van den Hoogen, Y. Th., van Beuzekom, A. A., van den Elst, H., van der Marel, G. A., van Boom, J. H., & Altona, C. (1988) *Nucleic Acids Res. 16*, 2971–2986.
Wagner, G., (1983) *J. Magn. Reson. 55*, 151–156.
Weiss, M. A., Patel, D. J., Sauer, R. T., & Karplus, M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 130–134.
Woodson, S. A., & Crothers, D. M. (1987) *Biochemistry 26*, 904–912.
Woodson, S. A., & Crothers, D. M. (1988a) *Biochemistry 27*, 436–445.
Woodson, S. A., & Crothers, D. M. (1988b) *Biochemistry 27*, 3130–3141.
Woodson, S. A., & Crothers, D. M. (1989) *Biopolymers 28*, 1149–1177.
Yoon, C., Privé, G. G., Goodsell, D. S., & Dickerson, R. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6332–6336.

Two-Dimensional ^1H and ^{31}P NMR Spectra and Restrained Molecular Dynamics Structure of a Mismatched GA Decamer Oligodeoxyribonucleotide Duplex†

Edward P. Nikonowicz and David G. Gorenstein*
Department of Chemistry, Purdue University, West Lafayette, Indiana 47907
Received March 29, 1990; Revised Manuscript Received June 12, 1990

ABSTRACT: Assignment of the ^1H and ^{31}P NMR spectra of a tandem G·A mismatched base pair decamer oligodeoxyribonucleotide duplex, d(CCAA*GA*TTGG)$_2$, has been made by two-dimensional ^1H–^1H and heteronuclear ^{31}P–^1H correlated spectroscopy. Unusual downfield ^{31}P resonances have been assigned by a pure absorption phase constant-time heteronuclear ^{31}P–^1H correlated spectrum to be associated with the phosphates on the 5′- and 3′-sides of the mismatched guanosine residue. $J_{H3'-P}$ coupling constants for each of the phosphates of the decamer were obtained from the ^1H–^{31}P *J*-resolved selective proton-flip 2D spectrum. The two most downfield-shifted ^{31}P resonances each appear to consist of two overlapping signals that can be resolved into two distinct doublets with different coupling constants in the *J*-resolved spectrum. This as well as the temperature dependence of the ^{31}P spectra demonstrates that two distinct conformations exist at lower temperatures. By use of a modified Karplus relationship, the C4′–C3′–O3′–P torsional angles (ϵ) were obtained. A linear correlation between ^{31}P chemical shifts and the measured coupling constants is quite good (only when the larger set of coupling constants of the two most downfield ^{31}P signals is included). The ^{31}P chemical shifts as well as the measured coupling constants tend to follow the positional variation seen in other duplexes of interior phosphates resonating more upfield than terminal residues and of interior phosphates exhibiting smaller coupling constants; however, this pattern is disrupted at the site of the mismatch. Modeling and initial NOESY distance restrained molecular mechanics energy minimization and restrained molecular dynamics support previous observations that the mismatched guanine and adenine bases are both in anti conformations. Most significantly, the ϵ backbone torsional angle variaions calculated from the NOESY distance restrained structures are in agreement with both the crystal structure values and the measured $J_{H3'-P}$ coupling constants.

The formation of base pair mismatches in DNA and the associated problems during replication, transcription, and gene activation if left unrepaired are current topics of concern (Modrich, 1987; Radman & Wagner, 1986). Their recognition and efficiencies of repair have been studied in bacteria (Fazakerley et al., 1986; Lu et al., 1984; Radman & Wagner, 1984), the latter having been found to be sequence dependent. Both X-ray crystallography and NMR spectroscopy have been used to probe the structural anomalies of mismatched base pairs (Brown et al., 1986; Dodgson & Wells, 1977a,b; Hare et al., 1986a; Kan et al., 1983; Kouchakdjian et al., 1988; Patel et al., 1984a,b; Prive et al., 1987). One of the first nonstandard types of base pairs to be structurally investigated in solution was the naturally occurring purine–pyrimidine base pair G·U found in tRNA (Hare & Reid, 1979; Johnston & Redfield, 1978). This G·U base pair was shown to be in a "wobble"-type pairing scheme. Later, Patel et al. examined the dynamics of the imino protons of four pseudo-self-complementary dodecamers containing pairs of well-separated mismatches of the types G·T, G·A, A·C, and T·C derived from d(CGX$_1$GAATTCX$_2$CG)$_2$, where X$_1$ and X$_2$ are the mismatched pair (Patel et al., 1984c, 1982a,b). The G·T was demonstrated to be wobble base paired; the G·A formed an anti–anti base-paired scheme through the adenine NH$_2$/N3 and guanine O6/NH1.

X-ray crystallography has also proved to be a useful tool for the elucidation of the structure of deoxyoligonucleotides, although crystal results are not always in good agreement with solution-state results (Joshua-Tor et al., 1988; Nikonowicz et al. 1990). The purine–purine mismatch G·A has been crystallized in two different sequences: the dodecamer d(CGCGAATTAGCG)$_2$ (Brown et al., 1986) and the decamer d(CCAAGATTGG)$_2$ (Prive et al., 1987). In the former case, the mismatch was determined to be in a G_{anti}–A_{syn} conformation and in the latter molecule the conformation was found

† Supported by the NIH (AI27744), the Purdue University Biochemical Magnetic Resonance Laboratory which is supported by the NIH (Grant RR01077 from the Biotechnology Resources Program of the Division of Research Resources), the NSF National Biological Facilities Center on Biomolecular NMR, Structure and Design at Purdue (Grants BBS 8614177 and 8714258 from the Division of Biological Instrumentation), and the National AIDS Research Center at Purdue (AI727713).

0006-2960/90/0429-8845$02.50/0

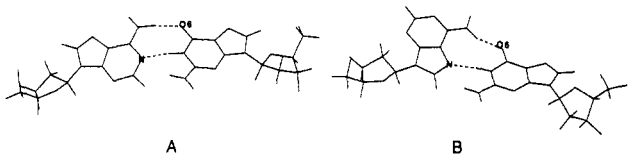

FIGURE 1: Nucleoside G·A base pair in either (a) G_{anti}–A_{anti} conformation or (b) G_{anti}–A_{syn} conformation.

to be G_{anti}–A_{anti} (see Figure 1).

Solution-state studies of the G·A mismatch in a dodecamer sequence have shown that this base pair is capable of adopting a G_{syn}–A_{anti} conformation at acidic pH (Gao & Patel, 1988). Using the tandem G·A decamer sequence d(CCAAGATTGG)$_2$, Kan et al. (1983) have demonstrated the presence of an NOE at A6H2 when the imino proton of the A·G base pair was irradiated at 1 °C, suggesting that the G and A were in an anti–anti conformation. Thus, while spectroscopic evidence has defined the base-pairing schemes possible for the G·A-type mismatch, a well-defined solution structure of the paired bases and the flanking sequences has not been established.

Much emphasis has been placed upon the use of two-dimensional ^{1}H NMR techniques in the past few years to extract structural information about oligonucleotides in solution. The majority of this information has been in the form of NOESY distance constraints of 5 Å or less (Hare et al., 1986c; Roy et al., 1987; Nikonowicz et al., 1989, 1990). Coupling constant information for the deoxyribose ring and hence torsion angle data are obtainable as well, but their use has been restricted to a small number of cases (Celda et al., 1989; Van De Ven & Hilbers, 1988; Zhou et al., 1988). This has been in large part due to the severe overlap in the H3′, H4′, and H5′/H5″ regions of the spectra of oligonucleotides of modest size (8–10 base pairs) and longer. This is unfortunate since the backbone structure of DNA can largely be defined by six torsion angles, four of which directly involve the sugar ring. The structural information gained by describing the entire backbone conformation would provide additional constraints that NOESY data do not directly afford.

Our laboratory has proposed that ^{31}P NMR spectroscopy is capable of providing information on the remaining two torsional angles involving the phosphate ester bonds. Theoretical studies (Gorenstein, 1984, 1987; Gorenstein et al., 1988) indicate that these two torsion angles (α, O3′–P–O5′–C5′; ζ, C3′–O3′–P–O5′) appear to be most important in determining ^{31}P chemical shifts. We have noted that ^{31}P chemical shifts can potentially provide a probe of the conformation of the phosphate ester backbone in nucleic acids and nucleic acid complexes (Gorenstein, 1978, 1984; Gorenstein & Goldfield, 1984). These studies suggested that the ^{31}P resonance of a phosphate diester in a g,g or g^-,g^- conformation should be several ppm upfield of the ^{31}P signal of an ester in a t,g or t,t conformation (Gorentein, 1984). [The notation for the P–O ester torsion angles follows the convention of Seeman et al. (1976) with the ζ, P–O3′ angle given first followed by the α, P–O5′ angle.] Additionally, it is now possible to directly measure the P–H3′ coupling constant (Sklenar & Bax, 1987). Use of a Karplus-type relationship provides for direct determination of the P–H3′ torsional angles (Lankhorst et al., 1984) and indirectly the C4′–C3′–O3′–P torsional angles (Dickerson, 1983). Thus both ^{31}P chemical shifts and ^{31}P coupling constants are now able to provide information on the deoxyribose phosphate backbone conformation of duplex DNA.

With the recent development of methodologies to assign individual ^{31}P resonances and measure P–H3′ coupling constants of oligonucleotides (Fu et al., 1988; Gorentein et al., 1988; Lai et al., 1984; Ott & Eckstein, 1985a,b; Petersheim et al., 1984; Schroeder et al., 1986; Shah et al., 1984a), we and others also have been able to begin to understand some of the factors apparently responsible for ^{31}P chemical shift variations in oligonucleotides (Chang et al., 1987, 1982; Lai et al., 1984; Ott & Eckstein, 1985b; Schroeder et al., 1986; Shah et al., 1984a). As mentioned above, one of the major contributing factors hypothesized to determine ^{31}P chemical shifts in oligonucleotides is the main-chain torsional angle of the individual phosphodiester groups. Phosphates located toward the middle of a B-DNA double helix assume the lower energy, stereoelectronically (Gorenstein, 1987) favored g^-,g^- conformation, while phosphodiester linkages located toward the two ends of the double helix tend to adopt a mixture of g^-,g^- and t,g^- conformations, where increased flexibility of the helix is more likely to occur. Because the g^-,g^- conformation is responsible for a more upfield ^{31}P chemical shift, while a t,g^- conformation is associated with a lower field chemical shift, internal phosphates in oligonucleotides would be expected to be upfield of those nearer the ends. Although several exceptions have been observed, this positional relationship appears to be generally valid for oligonucleotides where ^{31}P chemical shift assignments have been determined (Gorenstein, 1978; Gorenstein et al., 1976; Ott & Eckstein, 1985a,b; Schroeder et al., 1986, 1987).

Eckstein and co-workers (Connolly & Eckstein, 1984; Ott & Eckstein, 1985a,b) and our laboratory (Gorenstein et al., 1988; Schroeder et al., 1986, 1987) have recently suggested that ^{31}P chemical shifts are also sensitive to sequence-specific structural variations of the double helix, related to those proposed by Calladine (1982) and Dickerson (1983). The ^{31}P chemical shifts of duplex B-DNA phosphates correlate reasonably well with some aspects of the Dickerson/Calladine sum function for variation in the helical twist (or helical roll) of the individual base steps in the oligonucleotides. In the B-form DNA double helix with ~10 base pairs per turn (360°) of the helix, each base pair is rotated ca. +36° (helix twist) with respect to the nearest-neighbor base pair. However, analysis of X-ray crystallographic structures of a number of duplexes has revealed large variations in local helix structure, with helix twist varying from 25° to 45° (Dickerson, 1983). Correlations between experimentally measured P–O and C–O torsional angles and results from molecular mechanics/dynamics energy minimization calcultions show that these results are consistent with the hypothesis that sequence-specific variations in ^{31}P chemical shifts are attributable to sequence-specific changes in the deoxyribose phosphate backbone (Gorenstein et al., 1988).

In this paper, ^{31}P NMR and two-dimensional ^{1}H/^{1}H and ^{1}H/^{31}P NMR methods have been used to assign the ^{31}P signals and extract torsional angle information. Additionally, the NOESY-derived ^{1}H–^{1}H distances have been used in conjunction with restrained molecular dynamics calculations to initially define structures of the tandem G·A mismatched 10-mer duplex. A detailed structural analysis will be described in a subsequent paper.

Experimental Procedures

Synthesis and Sample Preparation. The self-complementary decamer duplex (dCCAAGATTG)$_2$ was synthesized by a manual modification of the phosphite triester method on a solid support (Lai et al., 1984; Schroeder et al., 1987; Shah et al., 1984a,b). The porous glass derivatized dG support was synthesized as described (Pon et al., 1988). After cleavage from the support and deprotection, the resulting oligonucleotide

was purified by reversed-phase HPLC as previously described (Gorenstein et al., 1984) followed by dialysis and lyophilization. The 10-mer was then detritylated with 40% acetic acid for 35 min and passed through a cation exchange column (Dowex 50X, K^+ form), dialyzed, lyophilized, and redissolved in 500 μL of 99.8% D_2O. The sample used in all experiments was prepared by dissolving approximately 510 OD units of the decamer in 600 μL of D_2O (99.995%) buffer containing 100 mM KCl, 0.1 mM NaN_3, and 100 mM phosphate buffer at pH* 7.1 (uncorrected pH meter reading).

NMR. The COSY, TOCSY, and NOESY spectra were recorded on a Varian VXR500 (500 MHz, ^1H) spectrometer at ambient temperature (20.5 ± 0.5 °C) according to the States et al. (1982) method. The TOCSY (Bax & Davis, 1985; Braunschweiler & Earnst, 1983) spectrum was acquired with 4096 points in the t_2 dimension and 512 t_1 increments. The mixing times of the TOCSY experiments were set to either 30 or 120 ms. The data were apodized with a shifted sine bell before Fourier transformation in both dimensions. The NOESY spectra were acquired with 4096 points in the t_2 dimension and 512 t_1 increments. Mixing times of 80–800 ms were used. The 400 ms mixing time NOESY data were treated with a shifted sine bell mutliplication in the t_1 and t_2 dimensions.

The ^{31}P 1D NMR spectra, the ^{31}P melting profiles, the $^1H/^{31}P$ correlation *P*ure *A*bsorption phase *C*onstant-time (PAC) spectra (Schroeder et al., 1986), and the $^1H/^{31}P$ heteronuclear 2D *J*-resolved spectrum were run on Varian XL-200A NMR spectrometer (200 MHz, ^1H). All experiments were run in either 99.998% D_2O or 90% H_2O/10% D_2O. Typical ^{31}P 1D NMR parameters were as follows: sweep width 200 Hz; acquisition time 2.98 s; block size 1K zero filled to 16K; pulse width 8.0 μs. Spectra were resolution enhanced with a combination of positive exponential and Gaussian apodization functions; the number of acquisitions was 128. The ^{31}P spectra were referenced to trimethyl phosphate (TMP) at 0.000 ppm, which is 3.456 ppm downfield of 85% phosphoric acid.

A PAC version (Fu et al., 1988; Jones et al., 1988) of the Kessler–Griesinger long-range heteronuclear correlation (COLOC) experiment (Kessler et al., 1984) was conducted on the decamer at ambient temperature. The PAC spectrum provides chemical shift correlation between the H3′, H4′, and H5′ protons of the deoxyribose rings and the three or four bond coupled phosphorus. Coupling occurs between the phosphorus and the H3′ proton on the 5′-side and between the H4′ and H5′ protons on the 3′-side of the dinucleotide subunit. The data set was acquired with 256 points in the ^{31}P dimension and 64 t_1 increments and then zero-filled to 1024 × 512. The data set was multiplied by a combination of an increasing exponential and a Gaussian function before Fourier transformation in both dimensions. ^1H NMR resonances are referenced to 3-(trimethylsilyl)propanesulfonic acid sodium salt (DSS), and ^{31}P NMR resonances are referenced again to an external sample of TMP.

The Bax–Freeman selective 2D J-resolved long-range correlation experiment with a Dante sequence for the selective 180° pulse (Sklenar & Bax, 1987) was performed on the decamer to correlate the ^{31}P chemical shift with the phosphorus–H3′ coupling constant. The ^{31}P–^1H *J*-resolved spectra were recorded at 9, 20, 30, 60, and 80 °C.

The data sets were acquired with 256 points in the ^{31}P dimension and 32 t_1 increments and then zero-filled to 512 × 128. Resolution enhancement of the Gaussian type was applied before Fourier transformation in both dimensions. *J* values were measured from peak center to peak center.

The observed three-bond coupling constant is used with a proton–phosphorus Karplus relationship to measure the H3′–C3′–O–P torsional angle θ from which we have calculated the C4′–C3′–O–P torsional angle ϵ ($= -120° - \theta$). The relationship $J = 15.3 \cos^2\theta - 6.1 \cos\theta + 1.6$, was determined by Haasnoot (Lankhorst et al., 1984; Sklenar & Bax, 1987).

NOESY Distance-Restrained Molecular Mechanics/Dynamics Calculations of Duplex. The molecular modeling program MIDAS operating on a Silicon Graphics Iris 3030 workstation was used to initially generate idealized Arnott B-DNA and A-DNA decamer duplexes $d(CCAAGATTGG)_2$. The two structures were then manually manipulated in order to accommodate the mismatched base pairs in the anti configuration. NOESY distance constraints were incorporated into the potential energy function in the molecular mechanics/dynamics programs AMBER (Weiner & Kollman, 1981) through addition of a flat-well potential (Nikonowicz et al., 1989). A total of 162 NOESY two-spin distance constraints were then energy refined until a rms gradient of 0.1 kcal/(mol·Å) was achieved or until the change in energy was less than 1.0×10^{-7} kcal/mol for successive steps. A total of 260 NOESY distance constraints derived from the hybrid matrix methodology (details presented in a subsequent paper) were used for the hybrid matrix/restrained MD calculations.

Results

Nonexchangeable Proton Assignments of the Decamer. Assignment of the proton signals of the decamer $d(CCAAGATTGG)_2$ was accomplished through analysis of two-dimensional COSY and NOESY spectra via a sequential assignment methodology (Broido et al., 1984; Feigon et al., 1983a,b; Hare et al., 1983; Scheek et al., 1984a; Schroeder et al., 1987). The TOCSY experiment, collected in D_2O, pH* 7.1, containing 0.1 M KCl (20 °C), confirmed the location the H6 resonances of the cytidines and thymidines through the off-diagonal cross-peaks of the scalar-coupled H5/H6 protons (cytidine) and CH3/H6 (thymidine) protons. Additionally, the intra sugar ring proton system connectivities, H1′–H4′, could be traced in the TOCSY spectrum. The 5′ and 5″ regions proved to be either too crowded (H4′/H3′–H5′,H5″) or weak (H1′–H5′,H5″) to unambiguously assign via the TOCSY and COSY.

The sequential assignment of the nonexchangeable proton resonances of the decamer duplex was accomplished by inspection of 2D NOESY spectra collected at mixing times of 200, 400, and 800 ms under the previous conditions, Figure 2.

The assignment procedure followed that expected for a right-handed B-DNA duplex wherein the purine H8 and pyrimidine H6 base protons give rise to cross-peaks from their own deoxyribose ring protons as well as those of the *i*−1 residue (Hare et al., 1983; Scheek et al., 1985b; Weiss et al., 1984) (region A, Figure 2). The sequential connectivities are contiguous through the complete decamer sequence in the 200 ms τ_m NOESY spectrum. However, at this mixing time the A4H1′–G5H8 cross-peak was found to be quite weak relative to the remaining peaks in the region, indicative of an anomalous structural feature. The C1H1′ and C1H5 protons were found to be degenerate; however, the remainder of the spectrum is well dispersed.

We note that the A6H8 proton is two to three times as broad as the other protons in the base region (20 vs 6–10 Hz). However, in the 8 °C ^1H spectrum all peaks in the base region have uniform line widths. We therefore believe this to be indicative of conformational flexibility of the mismatched

Table 1: Nonexchangeable ^{1}H Chemical Shifts (ppm) of d(CCAAGATTGG)$_2$

residue	H8/H6	H2/H5/CH3	H1′	H2′, H2″	H3′	H4′	H5′, H5″
C1	7.67	5.83	5.85	2.00, 2.42	4.58	4.03	3.67, 3.69
C2	7.52	5.66	5.18	2.02, 2.23	4.76	4.01	3.92, 3.95
A3	8.16	7.35	5.78	2.67, 2.76	4.99	4.31	3.89, 4.08
A4	8.00	7.34	5.39	2.25, 2.34	4.91	4.19	4.06, 4.02
G5	7.88		5.11	2.68, 2.64	4.96	4.23	4.03, 3.96
A6	8.28	7.87	6.31	2.70, 2.87	4.98	4.46	4.09, 4.13
T7	7.00	1.49	5.76	1.72, 2.37	4.73	4.11	3.90, 4.20
T8	7.28	1.61	5.58	1.97, 2.22	4.80	4.07	4.02, 4.00
G9	7.79		5.60	2.61, 2.67	4.92	4.29	3.94, 3.98
G10	7.723		6.05	2.44, 2.30	4.59	4.15	4.03, 4.05

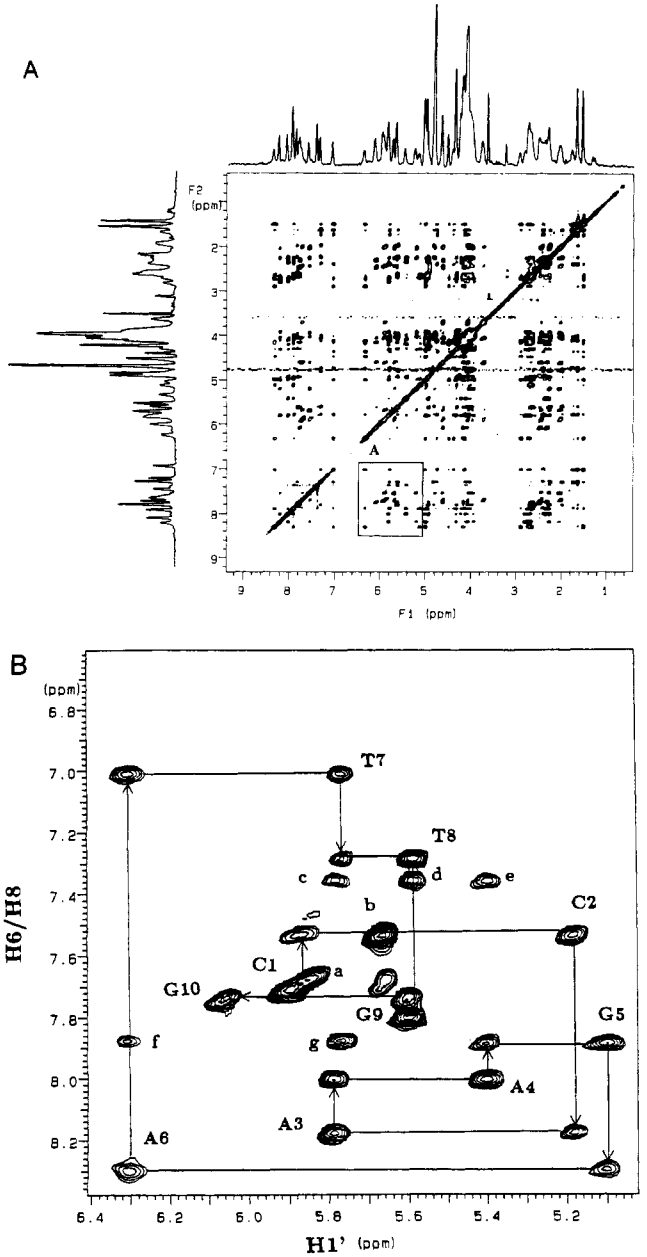

FIGURE 2: (A) Pure absorption phase ^{1}H NOESY NMR spectrum of duplex decamer, at 500 MHz. Region labeled A is expanded in (B): the sequential assignment of the base and deoxyribose H1′ protons is diagrammed. Note that peaks (a) C1H5–C1H6, (b) C2H5–C2H6, (c) A3H1′–A4H2, (d) T8H1–A4H2/G9H1′–A3H2, (e) A4H1′–A4H2, (f) A6H1′–A6H2, and (g) A6H2–T7H1′ are also present in the 200-ms NOESY spectrum. Peak d is composed of two interstrand cross-peaks.

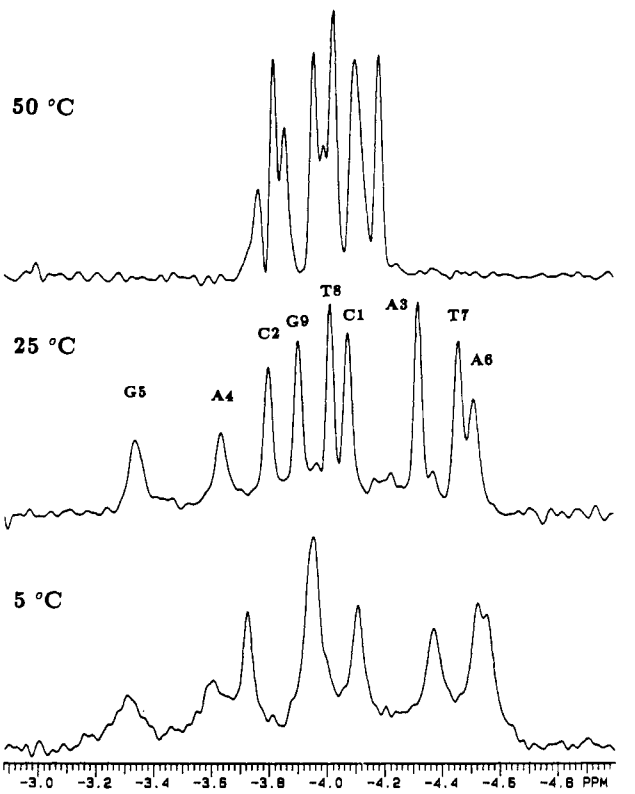

FIGURE 3: ^{31}P NMR spectra at various temperatures and phosphate assignments of the decamer (numbering corresponds to phosphate position from the 5′-end of the duplex).

adenosine base which is in intermediate to slow exchange at lower temperatures. Interestingly though, the H8 proton of G5 is not broadened, suggesting that the residue does not experience the effect of chemical exchange.

Intrastrand and interstrand cross-peaks for the H2 protons of A3, A4, and A6 are also present (Figure 2B). The H2 protons were assigned from the NOESY spectrum on the basis of their connectivities to the H1′ protons of the i and i+1 intrastrand residues and the H1′ protons of G9, T8, and A6 interstrand residues.

Identification of the H1′ resonances allows further sequence-specific assignment of protons H2′, H2″, H3′, and H4′. Stereospecific assignments of the H2′ and H2″ protons were made by analysis of cross-peak intensities at short mixing times (60 and 80 ms). The H2″ proton will generate a more intense intraresidue NOESY cross-peak to the H1′ proton than that produced by the interaction of H2′ and H1′ resulting from the former pair's closer spatial proximity (Feigon et al., 1983a,b; Scheek et al., 1984b).

Finally, the H5′/H5″ resonances were assigned (nonstereospecifically) through their NOESY connectivities with the H1′ and H3′ sugar protons and H6/H8 base protons. These interactions were observed by inspection of short (80 ms) and long (800 ms) mixing time NOESY spectra. The 800-ms NOESY contained the effect of spin diffusion which proved useful for clarification of ambiguities in the H6/H8–H4′,-

Table II: ^{31}P Chemical Shifts (ppm) for d(CCAAGATTGG)$_2$

position	peak[a]	^{31}P (ppm)[a]	20 °C[b]	H3′ (COLOC)[c]	H3′ (NOESY)[d]	H4′ (COLOC)[c]	H4′ (NOESY)[d]
CpC	6	−4.11	−4.11	4.57	4.58	4.02	4.03
CpA	3	−3.79	−3.72	4.76	4.76	4.31	4.31
ApA	7	−4.35	−4.39	5.00	4.99	4.19	4.19
ApG	2	−3.64	−3.60	4.91	4.91	4.23	4.23
GpA	1	−3.34	−3.41	4.97	4.96	4.46	4.45
ApT	9	−4.55	−4.55	4.99	4.98	4.12	4.11
TpT	8	−4.49	−4.49	4.52	4.53	4.08	4.07
TpG	5	−4.02	−4.01	4.80	4.80	4.30	4.29
GpG	4	−3.93	−3.96	4.92	4.92	4.16	4.15

[a] Peaks are numbered from lowest field. Chemical shifts of ^{31}P resonances derived from PAC experiment. [b] Chemical shifts of ^{31}P resonances observed at 20 °C in 1D spectrum. [c] H3′ and H4′ chemical shifts as observed in the PAC (pure absorption COLOC) spectrum. [d] H3′ and H4′ chemical shifts as observed in the NOESY spectrum.

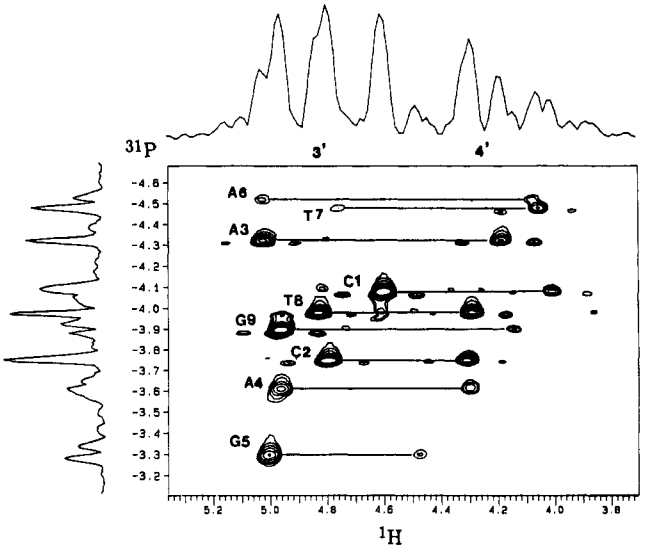

FIGURE 4: Two-dimensional ^{31}P–^{1}H PAC heteronuclear correlation NMR spectrum of duplex decamer at 200 MHz (^{1}H). The 1D decoupled ^{31}P NMR spectrum is shown along one axis, and the H3′, H4′, and H5′,5″ region of the proton spectrum is shown along the second axis.

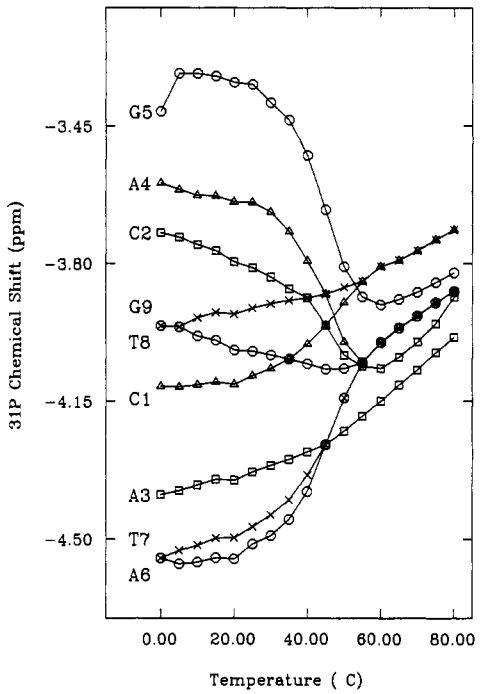

FIGURE 5: Temperature dependence of ^{31}P chemical shifts of duplex decamer.

H5′,H5″ region of the spectra. The proton assignments are listed in Table I.

^{31}P Resonance Assignments of the Decamer. The assignment of the resonances of the decamer ^{31}P spectrum (Figure 3) was accomplished via a pure absorption phase constant-time ^{1}H/^{31}P heteronuclear correlated spectrum, PAC (Fu et al., 1988). Rather than conventional 2D ^{1}H/^{31}P HETCOR NMR spectroscopy being used (Lai et al., 1984; Pardi et al., 1983), which suffers from poor sensitivity, a heteronuclear version of the "constant-time" coherence-transfer technique (Bax & Morris, 1981), referred to as COLOC and originally proposed for ^{13}C/^{1}H correlations (Kessler et al., 1984), was chosen. We have used a pure absorption phase version of the COLOC sequence (PAC) previously modified and optimized for the ^{31}P to 3′-proton of oligonucleotides in our laboratory (Fu et al., 1988; Nikonowicz et al., 1989) in conjunction with the 3′-proton assignments from the NOESY spectrum to assign ^{31}P resonances. The PAC spectrum of the decamer is shown in Figure 4. Table II lists the ^{31}P assignments of the nine phosphates in the decamer backbone. It can be seen in the PAC spectrum that the H3′ as well as the H4′ protons are well resolved. We note that although three of the most upfield H3′ resonances are nearly degenerate, ^{31}P assignments involving these protons can be resolved through the H4′ scalar connectivities.

^{31}P Melting Curve. Figure 3 shows representative examples of the ^{31}P spectra at various tempertures. The temperature dependence of the ^{31}P chemical shifts is shown in Figure 5. As can be seen, there is a general downfield shift of many of the resonances with increasing temperature as predicted (Gorenstein, 1981, 1984). However, it is interesting to note that three of the peaks actually move considerably upfield beginning at about 30 °C. Two of these resonances, labeled G5 and A4, are connected to the 5′-oxygens of the mismatched base pairs. The inflection point in the curves for all phosphates occurs at ~40 °C, indicating the melting of the duplex. We also point out that these two downfield-shifted peaks are broad relative to the remaining peaks in the ^{31}P spectrum. These broad peaks neither sharpen nor split into multiple resolvable peaks at lower temperatures (0–30 °C; however, see below for evidence of multiple peaks in the *J*-resolved spectra). At 35 °C these peaks begin to sharpen and move upfield; by 55 °C their chemical shifts begin to move downfield. Thus prior to melting these downfield-shifted phosphates are apparently undergoing a premelting transition involving chemical exchange between two different phosphate conformational states.

J-Correlated Spectra. It is our goal to use the NMR data from solution and computational methods to generate as valid a structure as possible for the decamer. Unfortunately, ^{1}H/^{1}H NOESY information has not been generally believed to provide any significant constraint on the conformation of the phosphate ester backbone (Van De Ven & Hilbers, 1988).

Our laboratory has been interested in applying the information provided by ^{31}P NMR to aid in the elucidation of the

Table III: ^{31}P-^{1}H Coupling Constants (J) and Calculated Torsion Angles

residue	^{31}P	J_{H-P} (Hz) 20 °C	9.5 °C	30 °C	ϵ, 20 °C	ζ, 20 °C
G5[a]	−3.37/−3.41	3.0/7.2	3.0/7.4	4.5	−175/−153	−102/−129
A4[a]	−3.60/−3.63	4.0/5.7	3.2/5.7	4.4	−170/−161	−110/−119
C2	−3.79	3.5	3.1	6.0	−172	−106
G9	−3.96	3.5	2.8	4.2	−172	−106
T8	−4.01	4.4	4.3	5.2	−165	−115
C1	−4.11	3.6	3.1	6.3	−162	−119
A3	−4.39	2.4	2.4	3.2	−180	−98
T7	−4.49	1.9	1.5	4.0	−184	−93
A6	−4.55	1.5	1.6	2.3	−188	−88

[a] Two sets of doublets were observed for both the G5 and A4 phosphates.

fine structure of DNA and DNA complexes. To this end, we have focused on understanding the origin of variations in ^{31}P chemical shifts and coupling constants. One parameter that has yet to be fully exploited in oligonucleotide NMR solution structural determinations is the backbone torsional angles dependent about the C–O–P–O–C bonds. Two of these torsional angles, ϵ (C4′–C3′–O–P) and β (C4′–C5′–O–P), are accessible through a ^{1}H–^{31}P J-resolved experiment. Unfortunately, due to the complexity of the 5′,5″ region, only the C4′–C3′–O–P torsional angle (ϵ) is easily accessible by this method. The heteronuclear proton-flip experiment (Bax & Freeman, 1982) as optimized for P–H3′ J correlations (Sklenar & Bax, 1987) was performed, without reverse detection, on the decamer at 9, 20, 30, 60, and 80 °C. We have additionally used this pulse sequence to derive qualitative information on the P–H5′ coupling constants. The J_{HP} couplings to the H4′ and H5′/H5″ regions are resolvable and quantitative at 80 °C where the proton lines become very sharp; however, the decamer is also in a single-strand form, thereby removing potential torsional angle variations which may exist in the duplex form. The H5′/H5″ region does yield some qualitative information at 20 °C which will be addressed under Discussion.

As seen in the ^{31}P melting curve, Figure 5, there appears to be little melting of the duplex (and phosphate ester backbone) at 20 °C. All ^{31}P resonances are resolved at this temperature as well as all ^{31}P–H3′ coupling constants. The 20 °C P–H3′ J correlation spectrum shown in Figure 6A was thus used to derive the ϵ torsion angles. Table III lists the P–H3′ coupling constants of the duplex at various temperatures. We note that at 20 °C the two most downfield-shfted ^{31}P resonances, those two which are 3′ to residues A4 and G5, each appear to consist of two overlapping resonances that are only resolved in the 2D J-spectrum. The two most downfield-shifted ^{31}P resonances each appear to consist of two overlapping signals that can be resolved into two distinct doublets with different coupling constants in the J-resolved spectrum. The overlapping signals show a ^{31}P chemical shift difference of about 0.05 ppm, with the more upfield of each in the pair having the smaller coupling constant. This demonstrates that two distinct phosphate ester conformations exist at lower temperatures for these two phosphates adjacent to the mismatch sites. It also explains why at lower temperatures the broad signals do not appear to resolve into separate slow-exchange signals because in the 1D spectra each of the broad resonances consists of two very nearly coincident signals in slow chemical exchange. At higher temperature where these two downfield signals sharpen, either the chemical shift difference between the overlapping signals is further reduced, or chemical exchange averages the overlapping signals.

Table III also lists the derived torsional angles at 20 °C. We have calculated H3′–C3′–O–P torsional angle θ using the coupling constants and the proton–phosphorus Karplus relationship of Lankhorst et al. (1984) (see Experimental Procedures) to determine the C4′–C3′–O–P torsional angle ϵ. Although it is possible to derive up to four different θ torsional angles (0–360°) from the same coupling constant, we initially assume that the value of θ closest to the undistorted crystallographically observed value of ca. 49° (ϵ = −169°; Saenger, 1984) is the correct value.

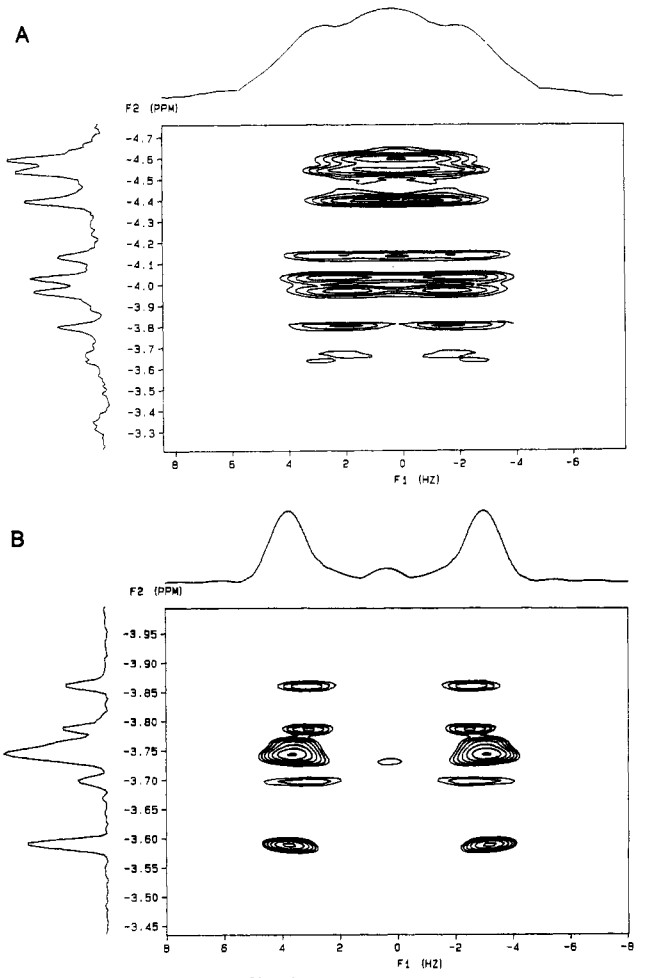

FIGURE 6: 2D J-resolved $^{31}P/^{1}H$ spectra of the decamer at (A) 20 and (B) 80 °C. The 1D decoupled ^{31}P NMR spectrum is also shown along one axis, and the H3′ coupled doublet projections are shown along the second dimension.

In addition to the torsion angle C4′–C3′–O3′–P (ϵ), we have calculated the torsion angle C3′–O3′–P–O5′ (ζ). A strong correlation (R = −0.92) between torsional angles ζ and ϵ in the crystal structures of a dodecamer was shown to exist by Dickerson, and the angles are related by the equation $\zeta = -317 - 1.23\epsilon$ [Dickerson, 1983; Dickerson & Drew, 1981; see also

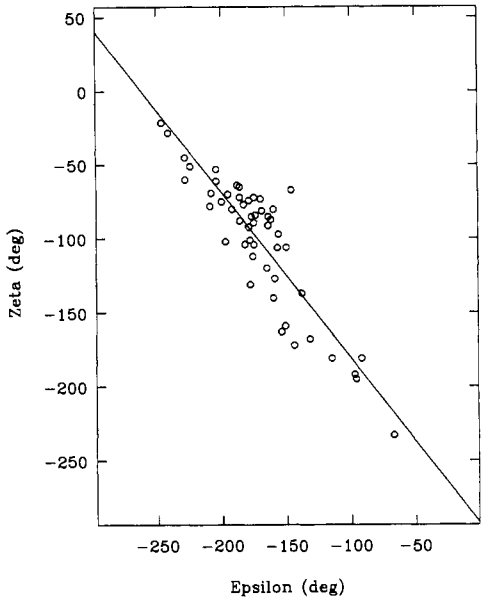

FIGURE 7: Plot of ϵ vs ζ for torsional angles derived from three mismatched crystal structures: d(CCAA*GA*TTGG)$_2$ (Prive et al., 1987), d(CGC*A*AATTC*G*CG)$_2$ (Hunter et al., 1987a), and d-(CGC*G*AATT*T*GCG)$_2$ (Hunter et al., 1987b).

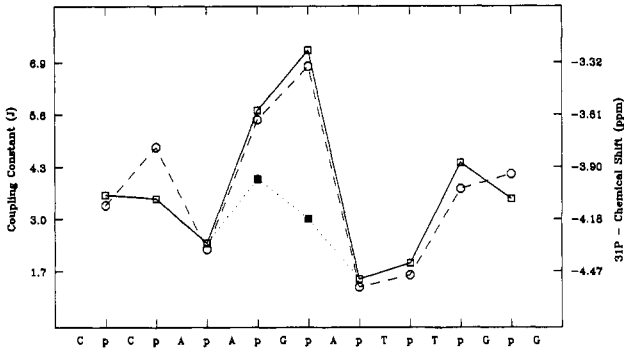

FIGURE 8: Plot of the ^{31}P chemical shift (O) vs phosphate position along the 5′ → 3′ direction of the sequence. Also plotted is the experimental $J_{H3'-P}$ coupling constant (□) of the decamer duplex [the coupling constants for the secondary resolved doublets for the A4p and G5p phosphates are also shown (■)].

Powers et al. (1990)]. As noted by Dickerson (Dickerson, 1983; Dickerson & Drew, 1981) when the P–O3′ (ζ) conformation is g^-, invariably the C–O3′ conformation (ϵ) is found to be t. This $\epsilon(t),\zeta(g^-)$ conformation is the most common backbone conformation [defined as the $B_I(t,g)$ conformation] (Dickerson, 1983; Dickerson & Drew, 1981). The other most common conformation for the ϵ,ζ pair is the g^-,t or B_{II} state. A "crankshaft" motion interconverts B_I and B_{II} conformations with only a modest movement of the phosphate. It is largely this variation in ϵ and ζ (as well as δ; Gorenstein et al., 1988, 1989; Nikonowicz et al., 1989; Schroeder et al., 1989) that allows the sugar–phosphate backbone to "stretch" or "contract" to allow for variations in the local structure of B-DNA.

A very similar correlation exists between ζ and ϵ in the crystal structures of various B-DNA base pair mismatches, including the decamer (Figure 7). The best fit to the data is given by the linear relationship $\zeta = -296 - 1.14\epsilon$ with a correlation coefficient of −0.90. We have assumed that the original Dickerson correlation of ζ and ϵ exists for the duplex structure in solution as well and have thus calculated ζ from ϵ and this relationship (Table III; using the slightly different correlation from Figure 7 increases all of the ζ values by 6°–8°). A comparison of the variation of the coupling constant (which is the direct experimental measure of the torsional angles) and ^{31}P chemical shifts for the decamer sequence is shown in Figure 8.

Table IV: Selected Distances Involving the Central G·A Mismatch

atom pair	80-ms NOESY[a]	200-ms MORASS[b]	B-DNA$_{syn}$ model[c]/MD[d]
G5H8–A4H2′	3.22	3.81	3.95/3.80
G5H8–A4H2″	3.17	2.89	2.56/2.93
G5H8–G5H3′	4.55	4.45	4.94/4.64
G5H1′–G5H4′	3.48	3.57	3.55/3.31
G5H1′–A6H8	3.61	3.86	5.04/3.76
A6H8–A6H1′		3.89	2.54/3.91
A6H1′–A6H4′	3.44	3.10	2.82/2.91
A6H2″–A6-8	3.16	3.40	4.22/3.71
A6H2″–A6-1′	2.61	2.39	2.44/2.46
A6H2″–A6H3′	3.16	2.72	2.73/2.76
A6H1′-T7H6	3.52	3.82	3.71/4.05
A6H2′–T7H6	3.46	3.90	3.88/3.75
A6H2″–T7H6	2.73	2.31	2.29/2.36
A6H2–T7H1′	3.64	3.28	3.33/3.29

[a] Constraint distances calculated from the 80-ms NOESY experiment with the "two-spin" approximation. [b] Constraint distances derived from the final iteration of the hybrid matrix/MD procedure (MORASS) with volumes from the 200-ms NOESY experiment. [c] Distances between specified proton pairs were from the energy-minimized model-built B-DNA structure with G_{anti} and A_{syn}. [d] Distances from final structure-derived after 40 ps of gas-phase dynamices with 200-ms MORASS constraints starting from the model-built B-DNA G_{anti}–A_{syn} structure.

When the larger of the two coupling constants from A4p and G5p are used, the correlation between $J_{H3'-P}$ coupling constants and ^{31}P chemical shifts is quite good. However, when the smaller of the two coupling constants is being considered, the above correlation is poorer than that which we have noted for a number of other oligonucleotide duplexes (Gorenstein et al., 1988, 1990) (see below; Figure 10).

Molecular Modeling of Decamer Duplex. Using the distances derived from the 2D NOESY spectrum of Figure 2, in conjunction with restrained molecular mechanics and dynamics calculations, we have derived structures for the duplex decamer. Distances were calculated by integrating the cross-peaks and utilizing the two-spin approximation at short mixing times (Wuthrich, 1986). The two-spin methodology assumes that, at very short mixing times in the linear regimes of the NOE buildup, the cross-peak intensity of a spin pair is inversely proportional to the sixth power of the distance between those two protons. Although volumes were measured for a complete set of mixing times (60, 80, 200, 400, and 800 ms), the volumes for a single 80-ms NOESY were used to measure the intra and interproton distances from the observed NOE cross-peaks (Table IV). [A complete hybrid relaxation matrix calculation analysis of the NOESY data will be presented in a subsequent paper with a detailed comparison of the structures derived from both the two-spin approximation and the complete relaxation matrix (Keepers & James, 1984) approach.]

Four alternative models for the decamer were considered, A_{syn}–G_{anti} and A_{anti}–G_{anti} for both A and B conformations of DNA. Using the molecular mechanics/dynamics program AMBER3 (Weiner & Kollman, 1981), we model built standard A- and B-DNA structures for the parent decamer duplex, d(CCAAGATTGG)$_2$. The mismatch adenosines on each strand were then rotated 180° about the C1′–N9 bond to produce the A_{syn}-oriented structures by use of the molecular modeling program MIDAS (Ferrin and Langridge, 1980) and the four different conformations energy minimized with AMBER. The NOESY-derived distances were then used in a distance-constrained energy minimization (AMBER) of the decamer duplex structures. Instead of a simple harmonic

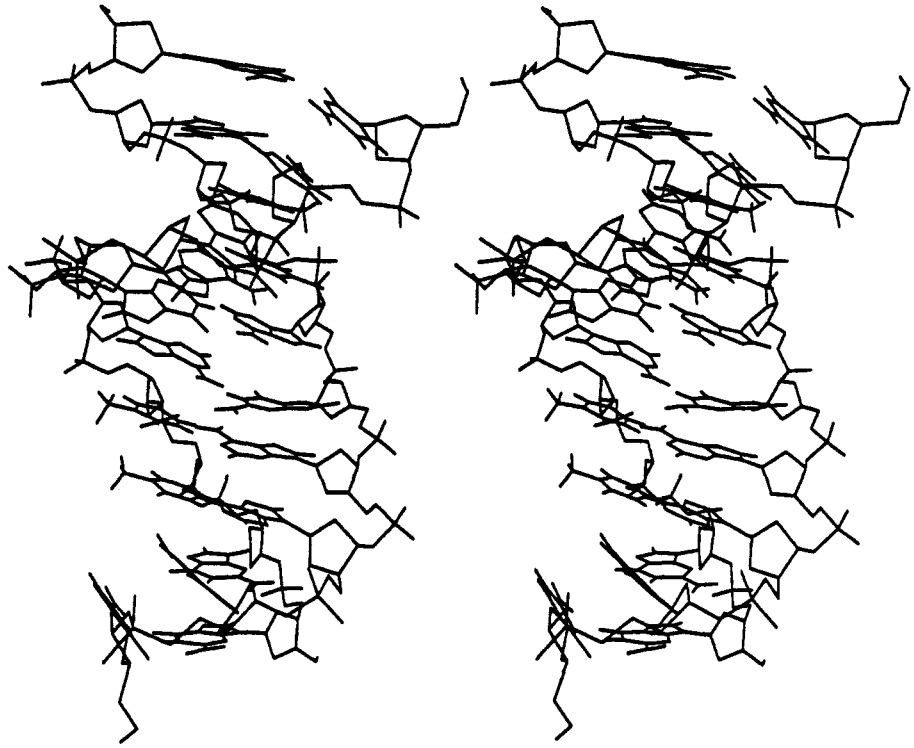

FIGURE 9: Stereoview of decamer after 40 ps of the hybrid matrix derived, distance-restrained MD refinement starting from model-built B-DNA with A_{syn}–G_{anti} for both mismatches. The final structure shown is A_{anti}–G_{anti}.

pseudo-energy error function to restrain the NMR-derived distances, we have modified AMBER so as to provide a flat-well harmonic function, which we believe better reflects the intrinsic accuracy of these NOESY distances restraints (Clore et al., 1985; Gorenstein et al., 1990). The left and right force constants in the flat-well harmonic potential for the NOESY distance constraint term were set to 10 kcal/(mol·Å^2) with permitted errors of ±15% in the NOESY distances. One of the refined structures calculated with distances derived from the 200-ms NOESY and the hybrid matrix methodology (details will be given in a subsequent paper) using the relaxation matrix program MORASS (by R. P. Meadows, C. B. Post, and D. G. Gorenstein, 1989) is shown in Figure 9. A summary of selected distances calculated from the 80- and 200-ms NOESY spectra is shown in Table IV.

Discussion

Spectral Assignments. We have studied the decamer sequence d(CCAAGATTGG) with particular emphasis on the two-dimensional proton and proton/phosphorus spectra. We have sequentially assigned all nonexchangeable protons of the decamer, including the 5′ and 5″ sugar ring protons (Table I), via the TOCSY spectrum and two-dimensional NOESY technique (Broido et al., 1984; Feigon et al., 1983a,b; Hare et atl., 1983; Scheek et al., 1984b; Wemmer et al., 1984). Sugar proton to base proton connectivities are consistent with that expected for a right-handed helix B-type DNA (Weiss et al., 1984).

The base protons have produced NOESY cross-peaks to their own H1′, H2′, H2″, and H3′ protons as well as to those protons of the 5′-neighboring nucleotide. The 4′-protons were assigned via their spatial connectivity with the previously assigned 1′- and 3′-protons. The chemical shifts of the protons in the base region are consistent with those earlier deduced by Kan et al. (1983) through comparison of the mismatched G·A decamer spectrum with that of the previously reported G·C parent decamer duplex (Kan et al., 1982). The relative regions of proton types are those typically expected in DNA, H6/H8–H1′–H3′–H4′,H5′–H2′/H2″; however, the H1′ protons are dispersed over a larger than typical range, 1.20 ppm. Not unexpectedly, this anomalous H1′ dispersion is centered about the site of the mismatch. While it is common for a guanine–adenine base step to produce an upfield GH1′ and a downfield AH1′ step, the G5H1′ proton is the most upfield shifted of the H1′ protons (5.11 ppm), nearly falling in the H3′ region, 4.23–4.99 ppm (Table I). This unusually high upfield shift of an H1′ proton has also been recently observed in a nontandem G·A mismatch dodecamer (Gao & Patel, 1988) in the G_{anti}–A_{anti} conformation, suggesting some common feature of the G at the mismatch site. The most downfield-shifted H1′ resonance is that of A6H1′, coming at 6.31 ppm.

G_{anti}–A_{anti} Nature at the Mismatch. The G·A-type mismatch has been reported to adopt three basic orientations, i.e., G_{anti}–A_{anti}, G_{anti}–A_{syn}, and G_{syn}–A_{anti}. Recently, the orientation of a nontandem mismatch was shown to be pH dependent in a dodecamer sequence in solution (Gao & Patel, 1988), going from G_{anti}–A_{anti} to G_{syn}–A_{anti} as the pH was decreased. The sequence studied here was first shown to be in a G_{anti}–A_{anti} orientation in solution by observation of an NOE at A6H2 upon irradiation of the imino proton of the GA base pair at 1 °C. Our results confirm that the initial observation at 1 °C, G_{anti}–A_{anti} (Kan et al., 1983), is also valid at 20 °C. The cross-peak intensity of A6H8–A6H1′ relative to that of H5–H6 of either C1 or C2 is quite weak. The greater intensity of this fixed-distance, 2.46 Å, proton pair even at short mixing times indicates a greater distance between the H8 and H1′ of A6, i.e., separated by more than 2.5 Å. (In the syn orientation, the H8–H1′ distance is ca. 2.5 Å.)

As stated above, the H8 proton of A6 appears broad relative to the other protons in the base region. This broadness is also seen in the NOESY spectrum cross-peaks which involve the A6H8 residue. It is likely that the broadening is due to intermediate exchange between multiple conformations although we have been unable to slow the exchange rate sufficiently in

order to observe these forms. Increasing the temperature causes the A6H8 to follow a pattern of line narrowing similar to that of the other base protons. Additionally, cross-peak intensities involving the A6H8 proton relative to each other and other cross-peaks in the spectrum at reduced temperature (9 °C) do not change from those measured at 20 °C.

Finally, the A6H2 resonance has cross-peaks to the A6H1′ as well as the T7H1′ (intrastrand) resonances. While the $H2_i$ to $H1'_{i+1}$ interaction is commonly seen in B-DNA, the intensity of this interaction (cross-peak) in the decamer is unusually strong (Figure 2B) and builds up quickly. These observations suggest a close spatial proximity between this proton pair, perhaps the result of an ApT dinucleotide unit with unusually short A6H2 to T7H1′ distance [confirmed in Nikonowicz et al. (1990)]. This further supports an $A6_{anti}$ conformation because the $H2_i$ to $H1'_{i+1}$ distance in the syn conformation is longer, >5 Å, than what should be observable in the NOESY spectrum.

^{31}P Assignments. The ^{31}P peaks were assigned via a 2D pure absorption phase constant-time (PAC) heteronuclear correlation NMR spectrum (Fu et al., 1988). Because the 3′- and 4′-resonances of the decamer were assigned via the TOCSY and NOESY experiments, it was possible to assign the ^{31}P peaks through their connectivities via scalar coupling to the 3′- and 4′-protons (Lai et al., 1984; Pardi et al., 1983). The most downfield-shifted peak is that phosphorus atom which is 3′ to residue G5. The adjacent upfield resonance is that of residue A4. It thus appears that the perturbation of the mismatch is felt most strongly in the backbone, namely, in the region of the phosphorus residues immediately 5′ to each of the mismatched bases. The ^{31}P spectrum is quite dispersed at 20 °C with a range of 1.2 ppm. This "spread" in chemical shifts is unusual for a duplex DNA fragment, most being generally less than 0.7 ppm. Even in the cases where there are two mismatch sites (nontandem), the chemical shifts of the ^{31}P spectrum are dispersed between 0.5 and 0.9 ppm (Kalnik et al., 1989; Patel et al., 1984a,b; Roongta et al., 1990). We also note that although the ^{31}P peaks of the phosphates attached to the 5′-oxygens of the mismatched bases integrate to a single phosphorus atom, they are quite broad. As noted earlier, these two most downfield-shifted ^{31}P resonances each appear to consist of two overlapping resonances that are only resolved in the 2D *J*-resolved spectrum. The broadness is largely attributable to the overlap of the signals although a contribution from intermediate chemical exchange is also possible. Presumably this is caused by chemical exchange between different conformations about the C–O–P bonds.

Positional and Sequence-Specific Variation of ^{31}P Chemical Shifts and Backbone Torsional Angles. As mentioned above, one of the factors that will affect ^{31}P chemical shifts is the degree of conformational constriant imposed by the duplex geometry (Connolly & Eckstein, 1984; Gorenstein, 1981; Ott & Eckstein, 1985a). Thus as one moves more interior to the duplex, there is generally an upfield shift of the ^{31}P resonances (Gorenstein et al., 1988; Gorenstein, 1981, 1984; Nikonowicz et al., 1989; Schroeder et al., 1989; Roongta et al., 1990). Base pairs closer to the ends of the duplex are less contrained to the stacked, base-paired geometry. This "fraying" at the ends imparts greater conformational flexibility to the deoxyribose phosphate backbone, and thus phosphates at the ends of the duplex will tend to adopt more of a mixture of g^-,g^- and t,g^- (ζ,α) conformations. Interior phosphates are more constrained to the polymer P–O g^-,g^- conformation.

This "positional" ^{31}P chemical shift effect is apparently superimposed on site- and sequence-specific effects (Connolly & Eckstein, 1984; Gorenstein et al., 1988; Ott & Eckstein, 1985a; Schroeder et al., 1986, 1987; Roongta et al., 1990). The minor groove clash steps (pyrimidine–purine base steps) have been generally associated with a relative downfield ^{31}P chemical shift. Our laboratory has shown that for some sequences ^{31}P chemical shifts correlate well (ca. −0.7) with the predicted Calladine helix twist parameter (Gorenstein et al., 1988). However, we have not found this to be universally true. In fact, there appears to be little or no correlation between predicted Calladine parameters and ^{31}P chemical shifts in this decamer as well as in another recently studied decamer (Powers et al., 1990). It is quite likely that this lack of correlation is a reflection of base pair geometry disortion at sites removed from the base pair mismatches. The failure of the Calladine rules to accurately predict local helical structure variation even for regular duplexes has also been attributed in part to their lack of consideration of electrostatic interactions, hydrogen bonding, and hydration forces (Prive et al., 1987).

A further variation in ^{31}P chemical shifts peculiar to the mismatch is that occurring around the site of the mismatch. While it may seem obvious that the mismatch will affect the chemical shift, it appears to do so in no predictable fashion. It might be expected that a downfield chemical shift perturbation would be observed for the G5p and A6p resonances as was observed for the A3′p resonance in a tridecamer extrahelical base duplex (Nikonowicz et al., 1989); however, this has proved to be only partially correct. Although the phosphorus resonance of the phosphate attached to the 3′-oxygen of G5 is shifted downfield, the phosphorus resonance 3′ to the mismatched A6 is not. Rather, it is the phosphorus resonance attached to the 3′-oxygen of A4. Thus the backbone distortions of both mismatched base pairs are evidenced on the 5′-side of the mismatched bases, pG5 and pA6.

The downfield shift of a single phosphorus resonance has been observed in other distorted duplex sequences as well. The extrahelical adenosine-containing duplexes d-(CGCGAAATTT*A*Cp*GCG) and d(CCGGAATTC*A*Cp*GG) previously studied [Hare et al. (1986b) and Kalnik et al. (1989a), respectively] both contained a downfield-shifted phosphorus peak, namely, that peak of a phosphate at a position one residue removed to the 3′-side of the extrahelical base. The extrahelical cytosine-containing duplex d-(CCGp**C*GAATTCCGG) also studied by Kalnik et al. (1989) contains no anomalously downfield-shifted ^{31}P peak; however, the phosphate attached to the 5′-oxygen of the extrahelical cytosine shifts downfield upon warming. This is unusual since warming has been observed to produce the opposite effect (i.e., an upfield movement of the affected ^{31}P resonance) for the extrahelical adenosine duplexes. Additionally, the ^{31}P assignments of several sequences containing nonstandard base pairs have been determined in our labortory (Roongta et al., 1990). These mismatches have been derived from the parent self-complementary dodecamer $d(CGX_1GAATTCX_2CG)_2$. In the A·G and C·A dodecamers ($X_1 \cdot X_2$) studied, it is the phosphorus attached to the 3′-oxygen of G2 which resonates downfield from the rest of the ^{31}P cluster. However in the T·G and U·G mismatches, it is the phosphorus resonance of the phosphate on the 5′-side of the mismatched T and U (p11) which becomes the most downfield-shifted peak. Finally, the G·G mismatch shows a unique downfield shift of the T8p peak (one residue removed to the 5′-side of the mismatched G10 residue). It is thus clear that there appear to be no rules which would allow the prediction of where distortion relief might occur along the phosphate backbone (as monitored by ^{31}P

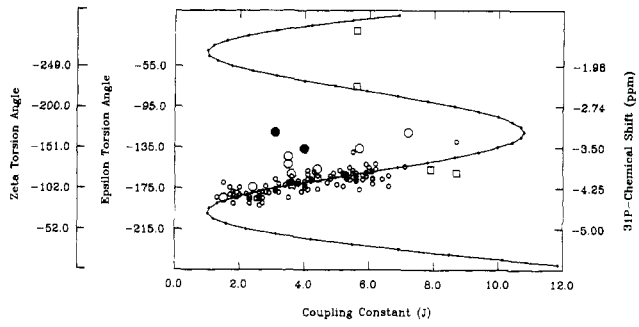

FIGURE 10: Plot of ^{31}P chemical shifts for the decamer sequence (○ and ●), oligonucleotide sequences previously studied in our laboratory (o) (Roongta et al., 1990), and the actinomycin D bound tetramer $d(CGCG)_2$ (□) (Delepierre et al., 1989) with their measured $J_{H3'-P}$ coupling constants [the two (●) data points correspond to the second 2D *J*-resolved signals for the A4 and G5 phosphates]. Also shown are the theortical ϵ and ζ torsion angles (solid curve) as a function of coupling constant derived from the Karplus relationship (ϵ) and the ϵ–ζ relation $\zeta = -317 - 1.23\epsilon$.

NMR) in molecules such as those described above.

Torsion Angles. Two of the more important parameters controlling ^{31}P chemical shifts in phosphate esters are the P–O torsional angles [in nucleic acids the P–O5′ (α) and P–O3′ (ζ) torsional angles (Gorenstein, 1987; Gorenstein & Kar, 1975) and the C–O5′ (β) and C–O3′ (ϵ) torsional angles (Giessner-Prettre et al., 1984; Ribas-Prado et al., 1979)], although the P–O torsional angle may be more important. Using the selective 2D *J*-resolved correlation experiment, we can measure the three-bond H3′–C3′–O–P coupling constant (listed in Table II) for each of the phosphates in the decamer (Gorenstein et al., 1988; Sklenar & Bax, 1987).

Our laboratory has postulated that the ^{31}P chemical shift observed in duplex B-DNA is to a large extent determined by the conformation about the backbone torsion angle ζ (C3′–O3′–P–O5′). We have previously demonstrated a correlation between the backbone torsion angle ϵ and ^{31}P chemical shifts in oligonucleotides using a number of DNA duplexes (Roongta et al., 1990; Gorenstein et al., 1988). Since a linear relation between ϵ and ζ is assumed, we have chosen to use the experimentally determined torsion angle ϵ. However, it has also been observed that in a few cases this correlation is not obeyed for phosphates adjacent to distorted DNA sites (Nikonowicz et al., 1989). Figure 10 depicts the torsion angle ϵ as a function of the H3′–P coupling constant [i.e., the Karplus relationship of Lankhorst et al. (1984)] and the observed variation of ^{31}P chemical shift with measured coupling constant for 12 DNA duplexes, including the dodecamer $d(CGCGAATTCGCG)_2$ (Ott & Eckstein, 1985a; Sklenar & Bax, 1987), the tridecamer $d(CGCAGAATTCGCG)_2$ (Nikonowicz et al., 1989), the decamer $d(CCAAGATTGG)_2$, and others from our labortory (Roongta et al., 1990).

A good linear correlation is seen between ^{31}P chemical shifts and P–H3′ coupling constants ($R = 0.82$). However, three of the nine resonances of the decamer are signficantly displaced to the downfield side of the curve (position C2 and the smaller coupling constants of A4 and G5) as well as two resonances from an extrahelical tridecamer. The remaining six phosphate positions of the decamer fall within two standard deviations (determined with the non-decamer data) of the curve. Several explanations can be offered for the deviations of the decamer from the overall correlation that we have previously observed between ^{31}P chemical shifts and phosphate ester torsional angles (Figure 10).

(1) The Karplus relationship provides for four different torsional angle solutions for each value of the coupling constant. Although all four values are shown in Figure 10, the limb which includes ϵ values between 0° and −270° is sterically inaccessible in nucleic acids (Saenger, 1984). We have assumed that the coupling constant, which can vary only between 1 and 11 Hz, all correspond to ϵ torsional angles on a single limb nearest the crystallographically observed average for nucleic acids. This need not always be true, especially since phosphate ester conformations corresponding to either of the other two solutions are observed crystallographically (Saenger, 1984). Indeed, it has recently been suggested (Delepierre et al., 1989) that the correlation between ^{31}P chemical shifts and ϵ/ζ torsional angles is not obeyed for a tetramer duplex–actinomycin D complex. A problem with the analysis offered by Delepierre et al. is that they fail to take into account the multiple solutions to the Karplus relationship. It is quite interesting that although the measured ^{31}P chemical shifts and coupling constants for the two phosphates at the binding site of the intercalation of the drug do not fit on the main limb of the Karplus curve, the values do fit quite well on the other two solutions of the curve [(□) Figure 10]. Indeed, crystal structures (Saenger, 1984) of related intercalator duplex complexes confirm that the phosphates often are constrained to conformations that best correspond to the other solutions. These intercalator drug–duplex complexes quite generally show large downfield shifts of the ^{31}P signals (Gorenstein & Goldfield, 1984). However as shown in Figure 10 the points for the decamer that significantly deviate from the curve do not fit any of the other solutions as well.

(2) Alternatively, perhaps the explanation for the "anomalous" behavior of these ^{31}P signals of phosphates neighboring the mismatch site is attributable to the observed chemical exchange behavior of these peaks. Note that all three of the decamer signals that do not fit the correlation very well (C2, A4, and G5 phosphates) are the furthest downfield-shifted signals which undergo a premelting transition (Figure 5C). Also, of the two different coupling constants measured for each of the downfield-shifted signals A4 and G5, the larger set of coupling constants better fits the curve in Figure 10. Indeed at higher temperatures, the predicted correlation is more closely approached as chemical exchange allows for averaging of the chemical shifts and coupling constants (plots not shown).

(3) Finally, perturbations in ^{31}P shifts can also arise from variations in the α (O3′–P–O5′–C5′) torsion angle or even O–P–O bond angles (Gorenstein, 1984). While the ζ torsional angle is found to be the most variable one in the B-form of the double helix, the other P–O torsional angle (α) is one of the most variable in the A-form of the duplex (Saenger, 1984). However, the unusual geometry of the mismatch could allow significant variation of the α and β (P–O5′–C5′–C4′) torsional angles as well. Unfortunately, as previously noted, these angles have in practice been unobtainable even in moderate-length duplexes. Broad, overlapping lines at low temperature and the inability to stereospecifically assign H5′ and H5″ are significant barriers to deriving useful quantitative data. However, in an effort to determine whether any outstanding anomalies may exist in the β torsion angles, the 2D J_{HP}-correlated spectra of the H5′/H5″ region were collected. These spectra were recorded for the duplexes $d(CGTGAATTCGCG)_2$ and $d(CCAAGATTGG)_2$ with the same pulse sequence as was used for the 3′-region. The former duplex was chosen because it contains a G·T mismatch yet fits the ^{31}P chemical shift/J_{HP} 3′ (ϵ torsion angle) correlation quite well. Imperfect π inversion pulses (generated by the Dante pulse train) due to tailing outside the 180° excited regions of interest produce a "center peak" in all spectra. At low tem-

perature (20 °C for the decamer and 30 °C for the dodecamer) the multiplet pattern was not resolved well enough so that accurate coupling constants could be obtained. However, it was possible to determine the difference (in hertz) between the two peaks centered on either side of $\omega = 0$ in the multiplet pattern. This difference in the decamer was ~2.6 Hz for the six most upfield resonances and ~2.2 Hz for the three remaining downfield resonances (±0.1 Hz). The G·T mismatched dodecamer (Roongta et al., 1990) had values of ~2.6 Hz for upfield resonances while the two most downfield resonances were centered about 2.9 Hz. Although there appears to be a slight variation in the differences described above, we are not able to unambiguously say that α and β have significantly contributed to the ^{31}P chemical shift dispersion. However, we can say that the β angle is to some degree distorted at the site of the mismatch and at residues 5′ to that site.

We have indicated earlier that ^{31}P chemical shifts are perturbed by factors other than torsional angle changes alone. As has previously been shown, ^{31}P chemical shifts are also very sensitive to bond angle distortions as well (Gorenstein, 1975, 1984). It is quite reasonable to assume that the unusual distortions from such structures as the mismatches and extrahelical base duplex not only display unusual torsional angle variations but also introduce some bond angle distortion to further widen the base step to accommodate the distorted geometry. Widening of the ester O–P–O bond angle is indeed expected to produce an *upfield* shift (Gorenstein, 1975, 1984) while narrowing of this bond angle causes a *downfield* shift, and it is possible this bond angle effect could account for the anomalous shifts. Thus, we conclude that for "normal" B-DNA geometry there is an excellent correlation between the phosphate resonances and the observed torsion angle while phosphates which are greatly distorted in their geometry may only show a modest correlation. Note, however, that a *combination* of both ^{31}P chemical shft and P–H3′ coupling constant can best be used to define the phosphate conformation.

^{31}P Melting Curves. As previously noted, the ^{31}P peaks of the main cluster move downfield (Figure 5A) as the sample is heated. In a right-handed B-type DNA duplex, the PO ester conformation is g^-,g^- (B_I); heating of the duplex to above the T_m produces increased conformational flexibility along the phosphate backbone. This is in agreement with previous observations (Gorenstein, 1984) that the non-gauche conformation of the PO ester resonates downfield of the g^-,g^- form. We have already pointed out that the two most downfield resonances move upfield upon heating and eventual melting of the duplex. The melting transition appears to be quite broad and would appear to have a range of ~15 °C.

J-Correlated spectra collected at 60 and 80°C indicate complete melting of the duplex as the 3′-coupling constants become nearly uniform for all resonances (5.3–6.9 Hz and 5.7–7.1 Hz at 60 and 80 °C, respectively; see Figure 6B). It was also possible to measure the H4′ to phosphorus coupling constants at elevated tempertures, ranging from 2.1 to 2.2 Hz. The H5′/H5″to P coupling constants however proved to be unobtainable due to the effects of the strong H5′–H5″ coupling and their nearly degenerate chemical shifts.

NOESY Distance-Restrained Structure. Four initial model-built structures for the tandem G·A mismatched decamer duplex were considered. Arnott B-DNA and A-DNA structures were generated with AMBER and were then manipulated with MIDAS to create structures where the adenine was either in a syn or anti conformation. These structures were then energy minimized with AMBER and used as four initial

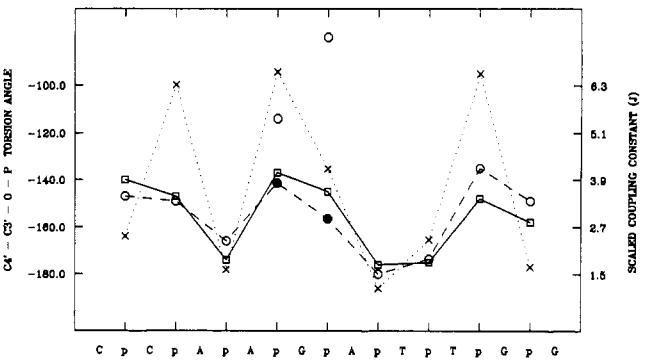

FIGURE 11: Plot of the ϵ torsional angle from the crystal structure (×) and final NOESY distance-restrained solution structure (□) vs phosphate position along the 5′ → 3′ strand for duplex decamer. The experimental $J_{H3'-P}$ coupling constant (○) vs phosphate position is also shown [the coupling constants for the additional resolved doublets for the A4p and G5p phosphates are also shown (●). The $J_{H3'-P}$ coupling constant vs sequence plot has been scaled to reflect the ϵ torsion angle variations. The solution structure ϵ torsional angles were calculated by averaging over 160 ps of the hybrid matrix distance-restrained/MD refinements [see text and Nikonowicz et al. (1990)].

model-built starting structures. The initial models were then reminimized with 162 NOESY two-spin-type distance constraints[1] and six hydrogen bonds, one for each imino proton of the three termial residues on either end. Each of the four minimized starting structures was then subjected to 6 ps of gas-phase restrained molecular dynamics with position coordinates stored every 50 fs [see Nikonowicz et al. (1990) for experimental approach]. The coordinates for the last 2 ps were position averaged and the averaged structures minimized with contraints. At this point all four structures contained the G_{anti}–A_{anti} conformation at the mismatched site. Interestingly, unconstrained molecular dynamics carried out on the original structures, without previous constrained minimization, maintained their original conformation at the site of the mismatch, i.e., $A_{syn} \rightarrow A_{syn}$ and $A_{anti} \rightarrow A_{anti}$. A detailed analysis of the decamer's structure will be presented in Nikonowicz et al. (1990).

Comparison of the Phosphate Backbone Torsional Angle Variations from Solution Coupling Constants, Restrained Molecular Dynamics Calculations, and the X-ray Crystal Structure. ^{1}H/^{1}H 2D NOESY spectra give no direct information on the sugar–phosphate conformation and NOESY distance-restrained structures have been suggested to be effectively disordered in this part of the structure (Van De Ven & Hilbers, 1988). The nonrandom variation in ^{31}P chemical shifts and $J_{H3'-P}$ coupling constants demonstrates clearly that the backbone is not disordered thoughout the sequence. Remarkably, as shown in Figure 11, the NOESY distance-restrained molecular dynamics (MD) calculations are able to reproduce the observed variation in the ϵ torsional angles for the decamer and rather accurately parallel the variation in ^{31}P chemical shifts and $J_{H3'-P}$ coupling constants. These variations were *not* reproduced in molecular dynamics calculations

[1] As noted earlier, the A4 and G5 ^{31}P resonances each appear to consist of two overlapping resonances that are only resolved in the 2D *J*-resolved spectrum. The multiplicity and broadness for these resonances in the 1D ^{31}P NMR spectrum is presumably attributable to chemical exchange between different conformations about the C–O–P bonds. While the A6H8 proton is broader than expected, no other evidence for this chemical exchange behavior between conformations is found in the ^{1}H NMR spectra. We have thus assumed that the distances derived from the 2D NOESY spectra best represent a "single" structure or an average from two or more rapidly interconverting similar conformations that predominantly or only differ in several P–O ester conformations.

without the hybrid matrix/NOESY distance restraints (imino H-bond restraints were added to prevent strand separation).

The ϵ torsional angles were calculated by averaging over the entire 40-ps dynamics of each of the four different hybrid matrix distance-restrained/MD refinements (a total of 160 ps). However the results were essentially the same for each of the 40-ps runs. The torsional angles from the MD simulations are plotted as a function of sequence in Figure 11. It is clear that a remarkably strong correlation exists between the pattern of the variation of the torsional angles derived from the restrained MD calculations and the experimentally measured coupling constants (the fit is poorer to the alternate conformation in Figure 11). These results provide further support for the reliability of the hybrid marix/restrained MD refinement methodology to accurately reproduce the solution structure. Similar correlations between the averaged torsional angles from restrained MD refinement of a canonical B-DNA decamer have been demonstrated (Powers et al., 1990).

It is important to note that the ϵ torsional angles in Figure 11 derived from the restrained MD calculations represent time averages during which large-amplitude fluctuations are observed on the picosecond time scale (Powers et al., 1990). The time course of the restrained MD calculated fluctuations about the six sugar–phosphate torsional angles in the decamer shows relatively small-amplitude fluctuations about the average B-DNA values for the 40-ps restrained molecular dynamics calculation except for the ϵ and ζ torsional angles (data not shown), consistent with the X-ray studies (Dickerson, 1983; Dickerson & Drew, 1981) of B-DNA. These torsional angles changes reflect a transition from the low-energy B_{I} conformation to the higher energy B_{II} conformation in a crankshaft motion. These transitions are short lived and relax back to the low-energy conformation.

Most remarkably, Figure 11 demonstrates that the variation in the backbone conformation in the crystal state also follows the same pattern of variation in the measured coupling constants and the averaged torsional angles derived from the restrained MD refinement. NMR studies have suggested that the duplex conformation in solution may not be identical with the static picture provided by X-ray diffraction in the crystal state (Sklenář et al., 1986; Gorenstein et al., 1989). This has raised the question whether some of the structural variations observed in the X-ray crystallographic studies are the result of less profound crystal packing forces (Dickerson et al., 1987; Jain & Sundaralingam, 1989). Indeed, Dickerson et al. (1987) have suggested that all of the sequence-specific variation in the phosphate ester B_{I} and B_{II} conformations arises from crystal packing forces. Similar conclusions have been reached in a Raman spectroscopy analysis of the backbone phosphate conformation in solution and the solid state (Benevides et al., 1988). Our own results clearly show that there are significant variations in the phosphate ester conformations in solution, comparable to those seen in the crystal state of the decamer (although as shown in Figure 11 the amplitude of the variation appears to be larger in the crystal state). Thus while it has been argued that variations in the backbone conformation only reflects crystal packing forces, our own results clearly demonstrate that similar variations of the backbone conformation are observed in both the crystal and solution state.

ACKNOWLEDGMENTS

We greatly appreciate the contributions of Dr. Claude Jones, Patricia Fagan, and Dr. Vikram Roongta.

Registry No. d(CCAAGATTGG), 86880-63-9; G, 73-40-5; A, 73-24-5.

REFERENCES

Bax, A., & Morris, G. A. (1981) *J. Magn. Reson. 42*, 501–505.

Bax, A., & Freeman, R. (1982) *J. Am. Chem. Soc. 104*, 1099–1100.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Benevides, J. M., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H., & Thomas, G. J., Jr. (1988) *Biochemistry 27*, 931–938.

Braunschweiler, L., & Earnst, R. R. (1983) *J. Magn. Reson. 71*, 521.

Broido, M. A., Zon, G., & James, T. L. (1984) *Biochem. Biophys. Res. Commun. 119*, 663–670.

Brown, T., Hunter, W. N., Kneale, G., & Kennard, O. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 2402–2406.

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

Celda, B., Widmer, H., Leupin, W., Chazin, W. J., Denny, W. A., & Wuthrich, K. (1989) *Biochemistry 28*, 1462–1471.

Cheng, D. M., Kan, L.-S., Miller, P. S., Leutzinger, E. E., & Ts'o, P. O. P. (1982) *Biopolymers 21*, 697.

Cheng, D. M., Kan, L., & Ts'o, P. O. P. (1987) *Phosphorus NMR in Biology* (Burt, C. T., Ed.) pp 135–147, CRC Press, Boca Raton, FL.

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

Connolly, B. A., & Eckstein, F. (1984) *Biochemistry 23*, 5523–5527.

Delepierre, M., Van Heijenoort, C., Igolen, J., Pothier, J., Le Bret, M., & Roques, B. P. (1989) *J. Biomol. Struct. Dyn. 7*, 557–589.

Dickerson, R. E. (1983) *J. Mol. Biol. 166*, 419–441.

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

Dickerson, R. E., Goodsell, D. S., Kopka, M. L., & Pjura, P. E. (1987) *J. Biomol. Struct. Dyn. 5*, 557–579.

Dodgson, J. B., & Wells, R. D. (1977a) *Biochemistry 16*, 2367–2374.

Dodgson, J. B., & Wells, R. D. (1977b) *Biochemistry 16*, 2374–2379.

Fazakerley, G. V., Quignard, E., Woisard, A., Guschlbauer, W., van der Marel, G. A., van Boom, J. H., Jones, M., & Radman, M. (1986) *EMBO J. 5*, 3697–3703.

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

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

Ferrin, T. E., & Langridge, R. (1980) *Computer Graphics 13*, 320.

Fu, J. M., Schroeder, S. A., Jones, C. R., Santini, R., & Gorenstein, D. G. (1988) *J. Magn. Reson. 77*, 577–582.

Gao, S., & Patel, D. J. (1988) *J. Am. Chem. Soc. 110*, 5178–5182.

Giessner-Prettre, C., Pullman, B., Ribas-Prado, F., Cheng, D. M., Iuorno, V., & Ts'o, P. O. P. (1984) *Biopolymers 23*, 377.

Gorenstein, D. G. (1975) *J. Am. Chem. Soc. 97*, 898–900.

Gorenstein, D. G. (1978) *Jerusalem Symposium, NMR in Molecular Biology* (Pullman, B., Ed.) pp 1–15, D. Riedel, Dordrecht, The Netherlands.

Gorenstein, D. G. (1981) *Annu. Rev. Biophys. Bioengin. 10*, 355.

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., & Kar, D. (1975) *Biochem. Biophys. Res. Commun. 65*, 1073–1080.
Gorenstein, D. G., & Goldfield, E. M. (1984) *Phosphorus-31 NMR: Principles and Applications* (Gorenstein, D. G., Ed.) Academic Press, New York.
Gorenstein, D. G., Findlay, J. B., Momii, R. K., Luxon, B. A., & Kar, D. (1976) *Biochemistry 15*, 3796–3803.
Gorenstein, D. G., Lai, K., & Shah, D. O. (1984) *Biochemistry 23*, 6717.
Gorenstein, D. G., Schroeder, S. A., Fu, J. M., Metz, J. T., Roongta, V. A., & Jones, C. R. (1988) *Biochemistry 27*, 7223–7237.
Gorenstein, D. G., Meadows, R. P., Metz, J. T., Nikonowicz, E., & Post, C. B. (1990) *Advances in Physical Chemistry* (Bush, C. A., Ed.) JAI Press, Greenwich, CT (in press).
Hare, D., Shapiro, L., & Patel, D. J. (1986a) *Biochemistry 25*, 7445–7456.
Hare, D., Shapiro, L., & Patel, D. J. (1986b) *Biochemistry 25*, 7456–7464.
Hare, D. R., & Reid, B. R. (1982) *Biochemistry 21*, 1835–1842.
Hare, D. R., Wemmer, D. E., Chou, S. H., Drobny, G., & Reid, B. (1983) *J. Mol. Biol. 171*, 319.
Hunter, W. N., Brown, T., & Kennard, O. (1987a) *Nucleic Acids Res. 15*, 6589–6606.
Hunter, W. N., Brown, T., Kneale, G., Anand, N. N., Rabinovich, D., & Kennard, O. (1987b) *J. Biol. Chem. 262*, 9962–9970.
Hurd, R. E., & Reid, B. R. (1979) *Biochemistry 18*, 4017.
Jain, S., & Sundaralingam, M. (1989) *J. Biol. Chem. 264*, 12780–12784.
Johnston, P. D., & Redfield, A. G. (1978) *Nucleic Acids Res. 5*, 3913.
Jones, C. R., Schroeder, S. A., & Gorenstein, D. G. (1988) *J. Magn. Reson. 80*, 370–374.
Joshua-Tor, L., Rabinovich, D., Hope, H., Frolow, F., Appella, E., & Sussman, J. L. (1988) *Nature 334*, 82–84.
Kalnik, M. W., Norman, D. G., Zagorski, M. G., Swann, P. F., & Patel, D. J. (1989) *Biochemistry 28*, 294–303.
Kan, L.-S., Cheng, D. M., Jayaraman, K., Leutzinger, E. E., Miller, P. S., & Ts'o, P. O. P. (1982) *Biochemistry 21*, 6723–6732.
Kan, L., Chandrasegaran, S., Pulford, S. M., & Miller, P. S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4263–4265.
Keepers, J., & James, T. (1984) *J. Magn. Reson. 57*, 404.
Kessler, H., Griesinger, C., Zarbock, J., & Loosli, H. R. (1984) *J. Magn. Reson. 57*, 331–336.
Kouchakdjian, M., Li, B. F. L., Swann, P. F., & Patel, D. J. (1988) *J. Mol. Biol. 202*, 139–155.
Lai, K., Shah, D. O., Derose, E., & Gorenstein, D. G. (1984) *Biochem. Biophys. Res. Commun. 121*, 1021.
Lankhorst, P. P., Haasnoot, C. A. G., Erkelens, C., & Altona, C. (1984) *J. Biomol. Struct. Dyn. 1*, 1387–1405.
Lu, A. L., Welsh, K., Clark, S., Su, S. S., & Modrich, P. (1984) *Cold Spring Harbor Symp. Quant. Biol.*, 589–596.
Modrich, P. (1987) *Annu. Rev. Biochem. 56*, 435–466.
Nikonowicz, E., Roongta, V., Jones, C. R., & Gorenstein, D. G. (1989) *Biochemistry 28*, 8714–8725.
Nikonowicz, E. P., Meadows, R. P., & Gorenstein, D. G. (1990) *Biochemistry 29*, 4193–4204.
Ott, J., & Eckstein, F. (1985a) *Biochemistry 24*, 253.
Ott, J., & Eckstein, F. (1985b) *Nucleic Acids Res. 13*, 6317–6330.
Pardi, A., Walker, R., Rapoport, H., Wider, G., & Wuthrich, K. (1983) *J. Am. Chem. Soc. 105*, 1652.
Patel, D. J. (1974) *Biochemistry 13*, 2396–2402.
Patel, D. J. (1979) *Acc. Chem. Res. 12*, 118–125.
Patel, D. J., Kolowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Dallas, C., Itakura, K., & Breslauer, K. J. (1982a) *Biochemistry 21*, 437–444.
Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Itakura, K., & Breslauer, K. J. (1982b) *Biochemistry 21*, 445–451.
Patel, D. J., Kozlowski, S. A., Ikuta, S., & Itakura, K. (1984a) *Biochemistry 23*, 3207–3217.
Patel, D. J., Kolowski, S. A., Ikuta, S., & Itakura, K. (1984b) *Biochemistry 23*, 3218–3226.
Patel, D. J., Kozlowski, S. A., Ikuta, S., & Itakura, K. (1984c) *Fed. Proc. 43*, 2663–2670.
Petersheim, M., Mehdi, S., & Gerlt, J. A. (1984) *J. Am. Chem. Soc. 106*, 439–440.
Pon, R. T., Usman, N., & Ogilvie, K. K. (1988) *Biotechniques 6*, 768–775.
Powers, R., Jones, C. R., & Gorenstein, D. G. (1990) *J. Biomol. Struct. Dyn.* (in press).
Prive, G. G., Heinemann, U., Chandrasegaran, S., Kan, L., Kopka, M. L., & Dickerson, R. E. (1987) *Science 238*, 498–504.
Quigley, G. J., Wang, A. H.-J., Ughetto, G., van der Marel, G., van Boom, J. H., & Rich, A. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 7204–7208.
Radman, M., & Wagner, R. (1984) *Curr. Top. Microbiol. Immunol. 108*, 23–28.
Radman, M., & Wagner, R. (1986) *Annu. Rev. Genet. 20*, 523–538.
Ribas-Prado, F., Giessner-Prettre, C., Pullman, B., & Daudey, J.-P. (1979) *J. Am. Chem. Soc. 101*, 1737.
Roongta, V. A., Jones, C. R., & Gorenstein, D. G. (1990) (submitted for publication).
Roy, S., Sklenar, V., Appella, E., & Cohen, J. S. (1987) *Biopolymers 26*, 2041–2052.
Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
Scheek, R. M., Boelens, R., Russo, N., Van Boom, J. H., & Kaptein, R. (1984) *Biochemistry 23*, 1371–1376.
Schroeder, S., Jones, C., Fu, J., & Gorenstein, D. G. (1986) *Bull. Magn. Reson. 8*, 137–146.
Schroeder, S., Fu, J., Jones, C., & Gorenstein, D. G. (1987) *Biochemistry 26*, 3812–3821.
Schroeder, S. A., Roongta, V., Fu, J. M., Jones, C. R., & Gorenstein, D. G. (1989) *Biochemistry 28*, 8292–8303.
Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Park Kim, J. J., & Rich, A. (1976) *J. Mol. Biol. 104*, 142–143.
Shah, D. O., Lai, K., & Gorenstein, D. G. (1984a) *Biochemistry 23*, 6717–6723.
Shah, D. O., Lai, K., & Gorenstein, D. G. (1984b) *J. Am. Chem. Soc. 106*, 4302.
Sklenar, V., & Bax, A. (1987) *J. Am. Chem. Soc. 109*, 7525–7526.
Sklenář, V., Miyashiro, H., Zon, G., Miles, H. T., & Bax, A. (1986) *FEBS Lett. 208*, 94–98.
States, D. J., Haberkorn, R. A., & Rueben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Van De Ven, F. J. M., & Hilbers, C. W. (1988) *Eur. J. Biochem. 178*, 1–38.
Weiner, P. K., & Kollman, P. A. (1981) *J. Comput. Chem. 2*, 287–303.

Weiss, M. A., Patel, D. J., Sauer, R. T., & Karplus, M. (1984) *J. Chem. Soc. 106*, 4269–4270.
Wemmer, D. E., Chou, S.-H., Hare, D. R., & Reid, B. R. (1984) *Biochemistry 23*, 2262–2268.
Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Zhou, N., Manogaran, S., Zon, G., & James, T. L. (1988) *Biochemistry 27*, 6013–6020.

^{1}H NMR Assignment and Melting Temperature Study of Cis-Syn and Trans-Syn Thymine Dimer Containing Duplexes of d(CGTATTATGC)·d(GCATAATACG)†

John-Stephen Taylor,*,‡,§ Daniel S. Garrett,‡ Ian R. Brockie,‡ Daniel L. Svoboda,‡ and Joshua Telser‖
Department of Chemistry, Washington University, St. Louis, Missouri 63130, and Amoco Technology Company, P.O. Box 400, Naperville, Illinois 60566

Received January 24, 1990; Revised Manuscript Received June 4, 1990

ABSTRACT: The preparation and spectroscopic characterization of duplex decamers containing site-specific cis-syn and trans-syn thymine dimers are described. Three duplex decamers, d(CGTATTATGC)·d(GCATAATACG), d(CGTAT[c,s]TATGC)·d(GCATAATACG), and d(CGTAT[t,s]TATGC)·d(GCATAATACG), were prepared by solid-phase phosphoramidite synthesis utilizing cis-syn and trans-syn cyclobutane thymine dimer building blocks (Taylor et al., 1987; Taylor & Brockie, 1988). NMR spectra (500 MHz 2D ^{1}H and 202 MHz 1D ^{31}P) were obtained in "100%" D_2O at 10 °C, and 1D exchangeable ^{1}H spectra were obtained in 10% D_2O at 10 °C. ^{1}H NMR assignments for H5, H6, H8, CH_3, H1′, H2′, and H2″ were made on the basis of standard sequential NOE assignment strategies and verified in part by DQF COSY data. Comparison of the chemical shift data suggests that the helix structure is perturbed more to the 3′-side of the cis-syn dimer and more to the 5′-side of the trans-syn dimer. Thermodynamic parameters for the helix ⇄ coil equilibrium were obtained by two-state, all or none, analysis of the melting behavior of the duplexes. Analysis of the temperature dependence of the T5CH_3 ^{1}H NMR signal gave $\Delta H = 44 \pm 4$ kcal and $\Delta S = 132 \pm 13$ eu for the trans-syn duplex. Analysis of the concentration and temperature dependence of UV spectra gave $\Delta H = 64 \pm 6$ kcal and $\Delta S = 178 \pm 18$ eu for the parent duplex and $\Delta H = 66 \pm 7$ kcal and $\Delta S = 189 \pm 19$ eu for cis-syn duplex. It was concluded that photodimerization of the dTpdT unit to give the cis-syn product causes little perturbation of the DNA whereas dimerization to give the trans-syn product causes much greater perturbation, possibly in the form of a kink or dislocation at the 5′-side of the dimer.

Cis-syn cyclobutane pyrimidine dimers are the major photoproducts produced upon exposure of DNA to sunlight [see Patrick and Rahn (1976) for a review] and have been correlated with mutation and skin cancer [see Brash (1988) for a recent discussion]. These photoproducts result from the photo [2 + 2] cycloaddition of the 5,6-double bond of two adjacent pyrimidine nucleotides in an anti glycosyl conformation. The cis-syn dimer of dTpdT[1] sequences is the major photoproduct of DNA, and as a consequence there is much interest in determining its structure–activity relationships. Cis-syn thymine dimers are the substrates for photolyases, pyrimidine dimer glycosylases, and the uvrABC excinuclease system (Friedberg, 1985). Cis-syn thymine dimers generally constitute blocks to replication, but recently it has been found that they can be bypassed by pol I of *Escherichia coli* in vitro in a nonmutagenic manner (Taylor & O'Day, 1990). The structural or physical properties of DNA containing the cis-syn dimer which contribute to its recognition by repair systems and its bypass by replication systems are unknown.

The cis-syn dimer of dTpdT itself has been extensively studied by NMR (Ogilvie, 1975; Liu & Yang, 1978; Rycyna & Alderfer, 1985, Kemmink et al., 1987c; Kan et al., 1988). The crystal structure of a phosphotriester derivative of the cis-syn dimer of dTpdT has been determined by X-ray crystallography and found to be somewhat different than what the solution-state NMR data indicate (Cadet et al., 1985; Hruska et al., 1986). NMR studies have also been reported on cis-syn dimer containing tri- and tetrathymidylates (Rycyna et al., 1988). Milligram quantities of d(GCGT[c,s]TGCG)·d(CGCAACGC) have been prepared by a direct photolysis route and the nonexchangeable ^{1}H NMR signals assigned (Kemmink et al., 1987a). In a separate study the exchangeable proton signals of the parent and cis-syn octamer duplexes were assigned and thermodynamic parameters for the helix ⇄ coil transition determined by analysis of the temperature dependence of one of the methyl proton signals (Kemmink et al., 1987b).

† This investigation was supported by USPHS Grant CA40463, awarded by the National Cancer Institute, DHHS. The assistance of the Washington University High-Resolution NMR Service Facility, funded in part through NIH Biomedical Research Support Shared Instrument Grant 1 S10 RR02004, and a gift from the Monsanto Co. are gratefully acknowledged.

* To whom correspondence should be addressed.
‡ Washington University.
§ Alfred P. Sloan Foundation Fellow, 1988–1990.
‖ Amoco Technology Co. Present address: Squibb Institute for Medical Research, P.O. Box 191, New Brunswick, NJ 08903.

[1] Abbreviations: COSY, correlated spectroscopy; 2D, two dimensional; DQF, double quantum filtered; EDTA, ethylenediaminetetraacetic acid; NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; TMP, trimethyl phosphate; dTpdT, thymidylyl-(3′→5′)-thymidine; T[c,s]T, cis-syn cyclobutane dimer of dTpdT; T[t,s]T, trans-syn cyclobutane dimer of dTpdT; TSP, 3-(trimethylsilyl)propionate.

0006-2960/90/0429-8858$02.50/0