Solution Structure of a DNA Dodecamer Containing the Anti-Neoplastic Agent Arabinosylcytosine: Combined Use of NMR, Restrained Molecular Dynamics, and Full Relaxation Matrix Refinement†,‡

Barry I. Schweitzer,*,§ Thomas Mikita,‖ Gregory W. Kellogg,‖,⊥,△ Kevin H. Gardner,‖ and G. Peter Beardsley§

Department of Pediatrics, Yale University School of Medicine, Department of Chemistry and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06510

Received March 16, 1994; Revised Manuscript Received June 17, 1994®

ABSTRACT: The effect of araC incorporation into the dodecamer duplex [d(CGCGAATT)(araC)d(GCG)]$_2$ was examined by comparing its nuclear magnetic resonance (NMR)-determined solution structure with that of the control duplex d[(CGCGAATTCGCG)]$_2$. ^{1}H and ^{31}P resonances in both duplexes were assigned using a combination of 2-D ^{1}H NMR and a 3-D ^{31}P–^{1}H heteroTOCSY-NOESY experiment. Proton–proton distances (determined from NOESY data) and sugar dihedral angles (from NOESY and COSY data) were used in restrained molecular dynamics simulations starting from canonical A- or B-form DNA models. Both the control and araC sets of simulations converged to B-type structures. These structures were subjected to full relaxation matrix refinement to produce final structures which were in excellent agreement ($R_{1/6} < 0.05$) with the observed NOE intensities. A detailed comparison of the final control and araC structures revealed a global similarity (overall RMSD ~1.3 Å), with significant differences localized at the araC site and neighboring bases. These included changes in sugar pucker, backbone torsion angles, base stacking, and other helical parameters. These findings are in general agreement with the previously published X-ray structure of a decamer duplex containing araC. One intriguing feature of the NMR solution structure not found in the crystal structure is the presence of an intramolecular hydrogen bond between the 2′ hydroxyl on the araC sugar and the 3′ phosphate group.

The nucleoside analog arabinosylcytosine (1-β-D-arabinofuranosylcytosine, araC,[1] Figure 1) is a potent agent in the treatment of various forms of human leukemia (Frei et al., 1969; Bodey et al., 1969). araC is converted by cellular kinases into the active metabolite araCTP (Coleman et al., 1975; Hande & Chabner, 1978), which in turn leads to suppression of replication (Graham & Whitmore, 1970; Heintz & Hamlin, 1983) and repair (Collins, 1977) DNA synthesis. Possible mechanisms by which DNA synthesis is suppressed by araCTP include competitive inhibition of various DNA polymerases (Furth & Cohen, 1968; Momparler, 1972; Yoshida et al., 1977; Dicioccio & Srivastava, 1977) and misincorporation into newly synthesized DNA at 3′ termini, followed by slow or absent addition of the next nucleotide (Momparler, 1972; Fridland, 1977; Cozzarelli, 1977). Incorporation of araCMP into cellular genomic DNA has been detected at internucleotide positions as well as at 3′ termini (Major et al., 1982; Kufe et al., 1984); furthermore, the total amount of araC misincorporated into DNA at internucleotide positions has been correlated with the cytotoxicity of the drug (Kufe et al., 1980), suggesting that misincorporation of araC into DNA is the lethal event in cells.

Our laboratory and others have investigated the mechanism(s) by which misincorporation of araC into DNA interferes with DNA synthesis. The development of methodology for chemical synthesis of DNA oligomers containing araC at specific sites (Beardsley et al., 1988) allows the use of such oligomers as substrates for various DNA polymerases in *in vitro* assays and enables one to study the effects of araC misincorporation into DNA separately from its effects as a competitive inhibitor. In such experiments, araC located at an internucleotide site in the template strand of a primer/template duplex was found to inhibit DNA polymerase function, significantly reducing the rate of replication bypass of the Klenow fragment, T4 polymerase, and polymerase α_2, but not of reverse transcriptase (Mikita & Beardsley, 1988). Depending on the polymerase examined, synthesis arrest occurred either after insertion of dG opposite araC, one nucleotide before the araC site, or one nucleotide after the araC site. Further studies have shown that the pattern of synthesis arrest can also be affected by the local sequence context (Mikita & Beardsley, unpublished results).

† Supported by National Institutes of Health (NIH) Grants F32 AI08472 (B.I.S.) and RO1 CA42300 (G.P.B.) and American Cancer Society Institutional Research Grant IN31-33 (B.I.S.). G.W.K. was supported by NIH Grant GM-41651 to Peter B. Moore (Department of Chemistry). NMR and additional computational facilities were supported by NIH Grant RR03475, NSF Grant DMB8610557, and American Cancer Society Grant RD259. K.H.G. was supported by a Predoctoral fellowship from the Howard Hughes Medical Institute.

‡ Coordinates of the final araC and control structures and the experimental distance and dihedral restraints have been deposited in the Brookhaven Protein Database with the IDENT codes 170D and 171D, respectively.

* To whom correspondence should be addressed at his present address: Walt Disney Memorial Cancer Institute, 12722 Research Parkway, Orlando FL 32826.

§ Department of Pediatrics, Yale University School of Medicine.

‖ Department of Molecular Biophysics and Biochemistry, Yale University.

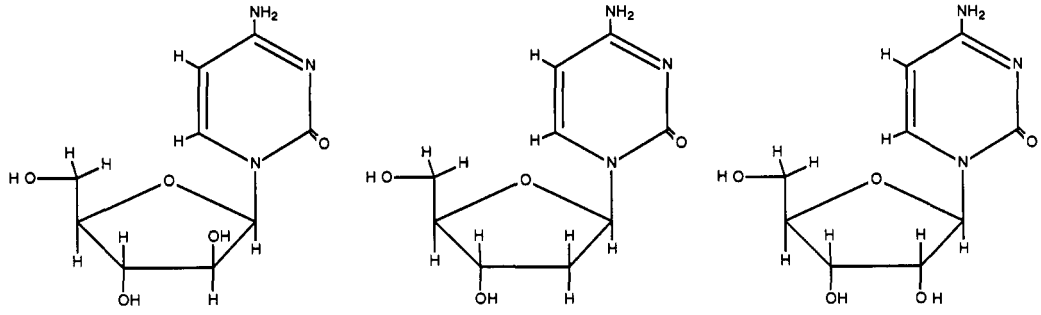
⊥ Department of Chemistry, Yale University.

△ Present address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115.

® Abstract published in *Advance ACS Abstracts,* September 1, 1994.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; DQF-COSY, double quantum filtered correlated spectroscopy; E-COSY, exclusive correlated spectroscopy; heteroTOCSY, heteronuclear total correlated spectroscopy; Ini-A, canonical A-DNA; Ini-B, canonical B-DNA; rMD, restrained molecular dynamics; RR, relaxation matrix refined; RMSD, root mean square deviation; araC, cytosine arabinoside; EDTA, ethylenediaminetetraacetic acid; IPSA, isolated spin-pair approximation.

0006-2960/94/0433-11460$04.50/0



A. B. C.

FIGURE 1: Chemical structure of cytosine arabinoside: (A) cytosine arabinoside; (B) deoxycytosine; (C) ribocytosine.

Additional experiments in our laboratory revealed that araC at the 3′ primer terminus greatly reduced the rate of addition of the next nucleotide by the Klenow fragment of *Escherichia coli* DNA polymerase I, T4 polymerase, HeLa cell polymerase α_2, and AMV reverse transcriptase. Polymerases possessing 3′–5′ exonuclease activity preferentially excised araCMP and replaced it with dCMP rather than extending an araCMP-terminated primer (Mikita & Beardsley, 1988). Using a preparation of human DNA polymerase α-primase complex devoid of exonuclease activity, Perrino and Mekosh (1992) demonstrated that extension from an araCMP 3′ terminus is approximately 2000-fold less efficient than extension from a dCMP 3′ terminus. They also found evidence that this effect was at least partially due to a decrease in the rate of phosphodiester bond formation rather than to an effect on binding of the DNA to the polymerase.

The above findings show that the seemingly subtle alteration in nucleoside structure present in araC produces profound biochemical effects, sparking considerable interest in determining the structural explanation for these observations. However, we have found it difficult to draw any firm conclusions about the structural basis for the mechanism of araC action from these studies. Of the four structures of DNA containing araC determined by X-ray crystallography by Wang and co-workers, one hexamer duplex was in a left-handed Z conformation (Teng et al., 1989), and two other hexamer duplexes were complexed with daunorubicin or doxorubicin (Zhang et al., 1992). A high-resolution X-ray structure of a B-type decamer duplex containing araC was described by Gao et al. (1991). In this structure, the araC substitution was found to produce small localized changes in the helix. However, it is becoming generally accepted that the structure of DNA can be altered by packing forces when examined in the crystal form; it is possible, therefore, that the araC X-ray structure does not accurately represent the structure present in solution.

The only oligonucleotide containing araC that has been analyzed by NMR in any detail was in a hairpin form (Pieters et al., 1990). This lack of a definitive NMR study on a DNA duplex containing araC prompted us to carry out the present study. For this study, we chose a familiar DNA sequence—the so-called Drew–Dickerson dodecamer—to place the araC substitution. This dodecamer has been the subject of numerous biophysical studies, including X-ray (Wing et al., 1980; Dickerson & Drew, 1981; Fratini et al., 1982), NMR (Sarma et al., 1982; Patel et al., 1983; Hare et al., 1983; Ott & Eckstein, 1985; Sklenar & Bax, 1987; Nerdal et al., 1989), and molecular dynamics simulations (Swaminathan et al., 1991; Kaluarachchi et al., 1991). Furthermore, the effects of base pair mismatches [e.g., G/T mismatch (Roongta et al., 1990)] and other substitutions [e.g., nebularine (Clore et al., 1988) and methylhypoxanthine (Yang et al., 1993)] have been extensively studied in this sequence by NMR. Using a novel three-dimensional ^{31}P–^{1}H heteroTOCSY-NOESY experiment (Kellog & Schweitzer, 1993), our laboratory completed the assignment of proton resonances for this dodecamer that began 10 years earlier (Hare et al., 1983).

In the present study, we used similar methodology to obtain complete proton and phosphorous assignments for the araC-containing dodecamer [d(CGCGAATT)(araC)d(GCG)]$_2$. A set of interproton distances and sugar dihedral angles derived from NOESY and COSY spectra, respectively, were used as the basis for structural refinement by restrained molecular dynamics and complete relaxation matrix calculations. A similar procedure was used to determine the solution structure of the control dodecamer [d(CGCGAATTCGCG)]$_2$. In both cases, convergence to a B-type structure was achieved starting from either canonical A- or B-type DNA. Although the resulting structures displayed overall similarity with the results of Gao et al.'s (1991) X-ray structure, the solution structure revealed some features which were not observed in the crystal structure, including a difference in the araC sugar pucker and an intramolecular hydrogen bond involving the araC 2′ hydroxyl group. These and other structural features of the araC-substituted dodecamer provide some possible insights into araC's mechanism of action.

MATERIALS AND METHODS

Sample Preparation. The DNA dodecamer [d(CGCGAATT)(araC)d(GCG)]$_2$ (also referred to herein as the araC duplex) and the control dodecamer [d(CGCGAATTCGCG)]$_2$ (also referred to herein as the control duplex) were synthesized on the Applied Biosystems automated DNA synthesis machine on a 10-μmol scale using standard solid-phase phosphoramidite chemistry. The araC phosphoramidite was prepared as previously described (Beardsley et al., 1988). The trityl on deprotected DNA was purified using a single reverse-phase column HPLC protocol, which includes the removal of failure sequences, on-column removal of the trityl group, and removal of nonfailure contaminating DNAs (Kellogg & Schweitzer, 1993). The purified (>98%) DNA was dialyzed twice against Chelex-100 (BioRad) and lyophilized to dryness. For NMR measurements in H_2O, the lyophilized DNA was dissolved in 0.4 mL of buffer containing 50 mM NaCl, 0.5 mM EDTA, 50 mM sodium phosphate, pH 7.0, and 10% D_2O. For measurements in D_2O, the DNA was repeatedly lyophilized to dryness, first from the aqueous buffer and then from 99.96% D_2O. Finally, 0.4 mL of 99.996% D_2O was added to give a duplex concentration of 5 mM.

NMR Spectroscopy. All NMR experiments were carried out on a Bruker AM-500 spectrometer (500.13 MHz for

protons, 202.45 MHz for phosphorus, other parameters as described in the text). The sample temperature was regulated at 303 K for all experiments. Six NOESY spectra for the control and araC DNA samples in D_2O with mixing times of 30, 50, 75, 100, 150, and 200 ms were each collected within separate 5-day periods, without removing the sample from the spectrometer. In each NOESY experiment, the mixing time was randomly varied over ±10% of the designated mixing time to eliminate zero-quantum coherence. The NOESY data were acquired in the phase-sensitive mode with time-proportional phase incrementation (TPPI) (Marion & Wüthrich, 1983) with 1024 complex points in t_2 and 512 points in t_1 and utilized a relaxation delay of 2.2 s (during which time the residual HDO peak was irradiated) and a spectral width of 4386 Hz. Three NOESY spectra for the control and araC DNA samples in H_2O with mixing times of 80, 120, and 200 ms were each collected within separate 3-day periods, without removing the sample from the spectrometer. The NOESY spectra in H_2O were acquired with a jump-return spin-echo pulse sequence for solvent suppression (Sklenar & Bax, 1987). These experiments were collected in the phase-sensitive mode of States et al. (1982) and 2048 complex points in t_2 and 256 complex points in t_1, a relaxation delay of 2.2 s, and a spectral width of 10 000 Hz. The exclusive COSY (E-COSY) spectrum (Griesinger et al., 1986, 1987) and the double quantum filtered (DQF)-COSY (Piantini et al., 1982) were acquired with 2048 complex points in t_2 and 800 points in t_1 and utilized a relaxation delay of 2.0 s (with no HDO irradiation) and a spectral width of 4386 Hz. The two- and three-dimensional ^{31}P–^1H heteroTOCSY-NOESY experiments were carried out as previously described (Kellogg, 1992; Kellogg et al., 1992; Kellogg & Schweitzer, 1993) with a heteroTOCSY mixing time of 125 ms (2-D) or 50 ms (3-D) and a NOESY mixing time of 300 ms. The time of the NOESY mixing period was subjected to a ±8% random variation. A homospoil pulse of 20 ms duration was applied at the beginning of this mixing period. In the 2-D experiment, the sweep widths were 252 (ω_1, ^{31}P) and 4386 Hz (ω_2, ^1H) with 60 and 2048 points acquired in the t_1 and t_2 dimensions, respectively. In the 3-D experiment, the sweep widths were 252 (ω_1, ^{31}P), 4386 (ω_2, ^1H), and 4386 (ω_3, ^1H) Hz with 48, 250, and 1024 points acquired in the t_1–t_3 dimensions, respectively.

NMR data were processed using the program Felix 2.05/2.10 (Hare Research, Inc.). In processing the NOESY data sets acquired in either H_2O or D_2O, a squared-sinebell shifted 90° was used as the window function in both dimensions. Data in both dimensions were zero-filled to 2048 real points, Fourier transformed, phased to pure absorption, and baseline-corrected with a fifth-order polynomial. NOE cross-peak volumes were obtained by cross-peak integration using the Felix software. Both the E-COSY and DQF-COSY data sets were apodized with unshifted sinebells in both dimensions, followed by zero-filling to 4096 real points in t_2 and 2048 real points in t_1 prior to Fourier transformation. The 3-D heteroTOCSY-NOESY data set was apodized with 60° shifted square sinebell functions in each dimension; in addition, the t_1 (^{31}P) dimension was zero-filled to 128 points, the t_2 (^1H) dimension was zero-filled to 512 points, and no zero-filling was used in the t_3 (^1H) dimension prior to Fourier transformation. ^{31}P resonances are referenced to an external sample of trimethyl phosphate. ^1H resonances are referenced to an external sample of sodium 2,2-dimethyl-2-silapentane-5-sulfonate in an identical buffer.

Determination of Interproton Distances. Initial buildup rates were calculated by fitting the cross-peak volumes at 30, 50, 75, 100, 150, and 200 ms mixing times to a linear equation. These rates (R_{ij}) were converted to distances (r_{ij}) using the cytosine H5–H6 distance of 2.45 Å as the reference distance (r_{ref}), the averaged rate calculated from well-resolved H5–H6 cross peaks (R_{ref}), and the relationship $r_{ij} = r_{ref}(R_{ref}/R_{ij})^{1/6}$. Cross-peak volumes obtained from the NOESY spectrum acquired in H_2O were corrected for the $\sin^3 \omega\tau_1$ offset dependence of the excitation profile in the jump-return experiment. Distances were calculated by fitting of the cross-peak volumes obtained at 80, 120, and 200 ms mixing times to a linear equation. These distances were given upper and lower bounds according to the following guidelines: distances (d) <3 Å had lower and upper bounds of 0.3 and 0.4 Å, respectively; $3.0 \leq d < 4.0$ Å had lower and upper bounds of 0.4 Å; $4.0 \leq d < 5.0$ Å had lower and upper bounds of 0.5 Å.

Structure Refinement Protocol. (1) Restraints. In addition to the experimentally determined interproton distance constraints and sugar ring dihedral angles used in restrained molecular dynamics and discussed in more detail below, four other types of restraints were used in calculating the structures presented in this study. *(a) Backbone torsion angles*: In order to preserve the right-handed character of the DNA duplex during the molecular dynamics calculations, the α, β, γ, ϵ, and ζ backbone torsion angles were restrained to a range covering both right-handed A- and B-DNA as well as satisfying the torsion angle values measured in both fiber diffraction and single crystal structures of all right-handed DNAs (Gronenborn & Clore, 1989; Baleja et al., 1990). *(b) Major groove distances*: To avoid the collapse of the major groove during the high-temperature dynamics phase of the simulation, C1′ atoms on opposite sides of the major groove were restrained to be greater than 16.0 Å apart during the calculations (Huang & Eisenberg, 1992). These distances are ca. 17.5 Å in canonical A-DNA and ca. 19 Å in canonical B-DNA (Arnott & Hukins, 1972, 1973). *(c) Base pair hydrogen bond distances*: Base pairs were kept hydrogen bonded in the Watson–Crick form by using distance restraints between the bases. These values were taken from X-ray analyses of ApU (Seeman et al., 1976) and GpC (Rosenberg et al., 1976). *(d) Base pair planarity*: The N1, N3, and C5 atoms of both cases of each base pair were constrained to be coplanar using the planarity energy term in X-PLOR. Portions of X-PLOR files containing the backbone torsion angles and the base pair hydrogen bond distances are given in supplementary material.

(2) Restrained Molecular Dynamics. Initial starting structures were generated by constructing the [d(CGCGAATTCGCG)]$_2$ duplex in canonical A- or B-form and then replacing the H2′ proton of the C9 residue with a hydroxyl group using the molecular editor function of QUANTA (version 3.1, Molecular Simulations, Inc.). These unminimized structures are designated Ini-A and Ini-B. All subsequent energy minimizations and molecular dynamics calculations were performed on a Silicon Graphics INDIGO R4000 XS24 using the program X-PLOR 3.1 (Brünger, 1993). All structures shown in this work were displayed using QUANTA 3.1. Nucleic acid parameters and force constants were obtained from the parameter file parallhdg.dna (X-PLOR 3.1). The parameters for the araC residue were deduced from ribose parameters. For the electrostatic component of the empirical energy function, the effect of solvent was approximated by a $1/r$ screening function (Brooks et al., 1993) and by reducing the net charge on the phosphate group to –0.32e (Tidor et al., 1982). All bond lengths were kept fixed with the SHAKE algorithm (Ryckaert et al., 1977). The protocol of restrained dynamics was similar to the one

employed by Mujeeb et al. (1993). Ini-A or Ini-B were first energy minimized with 500 cycles of conjugate gradient minimization with k_{NOE} at 0.5 kcal mol^{-1} A^{-2}, k_{CDIH} at 5.0 kcal mol^{-1} rad^{-2}, and k_{plan} at 0.5 kcal mol^{-1} A^{-2}. Using different random number seeds for each subsequent molecular dynamics run, initial velocities were assigned with a Maxwellian distribution at 0.4 K. The system was then slowly heated to 1000 K during the first 3 ps and then maintained at 1000 K for the next 10 ps. During this phase of the simulation, k_{CDIH} was constant at 50.0 kcal mol^{-1} rad^{-2}, and k_{plan} was constant at 15.0 kcal/mol^{-1} A^{-2}. K_{NOE} was slowly increased from 0.5 to a maximum value of 250 kcal mol^{-1} A^{-2} and then held constant at this level during the high-temperature phase. To ensure that the dynamics trajectory was mainly guided by experimental restraints, the relative weights of all force-field energy terms were reduced to 25% during the heating and high-temperature phases. Next, the system was slowly cooled to 300 K over 5 ps and then maintained at 300 K for an additional 12 ps. During the cooling and equilibrium phases, the weights of all force-field energy terms were restored to full scale, and k_{NOE} was slowly decreased to a final value of 25 kcal mol^{-1} A^{-2}. The coordinates of the last 4 ps were averaged and subjected to 500 cycles of restrained energy minimization to generate the final structures MD-A and MD-B. The 2-fold symmetry of the duplex was maintained by applying the noncrystallographic symmetry function of X-PLOR during this final minimization.

(3) Relaxation Matrix Refinement. In the final stage of structure determination, direct NOE refinements (Yip & Case, 1989) of structures MD-A and MD-B were carried out using the RELAX option of X-PLOR (Nilges et al., 1991). This option fits the coordinates of the input structures to the experimental volumes directly. Experimental volumes were taken from the 100- and 200-ms NOESY data sets. A total of 804 volumes were used in the control calculations, and 728 volumes were used in the araC calculations. All volumes were given a relative error of 15% based on the average difference between volumes for selected symmetry-related cross peaks. The theoretical volumes are calculated from a total relaxation matrix of the coordinate set. For these calculations, a uniform isotropic correlation time of 4.25 ns was used. With restraints of k_{NOE}(exp) = 0, k_{NOE} (hbond, grv) = 25, k_{CDIH} (backbone) = 0, and k_{CDIH} (sugar) = 50, MD-A and MD-B were first subjected to 50 steps of energy minimization with k_{relax} = 100. Minimized structures were further refined by dynamical simulated annealing (Nilges et al., 1988a,b; Holak et al., 1988) with k_{relax} = 400. Dynamical simulated annealing differs from restrained molecular dynamics in that all nonbonded interactions terms of the empirical energy function (i.e., Lennard-Jones, van der Waals, hydrogen-bonding, and electrostatic potentials) are replaced by a single hard-sphere repulsion term. In this procedure, the initial velocities were chosen from a Maxwellian distribution at 1000 K. The temperature of the system was then reduced by 25 K for another 30 steps of dynamics. This was repeated until a final temperature of 75 K was reached. During cooling, k_{NOE} (hbond, grv) was gradually decreased from 250 to 25 kcal/mol. The final step consisted of 150–200 steps of energy minimization with k_{relax} = 400 and the full nonbonded interaction terms included in the calculations. The 2-fold symmetry of the duplex was maintained by applying the noncrystallographic symmetry function of X-PLOR during this final minimization. *R* factors were calculated as the weighted average of the absolute value of the difference between the 1/6th power of the observed and calculated intensities (James, 1991; Thomas et al., 1991; Brünger, 1993).

Analysis of Helical Parameters. Analyses of helical parameters were performed with the program Dials-and-Windows (kindly provided by Dr. G. Ravishanker, Wesleyan University) (Ravishanker et al., 1989).

RESULTS

Resonance Assignments. The assignments of the nonexchangeable base protons and the H1′, H2′, and H2″ protons for the control dodecamer [d(CGCGAATTCGCG)]$_2$ and the araC dodecamer [d(CGCGAATT)(araC)(GCG)]$_2$ were made from NOESY and COSY spectra in a sequential manner (Hare et al., 1983; Scheek et al., 1983; Feigon et al., 1983). The sequential connectivity patterns for the control dodecamer were very similar to those given in Hare et al. (1983) (data not shown). For the araC dodecamer, assignments were made in a likewise manner by tracing connectivities between base protons (purine H8 and pyrimidine H6) and their own (*n*) and 5′-flanking (*n* – 1) sugar H1′ protons from the 5′ to the 3′ end of both duplexes; a similar connectivity tracing was carried out between base protons and their own (*n*) and 5′-flanking (*n* – 1) sugar H2′ and H2″ protons (supplementary material). The remaining H3′, H4′, and H5′/H5″ resonances are difficult to assign unambiguously from homonuclear experiments due to spectral overlap and crowding. These resonances as well as ^{31}P resonances were unambiguously assigned using a recently developed procedure involving two- and three-dimensional ^{31}P–^{1}H heteroTOCSY-NOESY experiments (Kellog & Schweitzer, 1993). In this procedure, H3′ and ^{31}P resonances are initially assigned in a sequential manner from 2-D heteroTOCSY and 2-D heteroTOCSY-NOESY experiments. A portion of the araC 2-D heteroTOCSY-NOESY spectrum is shown in Figure 2A; the control spectra are given in Kellogg and Schweitzer (1993). These assignments were confirmed, and the remaining backbone proton assignments were made using a 3-D ^{31}P–^{1}H heteroTOCSY-NOESY experiment. A representative slice from the 3-D heteroTOCSY-NOESY spectrum for the araC dodecamer is shown in Figure 2B. Assignment of exchangeable protons were made by analysis of the 200-ms NOESY spectrum acquired in 90% H_2O as previously described by Boelens et al. (1985); adenine H2 resonances were assigned from the cross peaks with the imino protons of complementary thymines in this spectrum (data not shown).

The ^{1}H resonance assignments for the araC dodecamer is given in Table 1. Corresponding assignments for the control dodecamer were not significantly different from those previously published (Hare et al., 1983; Kellogg & Schweitzer, 1993). Not including the araC residue, significant (≥0.1 ppm) differences in chemical shifts between the two DNAs exist at five resonances. Three of these differences occurred either at residues adjacent to [T(8) and G(10)] or opposite from [G(4)] the araC substitution. The largest difference is the 0.18 ppm downfield shift of the H1′ proton of G(4) in the araC dodecamer. Other shifts in the araC spectrum include a 0.17 ppm downfield shift of the H2″ resonance of T(8), a 0.15 ppm upfield shift of the H5′/H5″ resonances of C(1), a 0.11 upfield shift of the H2′ resonance of A(5), and a 0.1 ppm downfield shift of the H1 resonance of G(10). ^{31}P assignments for the araC dodecamer were obtained from 2-D heteroTOCSY and 3-D heteroTOCSY-NOESY experiments and are given in Table 2; the ^{31}P assignments for the control dodecamer in Table 2 were taken from Kellogg and Schweitzer (1993). Three out of four of the significant (≥0.1 ppm) ^{31}P chemical shift changes (Table 2) occur at or near the araC lesion site. The greatest change in phosphorus chemical shift is a 0.4 ppm

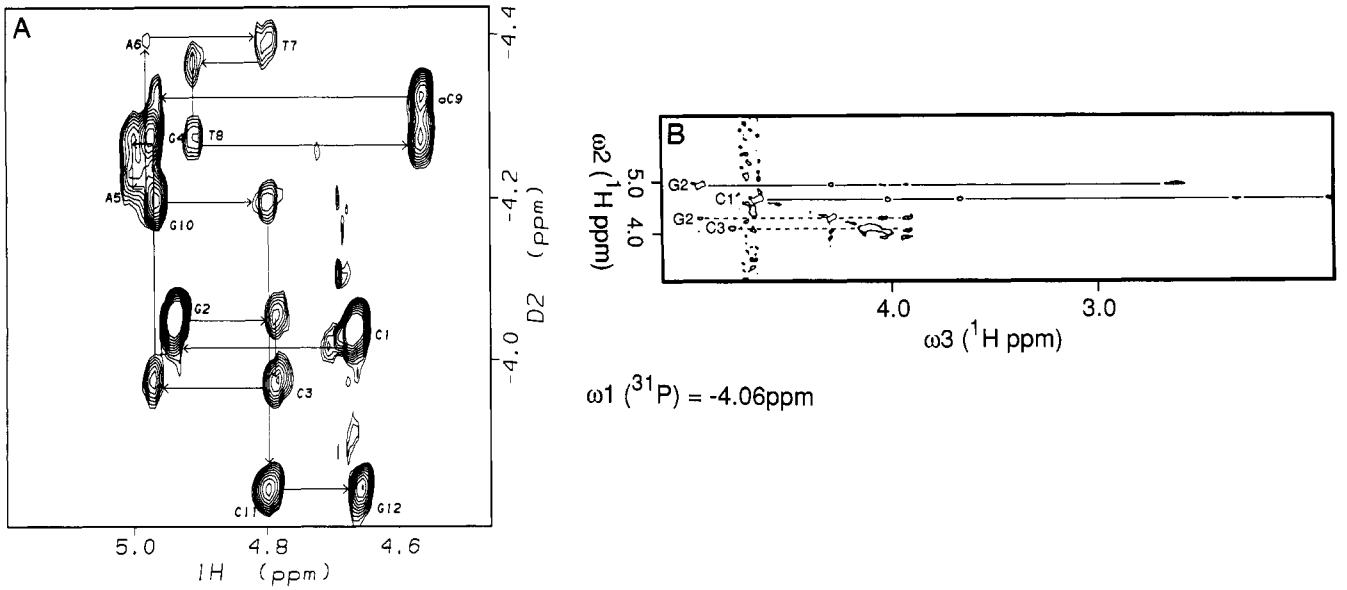


FIGURE 2: Phosphorus and backbone proton assignments in the araC duplex [d(CGCGAATT)(araC)d(GCG)]$_2$ using multidimensional heteroTOCSY-NOESY experiments. (A) Expanded H3′ region of the 2-D ^{31}P–^{1}H heteroTOCSY-NOESY spectrum of the araC duplex [d(CGCGAATT)(araC)d(GCG)]$_2$. The experiment was carried out as described in Materials and Methods and in Kellogg and Schweitzer (1993). The sequential connectivities are traced with a solid line. Each of the intraresidue ^{31}P–H3′ cross peaks is labeled with the residue number. (B) Cross-sectional plane perpendicular to ω1 (^{31}P) = –4.06 ppm in the 3-D ^{31}P–^{1}H heteroTOCSY-NOESY spectrum of the araC duplex [d(CGCGAATT)(araC)d(GCG)]$_2$. The experiment was carried out as described in Materials and Methods and in Kellogg and Schweitzer (1993). Solid horizontal lines connecting a set of deoxyribose correlations are drawn through the H3′ (ω2) frequencies of C1 and G2 and dashed horizontal lines through the H4′ frequencies of G2 and C3.

Table 1: Proton Chemical Shift Assignments for araC Dodecamer

residue	H8	H6	H5/methyl	H2	H1′	H2′	H2″	H3′	H4′	H5′/H5″ [a]	imino	amino
C1		7.61	5.90		5.76	1.91	2.37	4.67	4.03	3.70		na[b]
G2	7.93				5.84	2.66	2.75	4.95	4.33	4.07, 3.95	13.05	na
C3		7.29	5.37		5.62	1.90	2.34	4.80	4.14	4.12		8.42, 6.49
G4	7.85				5.63	2.64	2.78	4.99	4.32	4.10, 4.02	12.68	na
A5	7.99			7.19	5.94	2.58	2.86	5.01	4.43	4.20		na
A6	8.10			7.65	6.16	2.56	2.91	5.00	4.44	4.23		6.93, 5.89
T7		7.09	1.26		5.91	1.98	2.58	4.81	4.20	4.21	13.69	
T8		7.38	1.54		6.17	2.20	2.71	4.92	4.24	4.11	13.65	
araC9		7.46	5.50		5.61		4.13	4.58	4.03	4.34		8.45, 6.73
G10	7.87				5.91	2.66	2.78	4.97	4.36	4.15, 3.96	12.96	na
C11		7.31	5.42		5.74	1.89	2.31	4.80	4.12	4.09		8.50, 6.52
G12	7.93				6.16	2.61	2.37	4.67	4.17	4.06	na	na

[a] The H5′ and H5″ protons are not stereospecifically assigned. [b] na = not assigned due to broadening at 303 K.

Table 2: Phosphorus Chemical Shifts[a] in Control and araC Dodecamers

position	dinucleoside	control (ppm)	araC (ppm)
1	CpG	–4.04	–4.04
2	GpC	–4.16	–4.06
3	CpG	–3.98	–3.98
4	GpA	–4.11	–4.24
5	ApA	–4.24	–4.28
6	ApT	–4.37	–4.39
7	TpT	–4.37	–4.37
8	TpC	–4.23	–4.28
9	CpG	–3.90	–4.33
10	GpC	–4.04	–4.19
11	CpG	–3.91	–3.84

[a] Phosphorus chemical shifts are referenced to external trimethyl phosphate at 303 K.

upfield chemical shift at the phosphorus between araC(9) and G(10). Other significant shifts include a 0.15 ppm upfield shift at G(10), a 0.13 ppm upfield shift at G(4), and a 0.1 ppm downfield shift at G(2).

Sugar Torsion Angles. Sugar puckers were estimated using the procedure described by Kim et al. (1992). This approach involves the quantitative measurements of $J_{1'-2'}$ and $J_{1'-2''}$ via a high-resolution E-COSY experiment and qualitative estimation of $J_{2''-3'}$ and $J_{3'-4'}$ from either E-COSY or high-resolution DQF-COSY spectra. Additional constraints on the pseudorotation angle of the sugar ring can be obtained from NOESY-determined distances involving sugar protons, especially the H1′–H4′ distance that is highly dependent on the sugar pucker (Wüthrich, 1986; Hosur et al., 1986; Kim et al., 1992). While the individual values for the coupling constants and interproton distances may not be exactly correct, taken together they restrict the sugar conformation to a fairly narrow range. The J coupling measurements from the E-COSY spectrum of the control and araC dodecamers (supplementary material) indicate that in every residue except araC (which of course lacks the H2′ proton) $J_{1'-2'}$ is greater than $J_{1'-2''}$, thus limiting all the deoxyribose pseudorotation angles to 90–190° (Hosur et al., 1986, Kim et al., 1992).

While the values of $J_{2''-3'}$ and $J_{3'-4'}$ are difficult to determine accurately due to passive couplings, the intensity of these COSY cross peaks are directly related to the magnitude of the active J couplings (Hosur et al., 1986, Kim et al., 1992). From the equations and data presented in van Wijk et al. (1992), the relationship between the pseudorotation angle and these J values were calculated for the deoxyribose and 2′

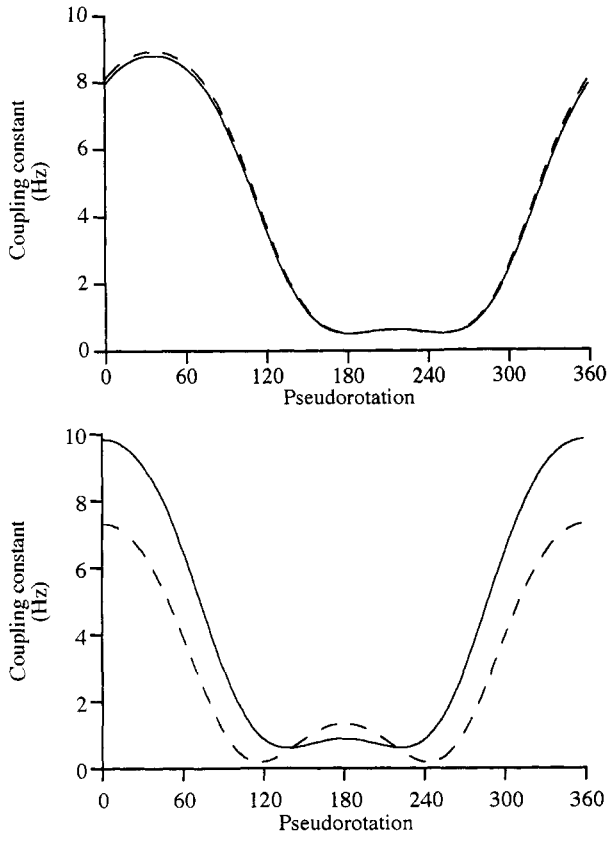


FIGURE 3: Three-bond vicinal J coupling constants in the deoxyribose (solid line) and arabinose sugars (dashed line) as a function of the pseudorotational phase angle P. (A, top) The generalized Karplus equation (van Wijk et al., 1992) was used to generate $J_{H3'-H4'}$ as a function of P. The coefficients for this equation were calculated from the data in van Wijk et al. (1992) and were identical for the deoxyribose and arabinose sugars. These coefficients were as follows: $C_0 = 4.294$; $C_1 = -1.06$; $C_2 = 3.767$; $C_3 = -0.54$; $S_2 = 0$. (B, bottom) The generalized Karplus equation (van Wijk et al., 1992) was used to generate $J_{H2''-H3'}$ as a function of P. The coefficients for this equation were calculated from the data in van Wijk et al. (1992). These coefficients were as follows: $C_0 = 5.237$; $C_1 = -1.06$; $C_2 = 4.566$; $C_3 = -1.06$; $C_2 = 3.704$; $C_3 = -0.54$; $S_2 = 1.548$ for the arabinose sugar.

arabinose sugar and are shown in Figure 3. In the H3′–H4′ region of a DQF-COSY spectrum (supplementary material) or E-COSY (both with and without ^{31}P decoupling) spectra (not shown) of the araC duplex, C(1), araC(9), and G(12) exhibit very strong cross peaks with the araC(9) being the strongest. It can be deduced from Figure 3A that the pseudorotation angles of these residues are limited to less than 120° or greater than 300°. Furthermore, only C(1), araC(9), and G(12) exhibit H2″–H3′ cross peaks in the araC dodecamer DQF-COSY spectrum (not shown), with C(1) and G(12) having cross peaks of weak intensity and araC(9) having a cross peak of weak to medium intensity. The existence of these cross peaks suggest a pseudorotation angle for C(1) and G(12) of greater than 60° and less than 100°, while according to Figure 3B, the araC(9) sugar should have a value of P less than 70°. For C(1) and G(12), pseudorotation angles greater than 200°, which have not be seen in nucleic acids but cannot be ruled out from Figure 3A, can be eliminated by the relative magnitude of $J_{1'-2'}$ and $J_{1'-2''}$. For araC(9) and G(12), a pseudorotation angle greater than 180° can also be ruled out from the H1′–H4′ interproton distances calculated from the NOESY data; this distance is greater than 3 Å for pseudorotation angles greater than 180° (Wüthrich, 1986). NOESY buildup calculations as well as the appearance of H1′–H4′ NOEs for araC(9) and G(12) even at short mixing times strongly suggest that the H1′–H4′ distance is less than 3 Å for both araC(9) and G(12). In summary, the NOESY and COSY data limit the value of P for C(1) and G(12) to the O4′-endo range (72–100°). The sugar pucker of araC(9) was given a broader range of 0–70°.

In the case of the control dodecamer, only C(1) and G(12) exhibit H2″–H3′ cross peaks (not shown) and strong H3′–H4′ cross peaks (supplementary material), thus placing their pseudorotation angles in the O4′-endo range. In both the control and araC duplexes, G(2), G(4), and G(10) have weak but detectable H3′–H4′ couplings in combination with nonexistent H2″–H3′ couplings, constraining P for these residues in the C2′-endo range (140–162°). Residues A(5), A(6), T(7), and T(8) in both the control and araC duplexes have medium intensity H3′–H4′ COSY cross peaks and nonexistent H2″–H3′ couplings; therefore, the pseudorotation angles of these residues should cover the C1′-exo sugar pucker region (P = 120–162°). Residues C(3) and C(11) in both the control and araC duplexes have strong H3′–H4′ COSY cross peaks and nonexistent H2″–H3′ cross peaks; the pseudorotation angles of these residues can therefore be constrained to a slightly lower range of 110–144°. These values for the pseudorotation angle ranges were converted into ranges for the individual sugar torsion angles, ν, using the relationship $\nu_j = T_m \cos[P + 144(j - 2)]$ where j is 0–4 (Altona & Sundaralingam, 1980). These torsion angles were included in the restrained molecular dynamics calculations described below.

Restrained Molecular Dynamics. The right-handed B-form structure of the araC dodecamer duplex was evident from the characteristic pattern of NOESY cross peaks. In addition, interstrand imino–imino and imino–H2 proton NOEs observed in NOESY spectra acquired in H_2O indicate that at least the internal 10 bases are stacked in both helices. Furthermore, interstrand imino proton and H2 proton to amino proton NOEs demonstrate Watson–Crick base pairing patterns. Qualitatively, therefore, introduction of an araC residue into the dodecamer sequence does not result in gross distortion of a right-handed, B-type DNA double helix. In order to obtain a more quantitiative comparison of the control and araC dodecamer structures, restrained molecular dynamics simulations were performed using experimentally determined interproton distance constraints and sugar dihedral constraints as well as the other restraints described above which preserve the right-handed nature and base pairing of the double helix. A total of 278 experimentally determined distances from the control DNA and 248 for the araC DNA were included in the restrained molecular dynamics calculations described below. At this stage of the calculations, an effort was made to eliminate those distances from interproton interactions where contamination from spin diffusion were most likely [e.g., H3′(i)–H8/H6(i+1)]. For both the control and araC calculations, three calculations were carried out from two starting structures, one in a canonical A-type conformation (designated Ini-A) and one in a canonical B-type conformation (designated Ini-B), by using different random number seeds for the assignments of the initial velocities. The two starting structures for the control simulations, Ini-A_{con} and Ini-B_{con} have an atomic rms deviation of 6.75 Å relative to each other. In the case of the control dodecamer, the three structures derived from Ini-A_{con} and Ini-B_{con} had atomic rms differences of 0.62 ± 0.27 and 0.66 ± 0.15 Å, respectively, and all six structures had an atomic rms difference of 0.58 ± 0.20 Å. The best fit superposition of these structures is shown in Figure 4A. In

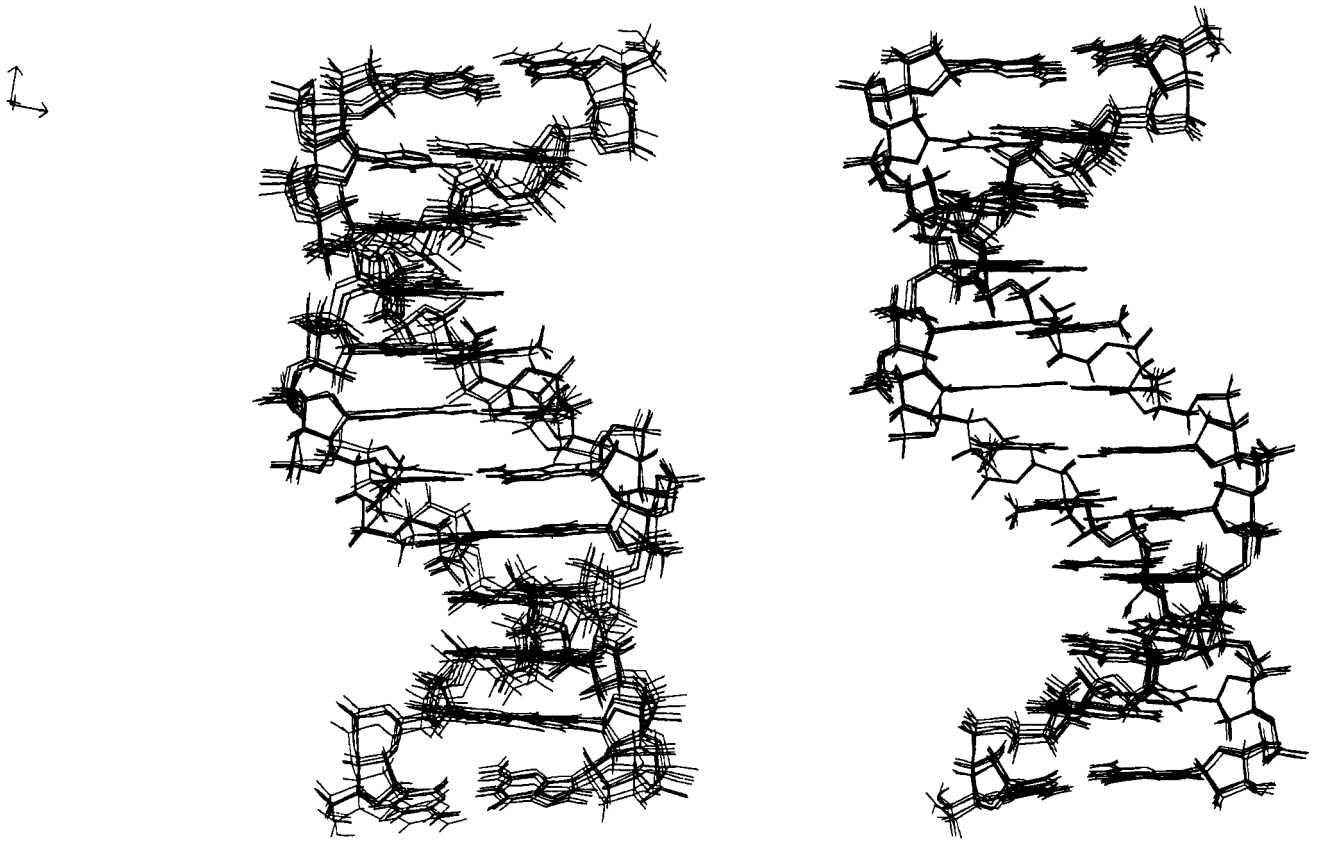


FIGURE 4: Best fit superposition of the three rMD structures generated from a canonical A starting structure and the three rMD structures generated from a canonical B starting structure. (A, left) View of the major groove along the helix axis of the six rMD_{con} structures. (B, right) View of the major groove along the helix axis of the six rMD_{ara} structures.

the case of the araC dodecamer, the three structures derived from Ini-A_{ara} and Ini-B_{ara} had atomic rms differences of 0.47 ± 0.05 and 0.45 ± 0.11 Å, respectively, and all six structures had an atomic rms difference of 0.48 ± 0.10 Å. The best fit superposition of these structures is shown in Figure 4B. It is obvious from these figures that both sets of calculations have converged to a B-type structure as predicted from the qualitative observations made above. Calculations in which the experimental distance were omitted did not converge (RMSD > 3 Å), demonstrating that the experimental restraints and not the empirical force field is driving this convergence.

The low RMSD between the six structures calculated by restrained molecular dynamics from the control data set indicated that convergence to essentially a single structure had been obtained; consequently, the coordinates of these six structures were averaged and subjected to 500 steps of restrained energy minimization to generate the structure rMD_{con}. The same procedure was followed to generate the structure rMD_{ara} from the six structures derived from the araC data set which had also demonstrated convergence to a single structure. Each of the resulting structures is clearly B-type in overall structure; both structures have a greater than 4.6-Å RMSD from canonical A-type DNA but less than a 3.2 Å RMSD from canonical B-type DNA (Table 3). In the case of the control dodecamer, energy-minimized starting structures, Ini-A and Ini-B, had values for the experimental distance energy of 508 and 326 kcal/mol, respectively, indicating significant violations of the experimentally determined distance constraints. Similarly, the violations of the experimentally determined dihedral restraints were also high, 2108 and 809 kcal/mol for Ini-A and Ini-B, respectively. After restrained molecular dynamics, the distance and dihedral energies were reduced to 6 and 86 kcal/mol, respectively. In the case of the araC dodecamer, the energy-minimized starting structures, Ini-A and Ini-B, had values for the experimental distance energy of 902 and 437 kcal/mol, respectively, while the violations of the experimentally determined dihedral restraints were 4733 and 2265 kcal/mol, respectively. The distance and dihedral energies for the araC structure obtained by restrained molecular dynamic were 9 and 54 kcal/mol, respectively.

Table 3: Atomic RMS Differences (Å) and *R* Factors

	rmsd (Å)							
	Ini-B_{min}[a]		rMD[b]		RR[c]		*R* factor[d]	
	control	araC	control	araC	control	araC	control	araC
Ini-A_{min}	6.75	6.75	4.66	5.12	4.69	5.10	0.076	0.079
Ini-B_{min}			3.10	3.02	3.17	3.08	0.073	0.071
rMD					0.59	0.67	0.061	0.060
RR							0.043	0.046

[a] Ini-A_{min} and Ini-B_{min}, energy-minimized canonical A- and B-type DNA dodecamer duplexes. [b] rMD, restrained energy-minimized average of six structures calculated by restrained molecular dynamics. [c] RR, relaxation matrix-refined structure. [d] *R* factor, 1/6th weighted residual as described in text.

Full Relaxation Matrix Refinement. In the final stage of the structural refinement protocol, spin diffusion was taken into account by refining back-calculated, full relaxation matrix-derived NOESY intensities directly against the experimental data. For this refinement, a repel energy function was used for dynamical simulated annealing, followed by energy minimization using Lennard-Jones potentials. This procedure gave structures with lower *R* factors than those derived from relaxation matrix refinement with 2 ps of restrained molecular dynamics with Lennard-Jones potentials

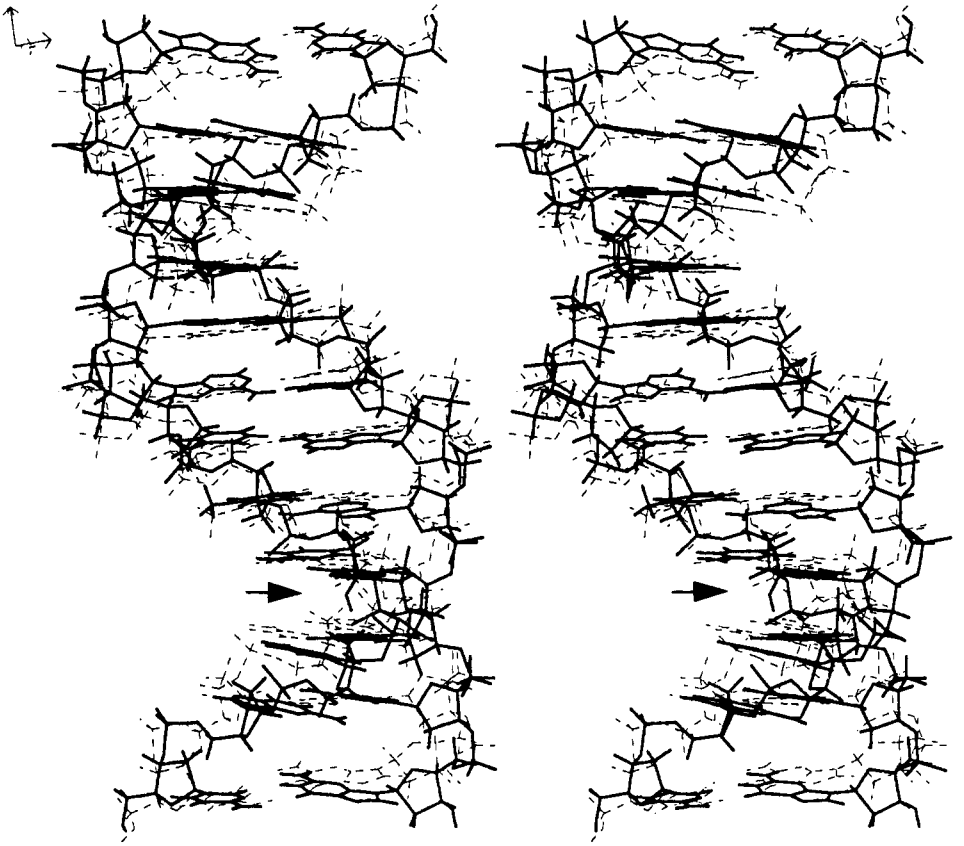


FIGURE 5: Stereoview of the best fit superposition of the relaxation matrix-refined structures for the control and araC dodecamers looking into the major groove. RR_{ara} is shown in thick lines; RR_{con} is shown in thin dashed lines. The arrows are pointing to the 2′ hydroxyl group on araC(9).

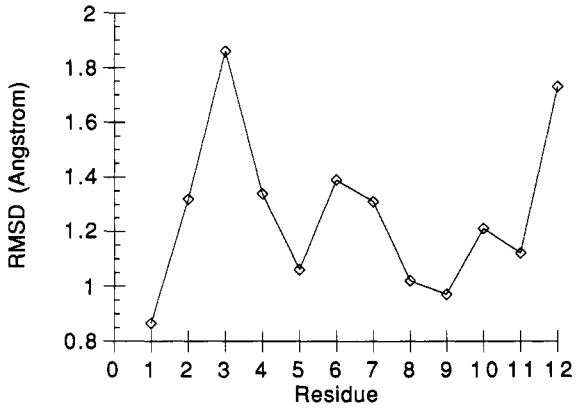


FIGURE 6: Root mean square difference (RMSD) between RR_{con} and RR_{ara} as a function of residue number.

at 300 K yet required far less computation time. Preliminary calculations indicated that taking fast local motions into account as described by Koning et al. (1991) did not result in improved *R* factors; consequently, an isotropic correlation time of 4.25 ns was calculated using a grid search to obtain a minimum value for the *R* factor, and this value was used in all relaxation matrix refinements. rMD_{con} and rMD_{ara} were each used as input structures for this refinement, yielding the final structures RR_{con} and RR_{ara}. Inspection of Table 3 shows that relaxation matrix refinement has resulted in structural shifts of 0.6–0.7 Å from the input structures. The best fit superposition of RR_{con} and RR_{ara} (Figure 5) reveals a global similarity between the two structures; incorporation of araC results in an overall structural change of 1.31 Å from the control. A residue by residue comparison (Figure 6) reveals that the greatest difference (1.8 Å) between these two structures is at residue 3, the residue immediately 5′ to the residue base paired with araC(9). The rms difference between the other residues of the two structures ranges from 0.8 to 1.4 Å. Interestingly, araC(9) and the corresponding dC(9) in the control structure had one of the lowest RMSD values.

Table 3 gives the *R* factors for the Ini-A, Ini-B, rMD, and RR structures from the control and araC calculations. These *R* factors reflect the degree to which the NOESY volumes back-calculated from the particular structure fit the experimental volumes. Due to the 6th power distance dependence of NOE intensities, the more familiar "crystallographic" *R* factor is dominated by errors in the shorter distances; the 1/6-weighted function used to calculate the *R* factors in Table 3 allows greater contribution of longer range NOE interactions to the *R* factor (James, 1991). As can be seen in Table 3, both the control and araC experimental volumes are better fit by the canonical B-type structure than the canonical A-type structure. After restrained molecular dynamics, all four MD structures show a significant reduction in the *R* factor, indicating that these structures are in better agreement with the experimental NOESY volumes than the starting structures. Relaxation matrix refinement resulted in further reduction in the *R* factor (Table 3); furthermore, the final control and araC structures had similar *R* factors, indicating that both sets of calculations had resulted in structures that were in excellent agreement with the experimental data. Using the more "traditional" crystallographic residual index, the final control and araC structures have $R_{x\text{-}ray}$ factors of 0.13 and 0.10, respectively.

Structural Analysis. (1) Helicoidal Parameters. Analysis of helical parameters of the relaxation matrix-refined structures RR_{con} and RR_{ara} was carried out with the Curves (Lavery & Sklenar, 1988, 1989) module of the structural analysis

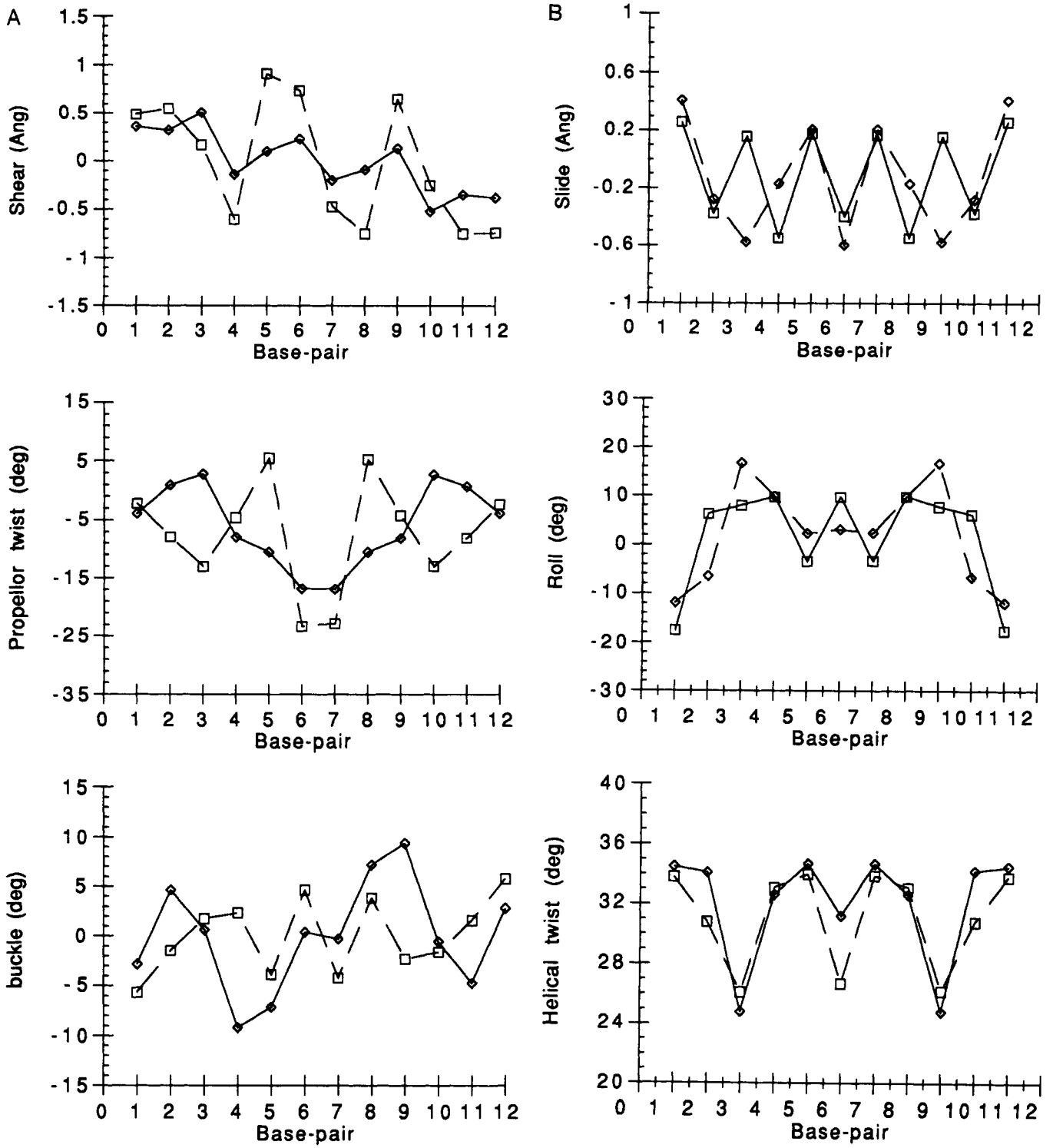


FIGURE 7: Comparison of helical parameters of RR_{con} and RR_{ara}. Parameter values were calculated using CURVES (Lavery & Sklenar, 1988, 1989). (A) Selected intrabase pair helical parameters as a function of base pair number for RR_{con} and RR_{ara}. (◇) RR_{con}; (□) RR_{ara}. (B) Selected interbase pair helical parameters as a function of base pair step for RR_{con} and RR_{ara}. (◇) RR_{con}; (□) RR_{ara}.

program Dials-and-Windows (Ravishanker et al., 1989). Among the *axis–base pair parameters*, *x*-axis displacements in both structures ranged from −1.52 to 2.0 Å, closer to that of canonical B-DNA (*x*-axis displacement = −0.7 Å) than to canonical A-DNA (*x*-axis displacement = −5.4 Å). The base pair inclination values for both structures were all negative as in B-type DNA (inclination = −5.9°). The largest deviation in this parameter from the canonical structure occurred in RR_{con} at base pair G(10)–C(15) and its symmetry-related base pair where the inclination was −10.9°; the value in RR_{ara} at the corresponding base pair was 5.2°.

The values for selected *intrabase pair parameters* as a function of base pair position are shown in Figure 7a. In RR_{ara}, base pairs 4 and 5, 6 and 7, and 8 and 9 exhibit rather large shears in opposite directions. With a positive shear indicating a shift toward the major groove and a negative shear indicating a shift toward the minor groove, these shear values indicate increased base-stacking interactions within these dinucleotide dimers. Shearing in the control structure is less pronounced, with the largest difference in shear values occurring at base pairs 3 and 4 and 9 and 10. Propeller twist values for RR_{con} and RR_{ara} show considerable deviation from

canonical B-DNA (propeller twist = 3.7°) and A-DNA (propeller twist = 13.7°) and are mostly negative. Interestingly, the largest differences in this parameter between the control and araC structures occur at the base pairs immediately 3′ and 5′ to the araC lesion site. For base pair buckle, however, the greatest difference between the two structures occurs at the araC site itself; in this case, base pairs 4 and 9 in RR_{con} have large negative and positive buckles, bending the center of these base pairs away from the center of the helix.

The values for selected *interbase pair parameters* as a function of base pair step are shown in Figure 7B. For both RR_{con} and RR_{ara}, the variations in roll and helical twist show a sequence dependence similar to the one found in the crystal structure of the Drew–Dickerson dodecamer and generally follow the trends predicted by Dickerson's (1983) sum functions. This suggests that these variations are driven by the need to relieve interstrand steric hindrance between purines at Pyr–Pur and Pur–Pyr steps (Callidine, 1982). RR_{ara} exhibits an alternating pattern of positive and negative slide; in the control structure, this pattern is disrupted at the C(3)–G(4) and C(9)–G(10) base pair steps. At this position, RR_{con} shows a slide value of −0.58 Å as compared with 0.16 Å for the araC structure. A relatively large negative slide like this at a Pyr–Pur step is usually not favored due to interstrand purine–purine clashes; however, this steric clash is relieved in the control structure by the large positive roll at these base pair steps. Positive roll opens the angle between base pairs toward the minor groove, resulting in a wider minor groove and bending toward the major groove. Thus, the positive roll at the 3–4 and 9–10 base pair steps and the positive roll values at the five other internal base steps in the control structure lead to a smooth bending of this helix towards the major groove. In the araC structure, all Pyr–Pur steps exhibit positive slide values, and all Pur–Pyr steps exhibit negative slide values. This relieves purine–purine clashes without the need for large positive rolls; consequently, the overall shape of RR_{ara} is straighter than RR_{con}.

Base Stacking. The stacking of the six unique base pair steps in the refined control and araC structures can be viewed in Figure 8. The stacking pattern observed in this figure is fairly typical of B-type double helices with a high degree of intrastrand base overlap at purine–pyrimidine steps and little intrastrand base overlap at pyrimidine–purine base steps. Purine–purine steps exhibit a modest degree of intrastrand base overlap, while pyrimidine–pyrimidine steps show virtually no intrastrand base overlap. Figure 8 also shows that the overall stacking patterns of the control and araC structures are very similar at every base pair step except the C(9)–G(10) step. In the control structure, there is no intrastrand base overlap at this step; however, there is a significant amount of interstrand overlap of the six-membered ring of the purine bases, similar to what is found in A-type DNA. This interstrand base overlap is almost completely abolished in the araC structure, resulting in a total lack of stacking interaction at this site. There are also some differences between the two structures at the C(11)–G(12) step, but this may be due to the lower number of restraints at the end of the helix.

Backbone Torsion Angles and Phosphate–Phosphate Distances. Values of the backbone torsion angles, α, β, γ, ϵ, and ζ as well as the glycosidic torsion angle χ and the pseudorotation angle P for RR_{con} and RR_{ara} are given in Table 4, Sections A and B, respectively. For the backbone torsion angles β, ϵ, and ζ, there is reasonably good agreement between the two structures. Exceptions to this pattern are angles β and ϵ of residue C(2) as well as the β torsion angle of residues C(9) and C(11). In addition to the agreement between the structures, these three torsion angles do not vary significantly with residue and are constrained to preferred orientations for right-handed DNA structures. These points are particularly noteworthy considering the fact that these angles were not directly constrained during relaxation matrix refinement. Backbone torsion angles α and γ, however, show considerably more variation from residue to residue as well as between the two structures. This variation appears to have a pattern insofar as a lower value of γ is accompanied by a higher value of α or vice versa. This strongly coupled "crankshaft" motion, where torsion about one of the sugar–phosphate bonds results in an equal but opposite rotation about the other bond has the net effect on keeping the sugar–phosphate backbone about the same length (Dickerson, 1983; Dickerson & Drew, 1981) and has been observed previously during molecular dynamics refinement (Kaluarachchi et al., 1991). Exceptions to this coupling of α and γ occur at araC(9), where a higher value of γ is not accompanied by a lower value of α, and at G(10), where a lower value of α in the araC structure is not accompanied by a higher value of γ. The uncoupling of these two torsion angles leads to changes in the interphosphate distances between residues 9 and 10 from 6.8 Å in the control structure to 6.1 Å in the araC structure and an apparently compensating change in the interphosphate distance between G(10) and C(11) from 6.3 Å in the control structure to 7.05 Å in the araC structure. The remaining interphosphate distances in the two structures do not exhibit differences >0.4 Å (supplementary material); thus, the overall lengths of the two helices are approximately equal.

Glycosidic Torsion Angles and Sugar Pucker. The glycosidic torsion angle χ was not directly restrained at any point during the calculations; nevertheless, the χ values in RR_{con} and RR_{ara} are quite similar except at residue 9, where the value of this torsion angle is 30° lower in RR_{ara} than in RR_{con}. The pseudorotation angles were constrained throughout the calculations with bounds ranging from ca. 30 to 70° instead of a rigid sugar conformation (Chary et al., 1988) or a dynamic equilibrium between C3′-endo and C2′-endo conformers (Rinkel & Altona, 1987). Solid-state ^2H NMR of highly solvated DNA duplexes have indicated that the deoxyribose sugar undergoes low amplitude motions (Kintanar et al., 1989; Huang et al., 1990; Alam et al., 1991), arguing against both the rigid and the two-state dynamic equilibrium models. Our laboratory has also shown that the magnetization transfer from phosphorus to protons in both the 3′ and 5′ directions that is observed in heteroTOCSY experiments is incompatible with pure C2′-endo or C3′-endo conformations (Kellogg & Schweitzer, 1993). Consequently, we feel that the use of fairly wide bounds for the sugar puckers in the calculations carried out in the present study represent a conservative approach. It is thus striking that the pseudorotation angles of the refined control and araC duplexes show little difference between the two structures except at residue 9 (Table 4). In the control structure, the pseudorotation angle at C(9) is lower than the average value of P for B-type DNA but is still in the C2′ endo family. The sugar of the araC(9) residue, on the other hand, is closer to the C3′-endo domain with a pseudorotation angle of 55. The only other significant difference in sugar pucker is seen at C(3), where the pseudorotation angle is 20° lower in the araC structure. The terminal residues C(1) and G(12) have sugar puckers in between that of A- and B-type DNA; this type of pucker at terminal residues is commonly observed in DNA solution structures. Finally, purine residues generally have a higher pseudorotation angle than pyrimidine residues,

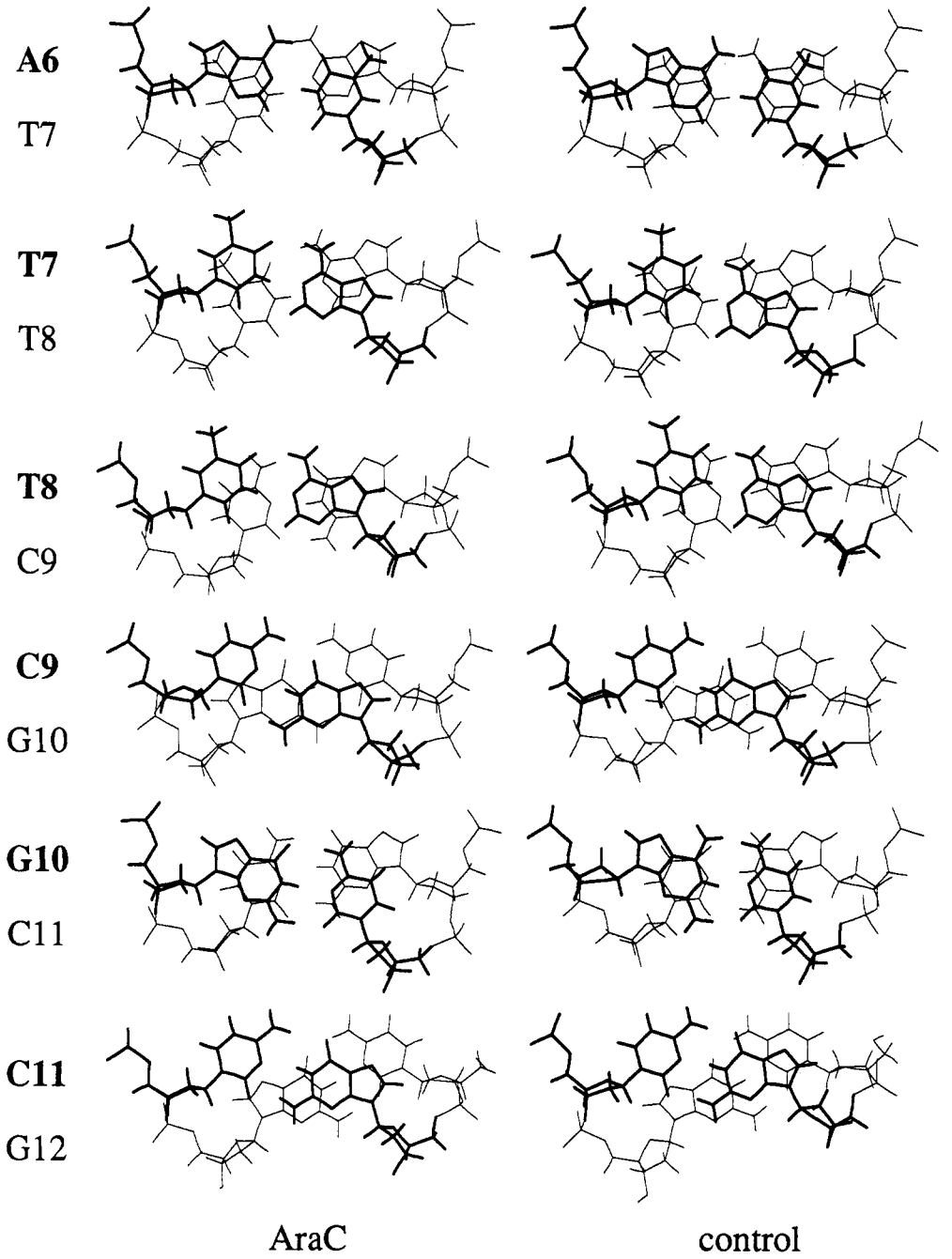


FIGURE 8: Base stacking in the RR_{con} (right) and RR_{ara} (left) structures. Base pair steps 6–11 are shown viewed down the global helix axis. Because of symmetry, base pair steps 1–5 are identical with base pair steps 7–11 and hence are not shown. The first base pair in the base step is shown in thick lines.

another common finding in solution structures of DNA.

Local Environment of the araC Lesion. It is clear from the above that most of the differences between the araC and control structures are located at or near the site of the araC lesion. These localized differences include changes in helical parameters, base stacking, and backbone, glycosidic, and sugar torsion angles. Further examination of the lesion site reveals some possible explanations for this myriad of structural perturbations. First, the O2′ oxygen of the arabinose sugar is less than 2 Å from the H8 proton of G(10); this close contact appears to be pushing the base of G(10) away from the araC-(9) base (Figure 9). This results in the loss of van der Waals interaction between the araC(9)-G and G(10)-C base pairs. Another close contact (<3 Å) is observed between the O2′ oxygen and the H6 proton of araC. A final observation is that the hydroxyl group of araC(9) is donating a hydrogen bond to either the O1P or O5′ oxygen of G(10). The distances between the araC(9) O2′ and the G(10) O5′ and O1P oxygens are both 3 Å, close enough for a hydrogen bond. The O2′ oxygen is also ~3 Å away from the N7 nitrogen of G(10), and as will be discussed further below, the possibility of a hydrogen bond with this acceptor group cannot be absolutely ruled out from the direct experimental data.

DISCUSSION

Previous Structural Studies on araC. araC has been the focus of a number of structural studies, both X-ray and NMR. An early X-ray crystallographic study on the araC mononucleoside indicated that the araC sugar possessed the C2′-endo conformation typical of canonical B-type DNA; interestingly, an intramolecular hydrogen bond between O2′H and O5′ was observed in the crystal which stabilized this sugar

Table 4: Backbone Torsion Angles,[a] Glycosidic Torsion Angle χ, and Pseudorotation Angle P of Relaxation Matrix-Refined Structures

residue	α	β	γ	δ	ϵ	ζ	χ	P
Section A: Control Structure RR_{con}								
C1	297	195	56	107	181	273	215	95
G2	282	202	80	138	143	267	258	151
C3	296	207	27	132	176	286	234	140
G4	287	191	62	134	167	262	254	144
A5	260	167	101	128	176	253	245	133
A6	338	183	12	129	192	260	229	136
T7	306	172	46	138	181	262	250	149
T8	289	186	55	125	176	266	246	129
C9	280	212	38	123	170	277	242	126
G10	301	187	62	136	163	263	251	147
C11	302	163	68	121	190	253	230	121
G12				100			219	82
mean	295	187	55	126	174	266	234	129
SD[b]	16	11	25	12	13	10	14	21
A-DNA[c]	285	208	45	83	178	313	206	13
B-DNA[d]	314	214	36	156	155	264	262	191
Section B: araC Structure RR_{ara}								
C1	308	178	68	107	183	267	216	96
G2	300	170	46	140	194	246	247	152
C3	303	197	29	119	186	275	240	120
G4	301	194	42	135	161	272	263	145
A5	321	165	56	125	173	242	243	129
A6	313	181	31	130	183	262	243	138
T7	281	185	62	132	180	261	247	139
T8	271	190	64	125	164	271	244	130
C9	285	183	79	89	181	280	212	55
G10	272	194	58	138	177	271	238	150
C11	291	195	66	122	175	259	245	125
G12				106			245	92
mean	295	185	55	122	178	264	240	123
SD[b]	16	11	16	15	10	12	14	29
A-DNA[c]	285	208	45	83	178	313	206	13
B-DNA[d]	314	214	36	156	155	264	262	191

[a] Backbone torsion angles are defined as $P(i)$–α–O5′–β–C5′–γ–C4′–δ–C3′–ϵ–O3′–ζ–$P(i+1)$. [b] SD, standard deviation. [c] Canonical A-DNA (Arnott & Hukins, 1972). [d] Canonical B-DNA (Arnott & Hukins, 1973).

conformation (Chwang & Sundaralingam, 1973). Several X-ray crystallography studies by Wang and co-workers have been carried out on DNA duplexes containing araC. In the first study, two hexamers which form left-handed Z structures, $[d(CG)r(CG)d(CG)]_2$ and $[d(CG)(araC)d(GCG)]_2$, were examined at 1.5-Å resolution (Teng et al., 1989). In both structures, the O2′ hydroxyl groups of the rC and araC residues formed intramolecular hydrogen bonds with the N2 amino group of the 5′ guanine residue; in addition, the O2′ hydroxyl group of araC formed hydrogen bonds with its own O5′ and O1P oxygens. The sugars of both the rC and araC residues adopted the C2′-endo conformational apparently to facilitate the formation of these hydrogen bonds. araC in a B-type DNA form was examined in a 1.6-Å resolution X-ray structure of the decamer $[d(CCAGGC)(araC)d(TGG)]_2$ (Gao et al., 1991). A comparison of this structure with the structure of the control duplex $[d(CCAGGCCTGG)]_2$ revealed a close similarity between the two (RMSD = 0.45 Å). The O2′ hydroxyl group of the araC residue was found to lie in the major groove of the helix in close contact with the C5 methyl and H6 atoms of its 3′ thymine neighbor. This close contact produced small localized changes in the helix, including a slight destacking between the two bases and an increase in buckle in the araC-dG base pair. The sugars of both araC and its counterpart dC in the control duplex adopted a C1′-exo (P = 111) conformation. Unlike the Z-DNA structure, the O2′ hydroxyl group of araC in the B-DNA structure did not form any intramolecular hydrogen bonds; however, it did appear to be hydrogen bonded to a water molecule, which in turn was hydrogen bonded to the 3′ and 5′ phosphate groups. This structure will be discussed in more detail below. In the most recent study, the structures of two hexamers, one with the sequence $[d(CG)(araC)d(GCG)]_2$ and formaldehyde cross-linked with daunorubicin and the other with the sequence $[d(CA)(araC)d(GTG)]_2$ and formaldehyde cross-linked with doxorubicin, were determined at high resolution (Zhang et al., 1992). In these B-type structures, the araC sugar again adopted a C1′-exo (P = 120) conformation with no intramolecular hydrogen bonding of its O2′ hydroxyl group. While there were numerous differences in these two hexamers from their corresponding control sequences, it is impossible to separate the effects of araC on these structures from those of the cross-linked anthracyclines.

The X-ray studies described above have provided useful information about the effects of araC substitutions on DNA structure. This information should be cautiously interpreted, however, since recent years have seen the accuracy of X-ray-determined nucleic acid structures called into question. A number of studies have shown disagreement between DNA structures derived from X-ray crystallography and those derived from NMR in solution (Rinkel et al., 1987; Sklenar & Bax, 1987; Joshua-Tor et al., 1988; Nikonowicz et al., 1990, Yang et al., 1993). One possible explanation for this discrepancy is that some features of X-ray-determined DNA structures are artifacts of crystal packing forces. For example, variations in local conformation have been found in different crystal forms of the same duplex (Jain & Sundaralingam, 1989; Shakked et al., 1989). Further, a number of DNA sequences which have been shown to crystallize as A-form DNA (Clark et al., 1990) have been found to have a B-form conformation in solution. Thus, in order for the effect of an araC substitution in a DNA duplex to be unambiguously established, a high-resolution NMR study is required.

While araC has been the focus of a number of NMR studies, no quantitative analysis of araC in a strictly double helical duplex form has been reported until now. In an early study, a coupling constant analysis by solution NMR on the araC mononucleoside demonstrated that the araC sugar adopted a conformation under neutral pH conditions that was 50% C2′-endo and 50% C3′-endo, similar to that for ribocytosine; furthermore, no evidence for the type of intramolecular hydrogen bonding observed in the X-ray structure of the mononucleoside was seen under neutral conditions in solution (Remin et al., 1976). Two more recent NMR studies have been carried out on the decamer $[d(CG)(araC)d(TAGCG)]_2$ (Pieters et al., 1989, 1990). In the first report, incorporation of araC was found to promote increased formation of a hairpin structure in which the central TA residues formed the loop of the hairpin (Pieters et al., 1989). Some formation of the duplex was observed, however, and a qualitative analysis of the data indicated that this duplex was B-type in structure. The only analysis of the B-form duplex given in this paper was a comparison of the chemical shifts of this duplex with the control sequence $[d(CGCTAGCG)]_2$ in duplex form, which revealed that the araC substitution produced significant chemical shift changes at the neighboring 5′ and 3′ residues. A more quantitative NMR and molecular dynamics analysis of the hairpin form of the araC-containing sequence showed it to have an unusual stacking arrangement in the loop region (Pieters et al., 1990). Based on a single coupling constant $J_{H1'-H2''}$, it was concluded that the araC sugar adopted a C2′-endo conformation; in this configuration, the O2′ hydroxyl group made intramolecular hydrogen bonds with the N2 amino group and N7 nitrogen of adenine in the loop. Unfortunately,

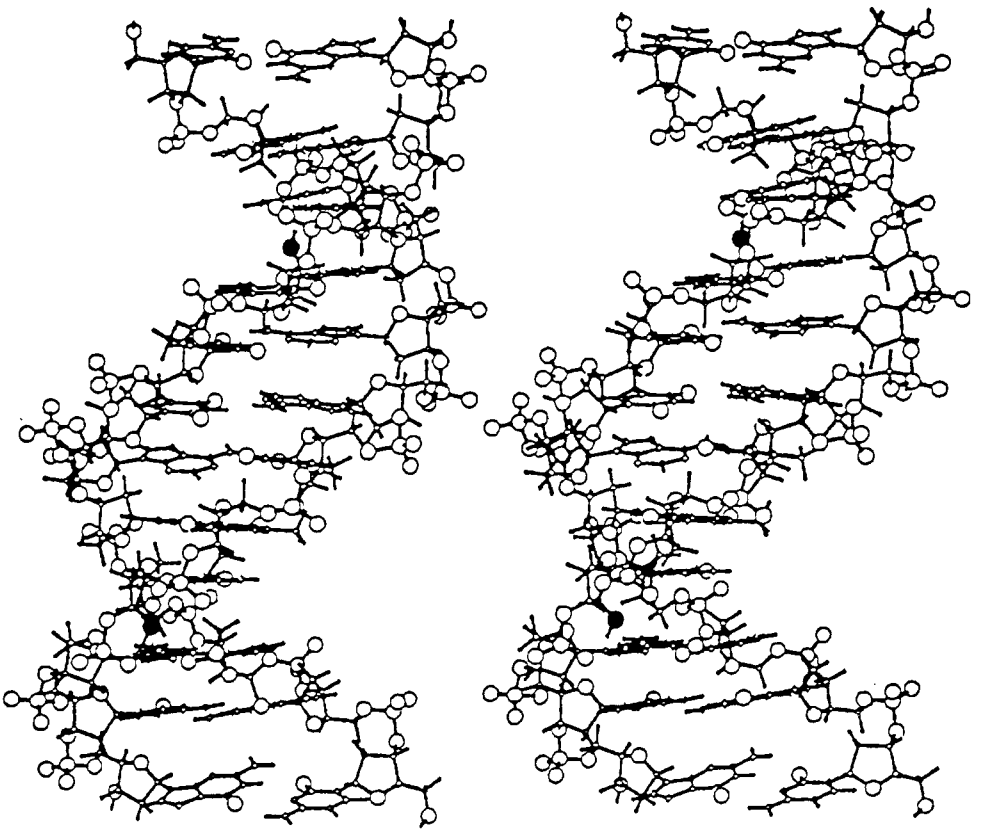


FIGURE 9: Stereoview ball-and-stick drawing of RR_{ara} looking into the minor groove. The O2′ oxygen of araC is shown as a filled circle.

no structure for the control hairpin was presented in this paper, making it impossible to determine the effect of the araC substitution on the overall structure.

Comparison of NMR-Determined Solution Structure of araC-Containing DNA with X-ray-Determined Crystal Structure. Before making a detailed comparison of the refined control and araC structures, it is necessary to assess the precision and accuracy of these structures. One way in which the precision of the structures can be gauged is by the rmsd among all the structures generated from a particular data set. The RMSDs for the six control and six araC structures were ≤0.6 Å, which is near the RMS fluctuations of the atoms about their average positions. This level of precision makes it possible for local structural variations in the structure to be fairly well determined (Gronenborn & Clore, 1989). As shown in Table 3, the *R* factors for the final structures are quite low, indicating that these structures are in very good agreement with the experimental NOE data. The low *R* factors also demonstrate that the level of precision has not been overestimated due to inappropriately restrictive bounds on the distance restraints.

The accuracy of the structures presented in this paper is somewhat harder to judge than the precision. Like most structures of DNA duplexes determined by NMR, the present structures were generated without long-range experimental restraints, explicit treatment of solvent, or a perfect force field. Additionally, some nonexperimental restraints (e.g., planarity, major groove distances) were used in the calculations. Finally, all NMR data represent time-averaged phenomena, yet this is not explicitly dealt with in our calculations. With these caveats in mind, the RMSD between structures derived from canonical A- or B-type DNA has sometimes been used as a measure of the "accuracy" of the NMR-derived model; a low RMSD suggests that sufficient conformational space has been sampled during the molecular dynamics and that convergence to the "true" structure has been obtained (Nilsson et al., 1986; Nilges et al., 1987a,b; Clore et al., 1988; Baleja et al., 1990). The RMSDs between structures derived from A- or B-type DNA in this study were <0.8 Å.

Previous work has also suggested that combining relaxation matrix refinement with restrained molecular dynamics as was done in the present study can help to more accurately determine helical parameters and backbone torsion angles even without any direct restraints on the backbone itself (Kaluarachchi et al., 1991). This approach has been shown to be superior to one utilizing IPSA-derived distances alone since spin diffusion is explicitly taken into account during the relaxation matrix refinement (Post et al., 1990). Increasing the number of NOE intensities used in the relaxation matrix refinement has also been shown to improve the likelihood of obtaining an "accurate" structure (Kaluarachchi et al., 1991). Furthermore, it has been demonstrated that using a larger number of constraints involving H5′/H5″ protons leads to better definition of the sugar–phosphate backbone portion of the structure (Kim et al., 1992). In the present study, our use of the 3-D heteroTOCSY-NOESY experiment allowed us to utilize a greater number of intensities from backbone proton NOEs than is usually reported with a duplex of this size. For example, out of a total of 728 intensities used to refine the rMD_{ara} structure, ca. 200 involved interactions with H4′, H5′, or H5″ protons, and 70 of these were sequential interactions.

Having considered the precision and accuracy of the solution structures derived in this study, it is now appropriate to discuss the effects of araC incorporation on the structure of the dodecamer and to compare these effects with those noted by Gao et al. (1991) in the crystal structure of an araC-containing B-type decamer. As in the crystal structure, incorporation of araC into the dodecamer did not produce a large change in the overall structure of the helix (RMSD = 1.3 Å). The most

significant change in the conformation of the araC residue itself observed in the crystal structure was a decrease in the glycosidic torsion angle from −114° to −125° (Gao et al., 1991). A similar change is observed in the solution structure, and in both cases, the change in glycosidic torsion angle may be due to close contact between O2′ on the araC sugar and the H6 of its own base. One change observed in the araC solution structure but not in the crystal structure is a decrease in the pseudorotation angle of the sugar at residue 9 from 125° in dC(9) of the control to 55° in the araC(9) sugar (Table 4). Modeling studies show that decreasing the pseudorotation angle brings the araC hydroxyl group closer to the 3′ phosphate backbone and thus may promote the formation of a hydrogen bond. This is reminiscent of the Z-RNA and araC-Z-DNA X-ray structures in which the ribose and arabinose sugars adopt *higher* pseudorotation angles to promote hydrogen bond formation (Teng et al., 1989). Additional NMR studies of araC in other sequence contexts will be necessary to determine whether the difference in the araC sugar pucker in the crystal and solution structures is due to sequence-specific or crystal packing effects.

In the crystal structure, the araC-G base pair has an increased buckle relative to the control decamer. This is not observed in the solution structure; in fact, the buckle of the araC(9)-G4 base pair is reduced relative to the control (Figure 7A). Again, this difference between the crystal and solution structures may be due to the different sequences of the DNAs used in each study. In the DNA used in the X-ray study, the methyl group of the thymine residue 3′ to the araC is in very close contact with the O2′ of araC. The adjustments needed to accommodate this contact create a "domino effect" of structural changes that are greatest at the araC base pair but also extend to the neighboring 3′ and 5′ base pairs (Gao et al., 1991). In the sequence used in the present study, there is a dG 3′ to the araC residue rather than a dT, but there is a similar close contact between the araC O2′ oxygen and the 3′ base (Figure 9). Like in the crystal structure, this close contact has the effect of pushing the 3′ base away from the araC residue. While this causes a loss of intrastrand stacking between the araC and dT bases in the crystal structure, it apparently also causes a loss of interstrand stacking between purine bases in the solution structure (Figure 8). The unusual A-like interstrand stacking observed at the C(9)–G(10) step in the control structure was also observed at the same step in the solution structure of the related dodecamer [d(CGCG-PATTCGCG)]$_2$ (P = nebularine) (Clore et al., 1988); indeed, pyrimidine–purine base pair steps often exhibit unusual helical parameters which are probably driven by the need to relieve interstrand steric hindrance between the purines (Calladine, 1982). Thus, this difference between the two araC-containing structures is most likely due to the different sequences employed or, more generally speaking, to changing from a pyrimidine–pyrimidine step to a pyrimidine–purine step.

Hydrogen Bonding in the araC Structure. As stated above, no intramolecular hydrogen bond involving the araC hydroxyl group was observed by Gao et al. (1991) in their crystal structure. It is not clear whether the different hydrogen bond arrangements observed in the crystal and solution structures are due to the different sequences, to an artifact of crystallization, or to the lack of explicit treatment of solvent in our *in vacuo* rMD simulations. No direct experimental constraints on the position of the araC hydroxyl group were used in the generation of the araC solution structure since the fast-exchange rate of hydroxyl protons usually precludes their observation by NMR even when they are involved in hydrogen bonds; consequently, the conformation of this hydroxyl group in the refined structure must have been indirectly determined by a combination of the empirical force field and the remaining experimental constraints. Thus, the actual position of the araC hydroxyl group in the araC dodecamer is subject to a certain degree of uncertainty. As stated above, there are several potential hydrogen bond acceptors within 3 Å of the O2′ oxygen; however, there are two pieces of indirect evidence which support a hydrogen bond between the 2′ hydroxyl group and the phosphate backbone. One possible indicator of this hydrogen bond is the significant upfield chemical shift of the ^{31}P resonance of the phosphate group 3′ to araC(9). Upfield ^{31}P chemical shifts have been previously shown to indicate possible changes in hydrogen bonding and/or solvation of phosphate diesters (Lerner et al., 1984). Furthermore, RR_{con} and RR_{ara} have virtually identical values for backbone torsion angles α and ζ at residue C(9), yet these two torsion angles appear to be the most important in determining ^{31}P chemical shifts (Gorenstein, 1984, 1987; Gorenstein et al., 1988; Roongta et al., 1990). This apparent paradox suggests that the ^{31}P chemical shift observed at this residue is not due to a change in backbone conformation. Work is in progress to determine $J_{H3'-P}$ values which will help to better define the backbone conformation. Additional support for an intramolecular hydrogen bond at this phosphate group comes from ^{31}P relaxation studies in which we have used modified heteroTOCSY pulse sequences to measure the T_1 and T_2 values for most of the ^{31}P resonances in the control and araC DNAs (Schweitzer & Kellogg, unpublished results). These studies suggest that the mobility of the phosphate 3′ to araC(9) may be reduced relative to the control; a hydrogen bond such as the one indicated in the araC structure at this site would be consistent with a reduction in mobility.

Biochemical Implications. As stated above, one motivation for studying the effect of araC incorporation on the structure of DNA was to better understand the structural underpinnings of the functional consequences that arise from araC misincorporation into DNA. These consequences include the severe inhibition of DNA polymerase-catalyzed primer extension when araC is the primer terminal residue (Mikita & Beardsley, 1988; Perrino & Mekosh, 1992), as well as interference with replication bypass when araC is in the template (Mikita & Beardsley, 1988). From their X-ray structure of araC incorporated into a B-type decamer duplex, Gao et al. (1991) suggested that the conformational perturbation at and near the araC site might interfere with polymerase function by decreasing binding to the enzyme. However, Perrino and Mekosh (1992) reported that araC at a 3′ primer terminus did not significantly reduce binding of a primer–template complex with human DNA polymerase α, and we have obtained the same result with the Klenow fragment of *E. coli* DNA polymerase I (Mikita & Beardsley, unpublished results). Furthermore, the structural perturbations accompanying araC misincorporation into DNA do not cause a large destabilization of the duplex; melting studies carried out on the duplex used in the present study indicated a reduction in melting temperature (T_m) of only 4 °C (Beardsley et al., 1988). In contrast, two dG-dT mispairs introduced into this same duplex drop the T_m by ~20 °C (Patel et al., 1982).

Of course, any extrapolation from a structure with araC locked into a duplex conformation to a possible structure of araC at a primer–template junction complexed with a polymerase should be made with a great deal of caution. However, a few observations are worth making. The loss of stacking or van der Waals interactions at the araC(9)–G(10)

step (Figures 8 and 9) could affect the stabilization of the incoming triphosphate when araC is in the template strand; this, in turn, would affect the rate of primer extension past this site, which is what we observe. The additional local helical distortions likely contribute to this lag in polymerase bypass as well. When araC is at the primer terminus, another factor in addition to the above is likely to be involved. Direct interference by the araC hydroxyl group almost certainly accounts for severe inhibition of DNA synthesis observed when arabinose is the primer terminal sugar (Mikita & Beardsley, 1988; Perrino & Mekosh, 1992). The orientation of the 2′ hydroxyl group as shown in our structure could interfere with polymerase-catalyzed addition in a number of ways, including altering the reactivity of the incoming triphosphate through hydrogen bonding, displacement of a critical solvent molecule, or steric interference with the incoming nucleotide or a catalytic group on the enzyme. Detailed pre-steady-state and steady-state experiments with DNA polymerase and araC are currently being undertaken in our laboratory to help distinguish these possibilities.

ACKNOWLEDGMENT

The authors wish to thank Dr. Suresh Srinivasta (ChemGenes, Inc.) for synthesis of the phosphoramidite derivative of araC, Mr. John Flory (Keck Biotechnology Resource Laboratory at Yale University School of Medicine) for synthesis of the araC-containing oligonucleotide, and Mr. Alex Szewczak for help in using X-PLOR.

SUPPLEMENTARY MATERIAL AVAILABLE

Two figures of expanded contour plots of the 200-ms mixing time NOESY of the araC duplex [d(CGCGAATT)(araC)-d(GCG)]$_2$, one showing sequential connectivities in the H8/H6–H1′ region, one showing sequential connectivities in the H8/H6–H2′/H2″ region; one figure showing the interphosphate distances for the refined control and araC structures; one table of $J_{H1\text{-}H2'}$ and $J_{H1\text{-}H2''}$ values determined from E-COSY experiments for the control and araC dodecamers; one table showing the backbone dihedral restraints; one table showing the base pair hydrogen bond distance restraints used in the restrained molecular dynamics (7 pp). Ordering information is given on any current masthead page.

REFERENCES

Alam, T. M., Orban, J., & Drobny, G. P. (1991) *Biochemistry 30*, 9229–9237.

Altona, C., & Sundaralingam, M. (1972) *J. Am. Chem. Soc. 94*, 8205–8212.

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

Arnott, S., & Hukins, D. W. (1973) *J. Mol. Biol. 81*, 93–105.

Baleja, J. D., Pon, R. T., & Sykes, B. D. (1990) *Biochemistry 29*, 4828–39.

Beardsley, G. P., Mikita, T., Klaus, M. M., & Nussbaum, A. L. (1988) *Nucleic Acids Res. 16*, 9165–9176.

Bodey, G., Freireich, E., Monto, R., & Hewlett, J. (1969) *Cancer Chemother. Rep. 53*, 59–66.

Boelens, R., Scheek, R. M., Dijkstra, K., & Kaptein, R. (1985) *J. Magn. Reson. 62*, 378–386.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 74*, 187–217.

Brünger, A. T. (1993) *X-PLOR 3.1: A System for Crystallography and NMR*, Yale University, New Haven, CT.

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

Chwang, A. K., & Sundaralingam, M. (1973) *Nature (London) 243*, 78–79.

Clark, G. R., Brown, D. G., Sanderson, M. R., Chwalinsky, T., Neidle, S., Veal, J. M., Jones, R. L., Wilson, W. D., Zon, G., Garman, E., & Stuart, D. I. (1990) *Nucleic Acids Res. 18*, 5521–5528.

Clore, G. M., Oschkinat, H., McLaughlin, L. W., Benseler, F., Scalfi Happ, C., Happ, E., & Gronenborn, A. M. (1988) *Biochemistry 27*, 4185–4197.

Coleman, C. N., Stoller, R. G., Drake, J. C., & Chabner, B. A. (1975) *Blood 46*, 791–803.

Collins, A. R. S. (1977) *Biochim. Biophys. Acta 478*, 461–473.

Cozzarelli, N. R. (1977) *Annu. Rev. Biochem. 46*, 641–684.

Dicioccio, R. A., & Srivastava, B. I. S. (1977) *Eur. J. Biochem. 79*, 411–418.

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

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

Fratini, A. V., Kopka, M. L., Drew, H. R., & Dickerson, R. E. (1982) *J. Biol. Chem. 257*, 14686–14707.

Frei, E., Bickers, J., Lane, M., Leary, W., & Talley, R. (1969) *Cancer Res. 29*, 1325–1352.

Fridland, A. (1977) *Biochemistry 16*, 5308–5312.

Furth, J. J., & Cohen, S. S. (1968) *Cancer Res. 28*, 2061–2067.

Gao, Y.-G., van der Marel, G. A., van Boom, J. H., & Wang, A. H.-J. (1991) *Biochemistry 30*, 9922–9931.

Gorenstein, D. G. (1984) *Phosphorus-31 NMR: Principles and Applications* (Gorenstein, D. G., Ed.) Academic Press, New York.

Gorenstein, D. G. (1987) *Chem. Rev. 87*, 1047–1077.

Gorenstein, D. G., Schroeder, S. A., Fu, J. M., Metz, J. T., Roongta, V. A., & Jones, C. R. (1988) *Biochemistry 27*, 7223–7237.

Graham, F., & Whitmore, C. (1970) *Cancer Res. 30*, 2636–2644.

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.

Gronenborn, A. M., & Clore, G. M. (1989) *Biochemistry 28*, 5978–5984.

Hande, K. R., & Chabner, B. A. (1978) *Cancer Res. 38*, 579–585.

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

Heintz, N. H., & Hamlin, J. L. (1983) *Biochemistry 22*, 3557–3562.

Holak, T. A., Engstrom, A., Kraulis, P. J., Lindeberg, G., Bennich, H., Jones, A., Gronenborn, A. M., & Clore, G. M. (1988) *Biochemistry 27*, 7620–7629.

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.

Huang, P., & Eisenberg, M. (1992) *Biochemistry 31*, 6518–6532.

Huang, W.-C., Orban, J., Kintanar, A., Reid, B. R., & Drobny, G. P. (1990) *J. Am. Chem. Soc. 112*, 9059–9068.

Jain, S., & Sundaralingam, M. (1989) *J. Biol. Chem. 264*, 12780–12784.

James, T. L. (1991) *Curr. Opin. Struct. Biol. 1*, 1042–53.

Joshua-Tor, L., Rabinovich, D., Hope, H., Frolow, F., Appella, E., & Sussman, J. L. (1988) *Nature 334*, 82–84.

Kaluarachchi, K., Meadows, R. P., & Gorenstein, D. G. (1991) *Biochemistry 30*, 8785–8797.

Kellogg, G. W. (1992) *J. Magn. Reson. 97*, 623–627.

Kellogg, G. W., & Schweitzer, B. I. (1993) *J. Biomol. NMR 3*, 577–595.

Kellogg, G. W., Szewczak, A. A., & Moore, P. B. (1992) *J. Am. Chem. Soc. 114*, 2727–2728.

Kim, S.-G., Lin, L.-J., & Reid, B. R. (1992) *Biochemistry 31*, 3564–3574.

Kintanar, A., Huang, W.-C., Schindele, D. C., Wemmer, D. E., & Drobny, G. P. (1989) *Biochemistry 28*, 282–293.
Koning, T. M. G., Boelens, R., van der Marel, G. A., van Boom, J. H., & Kaptein, R. (1991) *Biochemistry 30*, 3787–3797.
Kufe, D. W., Major, P. P., Egan, E. M., & Beardsley, G. P. (1980) *J. Biol. Chem. 255*, 8997–9000.
Lavery, R., & Sklenar, H. (1988) *J. Biomol. Struct. Dyn. 6*, 63–91.
Lavery, R., & Sklenar, H. (1989) *J. Biomol. Struct. Dyn. 6*, 655–667.
Lerner, D. B., Becktel, W. J., Everett, R., Goodman, M., & Kearns, D. R. (1984) *Biopolymers 23*, 2157–2172.
Major, P. P., Egan, E. M., Herrick, D., & Kufe, D. W. (1982) *Biochem. Pharmacol. 31*, 2937–2940.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Mikita, T., & Beardsley, G. P. (1988) *Biochemistry 27*, 4698–4705.
Mikita, T., & Beardsley, G. P. (1994) *Biochemistry 33*, 9195–9208.
Momparler, R. (1972) *Mol. Pharmacol. 8*, 362–370.
Mujeeb, A., Kerwin, S. M., Kenyon, G. L., & James, T. L. (1993) *Biochemistry 32*, 13419–13431.
Nerdal, W., Hare, D. R., & Reid, B. R. (1989) *Biochemistry 28*, 10008–10021.
Nikonowicz, E. P., Meadows, R. P., & Gorenstein, D. G. (1990) *Biochemistry 29*, 4193–4204.
Nilges, M., Clore, G. M., & Gronenborn, A. M., Brünger, A. T., Karplus, M., & Nilsson, L. (1987a) *Biochemistry 26*, 3718–3733.
Nilges, M., Clore, G. M., Gronenborn, A. M., Piel, N., & McLaughlin, L. W. (1987b) *Biochemistry 26*, 3734–3744.
Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988a) *FEBS Lett. 229*, 317–324.
Nilges, M., Gronenborn, A. M., Brünger, A. T., & Clore, G. M. (1988b) *Protein Eng. 2*, 27–38.
Nilges, M., Hazabettl, J., Brünger, A. T., & Holak, T. A. (1991) *J. Mol. Biol. 219*, 499–510.
Nilsson, L., Clore, G. M., Gronenborn, A. M., Brünger, A. T., & Karplus, M. (1986) *J. Mol. Biol. 188*, 455–475.
Ott, J., & Eckstein, F. (1985) *Biochemistry 24*, 2530–2535.
Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Dallas, J., Itakura, K., & Breslauer, K. J. (1982) *Biochemistry 21*, 437–444.
Patel, D. J., Kozlowski, S. A., & Bhatt, S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3908–3912.
Perrino, F. W., & Mekosh, H. L. (1992) *J. Biol. Chem. 267*, 23043–23051.
Piantini, U., Sorensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc. 104*, 6800–6801.
Pieters, J. M. L., de Vroom, E., van der Marel, G. A., van Boom, J. H., & Altona, C. (1989) *Eur. J. Biochem. 184*, 415–425.
Pieters, J. M. L., de Vroom, E., van der Marel, G. A., van Boom, J. H., Koning, T. M. G., Kaptein, R., & Altona, C. (1990) *Biochemistry 29*, 788–799.
Post, C. B., Meadows, R. P., & Gorenstein, D. G. (1990) *J. Biomol. Struct. Dyn. 8*, 253–294.
Ravishanker, R., Swaminathan, S., Beveridge, R. L., & Sklenar, H. (1989) *J. Biomol. Struct. Dyn. 6*, 669–699.
Remin, M., Darzynkiewicz, E., Ekiel, I., & Shugar, D. (1976) *Biochim. Biophys. Acta 435*, 405–416.
Rinkel, L. J., van der Marel, G. A., van Boom, J. H., & Altona, C. (1987) *Eur. J. Biochem. 163*, 275–286.
Roongta, V. A., Jones, C. R., & Gorenstein, D. G. (1990) *Biochemistry 29*, 5245–5258.
Rosenberg, J. M., Seeman, N. C., Day, R. O., & Rich, A. (1976) *J. Mol. Biol. 104*, 145–167.
Ryckaert, J. P., Ciccotti, G., & Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327–341.
Sarma, R. H., Warner, B. G., & Mitra, C. K. (1982) in *Biomolecular Stereodynamics* (Sarma, R. H., Ed.) Vol 1, pp 89–98, Adenine, New York.
Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H., & Kaptein, R. (1984) *Biochemistry 23*, 1371–1376.
Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Kim, J. J. P., & Rich, A. (1976) *J. Mol. Biol. 104*, 109–144.
Shakked, Z., Guerstein-Guzikevich, G., Eisenstein, M., Frolow, F., & Rabinovich, D. (1989) *Nature 342*, 456–460.
Sklenár, V., & Bax, A. (1987) *J. Am. Chem. Soc. 109*, 7525–7526.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Swaminathan, S., Ravishanker, G., & Beveridge, D. L. (1991) *J. Am. Chem. Soc. 113*, 5027–5040.
Teng, M.-K., Liaw, Y.-C., van der Marel, G. A., van Boom, J. H., & Wang, A. H.-J. (1989) *Biochemistry 28*, 4923–4928.
Thomas, P. D., Basus, V. J., & James, T. L. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1237–1241.
Tidor, B., Irikura, K., Brooks, B. R., & Karplus, M. (1983) *J. Biomol. Struct. Dyn. 1*, 231–252.
Van Wijk, J., Huckriede, B. D., Ippel, J. H., & Altona, C. (1992) *Methods Enzymol. 211*, 286–306.
Verlet, L. (1967) *Phys. Rev. 159*, 98–105.
Westhof, E., & Sundaralingam, M. (1980) *J. Am. Chem. Soc. 102*, 1493–1500.
Wing, R. M., Drew, H. R., Takano, T., Broka, C., Tamaka, S., Itakura, K., & Dickerson, R. E. (1980) *Nature (London) 287*, 755–758.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, pp 203–255, John Wiley & Sons, New York.
Yang, D., Gao, Y.-G., Robinson, H., van der Marel, G. A., van Boom, J. H., & Wang, A. H.-J. (1993) *Biochemistry 32*, 8672–8681.
Yip, P., & Case, D. A. (1989) *J. Magn. Reson. 83*, 643–648.
Yoshida, S., Yamada, M., & Masaki, S. (1977) *Biochim. Biophys. Acta 477*, 144–150.
Zhang, H., Gao, Y.-G., van der Marel, G. A., van Boom, J. H., & Wang, A. H.-J. (1993) *J. Biol. Chem. 268*, 10095–10101.