Biochemistry (submitted for publication).
Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1986) *Numerical Recipes*, pp 289–293, Cambridge University Press, Cambridge.
Redfield, C., & Dobson, C. M. (1988) *Biochemistry 27*, 122–136.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Torchia, D. A., Sparks, S. W., & Bax, A. (1989) *Biochemistry 28*, 5509–5524.
Wagner, G., Pardi, A., & Wüthrich, K. (1983) *J. Am. Chem. Soc. 105*, 5948–5949.
Wand, A. J., DiStefano, D. L., Feng, Y., Roder, H., & Englander, S. W. (1989) *Biochemistry 28*, 186–194.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Determination of the Conformation of d(GGAAATTTCC)$_2$ in Solution by Use of ^{1}H NMR and Restrained Molecular Dynamics†

Masato Katahira,‡ Hiromu Sugeta, and Yoshimasa Kyogoku*
Institute for Protein Research, Osaka University, Suita, Osaka 565, Japan
Satoshi Fujii
Faculty of Pharmaceutical Sciences, Osaka University, Suita, Osaka 565, Japan
Received January 4, 1990; Revised Manuscript Received April 18, 1990

ABSTRACT: The conformation of the putative bent DNA d(GGAAATTTCC)$_2$ in solution was studied by use of ^{1}H NMR and restrained molecular dynamics. Most of the resonances were assigned sequentially. A total of 182 interproton distance restraints were determined from two-dimensional nuclear Overhauser effect spectra with short mixing times. Torsion angle restraints for each sugar moiety were determined by qualitative analysis of a two-dimensional correlated spectrum. Restrained molecular dynamics was carried out with the interproton distances and torsion angles incorporated into the total energy function of the system in the form of effective potential terms. As initial conformations for restrained molecular dynamics, classical A-DNA and B-DNA were adopted. The root mean square deviation (rmsd) beteween these two conformations is 5.5 Å. The conformations obtained by use of restrained molecular dynamics are very similar to each other, the rmsd being 0.8 Å. On the other hand, the conformations obtained by use of molecular dynamics without experimental restraints or restrained energy minimization depended heavily on the initial conformations, and convergence to a similar conformation was not attained. The conformation obtained by use of restrained molecular dynamics exhibits a few remarkable features. The second G residue takes on the BII conformation [Fratini, A. V., Kopka, M. L., Drew, H. R., & Dickerson, R. E. (1982) *J. Biol. Chem. 257*, 14686–14707] rather than the standard BI conformation. There is discontinuity of the sugar puckering between the eighth T and ninth C. The minor groove of the oligo(dA) tract is rather compressed. As a result, d(GGAAATTTCC)$_2$ is bent.

It has been demonstrated that DNAs which contain an oligo(dA) tract are bent, by means of electrophoresis (Wu & Crothers, 1984; Hagerman, 1985, 1986; Koo et al., 1986; Haran & Crothers, 1989) and electron microscopy (Griffith et al., 1986). Several models for DNA bending have been proposed (Koo et al., 1986; Ulanovsky & Trifonov, 1987; Burkhoff & Tullius, 1987; Koo & Crothers, 1988; Calladine et al., 1988), but they are contradictory. Crystallographic studies (Nelson et al., 1987; Coll et al., 1987; DiGabriele et al., 1989) and spectroscopic studies (Roy et al., 1987; Kintanar et al., 1987; Katahira et al., 1988; Sarma et al., 1988; Patapoff et al., 1988; Leroy et al., 1988; Nadeau & Crothers, 1989; Celda et al., 1989) on DNAs containing the oligo(dA) tract were performed, but the correlation between the oligo(dA) tract and DNA bending remained unclear, although conformational deviations were reported in some of the papers.

To clarify the correlation, we have performed an NMR[1] study on the conformation of d(GGAAATTTCC)$_2$. When it was repeated in a synthetic polymer, macroscopic bending of the molecule was observed in an electrophoretic study (Hagerman, 1985). Thus, remarkable conformational features related to macroscopic bending are expected to be found in it with NMR spectroscopy. The conformation of d(GAAAATTTTC)$_2$, whose sequence is very similar to our molecule, has been already studied, and a conformational junction at the central AT sequence was suggested (Sarma et al., 1988). Our molecule has two base pairs at both ends, before and after the oligo(dA) tract. Therefore, it is expected

† This work was supported by a grant from the Ministry of Education, Science and Culture of Japan. M.K. was partly supported by grants from the Japan Society for the Promotion of Science for Japanese Junior Scientists and the CIBA-GEIGY Foundation for the Promotion of Science, Takarazuka, Japan.
‡ Present address: Department of Chemistry, University of Utrecht, Paduallaan 8, 3584 CH Utrecht, The Netherlands.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE spectroscopy; DQF-COSY, two-dimensional double quantum filtered correlated spectroscopy; rmsd, root mean square deviation.

0006-2960/90/0429-7214$02.50/0

that intrinsic conformational features at the junctions between the oligo(dA) tract and the preceding or following portions could also be detected in our case without being disturbed and hidden by the artificial end effect.

To determine the conformation of d(GGAAATTTCC)$_2$ in detail, we used restrained molecular dynamics (Kaptein et al., 1985; Clore et al., 1985), using the AMBER program (Weiner & Kollman, 1981), in which the information obtained with NMR spectroscopy is incorporated as effective potentials.

The numbering of residues is

	1	2	3	4	5	6	7	8	9	10	
5'	G	G	A	A	A	T	T	T	C	C	3'
3'	C	C	T	T	T	A	A	A	G	G	5'
	10	9	8	7	6	5	4	3	2	1	

Experimental Procedures

Sample Preparation. The DNA decamer d(GGAAATTTCC)$_2$ was synthesized on a 10-μmol scale with an automated version of the phosphoamidite coupling method. After deblocking of protection, the decamer was purified on a C-18 reversed-phase HPLC column and then converted into the sodium salt via the pyridine salt on an AG 50W ion exchange column (Uesugi et al., 1984). The decamer was annealed and then applied to a Sephadex G-50 column to remove single-stranded materials.

For measurement of NMR spectra in D_2O, a lyophilized sample was dissolved in 10 mM phosphate buffer (pH 7.0) containing 0.15 M NaCl. The solution was lyophilized and then dissolved in 99.96% D_2O (3.5 mM strands). For measurement in H_2O, a 9:1 H_2O:D_2O mixture was substituted for D_2O. DSS was used as an internal chemical shift reference.

NMR Spectroscopy. All NMR spectra were obtained with a JEOL GX-500 NMR spectrometer at 15 °C. Phase-sensitive NOESY (Jeener et al., 1979) and DQF-COSY (Rance et al., 1983) spectra were recorded by the methods of States et al. (1982). A phase-sensitive NOESY spectrum in H_2O was recorded with a 1-1 echo pulse (Sklenar et al., 1987). Various short mixing times, 50, 60, 90, 100, 120, and 150 ms, were used for the NOESY experiments in D_2O, and a mixing time of 50 ms was used for that in H_2O. The repetition delay was 2.0 s. The data sizes were 1024 complex points for t_2 and 256 complex points for t_1, and the t_1 data were zero-filled to 1024 points after measurement. To reduce t_1 noise, the first time domain data points were multiplied by a factor of 0.5 (Otting et al., 1986). Quantification of the intensities of cross peaks was carried out by volume integration.

Restrained Molecular Dynamics. All energy minimization and restrained molecular dynamics studies were carried out by use of the program AMBER (Weiner & Kollman, 1981) on an NEC-ACOS computer. The program AMBER was modified for restrained molecular dynamics. Effective interproton distance (E_{NOESY}) and torsion angle (E_{COSY}) restraint energy terms were incorporated in addition to the original empirical energy terms developed for nucleic acids (Weiner et al., 1986). The effective interproton distance restraint potential E_{NOESY} was given by

$$\begin{aligned} E_{NOESY} &= Wk_{NOESY}\,[r - (r_{NOESY} - a_1)]^2; && r < r_{NOESY} - a_1 \\ &= 0; && r_{NOESY} - a_1 < r < r_{NOESY} + a_2 \\ &= Wk_{NOESY}\,[r - (r_{NOESY} + a_2)]^2; && r_{NOESY} + a_2 < r \end{aligned} \quad (1)$$

where r and r_{NOESY} are the calculated and experimental interproton distances, a_1 and a_2 are the constants for defining the lower and upper limits, k_{NOESY} is an apparent force constant, and W is the weight of E_{NOESY} relative to the other energy terms. The effective torsion angle restraint potential E_{COSY} was given by

$$E_{COSY} = k_{COSY}(\phi - \phi_{COSY})^2 \quad (2)$$

where ϕ and ϕ_{COSY} are the calculated and experimental torsion angles and k_{COSY} is an apparent force constant.

With respect to the empirical electrostatic energy terms, the dielectric constant was set at $4r$, where r is the interatomic distance. This treatment reduces the electrostatic energy terms, and thus, the simulation without the explicit solvent molecules approximates the one under the effect of the solvent where the electrostatic force is shielded and weakened by the solvent. It is found that the dielectric constant $4r$ gives satisfactory results on correction for the omission of the solvent (Whitlow & Teeter, 1986). Moreover, to reduce the artificially strong repulsion among the negative charges of the phosphate groups, the phosphate groups were neutralized by placement of sodium ions (Singh et al., 1985). The sodium ions are placed 3.0 Å from each phosphorus atom on the bisection of the O–P–O angle in the initial conformations. While further treatment to keep the sodium ions near each phosphate group was not taken, the sodium ions remained near each phosphate group through the simulation. The cutoff distance of the nonbonded interaction was set at 15 Å. The temperature of the system was maintained at the intended value by gradually rescaling the velocities of the atoms. The time step of the integrator was 0.002 ps, and the nonbonded interaction lists were updated every 0.04 ps. The SHAKE algorithm (Rychaert et al., 1977) was used to maintain the lengths of bonds involving hydrogen atoms.

Results

Sequential Resonance Assignments. The resonances of d(GGAAATTTCC)$_2$ were assigned sequentially by means of two-dimensional NMR in the same way as our earlier works (Katahira et al., 1988, 1990a,b), by use of previously established methods (Reid et al., 1983; Scheek et al., 1983; Hare et al., 1983; Feigon et al., 1983; Clore & Gronenborn, 1983; Clore et al., 1984; Weiss et al., 1984). Panels a and b of Figure 1 show expansion of the NOESY spectrum with a mixing time of 150 ms in D_2O, indicating the sequential assignments of H1′ and H6/H8 through the H1′(i–1)–H6/H8(i)–H1′(i) connectivities and those of H2′ and H6/H8 through the H2′(i–1)–H6/H8(i)–H2′(i) connectivities, respectively. H2″'s were assigned in the same way (data not shown). In Figure 1b, the assignments of methyl protons of thymines through the H6/H8(i–1)–TCH$_3$(i)–TH6(i) connectivities are also shown. Figure 2a shows the assignments of H3′ and H4′ on the basis of the NOESY cross peaks between H8(i) and H3′(i)/H4′(i). The assignments were confirmed by the NOESY cross peaks between H1′(i) and H3′(i)/H4′(i) indicated in Figure 2b. All these assignments are consistent with the DQF-COSY spectrum.

The imino proton of the thymine in an A-T base pair is expected to be close to the H2 of the adenine in the same base pair and to the H2 of an adenine on the 3′ side of right-handed DNA. Such connectivities were observed in the NOESY spectrum in H_2O; thus, the imino protons of thymines and H2's of adenines were assigned sequentially (data not shown). The imino protons of guanines were assigned through T8 imino–G2 imino and G2 imino–G1 imino connectivities, and those of C9 amino protons were assigned through the G2 imino–C9 amino connectivities. The lower field resonance of two separated C9

Table I: Assignments (ppm) of Imino, Amino, and Nonexchangeable Protons of d(GGAAATTTCC)$_2$ at 15 °C

residue	H6/H8[a]	H2/H5/CH$_3$[b]	H1′	H2′	H2″	H3′	H4′	imino/amino[c]
G1	7.78		5.57	2.42	2.64	4.78	4.17	12.85
G2	7.77		5.29	2.59	2.59	4.95	4.28	12.77
A3	8.13	7.22	5.89	2.67	2.89	5.05	4.43	
A4	8.07	7.03	5.93	2.61	2.92	5.05	4.47	
A5	8.09	7.57	6.15	2.51	2.95	5.01	4.48	
T6	7.11	1.21	5.94	2.01	2.61	4.86	4.24	13.73
T7	7.44	1.49	6.14	2.20	2.66	4.91	4.24	14.00
T8	7.40	1.60	6.12	2.16	2.60	4.92	4.23	13.78
C9	7.53	5.66	5.97	2.13	2.44	4.84	4.16	8.45,[d] 6.97[e]
C10	7.50	5.50	6.17	2.26	2.26	4.55	3.97	

[a] H6 of pyrimidine bases or H8 of purine bases. [b] H2 of adenines, H5 of cytosines, or methyl protons of thymines. [c] H1 of guanines, H3 of thymines, or NH$_2$ of cytosines. [d] The hydrogen-bonded amino proton. [e] The non-hydrogen-bonded amino proton.

Table II: Intraresidue Nonexchangeable Proton Distance Restraints (Å) for d(GGAAATTTCC)$_2$

	H1′–H6/H8	H2′–H6/H8	H2″–H6/H8	H1′–H2′	H1′–H2″	H1′–H3′	H1′–H4′	H2′–H3′	H2′–H4′	H2″–H3′	H2″–H4′	H3′–H4′
G1				2.7	2.5	3.3	3.0	2.5		2.7	3.1	
G2	3.5					3.3	2.9					2.6
A3	3.5			2.8	2.5		3.0		3.4		3.4	2.6
A4					2.5		2.9		3.6			2.7
A5	3.4	2.6		2.8	2.5	3.3	3.0	2.6	3.4	2.8		2.7
T6	3.4	2.6	3.2	2.9		3.3		2.6	3.4			
T7												
T8												
C9				2.9	2.4	3.3	2.7	2.5	3.3	2.9		
C10	3.1					3.3	2.8					2.8

Table III: Interresidue Nonexchangeable Proton Distance Restraints (Å) for d(GGAAATTTCC)$_2$

	H1′–H6/H8	H2′–H6/H8	H2″–H6/H8	H3′–H6/H8	H1′–TCH$_3$	H2′–TCH$_3$	H2″–TCH$_3$	H3′–TCH$_3$	H1′–CH5
G1–G2									
G2–A3	2.8			3.8					
A3–A4									
A4–A5									
A5–T6	3.1	3.0	2.8	3.6	3.7	3.1	3.2	3.9	
T6–T7	3.1	2.9			3.6	3.0	2.9	3.8	
T7–T8								3.8	
T8–C9	3.1		2.7						3.7
C9–C10									4.1

	H2′–CH5	H2″–CH5	AH2–H1′	H6/H8–TCH$_3$	H6/H8–CH5	TCH$_3$–CH5	H6/H8–H6/H8	AH2–H1′[a]
G1–G2								
G2–A3							4.5	
A3–A4								
A4–A5								
A5–T6			3.6	3.0			4.3	
T6–T7				3.0			4.3	
T7–T8								
T8–C9	2.9	3.3			3.4	3.4		
C9–C10	3.2	3.3						
A3–C9[a]								3.8
A5–T7[a]								3.3

[a] Interstrand.

amino resonances was regarded due to a hydrogen-bonded proton and the higher field one due to a non-hydrogen-bonded one as usual. The assignments are listed in Table I.

Interproton Distance Restraints. Interproton distances were calculated on the basis of the intensities of the cross peaks in the NOESY spectrum with the 50-ms mixing time, by use of

$$r_{ij} = (a_{ref}/a_{ij})^{1/6} r_{ref} \quad (3)$$

where r_{ij} is the distance between protons i and j, r_{ref} is an internal reference distance, which is fixed geometrically, and a_{ij} and a_{ref} are the integrated volumes of corresponding cross peaks in the NOESY spectrum. For the NOESY spectrum *in* H_2O with a 1–1 echo pulse, the offset dependence of the 1–1 echo excitation was taken into account (Sklenar et al., 1987), and integrated volumes were calibrated. For interproton distances involving methyl protons, the distance between CH$_3$ and H6 of a thymine (2.8 Å) was used as the internal reference. For the others, the distance between H5 and H6 of a cytosine (2.46 Å) was used. The validity of eq 3 rests on the assumption that the initial rate condition is approximately valid and that the effective correlation times of the $i - j$ and reference vectors are about the same.

Table IV: Interstrand Proton Distance Restraints (Å) Involving Exchangeable Protons for d(GGAAATTTCC)$_2$

T6H3–A5H2	2.7	T8H3–A3H2	2.9
T7H3–A4H2	2.9	T8H3–A4H2	3.3
T7H3–A5H2	3.3	G2H1–C9amino[a]	2.8

[a] The hydrogen-bonded amino proton.

Interproton distances were calculated on the basis of the NOESY spectra with a series of short mixing times, 50, 60, 90, 100, 120, and 150 ms, by use of eq 3, and the mixing time

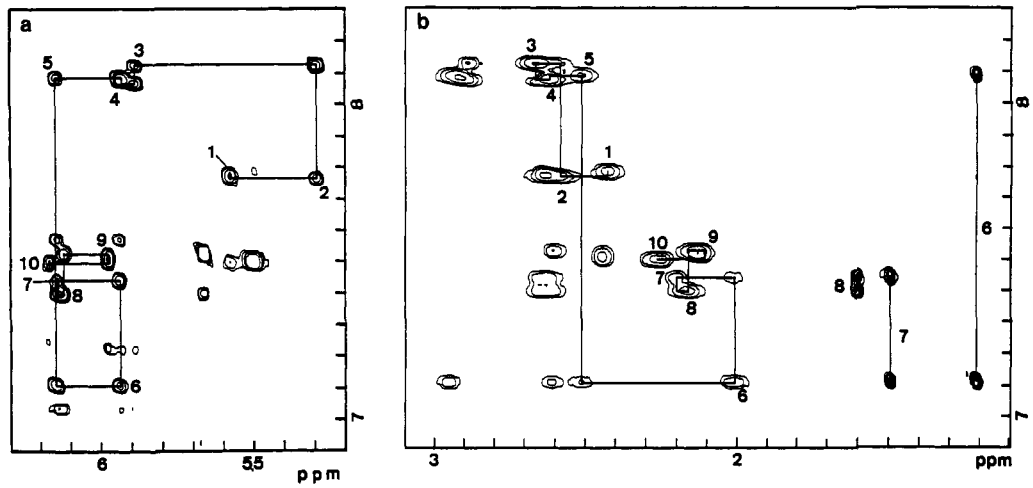

FIGURE 1: Expansion of the NOESY spectrum with the 150-ms mixing time in D_2O. (a) The H1′(i−1)–H6/H8(i)–H1′(i) and (b) H2′(i−1)–H6/H8(i)–H2′(i) connectivities are shown. The numbers denote the intraresidue cross peaks. The H6/H8(i−1)–TCH$_3$(i)–TH6(i) connectivities are also shown in (b), and the numbering of methyl protons is indicated.

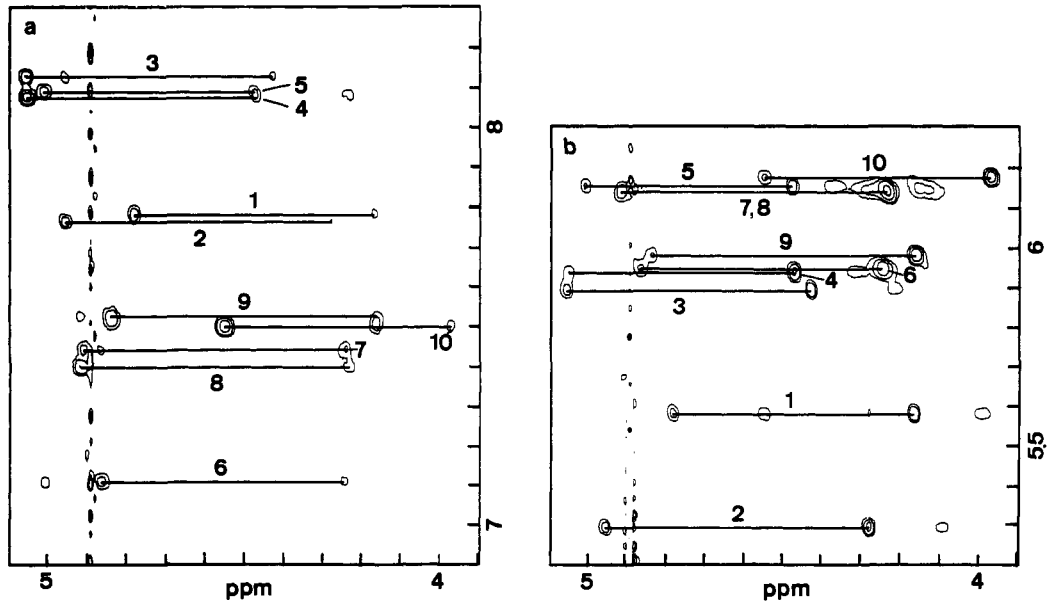

FIGURE 2: Expansion of the NOESY spectrum with the 150-ms mixing time in D_2O. (a) The H8(i)–H3′(i) and H4′(i) and (b) H1′(i)–H3′(i) and H4′(i) cross peaks are connected by horizontal lines. The numbering of each residue is indicated.

dependency was checked. For any kind of interproton distance listed in Tables II–IV, values obtained from the 50- and 60-ms NOESY spectra were nearly the same (difference <0.1 Å). The values obtained from the 90-ms NOESY spectrum were a little different from those obtained from the 50- and 60-ms NOESY spectra (difference 0.1–0.2 Å). The values obtained from the 100-, 120-, and 150-ms NOESY spectra were more different from those obtained from the 50- and 60-ms NOESY spectra. These results strongly suggest that interproton distances obtained from the 50-ms NOESY spectrum are little affected by spin diffusion and therefore are rather accurate. This is supported by the results of a study on a similar oligonucleotide, d(GAAAATTTTC)$_2$, i.e., that for such interproton distances as those listed in Tables II–IV, NOEs are primary direct ones when the mixing time is 50 ms (Sarma et al., 1988).

The effect of variation in the effective correlation time should also be considered. Sequence-dependent variation in the effective correlation times for the geometrically fixed distances, H2′–H2″, H5–H6 of cytosine, and CH$_3$–H6 of thymine vectors, was not detected. Moreover, the effective correlation time of H2′–H2″ vectors was not so different from that of H5–H6 vectors. This is consistent with the results of recent studies, in which no or only a slight difference was observed (Reid et al., 1988; Clore et al., 1988). This was also supported by a study involving complete relaxation matrix simulation, where all experimental NOEs involving H5–H6 and base–sugar protons could be reproduced equally with one correlation time (Sarma et al., 1988). Therefore, we used the H5–H6 distance as an internal reference in eq 3, except for the case of when methyl protons were concerned.

The obtained interproton distances, 91 in total (182 in the duplex), are summarized in Tables II–IV. They involve the error resulting from spin diffusion discussed above and inaccuracy of volume integration. Therefore, we set the upper and lower limits around the obtained interproton distances, between these limits the restraint energy term still being set at zero (see eq 1). a_1 and a_2 determining these limits were set on the basis of the results of a previous simulation study on spin diffusion (Clore & Gronenborn, 1985) and the precision of volume integration as follows: 0.20 and 0.25 Å, respectively, for r_{ij} of 2.5–3.0 Å; 0.25 and 0.30 Å for 3.0–3.5 Å, 0.30 and 0.35 Å for 3.5–4.0 Å, and 0.40 and 0.50 Å for 4.0–4.5 Å. As pointed out already, the obtained interproton distances tend

to be underestimates (Nilsson et al., 1986), so a_2 was set so as to be larger than a_1.

The obtained interproton distances and the setting of a_1 and a_2 can be justified partly by use of interproton distances that are not fixed geometrically but have a very limited range (Nilges et al., 1987). The ranges of the H1′–H2″ and H2′–H3′ distances are 2.2–2.3 and 2.2–2.4 Å, respectively (Wuthrich, 1986; Chary & Modi, 1988; van de Ven & Hilbers, 1988). The lower and upper limits set up for the corresponding distances in Table II on the basis of eq 1 cover these limited ranges fully, which justifies the obtained interproton distances and the setting of a_1 and a_2.

k_{NOESY} in eq 1 was fixed at 4 kcal mol^{-1} Å^{-2}. The weight, W, in eq 1 was raised from 0.25 to 8.35 in the course of restrained molecular dynamics simulation (see below), and thus, the effective force constant was 33.4 kcal mol^{-1} Å^{-2} at the most. This is rather moderate in comparison with force constants used in AMBER for single- and double-bond stretching, 300 and 600 kcal mol^{-1} Å^{-2}, and that used in the other study, 200 kcal mol^{-1} Å^{-2} (Gronenborn & Clore, 1989).

Torsion Angle Restraints for Sugars. Rough comparison of the relative intensities of DQF-COSY cross peaks enables one to determine the sugar conformation (Hosur et al., 1986). Judging from the point resolution of the DQF-COSY spectrum and the line width of resonances, the minimum value of the corresponding coupling constant necessary for the appearance of cross peaks was estimated to be about 3 Hz. Figure 3 shows the H1′–H2′/H2″ and H3′–H2′/H2″ cross-peak regions of the DQF-COSY spectrum. For G1 and A3 to T8, the H1′–H2′, H1′–H2″, and H3′–H2′ cross peaks are visible, but the H3′–H2″ cross peak is absent. For C9, the H1′–H2′, H1′–H2″, H3′–H2′, and H3′–H2″ cross peaks are all visible. Moreover, the H3′–H4′ cross peak is visible for all these residues (data not shown). On the basis of the correlation between the sugar conformation and the coupling constant (Hosur et al., 1986), it is concluded that the sugar puckering of G1 and of A3 to T8 is around C1′-exo and that of C9 is around O4′-endo. For G2 and C10, Figure 3 is not so informative because of the overlapping of the H2′ and H2″ resonances, so the relative intensities of the H3′–H4′ cross peaks were used. The intensity for G2 is comparable to those for G1 and A3 to T8. The intensity of C10 is comparable to that for C9 and rather stronger than those for the others. Thus, it is concluded that the sugar puckering of G2 is around C1′-exo and that of C10 around O4′-endo. Then the restraints for the main-chain torsion angle, δ, and furanose endocyclic torsion angles, ν_0 to ν_4, were given as follows (Altona & Sundaralingam, 1972); δ, ν_0, ν_1, ν_2, ν_3, and ν_4 are 121°, −35.2°, 35.2°, −21.7°, 0°, and 21.7° for G1 to T8 and 95°, −35.2°, 21.7°, 0°, −21.7°, and 35.2° for C9 and C10 [for the nomenclature, see Saenger (1984)]. A value of 37.0° is used as the amplitude of endocyclic torsion angles. k_{COSY} in eq 2 was set at 16.2 kcal mol^{-1} rad^{-2}. During the course of restrained molecular dynamics simulation, the sugar puckering of G1 to T8 fluctuated between O4′-endo and C2′-endo and that of C9 and C10 between C4′-exo and C1′-exo. This variation in the sugar puckering is of the same order of magnitude as the accuracy of the sugar puckering determined from the DQF-COSY spectrum.

Procedure for Structure Determination. A procedure similar to that for structure determination involving the use of CHARMM (Nilges et al., 1987) was carried out. Restrained molecular dynamics studies were started from two quite different initial conformations, classical A-DNA (IniA) and B-DNA (IniB) (Arnott & Hukins, 1972). Each conformation

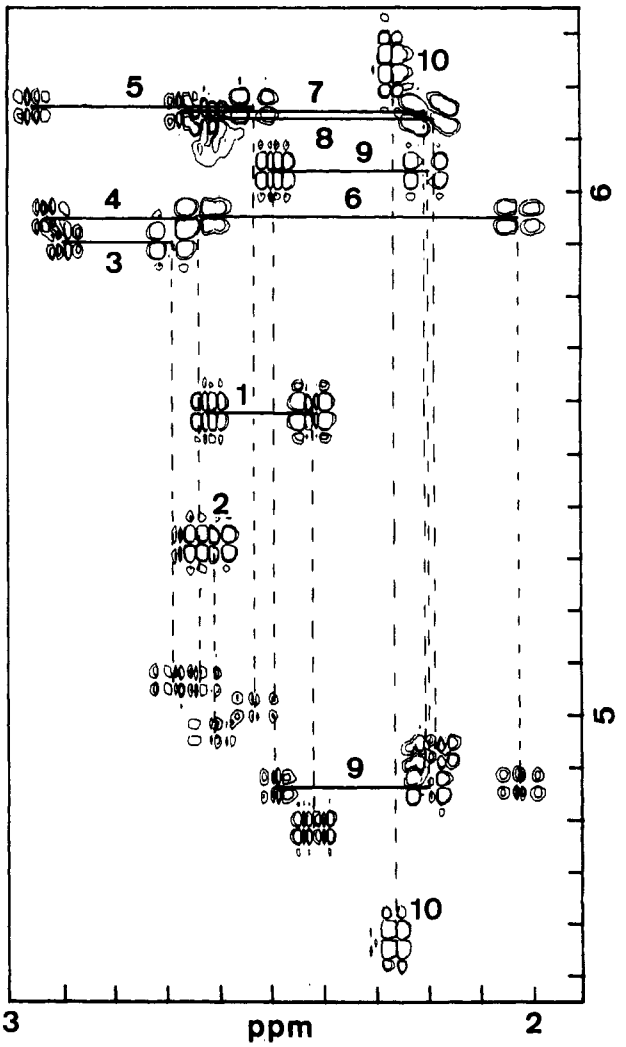

FIGURE 3: Expansion of the DQF-COSY spectrum. The H1′–H2′ and H1′–H2″ cross peaks are connected by horizontal lines (upper). The numbering of each residue is indicated. The H3′–H2′ and H2″ regions are shown (lower). The H3′–H2′ and H3′–H2″ cross peaks of C9 are connected by a horizontal line. C9 and C10 are labeled. Vertical dotted lines connect the H1′–H2′ and H3′–H2′ cross peaks for reference. For C9, the H1′–H2″ and H3′–H2″ cross peaks are also connected by a vertical dotted line.

was then subjected to the following steps: (1) equilibration for 18.4 ps during which time the molecule was heated from 10 to 400 K, following by cooling down to 300 K with an increase in W in eq 1 from 0.25 to 6.25; (2) restrained molecular dynamics for 16 ps at 300 K with W set at 6.25; and (3) an additional restrained molecular dynamics for 33.6 ps at 300 K with W set at 8.35. After long time simulation, conformational fluctuation became small. In fact, the rmsd between any two snap shots during the final 10-ps period (58–68 ps) was only about 0.5 Å. Therefore, it is concluded that the molecule reaches a stable conformation. Thus, (4) the coordinates of conformations stored per every 0.1 ps during the final 2 ps were averaged, and restrained molecular dynamics conformations were obtained, RDA and RDB, respectively.

Two other sets of calculations were carried out for comparison. In the second set, free molecular dynamics studies (where E_{NOESY} and E_{COSY} were ignored during simulation) were carried out. The averaged conformations were obtained in the same way as in the case of restrained molecular dynamics, FDA and FDB, respectively. In the third set, the initial structures were subjected to restrained energy minimization with W set at 8.35, and thus, restrained energy

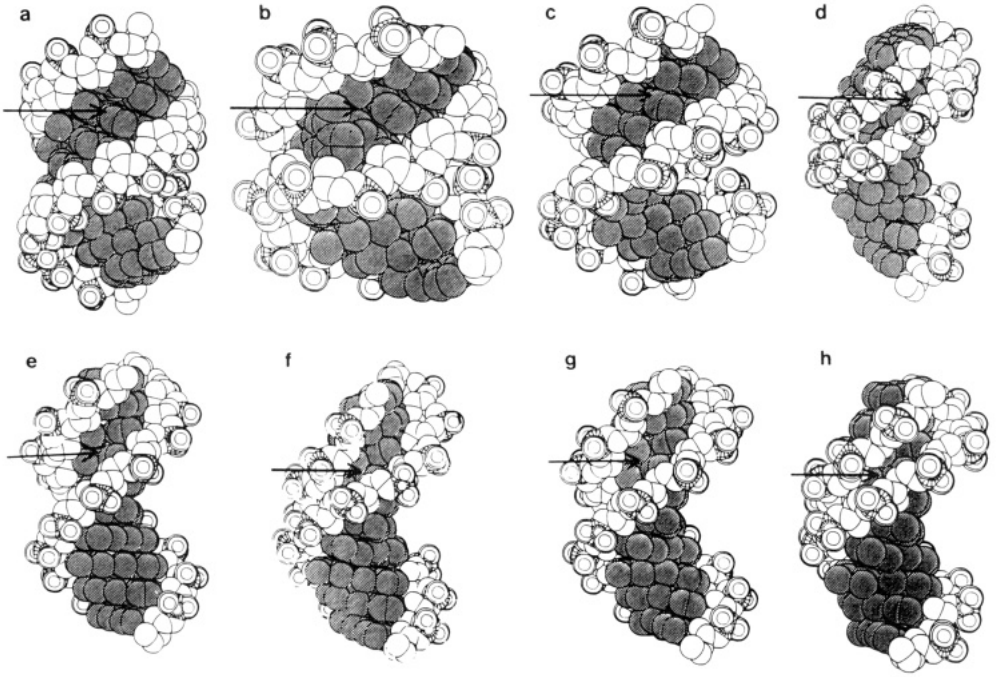

FIGURE 4: Views of the conformations obtained by various methods: (a) IniA, (b) FDA, (c) RMA, (d) RDA, (e) IniB, (f) FDB, (g) RMB, and (h) RDB. The minor grooves are indicated by arrows.

Table V: Atomic Overall rmsd (Å) between Initial (IniA, IniB), Free Molecular Dynamics (FDA, FDB), Restrained Energy Minimized (RMA, RMB), and Restrained Molecular Dynamics (RDA, RDB) Conformations

	IniB	FDA	FDB	RMA	RMB	RDA	RDB
IniA	5.5	3.4	6.1	1.7	5.5	5.8	6.1
IniB		6.3	1.8	5.1	1.0	1.7	1.9
FDA			6.6	3.5	6.3	6.4	6.6
FDB				5.5	1.5	1.7	1.8
RMA					4.9	5.1	5.4
RMB						1.2	1.5
RDA							0.8

minimized conformations were obtained, RMA and RMB, respectively.

The conformations obtained are illustrated in Figure 4. The rmsd between the conformations are given in Table V, and the energies of individual conformations are listed in Table VI.

Discussion

Comparison of Various Obtained Conformations. The rmsd between two initial conformations, A and B, is 5.5 Å. That between two restrained molecular dynamics conformations is very small, 0.8 Å. On the other hand, that between two restrained energy minimized conformations is 4.9 Å, just a little smaller than the initial one. That between two free molecular dynamics conformations, 6.6 Å, is, if anything, larger. These results mean that when restrained molecular dynamics was carried out, a converged conformation could be obtained independent of the initial conformations. When restrained energy minimization or free molecular dynamics was carried out, however, the conformations obtained depended heavily on the initial conformations, and it was impossible to obtain a converged conformation. This situation is apparent also in Figure 4. As to the global shape of the molecule, "fat" or "slim", and the width of the minor groove indicated by the arrow, FDA and RMA resemble IniA, and FDB and RMB resemble IniB. On the other hand, RDA and RDB (especially RDA) do not resemble IniA and IniB, respectively, but resemble each other (notice the extremely narrow minor grooves in RDA and RDB).

The results in Table V and Figure 4 suggest that the conformations obtained by use of restrained energy minimization are ones trapped in local minima of the total energy (the empirical energy + E_{NOESY} + E_{COSY}) function. This was confirmed by the fact that the total energy of RMA or RMB

Table VI: Individual Energy Terms (kcal/mol) for Initial (IniA, IniB), Free Molecular Dynamics (FDA, FDB), Restrained Energy Minimized (RMA, RMB), and Restrained Molecular Dynamics (RDA, RDB) Conformations

conformation	total[a]	empirical[a]	bond	angle	torsion	electrostatic	van der Waals	hydrogen bonding	restraints	
									E_{NOESY}[b]	E_{COSY}
IniA	3331	1048	45	443	206	−451	805	−4	1604	679
IniB	1767	524	46	95	262	−455	577	−1	832	411
FDA[c]	−96	−214	15	93	269	−432	−148	−11	107	11
FDB[c]	−242	−337	10	94	207	−440	−197	−11	91	4
RMA	−191	−298	12	83	210	−424	−167	−12	103	4
RMB	−255	−349	10	85	197	−441	−188	−12	89	5
RDA[c]	−279	−368	11	97	209	−463	−211	−11	79	10
RDB[c]	−284	−376	10	94	206	−458	−216	−12	85	7

[a] The total energy includes the restrained energy terms, whereas the empirical energy does not. [b] The weight, W, of E_{NOESY} is set at 8.35. [c] For average free and restrained molecular dynamics conformations, the energies are those obtained after the average structures had been subjected to restrained energy minimization with the same restraints as used to obtain RMA and RMB. This procedure is used to prevent artificial contact of atoms resulting from averaging and to make it meaningful to compare the energies of the various obtained conformations. It should be stressed that this procedure results in only very small atomic rms shifts (<0.2 Å).

Table VII: Values (deg) of Backbone and Glycosyl Torsion Angles (α–ζ, χ), Pseudorotation Angle (P), and Propeller Twist Angle for RDB

	α	β	γ	δ	ϵ	ζ	χ	P	propeller twist
G1			63	113	−163	−77	−133	119	10
G2	−81	169	54	137	−83	155	−92	137	11
A3	−75	127	57	117	−172	−96	−130	124	24
A4	−69	173	65	123	179	−96	−111	132	19
A5	−69	170	66	116	177	−98	−114	131	17
T6	−63	171	61	125	176	−99	−120	137	17
T7	−66	163	76	121	−177	−92	−147	126	19
T8	−63	−178	60	128	175	−93	−105	136	24
C9	−65	171	64	105	−179	−87	−126	107	11
C10	−75	173	60	97			−133	90	10

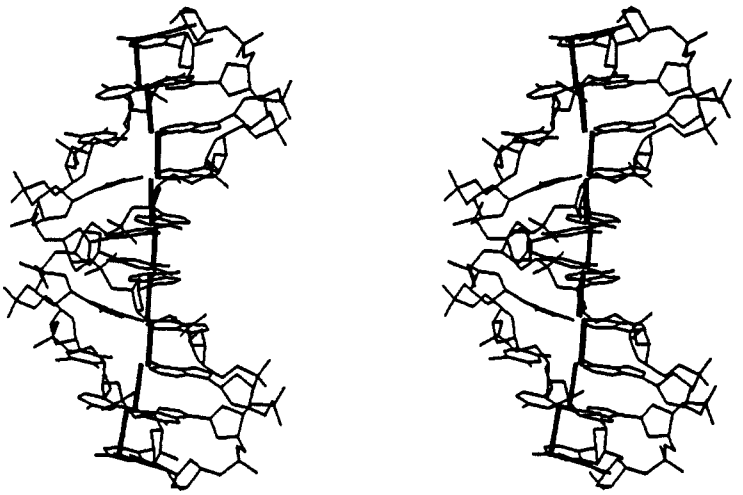

FIGURE 5: Stereoview of RDB. The local helical axes are indicated by heavy lines.

is higher than those of RDA and RDB (Table VI). The results in Table V and Figure 4 suggest that the conformations obtained by use of free molecular dynamics are also ones trapped in local minima of the empirical energy function. As a reason for this, it is suggested that the molecule could not reach the global minimum of the empirical energy function in the simulation for 68 ps, which is a practically available computational time, because its configuration space is too large and it would take too much time to search all the space with our computer. In contrast, the results in Table V and Figure 4 and the fact that the total energies of RDA and RDB are nearly the same and that they are lower than those of the other conformations suggest that the conformations obtained by use of restrained molecular dynamics are ones corresponding to around the global minimum of the total energy function. The total energy (empirical energy + E_{NOESY} + E_{COSY}) is very large for conformations deviating from experimental restraints, because E_{NOESY} and E_{COSY} are very large for such conformations (see eqs 1 and 2). So they are energetically very unstable. It is suspected that the configuration space giving such conformation is, in practice, excluded and the effective configuration space becomes quite smaller, and thus, the molecule could reach the global minimum in the limited simulation time. This means that convergence is due to incorporation of the experimental restraints into the total energy of the system in the form of effective potentials. The total energies of RDA and RDB are nearly the same as each other and lower than those of other conformations, which supports that these two conformations correspond to around the global minimum of the total energy function. Thus, conformations obtained by use of restrained molecular dynamics are suspected to be the most probable ones.

Bending Detected in Conformations Obtained by Use of Restrained Molecular Dynamics. The backbone and glycosyl torsion angles, the pseudorotation angle, and the propeller twist angle of RDB are listed in Table VII. The conformations obtained by use of restrained molecular dynamics are very similar to each other, and there are no remarkable differences in the conformational parameters. Therefore, the following features apply to RDA, too. The torsion angles, α, β, γ, and δ, come into the range corresponding to standard g^-, t, g^+, and anti positions. It is remarkable that G2 takes on the BII conformation around C3′–O3′–P bonds (ϵ:g^-, ζ:t) (Fratini et al., 1982), although the other residues take on the standard BI conformation (ϵ:t, ζ:g^-). Judging from the pseudorotation angles, the imposed restraints for sugar puckering are conserved in RDB, although the restraints are not so strong as to fix the sugar conformation rigidly, and in fact, it fluctuated to some extent in the course of restrained molecular dynamics as mentioned above. The sugar puckering of the residues from G1 to T8 is C1′-exo and that of C9 and C10 is O4′-endo. The combination of the occurrence of the BII conformation at G2 and the discontinuity of sugar puckering between T8 and C9 causes bending between the oligo(dA) tract and the other portions, as indicated in Figure 5. It must be difficult to detect such bending in previously studied $d(GAAAATTTTC)_2$, of which "theater portions" consist of just one residue at both ends of the sequence respectively, which must suffer from an artificial end effect.

The propeller twist angles of the oligo(dA) tract are relatively large, about 20°. This value is comparable to those observed for the oligo(dA) tract in crystals (Nelson et al., 1987; Coll et al., 1987) and is in good agreement with that found for the oligo(dA) tract in solution on the basis of the results of semiquantitative NMR analysis (Katahira et al., 1988, 1990a). The expected compression of the minor groove of the oligo(dA) tract corresponding to large propeller twist angles (Fratini et al., 1982) can be observed clearly in Figure 4h (and Figure 4d). As a result of compression of the minor groove (Drew & Travers, 1985; Satchwell et al., 1986; Travers &

Klug, 1987), the molecule is bent a little toward the minor groove within the oligo(dA) tract, as indicated in Figure 5. This kind of bending must also be indispensable for explaining DNA bending systematically (Katahira et al., 1990a). The direction of bending between the oligo(dA) tract and the other portions, and within the oligo(dA) tract, is the same; thus, the bending is additive, the total degree of bending being about 20°.

Concluding Remarks. We have shown that restrained molecular dynamics in conjunction with the modified AMBER program works very well for determining the conformation of DNA on the basis of the results of NMR experiments. A putative bent DNA was studied. Relatively remarkable bending between the oligo(dA) tract and the other portions and slight bending within the oligo(dA) tract were found. The bending is additive, the total degree of bending being about 20°.

ACKNOWLEDGMENTS

We thank Drs. Weiner and Kollman, University of California, for sending us the AMBER program.

Registry No. d(GGAAATTTCC), 117940-93-9.

REFERENCES

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

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

Burkhoff, A. M., & Tullius, T. D. (1987) *Cell 48*, 935–943.

Calladine, C. R., Drew, H. R., & McCall, M. J. (1988) *J. Mol. Biol. 201*, 127–137.

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

Chary, K. V. R., & Modi, S. (1988) *FEBS Lett. 233*, 319–325.

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

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

Clore, G. M., Lauble, H., Frenkiel, T. A., & Gronenborn, A. M. (1984) *Eur. J. Biochem. 145*, 629–636.

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

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

Coll, M., Frederick, C. A., Wang, A. H. J., & Rich, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8385–8389.

DiGabriele, A. D., Sanderson, M. R., & Steiz, T. A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1816–1820.

Drew, H. R., & Travers, A. A. (1985) *J. Mol. Biol. 186*, 773–790.

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

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

Griffith, J., Bleyman, M., Rauch, C. A., Kitchin, P. A., & Englund, P. T. (1986) *Cell 46*, 717–724.

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

Hagerman, P. J. (1985) *Biochemistry 24*, 7033–7037.

Hagerman, P. J. (1986) *Nature 321*, 449–450.

Haran, T. E., & Crothers, D. M. (1989) *Biochemistry 28*, 2763–2767.

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

Hosur, R. V., Ravikumar, M., Chary, K. V. R., Sheth, A., Govil, G., Zu-kun, T., & Todd Miles, H. (1986) *FEBS Lett. 205*, 71–76.

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

Kaptein, R., Zuiderweg, E. R. P., Scheek, R. M., Boelens, R., & van Gunsteren, W. F. (1985) *J. Mol. Biol. 182*, 179–182.

Katahira, M., Sugeta, H., Kyogoku, Y., Fujii, S., Fujisawa, R., & Tomita, K. (1988) *Nucleic Acids Res. 16*, 8619–8632.

Katahira, M., Sugeta, H., & Kyogoku, Y. (1990a) *Nucleic Acids Res. 18*, 613–618.

Katahira, M., Lee, S. J., Kobayashi, Y., Sugeta, H., Kyogoku, Y., Iwai, S., Ohtsuka, E., Benevides, J. M., & Thomas, G. J., Jr. (1990b) *J. Am. Chem. Soc. 112*, 4508–4512.

Kintanar, A., Klevit, R. E., & Reid, B. R. (1987) *Nucleic Acids Res. 15*, 5845–5862.

Koo, H. S., & Crothers, D. M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1763–1767.

Koo, H. S., Wu, H. M., & Crothers, D. M. (1986) *Nature 320*, 501–506.

Leroy, J. L., Charretier, E., Kochoyan, M., & Gueron, M. (1988) *Biochemistry 27*, 8894–8898.

Nadeau, J., & Crothers, D. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2622–2626.

Nelson, H. C. M., Finch, J. T., Luisi, B. F., & Klug, A. (1987) *Nature 330*, 221–226.

Nilges, M., Clore, G. M., Gronenborn, A. M., Brunger, A. T., Karplus, M., & Nilsson, L. (1987) *Biochemistry 26*, 3718–3733.

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

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

Patapoff, T. W., Thomas, G. A., Wang, Y., & Peticolas, W. L. (1988) *Biopolymers 27*, 493–507.

Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wuthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.

Reid, B. R., Banks, K. M., Flynn, P., Nerdal, W., & Chou, S. H. (1988) *Proceeding of the XIII International Conference on Magnetic Resonance in Biological Systems*, L13, University of Wisconsin, Madison, WI.

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

Roy, S., Borah, B., Zon, G., & Cohen, J. S. (1987) *Biopolymers 26*, 525–536.

Rychaert, J. P., Cicotto, G., & Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327–337.

Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.

Sarma, M. H., Gupta, G., & Sarma, R. H. (1988) *Biochemistry 27*, 3423–3432.

Satchwell, S. C., Drew, H. R., & Travers, A. A. (1986) *J. Mol. Biol. 191*, 659–675.

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

Singh, U. C., Weiner, S. J., & Kollman, P. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 755–759.

Sklenar, V., Brooks, B. R., Zon, G., & Bax, A. (1987) *FEBS Lett. 216*, 249–252.

States, D. J., Haberkorn, R. A., & Rubern, D. J. (1982) *J. Magn. Reson. 48*, 286–292.

Travers, A., & Klug, A. (1987) *Nature 327*, 280-281.
Uesugi, S., Ohkubo, M., Ohtsuka, E., Kobayashi, Y., & Kyogoku, Y. (1984) *Nucleic Acids Res. 12*, 7793-7810.
Ulanovsky, L. E., & Trifonov, E. N. (1987) *Nature 326*, 720-722.
van de Ven, F. J. M., & Hilbers, C. W. (1988) *Eur. J. Biochem. 178*, 1-38.
Weiner, P., & Kollman, P. A. (1981) *J. Comput. Chem. 2*, 287-303.
Weiner, S. J., Kollman, P. A., Nguyen, D. T., & Case, D. A. (1986) *J. Comput. Chem. 7*, 230-252.
Weiss, M. A., Patel, D. J., Sauer, R. T., & Karplus, M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 130-134.
Whitlow, M., & Teeter, M. M. (1986) *J. Am. Chem. Soc. 108*, 7163-7172.
Wu, H. M., & Crothers, D. M. (1984) *Nature 308*, 509-513.
Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Ultraviolet-Induced Thymine Hydrates in DNA Are Excised by Bacterial and Human DNA Glycosylase Activities†

Tapan Ganguly, Kim M. Weems, and Nahum J. Duker*
Department of Pathology and Fels Institute for Cancer Research and Molecular Biology, Temple University School of Medicine, Philadelphia, Pennsylvania 19140
Received March 6, 1990; Revised Manuscript Received April 30, 1990

ABSTRACT: Ultraviolet irradiation of DNA results in various pyrimidine modifications. We studied the excision of an ultraviolet thymine photoproduct by *Escherichia coli* endonuclease III and by a preparation of human WI-38 cells. These enzymes cleave UV-irradiated DNA at apyrimidinic sites formed by glycosylic removal of the photoproduct. Poly(dA-[^{3}H]dT)·poly(dA-[^{3}H]dT) was UV irradiated and incubated with purified *E. coli* endonuclease III. ^{3}H-Containing material was released in a manner consistent with Michaelis-Menten kinetics. This ^{3}H-labeled material was determined to be a mixture of thymine hydrates (6-hydroxy-5,6-dihydrothymine), separable from unmodified thymine by chromatography in three independent systems. Both *cis*-thymine hydrate and *trans*-thymine hydrate were chemically and photochemically synthesized. These coeluted with the enzyme-released ^{3}H-containing material. No thymine glycol was released from the UV-irradiated polymer. Similar results were obtained with extracts of WI-38 cells as the enzyme source. The release of thymine hydrates by both glycosylase activities was directly proportional to the amount of enzyme and the irradiation dose to the DNA substrate. These results demonstrate the modified thymine residues recognized and excised by endonuclease III and the human enzyme to be a mixture of *cis*-thymine hydrate and *trans*-thymine hydrate. The repairability of these thymine hydrates suggests that they are stable in DNA and therefore potentially genotoxic.

Endonuclease III of *Escherichia coli* is involved in the initiation of the repair of various pyrimidine damages in DNA. These include photochemical alterations resulting from UV[1] irradiation (Radman, 1976; Gates & Linn, 1977). Chemical oxidation, X-irradiation, and γ-irradiation of DNA also result in base damages that are substrates for this enzyme (Gates & Linn, 1977; Demple & Linn, 1980, 1982; Katcher & Wallace, 1983; Breimer & Lindahl, 1984, 1985). An analogous activity, present in cultured human cells, was originally found to incise UV-irradiated DNA (Brent, 1972; Duker & Teebor, 1975). This human enzyme, like *E. coli* endonuclease III, recognizes monomeric pyrimidine lesions in UV-irradiated DNA and also base damages in OsO_4-oxidized or X-irradiated DNAs (Brent, 1973, 1976, 1983; Doetsch et al., 1987). Studies using DNA sequencing techniques and chemical analyses of reaction products indicate the two enzymes to have an identical range of modified pyrimidine substrates (Boorstein et al., 1989; Doetsch et al., 1986, 1987, 1988; Higgins et al., 1987; Gallagher et al., 1989a,b; Lee et al., 1988; Weiss et al., 1989). Common structural features of these bases include saturation of the 5,6-double bond or loss of ring planarity (Breimer & Lindahl, 1985; Teoule, 1987; Teebor et al., 1988). It has been recently established that both enzymes remove cytosine hydrate (6-hydroxy-5,6-dihydrocytosine) and uracil hydrate (6-hydroxy-5,6-dihydrouracil) from UV-irradiated DNA (Boorstein et al., 1989; Weiss et al., 1989; Ganguly & Duker, 1990). This type of enzyme, sometimes termed a "redoxyendonuclease", is widely conserved and present in a wide variety of species. These enzymes contain two activities that act sequentially: a DNA glycosylase, which releases the modified pyrimidines, and a DNA apurinic/apyrimidinic site endonuclease, which incises DNA at the resultant apyrimidinic sites by catalysis of β-elimination. They are well discussed in a number of recent reviews (Weiss & Grossman, 1987; Sancar & Sancar, 1988; Wallace, 1988).

The DNA glycosylase activity of endonuclease III releases 5,6-saturated thymine and its derivatives, such as thymine glycol (5,6-dihydroxy-5,6-dihydrothymine), 5-hydroxy-5-methylhydantoin, methyltartronylurea, and urea, as free bases from oxidized DNA (Demple & Linn, 1980; Katcher &

†This work was supported by U.S. Public Health Service Grants CA-24103 from the National Cancer Institute and AG-00378 from the National Institute of Aging.
*To whom correspondence and request for reprints should be addressed.

[1] Abbreviations: UV, ultraviolet; HPLC, high-performance liquid chromatography; TLC, thin-layer chromatography; EDTA, ethylenediaminetetraacetic acid; assay buffer, 40 mM KH_2PO_4, pH 7.4, and 1 mM EDTA; 6-4 photoproduct, 6-(1,2-dihydro-2-oxo-4-pyrimidinyl)-5-methyl-2,4(1*H*,3*H*)-pyrimidinedione.

0006-2960/90/0429-7222$02.50/0