., Macura, S., & Markley, J. L. (1989) *J. Magn. Reson. 82*, 518–528.
Flynn, P. F., Kintanar, A., Reid, B. R., & Drobny, G. (1988) *Biochemistry 27*, 1191–1197.
Franklin, R., & Gosling, R. G. (1953) *Nature 171*, 740–742.
Gochin, M., & James, T. L. (1990) *Biochemistry 29*, 11171–11180.
Gochin, M., Zon, G., & James, T. L. (1990) *Biochemistry 29*, 11161–11171.
Griesinger, C., Sorensen, O. W., & Ernst, R. R. (1986) *J. Chem. Phys. 85*, 6837–6852.
Griesinger, C., Sorensen, O. W., & Ernst, R. R. (1987) *J. Magn. Reson. 75*, 474–492.
Hagerman, P. J. (1986) *Nature 321*, 449–450.
Hagerman, P. J. (1988) in *Unusual DNA Structures* (Wells, R. D., & Harvey, S. C., Eds.) pp 225–236, Springer-Verlag, New York.
Happ, C. S., Happ, E., Nilges, M., Gronenborn, A. M., & Clore, G. M. (1988) *Biochemistry 27*, 1735–1743.
Haran, T. E., & Crothers, D. M. (1989) *Biochemistry 28*, 2763–2767.
Hare, D. R., & Reid, B. R. (1986) *Biochemistry 25*, 5341–5350.
Hare, D. R., Wemmer, D. E., Chou, S.-H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319–336.
Hare, D. R., Shapiro, L., & Patel, D. J. (1986a) *Biochemistry 25*, 7445–7456.
Hare, D. R., Shapiro, L., & Patel, D. J. (1986b) *Biochemistry 25*, 7456–7464.
Havel, T. F., & Wüthrich, K. (1985) *J. Mol. Biol. 21*, 5129–5135.
Havel, T. F., Kuntz, I. D., & Crippen, G. M. (1983) *Bull. Math. Biol. 45*, 665–720.
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. (1988) *Magn. Reson. Chem. 26*, 927–944.
Hyberts, S. G., & Wagner, G. (1989) *J. Magn. Reson. 81*, 418–422.
Katahira, M., Sugeta, H., & Kyogoku, Y. (1990) *Nucleic Acids Res. 18*, 613–618.
Kim, S.-G., & Reid, B. R. (1991) *J. Cell. Biochem. 15G*, 77.
Kim, S.-G., & Reid, B. R. (1992) *J. Magn. Reson.* (in press).
Kim, S.-G., Lin, L.-J., & Reid, B. R. (1992) *Biochemistry 31*, 3564–3574.
Kintanar, A., Klevit, R. E., & Reid, B. R. (1987) *Nucleic Acids Res. 15*, 5845–5862.
Kirkpatrick, S., Gelatt, C. D., Jr., & Vecchi, M. P. (1983) *Science 220*, 671–680.
Lane, A. N., Jenkins, T., Brown, T., & Neidle, S. (1991) *Biochemistry 30*, 1372–1385.
Levitt, M. H., Freeman, R., & Frenkel, T. (1982) *J. Magn. Reson. 47*, 328–330.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Marion, D., Genest, M., & Ptak, M. (1987) *J. Biophys. Chem. 28*, 235–244.
Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953) *J. Chem. Phys. 21*, 1087–1092.
Metzler, W. J., Wang, C., Kitchen, D. B., Levy, R. M., & Pardi, A. (1990) *J. Mol. Biol. 214*, 711–736.
Nerdal, W., Hare, D. R., & Reid, B. R. (1988) *J. Mol. Biol. 201*, 717–739.
Nerdal, W., Hare, D. R., & Reid, B. R. (1989) *Biochemistry 28*, 10008–10021.
Nilsson, L., Clore, G. M., Gronenborn, A. M., Brünger, A. T., & Karplus, M. (1986) *J. Mol. Biol. 188*, 455–475.
Orbans, L. P. M., & Altona, C. (1986) *Eur. J. Biochem. 160*, 141–148.
Otting, G., Widmer, H., Wagner, G., & Wüthrich, K. (1985) *J. Magn. Reson.* 66, 187–193.
Pardi, A., Hare, D. R., & Wang, C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8785–8789.
Prive, G. G., Yanagi, K., & Dickerson, R. E. (1991) *J. Mol. Biol. 217*, 177–199.
Reid, B. R. (1987) *Q. Rev. Biophys. 20*, 1–34.
Reid, B. R., Banks, K., Flynn, P., & Nerdal, W. (1989) *Biochemistry 28*, 10001–10007.
Remerowski, M. L., Glaser, S. J., & Drobny, G. P. (1989) *Mol. Phys. 68*, 1191–1218.
Rinkel, L. J., & Altona, C. (1987) *J. Biomol. Struct. Dyn. 4*, 621–649.
Scheek, R. M., Russo, N., Boelens, R., Kaptein, R., & van Boom, J. H. (1983) *J. Am. Chem. Soc. 105*, 2914–2916.
Schmitz, U., Zon, G., & James, T. L. (1990) *Biochemistry 29*, 2357–2368.
Shakked, Z., Guerstein-Guzikevich, G., Eisenstein, M., Frolow, F., & Rabinovich, D. (1989) *Nature 342*, 456–460.
Sletten, E., Nerdal, W., & Reid, B. R. (1991) *J. Cell. Biochem. 15G*, 81.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Tropp, J. (1980) *J. Chem. Phys. 72*, 6035–6043.
Tropp, J., & Redfield, A. G. (1981) *Biochemistry 20*, 2133–2140.
van de Ven, F. J. M., & Hilbers, C. W. (1988) *Eur. J. Biochem. 178*, 1–38.
Wüthrich, K. (1986) in *NMR of Proteins and Nucleic Acids*, pp 203–255, Wiley, New York.
Yanagi, K., Prive, G. G., & Dickerson, R. E. (1991) *J. Mol. Biol. 217*, 201–214.
Yip, P., & Case, D. A. (1989) *J. Magn. Reson. 83*, 643–648.
Zhurkin, V. B., Poltev, V. I., & Florentev, V. L. (1980) *Mol. Biol. 14*, 1116–1130.