# Refinement of the Solution Structure of the DNA Decamer 5′d(CTGGATCCAG)$_2$: Combined Use of Nuclear Magnetic Resonance and Restrained Molecular Dynamics†

Michael Nilges, G. Marius Clore,* and Angela M. Gronenborn*

*Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, FRG*

Norbert Piel‡

*Max-Planck-Institut für Experimentelle Medizin, D-3000 Gottingen, FRG*

Larry W. McLaughlin

*Department of Chemistry, Boston College, Chestnut Hill, Massachusetts 02167*

*Received January 26, 1987*

ABSTRACT: The solution structure of the self-complementary DNA decamer 5′d(CTGGATCCAG)$_2$ comprising the specific target site for the restriction endonuclease *Bam*H1 is investigated by using nuclear magnetic resonance sectroscopy and restrained molecular dynamics. With the exception of the H5′/H5″ sugar proton resonances, all the nonexchangeable proton resonances are assigned sequentially by using pure-phase absorption two-dimensional nuclear Overhauser enhancement spectroscopy. From the time dependence of the nuclear Overhauser effects a set of 160 approximate interproton distances is determined and used as the basis of a structure refinement employing restrained molecular dynamics in which the interproton distances are incorporated into the total energy function of the system in the form of an effective potential term. Two restrained dynamics simulations are carried out, starting from classical B- and A-DNA [atomic root mean square (rms) difference 5.7 Å]. In both cases convergence is achieved to very similar B-type structures with an atomic rms difference of 0.9 Å which is comparable to the rms fluctuations of the atoms about their average positions. In addition, the rms difference between the experimental and calculated values of the interproton distances for both average restrained dynamics structures is ~0.3 Å. These results suggest that the converged restrained molecular dynamics structures represent reasonable approximations of the solution structure. The average restrained dynamics structures exhibit clear sequence-dependent variations of torsion angles and helical parameters. In addition, the structures exhibit a small bend of around 10–20° at the second (TpG) and eighth (CpA) base pair steps. This can be attributed to the positive base roll angles and large base pair slide values at the two Pyr-Pur steps. The central core of the decamer comprising the six-base recognition site for *Bam*H1 (GGATCC), however, is straight.

As part of a study on the effects of base sequence on nucleic acid structure in solution we present a combined nuclear magnetic resonance (NMR)[1] and restrained molecular dynamics study on the self-complementary DNA decamer 5′d(CTGGATCCAG)$_2$ comprising the target site GGATCC for the restriction endonuclease *Bam*H1. First, all nonexchangeable proton resonances (with the exception of the H5′/H5″ sugar proton resonances) are assigned in a sequential manner by means of pure-phase absorption two-dimensional nuclear Overhauser enhancement spectroscopy (NOESY). From the time dependence of the NOE cross-peak intensities a set of 160 approximate interproton distances is derived and then used as the basis for a structure refinement by restrained molecular dynamics (Kaptein et al., 1985; Clore et al., 1985, 1986; Brunger et al., 1986). As in our two previous studies on two DNA hexamers (Nilsson et al., 1986; Nilges et al., 1987), convergence is achieved by starting from two quite different initial structures, namely, classical B- and A-DNA, which in this case have an atomic rms difference of 5.7 Å. The converged structures are of the B type and have an atomic rms differences of 0.9 Å which is comparable to the rms fluctuations of the atoms about their average positions. Finally, the converged structures are analyzed and shown to display sequence-dependent variations in the values of the torsion angles and helical parameters. In this respect, we note that although there have been other NMR studies on oligonucleotides containing the *Bam*H1 recognition site, in particular on the self-complementary hexamer 5′d(GGATCC)$_2$ (Sarma et al., 1985) and dodecamer 5′d(GGATCCGGATCC)$_2$ (Kumar et al., 1985), these have been limited to assignment of proton resonances and the structural conclusions have been restricted to a qualitative interpretation of absolute value NOESY spectra (viz., the distinction between B- and A-DNA).

† This work was supported by the Max-Planck Gesellschaft and Grant Cl 86/1-1 of the Deutsche Forschungsgemeinschaft (G.M.C. and A.M.G.) and by Grant NSF-OMB/8518940 of the National Science Foundation (L.W.M.). M.N. thanks the Max-Planck Gesellschaft for a Max-Planck predoctoral fellowship.

‡ Present address: Bayer AG, Abteilung ZFBT, D-5090 Leverkusen, FRG.

[1] Abbreviations: NMR, nuclear magnetic resonance spectroscopy; NOE, nuclear Overhauser enhancement or effect; NOESY, two-dimensional nuclear Overhauser enhancement spectroscopy; rms, root mean square; RD, restrained dynamics; EDTA, ethylenediaminetetraacetic acid.

0006-2960/87/0426-3734$01.50/0 

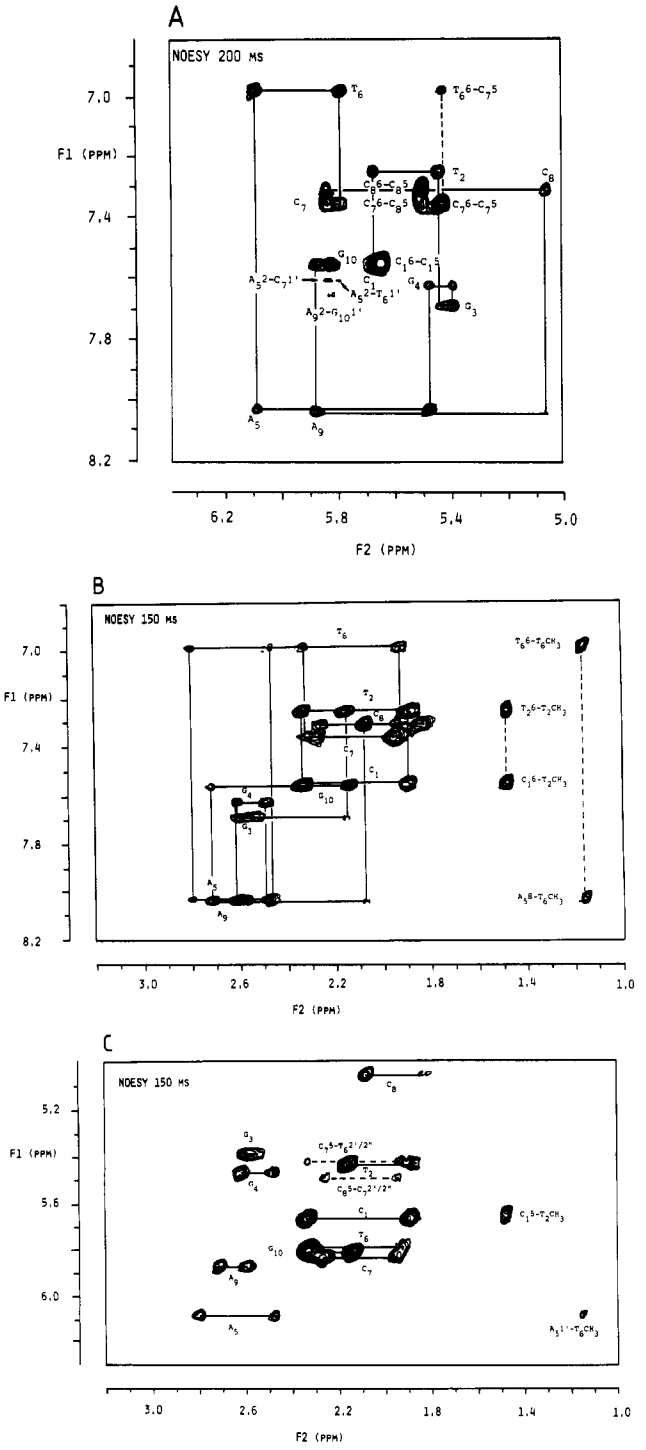


FIGURE 1: Pure-phase absorption NOESY spectra of the decamer. (A), (B), and (C) show the H8/H6($F1$ axis)–H1′/H5($F2$ axis), H8/H6($F1$ axis)–H2′/H2″/$CH_3$($F2$ axis), and H1′/H5($F1$ axis)–H2′/H2″/$CH_3$($F2$ axis) regions of the NOESY spectrum, respectively. Mixing times are indicated in the figure. Apodization was carried out by multiplying the time domain data with a sine-squared bell shifted by $\pi/4$ in both the $t_1$ and $t_2$ dimensions.

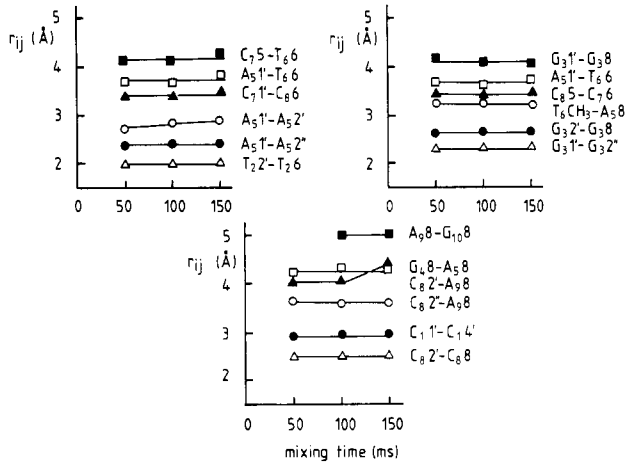


FIGURE 2: Examples of the dependence of the calculated values of $r_{ij}$ on mixing time determined by using eq 2.

Table I: Proton Resonance Assignments of the Decamer at 20 °C

| | chemical shift (ppm) | | | | | | |
|---|---|---|---|---|---|---|---|
| residue | H8/H6 | H5/CH3/ H2 | H1′ | H2′ | H2″ | H3′ | H4′ |
| C1 | 7.58 | 5.65 | 5.08 | 1.91 | 2.34 | 4.50 | 3.93 |
| T2 | 7.27 | 1.50 | 5.45 | 1.89 | 2.17 | 4.69 | 3.97 |
| G3 | 7.70 | | 5.41 | 2.54 | 2.60 | 4.86 | 4.20 |
| G4 | 7.63 | | 5.49 | 2.48 | 2.63 | 4.52 | 4.27 |
| A5 | 8.04 | 7.63 | 6.09 | 2.48 | 2.80 | 4.89 | 4.32 |
| T6 | 7.00 | 1.16 | 5.79 | 1.93 | 2.34 | 4.71 | 4.44 |
| C7 | 7.37 | 5.44 | 5.84 | 1.95 | 2.28 | 4.70 | 4.46 |
| C8 | 7.31 | 5.51 | 5.07 | 1.84 | 2.09 | 4.65 | 3.91 |
| A9 | 8.05 | 7.67 | 5.88 | 2.60 | 2.73 | 4.51 | 4.24 |
| G10 | 7.57 | | 5.82 | 2.15 | 2.34 | 4.51 | 3.97 |

## EXPERIMENTAL PROCEDURES

*Sample Preparation.* The DNA decamer 5′d-$(CTGGATCCAG)_2$ was synthesized on a solid support of controlled-pore glass containing a long-chain alkylamine (CPG-LCAA) by using 1-hydroxybenzotriazole-activated nucleotides as described previously (Piel et al., 1985; Marugg et al., 1983, 1984). After deprotection of the phosphate and nucleobase residues, the decamer was isolated as the terminal 9-phenyl-9-xanthenyl derivative on a 9.4 × 250 mm column of ODS-Hypersil (McLaughlin & Piel, 1984). Following removal of the 9-phenyl-9-xanthenyl group, the isolated oligodeoxynucleotide was eluted as a single peak from both anion-exchange and reverse-phase HPLC columns. Wandering spot analysis confirmed both the nucleoside composition and sequence (Wu et al., 1984).

After desalting and extensive lyophilizaton, the decamer (final concentration 3.4 mM) was taken up in 99.96% $D_2O$ containing 300 mM KCl, 50 mM potassium phosphate, pH* 6.5 (meter reading uncorrected for the isotope effect on the glass electrode), and 0.02 mM EDTA. The temperature used for all NMR experiments was 20 °C. Under these conditions of ionic strength and temperature, the decamer was entirely double stranded as judged from thermal denaturation studies and was of the B type as judged from its circular dichroism spectrum (unpublished data).

*NMR Spectroscopy.* All NMR spectra were recorded on a Bruker AM500 spectrometer equipped with an ASPECT 3000 computer and digital phase shifters. Quadrature detection was used with the carrier placed at the position of the residual HOD resonance. Chemical shifts are expressed relative to sodium 4,4-dimethyl-4-silapentane-1-sulfonate.

Two-dimensional NOESY spectra (Jeener et al., 1979) were recorded as pure-phase absorption spectra by using the time proportional phase incrementation method (Redfield & Kuntz, 1975; Bodenhausen et al., 1980) as described by Marion and Wüthrich (1983). Apropriate phase cycling was used to eliminate axial peaks and peaks due to multiple quantum coherence transfer; in addition, a 10% random variation in the mixing time was used to eliminate zero quantum coherence transfer (Macura et al., 1981). A total of 256 transients were collected for each of 512 increments with a relaxation delay of 1 s between sucessive transients. The spectral width em-

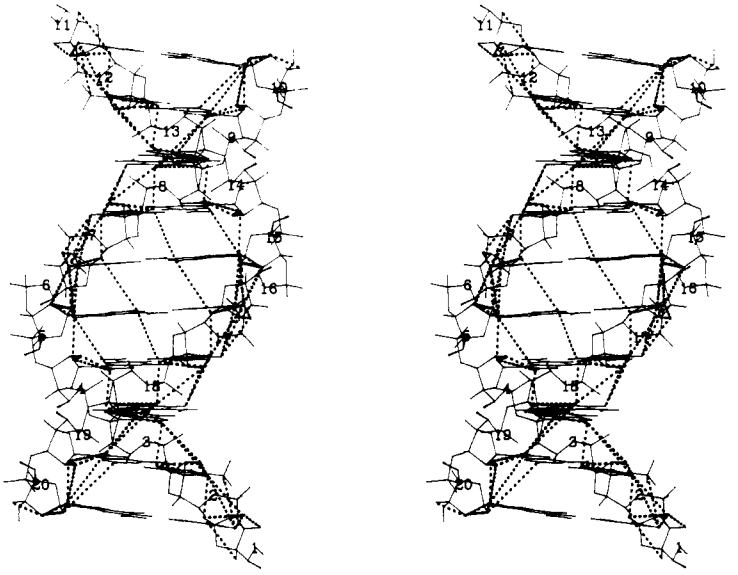

FIGURE 3: Stereoview of the interproton distance restraints as dashed lines on a classical B-DNA framework. Note that the values of the distance restraints in this figure are those found in classical B-DNA and *not* the experimental values.

Table II: $(\langle r^{-6}\rangle)^{-1/6}$ Mean Interproton Distances Calculated from Time-Dependent NOE Measurements[a]

| | Intranucleotide | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $r_{ij}$ (Å) | | | | | | | | | |
| proton | C1 | T2 | G3 | G4 | A5 | T6 | C7 | C8 | A9 | G10 |
| sugar–sugar | | | | | | | | | | |
| H1′–H2′ | 2.2 | 2.4 | 2.7 | 2.6 | 2.8 | | 2.2 | 3.1 | 2.5 | |
| H1′–H2″ | 2.0 | 2.2 | 2.3 | 2.4 | 2.5 | | 2.1 | 2.3 | 2.4 | 2.0 |
| H1′–H4′ | 2.9 | | | | | | | | | 2.5 |
| H2′–H3′ | 2.3 | | | | | | | | | |
| sugar–base | | | | | | | | | | |
| H1′–H8/H6 | | 3.3 | 4.1 | 4.4 | 4.2 | 3.9 | 3.1 | 3.6 | 3.7 | 3.2 |
| H2′–H8/H6 | 2.3 | 2.0 | 2.6 | 2.7 | 2.5 | 2.5 | 2.1 | 2.5 | 2.3 | 2.5 |
| H3′–H8/H6 | | 3.5 | | | | | | 3.8 | | |

| | | Internucleotide (Intrastrand) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $r_{ij}$ (Å) | | | | | | | | |
| proton of 5′-residue | proton of 3′-residue | $C_1pT_2$ | $T_2pG_3$ | $G_3pG_4$ | $G_4pA_5$ | $A_5pT_6$ | $T_6pC_7$ | $C_7pC_8$ | $C_8pA_9$ | $A_9pG_{10}$ |
| H1′ | H8/H6 | 3.3 | 3.8 | 3.6 | 3.5 | 3.6 | 3.7 | 3.4 | 4.0 | 3.4 |
| H2′ | H8/H6 | | 4.4 | | | 3.4 | | 2.4 | 4.1 | 3.1 |
| H2″ | H8/H6 | 2.4 | 3.4 | | | 3.1 | 3.2 | | 3.6 | 3.1 |
| H1′ | H5/CH3 | | | | | 4.0 | | | | |
| H2′ | H5/CH3 | | | | | | | 3.7 | | |
| H2″ | H5/CH3 | | | | | | 2.9 | 3.1 | | |
| H8/H6 | H8/H6 | 4.9 | 5.1 | | 4.3 | 4.8 | 5.6 | | 3.4 | 5.0 |
| H8/H6 | H5/CH3 | 3.0 | | | | 3.3 | 4.1 | 3.4 | | |
| H5/CH3 | H5/CH3 | | | | | | 4.5 | | | |
| H2 | H1′ | | | | | 4.1 | | | | 4.5 |

Internucleotide (Interstrand)

A5(H2)–C17(H1′)/A15(H2)–C7(H1′) $r_{ij}$ = 4.0 Å

[a] When the interproton distances were calculated by using eq 2, the H2′–H2″ initital cross-peak buildup rate and distance were used for all sugar–sugar and sugar–base (with the exception of the sugar–H1′ base) distances and the C(H5)–C(H6) initial cross-peak buildup rate and distance were used for all base–base and H1′–sugar base distances (see text). The estimated errors in the distances are as follows: −0.2/+0.2 Å for r ≤ 2.0 Å; −0.2/+0.4 Å for 2.0 Å < r ≤ 2.5 Å; −0.3/+0.5 Å for 2.5 Å < r ≤ 3.3 Å; and −0.5/+0.7 for 3.3 Å < r ≤ 6 Å.

ployed was 5000 Hz. A square 1K × 1K frequency matrix was obtained by zero filling in the $t_1$ dimension to give a digital resolution of 4.88 Hz per point in both dimensions. To reduce $t_1$ noise, the first time domain data point were multiplied by a factor of 0.5 (Otting et al., 1986). An initial phase correction was carried out during transformation with a final adjustment after completion of the two-dimensional transform. These manipulations were followed by base line correction (Pearson, 1977) and finally symmetrization (Bauman et al., 1981). NOESY spectra were recorded at four mixing times: 50, 100, 150, and 200 ms. Quantification of cross-peak intensities was carried out on a Vax 11/780 by determining the volume of each cross-peak by two-dimensional integration using a modified version of the Groningen 2D NMR processing program (Boelens, Kaptein, and Scheek, unpublished data).

*Molecular Dynamics.* All energy minimization and molecular dynamics calculations were carried out by using the program CHARMM (Brooks et al., 1983), optimized for the CRAY computer (Brunger, unpublished data), as described in our two previous restrained molecular dynamics studies on oligonucleotides (Nilsson et al., 1986; Nilges et al., 1987). The effective potential $E_{NOE}$ representing the interproton distance restraints was added to the total energy function of the system in the form of a skewed biharmonic effective potential [Clore et al., 1985; cf. eq 1 and 2 of the preceding paper (Nilges et al., 1987)].

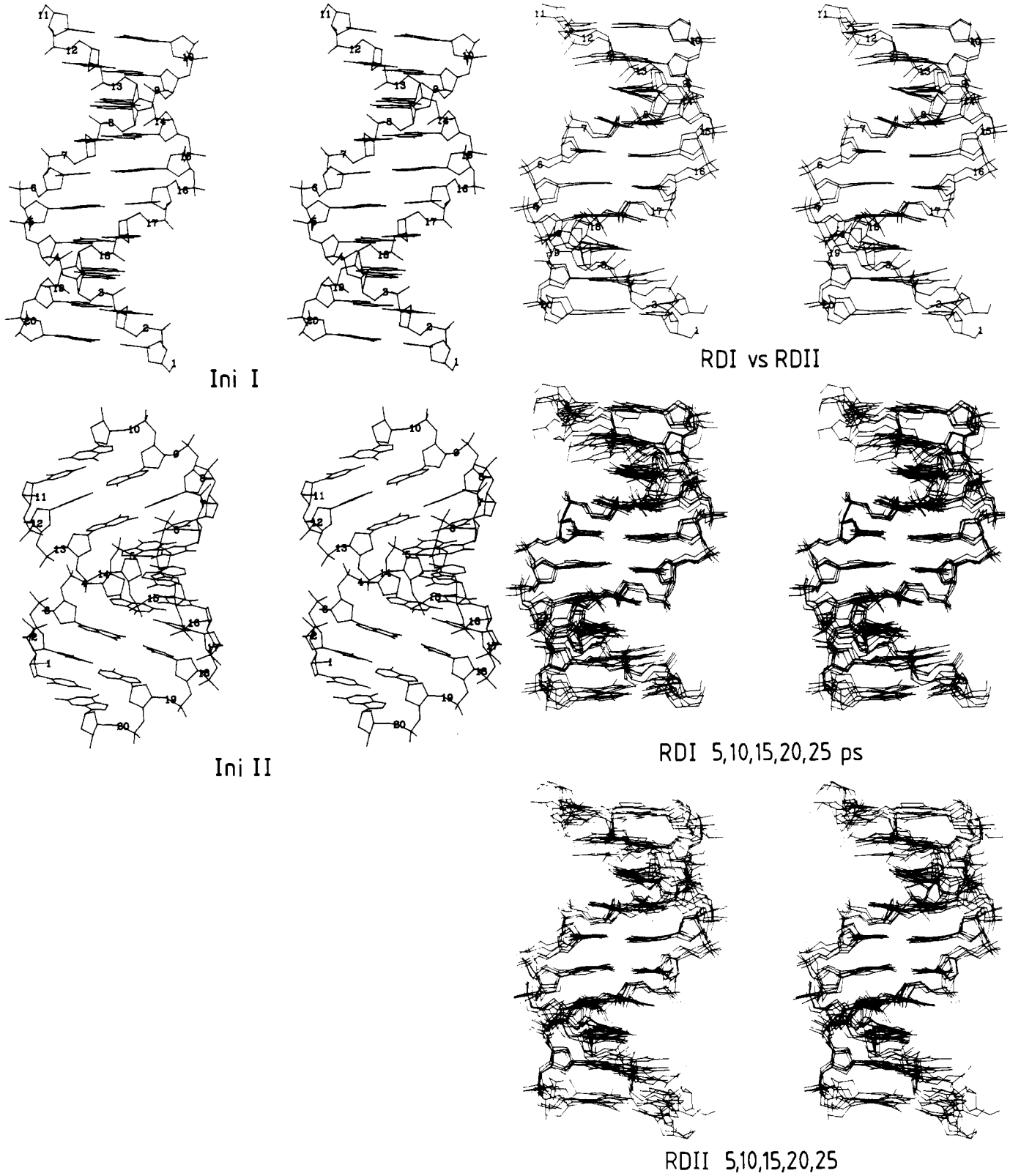


FIGURE 4: Stereoviews along the helix axis of the initial structures IniI and IniII, the best fit superposition of the two average restrained dynamics structures RDI and RDII, and the best fit superposition of the structures at 5, 10, 15, 20, and 25 ps of the second dynamics run for the restrained dynamics structures RDI and RDII.

Analysis of helical parameters was carried out by using modified versions of the AHELIX (written by J. Rosenberg) and BROLL and CYLIN (written by R. E. Dickerson) programs adapted to deal with dynamics trajectories (Nilges et al., 1987).

## Results and Discussion

*Resonance Assignment and Interproton Distances.* The sequential assignment of resonances was accomplished by means of NOESY spectroscopy to demonstrate through-space connectivities <5 Å as described previously (Reid et al., 1983; Scheek et al., 1983; Hare et al., 1983; Feigon et al., 1983; Clore & Gronenborn, 1983; Weiss et al., 1984). This involves establishing intranucleotide connectivities between sugar protons (e.g., H1′–H2′, H1′–H2″, H1′–H4′) and between base and sugar protons (e.g., H1′/H2′/H3′–H8/H6) and internucleotide connectivities of the type H1′/H2′/H2″(*i*)–H8/

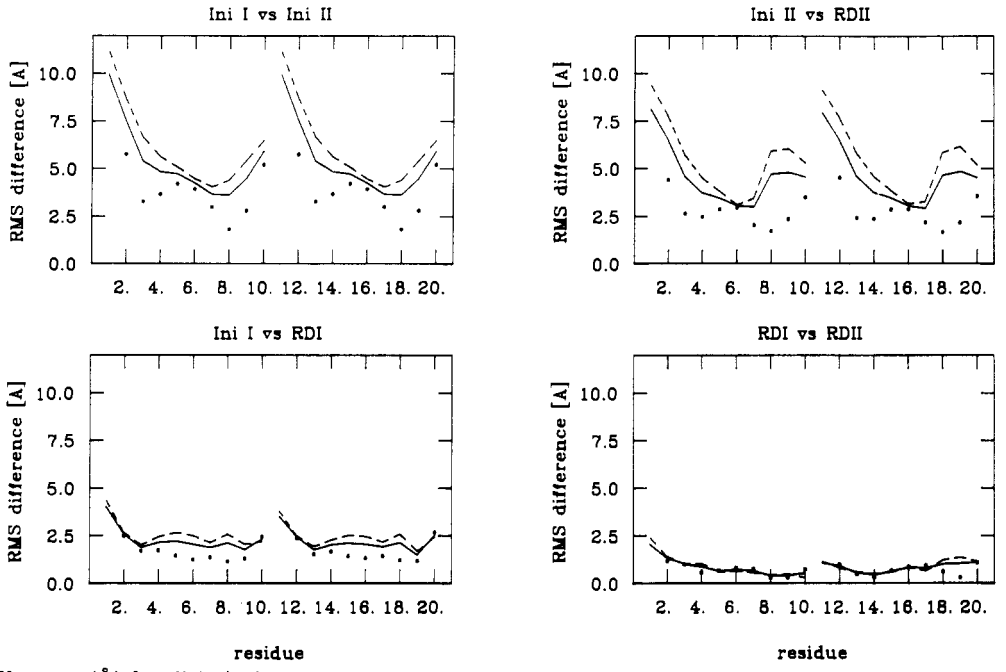


FIGURE 5: rms differences (Å) for all (—), the sugar–phosphate backbone (- - -), and the base (···) atoms as a function of residue number for various pairs of structures involving the initial (IniI, IniII) and average restrained dynamics (RDI, RDII) structures.

Table III: Atomic rms Differences between Initial (IniI, IniII) and Average Restrained Dynamics (RDI, RDII) Structures

| | overall rms difference (Å) | | |
|---|---|---|---|
| | IniII | RDI | RDII |
| IniI | 5.7 | 2.3 | 1.9 |
| IniII | | 4.3 | 4.8 |
| FDI | | | 0.9 |

Table IV: rms Differences of the Interproton Distances for Initial (IniI, IniII) and Average Restrained Dynamics (RDI, RDII) Structures

| | rms differences of interproton distances (Å) | | |
|---|---|---|---|
| | all (160) | intraresidue (82) | interresidue (78) |
| IniI | 0.56 | 0.39 | 0.70 |
| IniII | 0.87 | 0.79 | 0.96 |
| RDI | 0.32 | 0.29 | 0.35 |
| RDII | 0.32 | 0.29 | 0.34 |

H6($i$ + 1), H8/H6($i$)–H5/$CH_3$ (i + 1), and H8/H6($i$)–H8/H6 ($i$ + 1). Some examples of NOESY spectra are shown in Figure 1, and the list of assignments is given in Table I.

In order to determine interproton distances, the cross-peak intensities were measured as a function of mixing time. Relative cross-relaxation rates for the fixed distance reference vectors C(H5)–C(H6) (2.5 Å), T($CH_3$)–T(H6) (2.7 Å), and H2′-H2″ (1.8 Å) were then determined from the initial buildup rates of the corresponding cross-peaks. As $\omega\tau_c \gg 1$ (where $\omega$ is the spectrometer frequency and $\tau_c$ the correlation time), ratios of effective correlation times were calculated from (Solomon, 1955)

$$\frac{s_{ij}r_{ij}^6}{s_{kl}r_{kl}^6} \sim \frac{\tau_{\text{eff}}(ij)}{\tau_{\text{eff}}(kl)} \qquad (1)$$

where $s_{ij}$ and $s_{kl}$ are relative cross-relaxation rates between protons $i$ and $j$ and between protons $k$ and $l$, respectively, $r_{ij}$ and $r_{kl}$ the corresponding distances, and $\tau_{\text{eff}}(ij)$ and $\tau_{\text{eff}}(kl)$ the corresponding effective correlation times. As in previous cases (Gronenborn et al., 1984; Clore & Gronenborn, 1984; Nilges et al., 1987), no residue to residue variation in effective correlation times could be detected and the effective correlation time of the H2′–H2″ sugar vector was significantly shorter than that of the other two base vectors, in this instance by a factor of 3. Consequently, we used the same choice of reference distance in calculating unknown distances that we have discussed in detail previously (Gronenborn et al., 1984; Gronenborn & Clore, 1985): namely, the H2′–H2″ vector was used in the calculation of all sugar–sugar and sugar–base (with the exception of sugar H1′–base) distances, and the H5–H6 (or $CH_3$–H6) vector was used in the calculation of the sugar H1′–base and base–base distances. As no significant departures from the initial rate approximation $a_{ij}(t) \sim s_{ij}t$ (where $a_{ij}(t)$ is the cross-peak intensity at time $t$; Wagner & Wüthrich, 1979;: Dobson et al., 1982; Clore & Gronenborn, 1985) were apparent up to mixing times of 150 ms for all cross-peaks except the H2′–H2″ cross-peaks, interproton distances were calculated at each mixing time from (Clore & Gronenborn, 1985)

$$r_{ij}(t) = [s_{\text{ref}}t/a_{ij}(t)]^{1/6}r_{\text{ref}} \qquad (2)$$

where $s_{\text{ref}}$ and $r_{\text{ref}}$ are the relative cross-relaxation rate and distance of the appropriate reference vector, respectively. Some typical plots of calculated values of $r_{ij}$ as a function of mixing time are shown in Figure 2. (Note that, in this representation, a departure from the initial rate approximation is manifested by an increase in the calculated value of $r_{ij}$). A summary of the calculated interproton distances (taken as the average of the values calculated at 50, 100, and 150 ms) is given in Table II. On the basis of our previous calculations (Clore & Gronenborn, 1985) and taking into account the errors involved in determining cross-peak intensities by volume integration, we have estimated the errors in the interproton distances as follows: for $r_{ij} \leq 2.0$ Å, the errors are −0.2/+0.2 Å; for 2.0 Å $< r_{ij} \leq 2.5$ Å, they are −0.2/+0.4 Å; for 2.5 Å $< r_{ij} \leq 3.3$ Å, they are −0.3/+0.5 Å; and for 3.3 Å $< r_{ij} \leq 6.0$ Å, they are −0.5/+0.7 Å. These error estimates are sufficiently generous to ensure that errors arising from variations in effective correlation times have a negligible effect on the end results. A stereoview of this distance set, comprising 160 distances, superimposed on a classical B-DNA framework is shown in Figure 3.

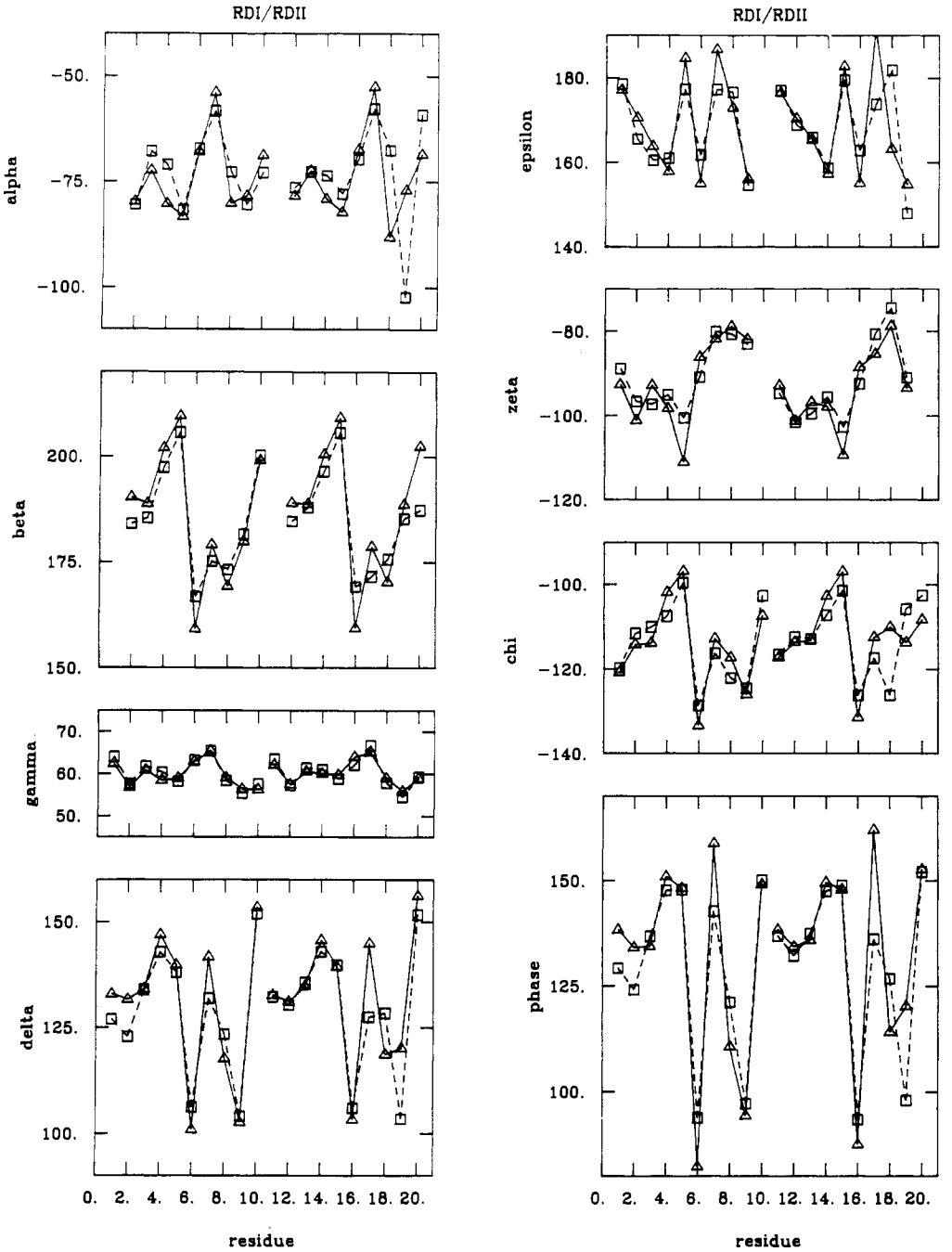


FIGURE 6: Variation in the backbone and glycosidic bond torsion angles as well as the phase angle describing the sugar pucker for the two restrained dynamics structures RDI (□) and RDII (△). The phase angle is calculated as described by Cremer and Pople (1975) with the apex at atom 3 and O4′ = atom 0, C1′ = atom 1, and so on.

Table V: Individual Energy Terms for Initial (IniI, IniII) and Average Restrained Dynamics (RDI, RDII) Structures

| | energy (kcal/mol) (number of terms) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| structure | total[a] | potential[a] | bond (680) | angle (1230) | improper (286) | torsion (580) | electrostatic | van der Waals | hydrogen bonding | restraints[b] (160) |
| IniI | 321 | −30 | 20 | 163 | 0.08 | 311 | −327 | −132 | −65 | 351 |
| IniII | 970 | 218 | 11 | 174 | 0.10 | 341 | −193 | −53 | −62 | 752 |
| RDI[c] | −316 | −398 | 12 | 179 | 9.4 | 243 | −426 | −328 | −87 | 82 |
| RDII[c] | −302 | −382 | 11 | 176 | 9.7 | 245 | −403 | −334 | −87 | 80 |

[a] The total energy includes the restraints energy whereas the potential energy does not. [b] The restraints scale factor $S$ in eq 2 of Nilges et al. (1987) used in calculating the restraints energy is 4. Thus error estimates in the interproton distances of 0.2, 0.3, 0.4, 0.5, and 0.7 Å correspond to force constants of 29.8, 13.2, 7.5, 4.8, and 2.4 kcal/mol, respectively. [c] The energies for the restrained dynamics structures are those obtained after subjecting the average structures to 500 cycles of restrained energy minimization constrained to their original average structures by weak harmonic constraints (Bruccoleri & Karplus, 1986). This procedure is used to correct minor distortions in bond lengths and angles produced by the averaging procedure and results in only very small atomic rms shifts (≲0.1 Å) (1983).

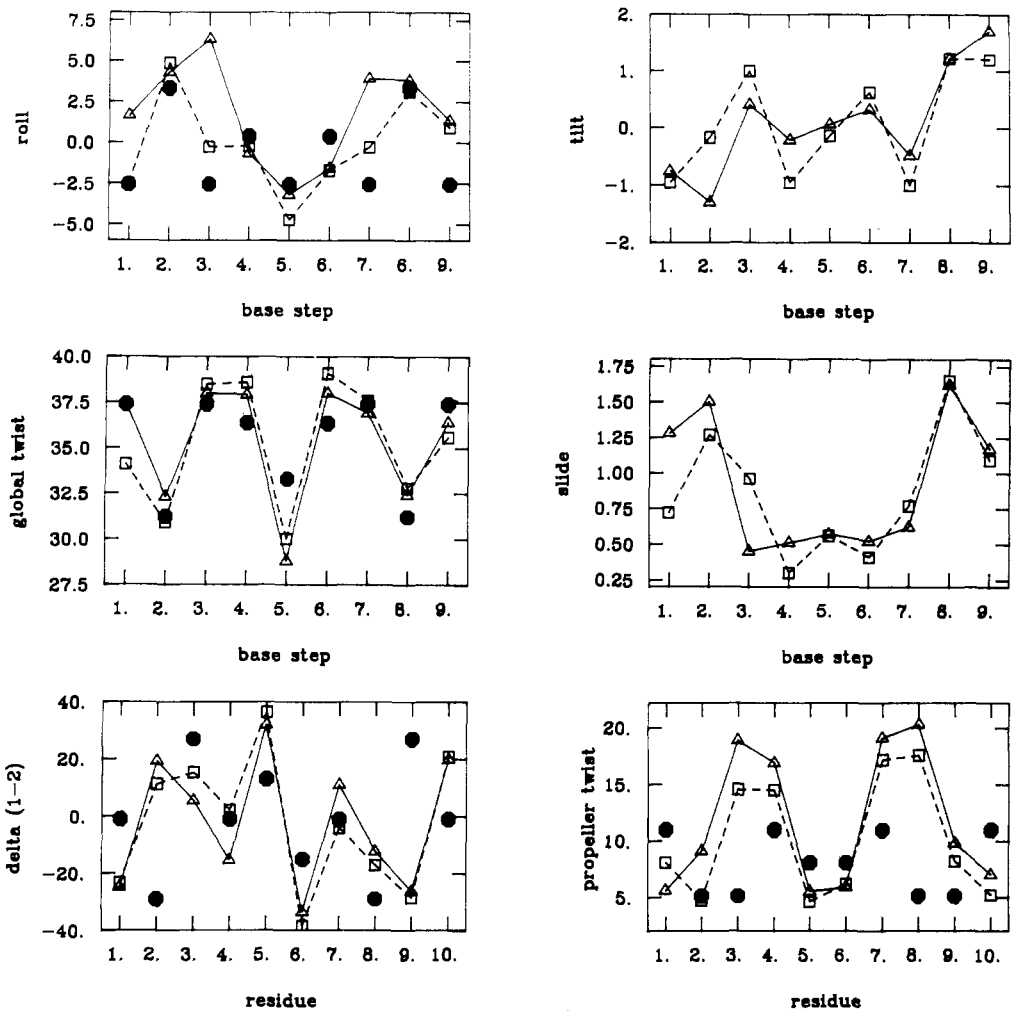


FIGURE 7: Variation in base roll, base tilt, global helical twist, and propeller twist angles, in base pair slide, and in the difference of the C4′–C3′ bond torsion angles $\delta$ at the two ends of a base pair ($\Delta\delta_{1-2}$) for the average restrained dynamics structures RDI (□) and RDII (△). The closed circles (●) represent the best fits of Dickerson's (1983) sum functions ($\sum_1$–$\sum_4$) to the data. $\sum_1$ is the sum function for global helical twist, $\sum_2$ for base roll, $\sum_3$ for $\Delta\delta_{1-2}$, and $\sum_4$ for propeller twist. The terms for $\sum$ are +1, −2, and +1 for x-Pur-Pyr-x and +2, −4, and +2 for x-Pyr-Pur-x; for $\sum_2$, they are +1, −2, and +1 for x-Pur-Pyr-x and −2, +4, and −2 for x-Pyr-Pur-x; for $\sum_3$, they are +1 and −1 for Pur-Py r and −2 and +2 for Pyr-Pur; and for $\sum_4$, they are −1 and −1 for Pur-Pyr and −2 and −2 for Pyr-Pur. The best fits are calculated by using the equation $y = S + T\sum$, where $y$ is the experimental value. In the case of roll and global helical twist, the fit shown represents the fit to all base pairs. In the case of $\Delta\delta_{1-2}$ and propeller twist, the fit shown represents the fit to all base pairs excluding base pairs 2 and 9 and base pairs 3 and 8, respectively.

*Structure Refinement.* In order to obtain an approximate picture of the decamer in solution, restrained molecular dynamics calculations (Kaptein et al., 1985; Clore et al., 1985, 1986; Brunger et al., 1986; Nilsson et al., 1986), incorporating the experimental interproton distances into the total energy function of the system in the form of an effective potential, were carried out, starting from two different initial structures, namely, classical B- (IniI) and classical A- (IniII) DNA (see Figure 4). The atomic rms difference between these two initial structures was 5.7 Å. Each structure was then subjected to the following steps: (i) 500 cycles of restrained energy minimization with the restraints scale factor $S$ [cf. eq 2 of preceding paper (Nilges et al., 1987)] set to 0.25; (ii) 1 ps of equilibration during which time the structure was heated up from 200 K to 300 K in steps of 10 K every 0.1 ps and $S$ was increased from 0.25 to 2.75 in steps of 0.25 every 0.1 ps; (iii) 15 ps of restrained dynamics (known as the first dynamics run) with the initial velocities assigned at 300 K and $S$ set to 3; and (iv) 28 ps of restrained dynamics (known as the second dynamics run) with the initial velocities reassigned at 300 K and $S$ set to 4. The temperature remained stable during the second dynamics run. The average restrained dynamics structures RDI and RDII were then obtained by averaging the coordinate trajectories from 5 to 28 ps of the second dynamics run.

The atomic rms differences between the structures is given in Table III, the rms differences between the calculated and experimental interproton distances in Table IV, and the energies of the initial and restrained dynamics structures in Table V. Stereoviews of the initial structures, the superposition of the average restrained dynamics structures, and the superposition of snapshots of the second dynamics run for each restrained dynamics structure are shown in Figure 4. The atomic rms differences as a function of residue number are shown in Figure 5.

It is clear from the data in Tables III–V and Figures 4 and 5 that convergence to essentially the same structure, both globally and locally, has been achieved, starting from both initial structures. The atomic rms difference between the average restrained dynamics structures is 0.9 Å, which is comparable to the rms fluctuation of the atoms about their average positions (see Figure 4), and the rms difference in the interproton distances (~0.3 Å) is within the distance errors specified. The extent of convergence can also be assessed by a comparison of the plots of backbone torsion angles (Figure 6) and helical parameters (Figure 7) as a function of residue number for both average restrained dynamics structures.

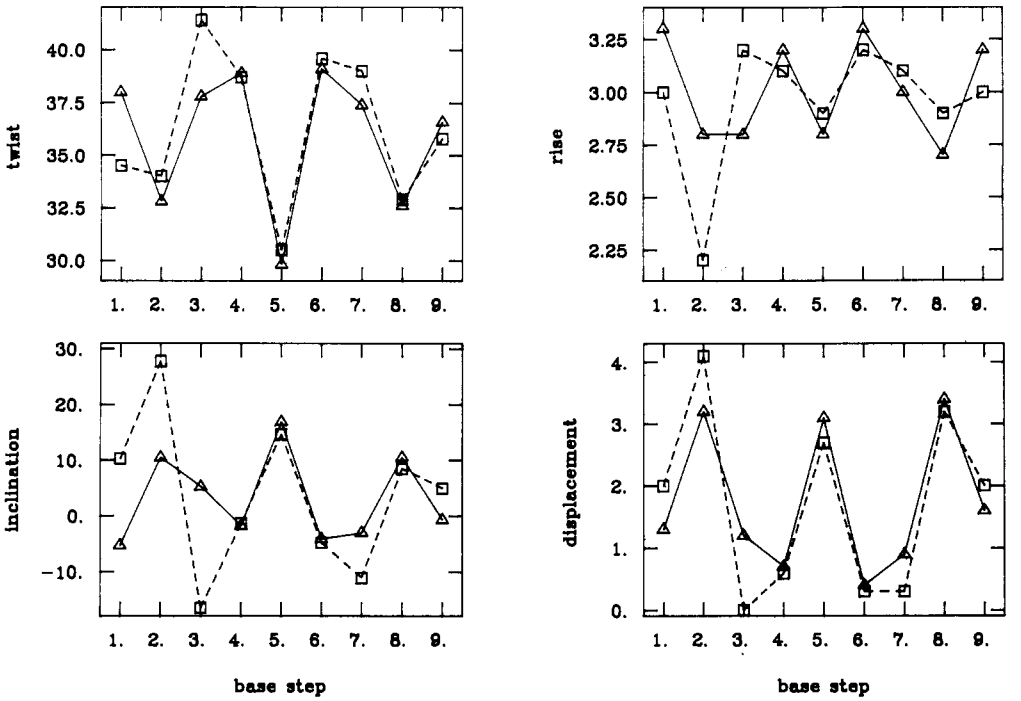


FIGURE 8: Variation in the local helical parameters (twist, rise, base pair inclination, and base pair displacement) as a function of residue number for the average restrained dynamics structure RDI (□) and RDII (△).

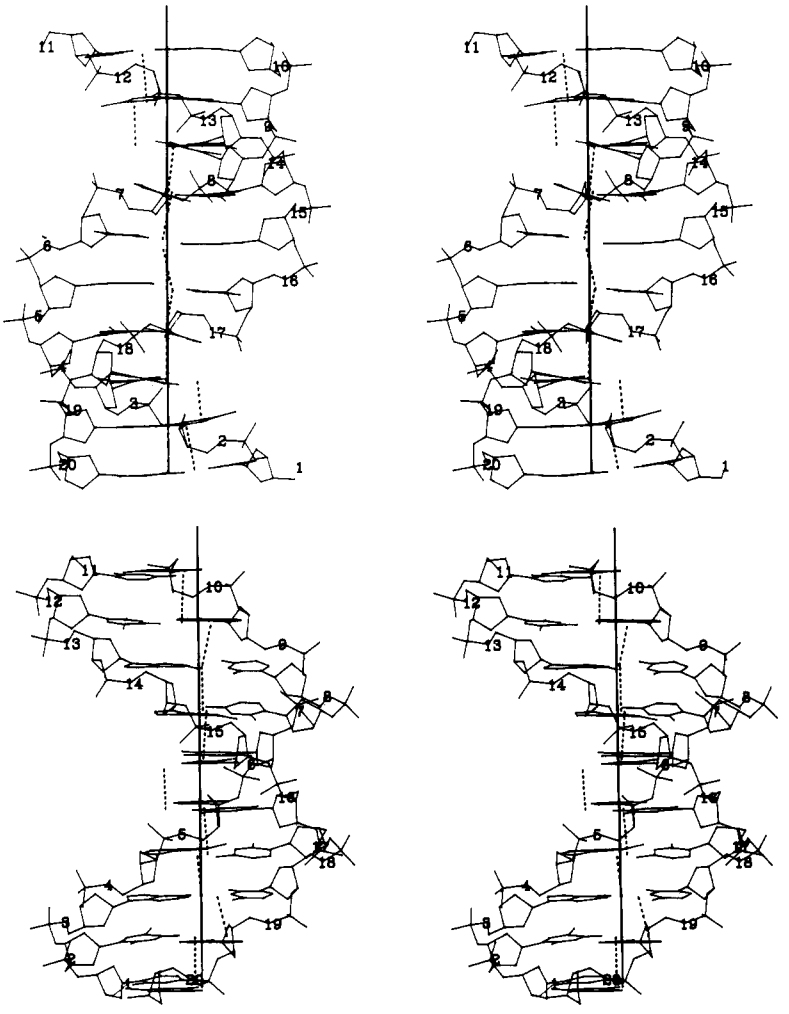

FIGURE 9: Two stereoviews of the average restrained dynamics structure RDI with the global helix axis (—) and the local helix axes (- - -) superimposed.

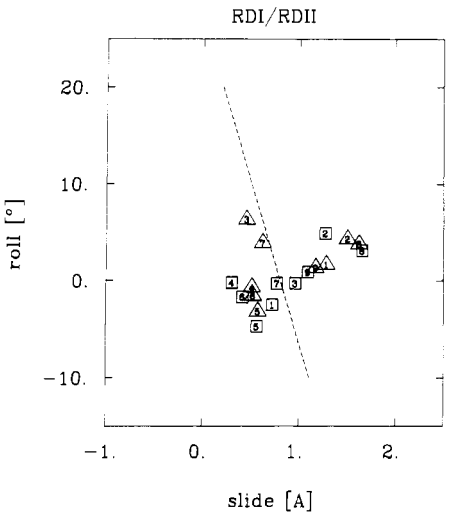


FIGURE 10: Roll–slide diagram for the two average restrained dynamics structures RDI (□) and RDII (△). The base pair steps are numbered inside the symbols. The dashed line from roll, slide = −10°, 1 Å, to +20°, 0.2 Å, represents the break between A- and B-type geometries, which lie to the right and left, respectively, of the line (Calladine & Drew, 1984).

*Structural Features of the Average Restrained Dynamics Structures.* The good convergence of the two restrained dynamics simulations starting from A- and B-DNA suggests that the two very similar average restrained dynamics structures, RDI and RDII, provide a reasonable representation of the conformational space sampled by the decamer in solution. It is therefore of interest to analyze the structures of RDI and RDII in detail, particularly with a view to examining possible sequence-dependent structural features.

The variation in torsion angle values as a function of residue is clearly symmetric (Figure 6) as expected given the symmetry of the NOE restraints. The degree of symmetry, although significantly better than that seen in the crystal structure of the self-complementary dodecamer (Dickerson & Drew, 1981), is not quite as good as that in the hexamer (Nilges et al., 1987). This is probably due to the increased length of the decamer so that a much longer period of restrained dynamics may be required to average out all fluctuations. As in the case of the hexamer 5′d(GCATGC)$_2$ (Nilges et al., 1987), the agreement in the values of the $\alpha$, $\beta$, $\gamma$, $\epsilon$, and $\xi$ torsion angles for the two restrained dynamics structures is reasonable despite the absence of any measured interproton distances directly related to these angles. This reaffirms the view that the positioning of the nucleotide units relative to each other achieved by the NOE restraints is sufficient, in the presence of the empirical energy function, to localize these backbone torsion angles to a relatively narrow region of conformational space within the confines of the range of values that can be adopted by A- and B-DNA.

All the sugar residues with the exception of those for residues T6/T16 and A9/A19 have sugar puckers in the C1′-exo to C2′-endo range associated with values of 120–150° for the C4′–C3′ bond torsion angle $\delta$ and values of −100° to −120° for the glycosidic bond torsion angle $\chi$. In the case of residues T6/T16 and A9/A19, however, the sugar pucker conformation is O1′-endo with values around 100° for $\delta$ and is correlated with more negative values of $\chi$ (−125° to −135°) (see Figure 6). These A-like features for these four residues are reflected in a number of other structural features associated with the heteropolymer base pair steps 2 (Pyr-Pur), 5 (Pur-Pyr), and 8 (Pyr-Pur) which were also observed in the restrained dynamics structure of the hexamer 5′d(GCATGC)$_2$ (Nilges et al., 1987). For all these three base pair steps, the local helix axis is displaced into the major groove, the local base pair inclination is increased, and the local helical rise, local helical twist, and global helical twist are reduced compared to other base pair steps in the sequence (see Figures 7–9). The roll–slide values, however, are B-like for step 5 but A-like for steps 2 and 8 (Figure 10). Similarly, the stacking pattern of the bases is B-like for step 5, which shows only intrastrand overlap of the bases, but A-like for steps 2 and 8, which exhibit interstrand base overlap (Figure 11). These two additional A-like features for the Pyr-Pur steps 2 and 8 result in bending of the ends of the decamer with respect to the straight central base pair steps 3–7 such that the bend angle (i.e., the angle between the best helix axis for steps 3–7 and the local helix axes for steps 1–9) is around 17° and 10° for RDI and RDII, respectively (see Figure 9). The additional 7° of bending in RDI arises from the difference in the values of the roll angles at the adjacent homopolymer steps 3–7 between the two average restrained dynamics structures, with RDI having values approximately 7° larger.

The central five base pair steps 3–7 are entirely straight. The local helix axes for steps 3, 4, 6, and 7 coincide with the global helix axis, and the local helix axis for step 5, although displaced, is parallel to the global helix axis (Figure 9).

The global helical twist is well predited by Dickerson's (1983) sum function $\sum_1$ (Figure 7) based on Calladine's (1982) rules. Fitting this sum function to the experimental data by means of the regression line $y = S + T\sum$ yields values of $S$ and $T$ of 35.3° and 1.05°, respectively, which are similar to those found in the crystal structure of the B-DNA dodecamer (35.6° and 2.1°, respectively; Dickerson, 1983) and the solution structure of the hexamer (34.9° and 0.9°, respectively; Nilges et al., 1987). The correlations of roll, $\Delta\delta_{1-2}$, and propeller twist with the appropriate sum function, however, are poor but can be greatly improved by omitting the "bad" steps, residues 3 and 7 for roll, 2 and 9 for $\Delta\delta_{1-2}$, and 3 and 8 for propeller twist (Figure 7). In the case of propeller twist this may be due to the alleviation of steric clash between the purines G3/G13 and A19/A9 on opposite strands by means of an increase in roll and slide and consequent bending at base pair steps 2 and 8. As a result, base pairs 3 and 8 can be highly propeller twisted with a concomitant improvement in base stacking with base pairs 4 and 7, respectively, without inducing steric clash (see Figure 11).

Considering both the average restrained dynamics structure of the decamer presented in this paper and that of the hexamer 5′d(GCATGC)$_2$ described previously (Nilges et al., 1987), certain common features emerge. In both cases the Pyr-Pur steps exhibit large roll and slide values and are responsible for bending of the DNA. The bend angle induced at these steps, though of the same order of magnitude, is slightly smaller in the decamer (10–17°) than in the hexamer (24–26°). The occurrence of these A-like features in two B-DNA oligonucleotides is in accordance with the strong bistability of Pyr-Pur steps proposed by Calladine and Drew (1984). Whereas the overall structure of the decamer is straight, that of the hexamer is very bent with a radius of curvature of approximately 20 Å. A possible explanation for this is that in the decamer the two Pyr-Pur steps are separated by five straight base pair steps whereas in the hexamer they are separated by only one base pair step. Such sequence-dependent variations in the structure of a DNA fragment at various steps in the sequence may play a role in the recognition process by restriction endonucleases. If one considers the decamer and the hexamer, the only difference between the *Bam*H1 recognition sequence GGATCC and the *Sph*1 recognition sequence

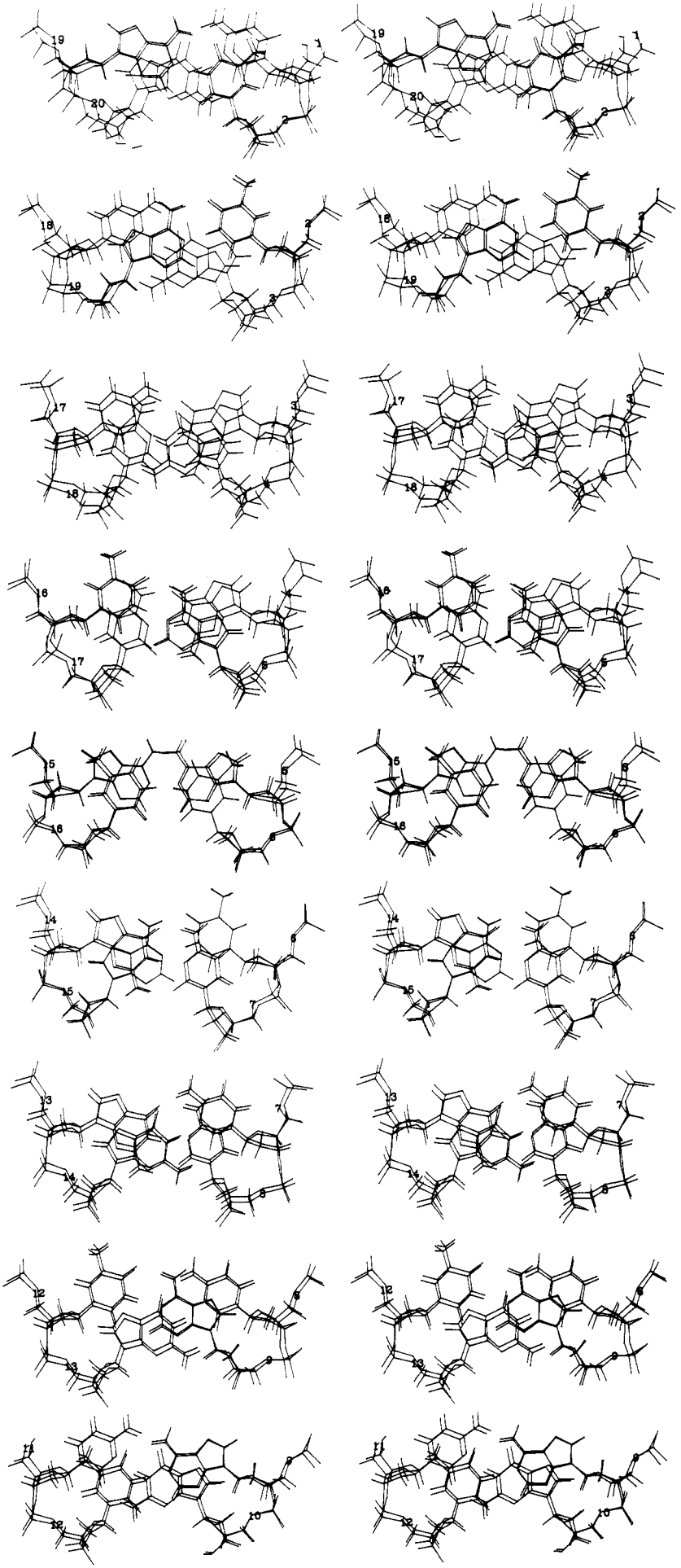

FIGURE 11: Stereoviews of the best fit superposition of the nine individual base pair steps of the two average restrained dynamics structures, RDI and RDII, viewed down the helix axis.

GCATGC is the exchange between the G and C at the symmetrically related positions 2 and 5 of the hexanucleotide. Thus a difference of only two nucleotides is all that is required to change the specificity for *Bam*H1 into that for *Sph*1. At the same time this same change is all that is required to change a straight piece of DNA (viz., GGATCC) into a bent one (viz., GCATGC). Similarly, the exchange of G and C at positions 2 and 5 of the *Bam*H1 sequence to A and T, respectively, converts the hexanucleotide into the *Eco*RI recognition site GAATTC. The structure of this hexanucleotide in solution would be expected to be similar to that of GGATCC as this alteration, in contrast to the one above, does not involve a purine for pyrimidine exchange. Indeed, the structure of the GAATTC segment in the crystal structure of the dodecamer 5'd$(CGCGAATTCGCG)_2$ (Dickerson & Drew, 1981) is similar to that of the GGATCC segment in the decamer: both are essentially straight and exhibit approximately the same pattern of variations in helical twist. This tentatively suggests two complementary mechanisms governing specificity, the first based upon the general shape of the specific DNA target site (e.g., straight vs. bent) and the second based upon specific hydrogen-bonding interactions.

## ACKNOWLEDGMENTS

We thank the Max-Planck-Institut für Plasma Physik (Garching) for computing facilities on the CRAY 1 computer.

## REFERENCES

Arnott, S., & Hukins, D. W. L. (1972) *Biochem. Biophys. Res. Commun. 47*, 1504–1509.

Bauman, R., Wider, G., Ernst, R. R., & Wüthrich, K. (1981) *J. Magn. Reson. 44*, 402–406.

Bodenhausen, G., Vold, R. L., & Vold, R. R. (1980) *J. Magn. Reson. 37*, 93–106.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.

Bruccoleri, R. E., & Karplus, M. (1986) *J. Comput. Chem. 7*, 175–185.

Brunger, A. T., Clore, G. M., Gronenborn, A. M., & Karplus, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3801–3805.

Calladine, C. R. (1982) *J. Mol. Biol. 161*, 343–352.

Calladine, C. R., & Drew, H. R. (1984) *J. Mol. Biol. 178*, 773–782.

Clore, G. M., & Gronenborn, A. M. (1983) *EMBO J. 2*, 2109–2115.

Clore, G. M., & Gronenborn, A. M. (1984) *FEBS Lett. 172*, 219–225.

Clore, G. M., & Gronenborn, A. M. (1985) *J. Magn. Reson. 61*, 158–164.

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

Clore, G. M., Brunger, A. T., Karplus, M., & Gronenborn, A. M. (1986) *J. Mol. Biol. 191*, 523–551.

Cremer, D., & Pople, J. A. (1975) *J. Am. Chem. Soc. 97*, 1358–1367.

Dickerson, R. E., & Drew, H. R. (1981) *J. Mol. Biol. 149*, 761–786.

Dobson, C. M., Olejniczak, E. T., Poulsen, F. M., & Ratcliffe, R. G. (1982) *J. Magn. Reson. 48*, 87–110.

Feigon, J., Denny, W. A., Leupin, W., & Kearns, D. R. (1983) *Biochemistry 22*, 5930–5942.

Gronenborn, A. M., & Clore, G. M. (1985) *Prog. Nucl. Magn. Reson. Spectrosc. 17*, 1–33.

Gronenborn, A. M., Clore, G. M., & Kimber, B. J. (1984) *Biochem. J. 221* 723–736.

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

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.

Kumar, R. M., Hosur, R. V., Roy, K. B., Miles, H. T., & Govil, G. (1985) *Biochemistry 24*, 7703–7711.

Macura, S., Huang, Y., Suter, D., & Ernst, R. R. (1981) *J. Magn. Reson. 43*, 259–281.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Marugg, J. E., McLaughlin, L. W., Piel, N., Tromp, M., van der Marel, G. A., & van Boom, J. H. (1983) *Tetrahedron Lett. 24*, 3989.

Marugg, J. E., Piel, N., McLaughlin, L. W., Tromp, M., Veeneman, G. H., van der Marel, G. A., & van Boom, J. H. (1984) *Nucleic Acids Res. 12*, 8639–8651.

McLaughlin, L. W., & Piel, N. (1984) in *Oligonucleotide Synthesis: A Practical Approach* (Gait, M. J., Ed.) pp 117–133, IRL Press, Oxford.

Nilges, M., Clore, G. M., Gronenborn, A. M., Brunger, A. T., Karplus, M., & Nilsson, L. (1987) *Biochemistry* (preceding paper in this issue).

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

Otting, G., Widmer, W., Wagner, G., & Wüthrich, K. (1986) *J. Magn. Reson. 66*, 187–193.

Pearson, G. A. (1977) *J. Magn. Reson. 27*, 265–272.

Piel, N., Benseler, F., Graeser, E., & McLaughlin, L. W. (1985) *Bioorg. Chem. 13*, 323–335.

Redfield, A. G., & Kuntz, S. D. (1975) *J. Magn. Reson. 19*, 250–254.

Reid, D. G., Salisbury, S. A., Bellard, S., Shakked, Z., & Williams, D. H. (1983) *Biochemistry 22*, 2019–2025.

Sarma, M. H., Dhingra, M. M., Gupta, G., & Sarma, R. H. (1985) *Biochem. Biophys. Res. Commun. 131*, 269–276.

Scheek, R. M., Russo, N., Boelens, R., Kaptein, R., & van Boom, J. H. (1983) *J. Am. Chem. Soc. 105*, 2914–2916.

Solomon, I. (1955) *Phys. Rev. 99*, 559–565.

Tidor, B., Irikura, K., Brooks, B. R., & Karplus, M. (1983) *J. Biomol. Struct. Dyn. 1*, 231–252.

Wagner, G., & Wüthrich, K. (1979) *J. Magn. Reson. 33*, 675–680.

Weiss, M. A., Patel, D. J., Sauer, R. T., & Karplus, M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 130–134.

Wu, R., Wu, N.-H., Hanna, Z., Georges, F., & Narang, S. (1984) in *Oligonucleotide Synthesis: A Practical Approach* (Gait, M. J., Ed.) pp 135–151, IRL Press, Oxford.