Solution Structure of the Parallel-Stranded Duplex Oligonucleotide α-d(TCTAAAC)–β-d(AGATTTG) via Complete Relaxation Matrix Analysis of the NOE Effects and Molecular Mechanics Calculations

G. Lancelot,* J.-L. Guesnet, and F. Vovelle
Centre de Biophysique Moléculaire, CNRS, 45071 Orléans Cédex 2, France
Received September 13, 1988; Revised Manuscript Received March 9, 1989

ABSTRACT: The solution structure of the duplex formed by the association of the unnatural oligonucleotide α-d(TCTAAAC) with its natural and parallel complementary sequence β-d(AGATTTG) was investigated by nuclear magnetic resonance spectroscopy and constrained molecular mechanics calculations. The structure was refined on the basis of interproton distances determined by NOE measurements for a series of mixing times. The NOE values were converted to distances by using the complete 134 × 134 relaxation matrix including all proton dipole–dipole interactions and spin diffusion. The computation of the relaxation matrix requires the Cartesian coordinates of the oligonucleotide, which are not known, a priori. To avoid this ambiguity, we used an iterative procedure in which the new distance constraints are obtained by using the complete relaxation matrix calculated from the previous structure. After three iterations, the process converged. The unnatural duplex α-d(TCTAAAC)–β-d(AGATTTG) adopts in solution a right-helical structure with Watson–Crick base pairing, an anti conformation on the glycosyl linkage on the β-strand, a syn conformation on the α-strand, and a 3′-exo conformation of the deoxyriboses for both sugar anomers. The three-dimensional structure obtained allowed us to describe the local heterogeneity of the duplex.

The binding of an oligonucleotide to its complementary sequence is a highly specific process governed by stacking interactions between base pairs and by hydrogen-bond formation between complementary bases. We have already shown that covalent attachment of an intercalating agent to an oligodeoxynucleotide strongly increases the stability of hybrid formation (Asseline et al., 1984).

However, the use of these molecules in living organisms is strongly limited by their fast degradation by endonucleases. To increase the resistance of such compounds, α-oligonucleotides were synthesized, and it was shown that their rate of hydrolysis was reduced 300–500 times for both exo- and endonucleases compare to the β-analogues (Lancelot et al., 1987; Thuong et al., 1987). Moreover, the strength of the association of the α-strand with its complementary β parallel strand was found equivalent to that between their antiparallel complementary β-analogues (Lancelot et al., 1987; Morvan et al., 1987).

NOE[1] data provided evidence that the oligonucleotides α-d(TCTAAAC)–β-(AGATTTG) form a stable right-helical duplex with anti conformation on the glycosyl linkages of the β-strand residues, Watson–Crick base pairing, and sugars in 3′-exo conformation (Lancelot et al., 1987).

Recent progress in DNA synthesis combined with analysis of NMR data and molecular dynamics or restrained molecular dynamics have enabled detailed structural studies of oligonucleotides in solution (Broido et al., 1985; Jamin et al., 1985; Suzuki et al., 1986; Zhou et al., 1987; Nerdal et al., 1988; Scalfi Happ et al., 1988). In the present study, we report on a detailed conformational investigation of the unnatural duplex α-d(TCTAAAC)–β-d(AGATTTG) in solution.

The previous assignment of the nonexchangeable protons (Lancelot et al., 1987) has been completed by the assignment of the H5′–H5″ proton resonances. All the nonoverlapping cross-peak areas were quantified for several mixing times. An iterative procedure based on the theoretical 2D NOE maps calculated by taking into account the complete relaxation matrix of the oligonucleotide and the energy-minimized structure computed by using the molecular mechanics program AMBER allowed us to determine the solution structure of our unnatural duplex.

Experimental Procedures

The synthesis and the purification of the oligonucleotides have been described before (Lancelot et al., 1987).

NMR experiments were carried out with a Bruker AM-300 spectrometer and processed on an Aspect 3000 computer. The NOESY data were obtained at 22 °C from degassed solutions (2.5 mM in duplex) contained in sealed tubes. Two-dimensional data sets for COSY DQF and NOESY spectra were collected in the phase-sensitive mode with the time-proportional phase incrementation scheme (Bodenhausen et al., 1980; Drobny et al., 1979; Marion & Wüthrich, 1983; Marion & Lancelot, 1984).

Typically, 4096 or 8192 complex t_2 data points were collected for each 256 or 384 t_1 values in NOESY or COSY DQF experiments. The 512 × 4096 or 1024 × 8192 data matrices were resolution-enhanced by a Gaussian window function in direction 2 and by a shifted squared sine-bell window function in direction 1 and then Fourier transformed and phase adjusted. The t_1 and t_2 ridges were decreased by using the suppression procedure already described (Otting et al., 1986). The good resolution in direction 2 (0.68 Hz/point) allowed us to nearly avoid the overlapping of the aromatic resonances. The window functions were optimized to minimize the side lobe of the resonances induced by the partial truncation of the FID in direction 1. Measurements of the volume area and intensities of cross-peaks were compared on several matrix data. It was found that area measurement of cross-peaks on the spectra in direction 1 were the most reproducible when smaller overlap occurred.

[1] Abbreviations: NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; COSY DQF, 2D double quantum filter correlated spectroscopy; AMBER, assisted model building with energy refinement.

0006-2960/89/0428-7871$01.50/0

Theory

The relaxation behavior of a system of protons is governed by (Noggle & Schirmer, 1971)

$$[dMz/dt] = [R][Mz - Mz^e]$$

where $[Mz - Mz^e]$ is the deviation of the magnetization from thermal equilibrium and [R] the relaxation matrix.

The resolution of this differential equation system gives the time evolution of the magnetization of each spin i. $[Mz - Mz^e] = [U][K][e^{\lambda t}]$, where [U] is the matrix of the eigenvectors of [R], [K] a diagonal matrix, and $[e^{\lambda t}]$ a column matrix where the element of row i is $e^{\lambda_i t}$ (λ eigenvalues of [R]). The elements of [K] are computed by using the initial conditions Mz(0).

At time τ_m, the 2D NOE cross-peak η_{ij} is given by $Mz(\tau_m)$ with $Mz_i(0) = 1$ and $Mz_j(0) = 0$ if $j \neq i$ and $Mz_i^e = 0$ for each i.

For a two-spin system without coupling, the dipolar Hamiltonian is

$$H_d = \frac{\mu_0 \hbar \gamma^2}{4\pi}\left[\frac{\vec{I}\cdot\vec{S}}{r^3} - \frac{3(\vec{I}\cdot\vec{r})(\vec{S}\cdot\vec{r})}{r^5}\right]$$

and the value of the matrix elements $\langle\varphi|H_d|\psi\rangle$ will vary randomly since H_d, which can be written $\sum_{q=-2}^{+2} F_q(r,\Theta,\varphi)A_q$, depends on the fluctuating functions (Abragam, 1978)

$$F_0 = \frac{1 - 3\cos^2\Theta(t)}{r^3(t)}$$

$$F_{\pm 1} = \frac{\sin\Theta(t)\cos\Theta(t)\, e^{\pm i\varphi(t)}}{r^3(t)}$$

$$F_{\pm 2} = \frac{\sin^2\Theta(t) e^{\pm 2i\varphi(t)}}{r^3(t)}$$

The autocorrelation and the cross correlation functions must be computed by using total description motion (amplitude, direction, and frequency). Since the mode of motion of the interproton vectors (aromatic proton–sugar proton or sugar proton–sugar proton) is generally unknown and the terms of the relaxation matrix cannot be rigorously calculated, we have computed the matrix elements in three cases:

(i) The assumption that each interproton distance is not dependent upon time and possessed an isotropic motion is described by the spectral density $J(\omega) = \tau_c/(1 + \omega^2\tau_c^2)$.

(ii) The assumption that some interproton vectors are moving with an isotropic internal motion is described by the spectral density that has been evaluated by Woessner (1962).

(iii) In the assumption of an isotropically internal motion, the search for the τ_c effective allowed the best fit of the data.

The total spin–lattice relaxation R_{ii} and the cross-relaxation rates R_{ij} may be written as functions of the zero-, single-, and double-quantum transitions probabilities.

$$R_{ii} = \sum W_0^{ij} + 2\, W_1^{ij} + W_2^{ij}$$

$$R_{ij} = W_2^{ij} - W_0^{ij}$$

The elements W can be expressed as a function of the spectral densities

$$W_0^{ij} = kJ(0)r_{ij}^{-6}$$

$$W_1^{ij} = 1.5kJ(\omega)r_{ij}^{-6}$$

$$W_2^{ij} = 6kJ(2\omega)r_{ij}^{-6}$$

where $k = \mu^2\gamma^4\hbar^2/16\pi^2 = 5.696 \times 10^{-49}$ m^6 s^{-2} in the SI system and where $J(\omega)$ is the spectral density function.

The internal motion of the protons of the methyl group is difficult to compute rigorously since the distances between a nucleus and each of the three protons is an unknown function of time, and generally the NOE data relative to the methyl group are not simulated.

We have tentatively simulated their NOE connectivities by computing the magnetization transfer of each of the three protons for each of the 12 positions of the C–H bond obtained by rotating the protons around the C6–C7 bond by 10° steps. The NOE data were then compared to the average value of the 12 computed data by assuming an isotropic motion of the C7–H protons around the C6–C7 bond.

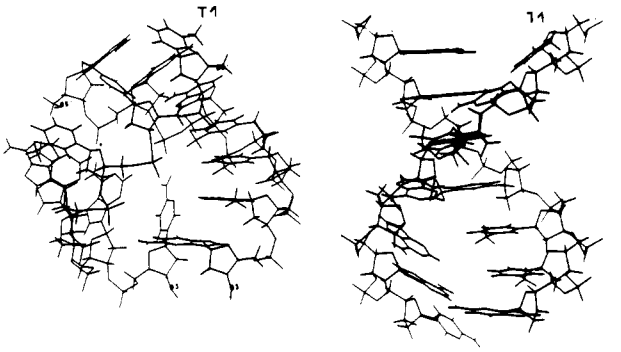

FIGURE 1: View of the two A (left) and B (right) starting conformations of the unnatural duplex α-d(TCTAAAC)–β-d(AGATTTG).

Molecular Mechanics

The calculations were carried out on the heptamer α-d(TCTAAAC)–β-d(AGATTTG) by using the molecular mechanics program AMBER (Weiner & Kollman, 1981). The energy function includes harmonic bond stretching and bond angle terms, torsional energy functions, and Lennard-Jones and electrostatic terms for nonbonded interactions plus a 10–12 function for hydrogen bonds. The various force field parameters and charges used in the calculation were reported previously (Weiner et al., 1984). No counterions were included in the calculation, and an all-atom model, in which all hydrogen atoms are included, was explicitly used. The dielectric constant was set equal to $4r_{ij}$ the interatomic distance between atoms i and j (Weiner et al., 1984). The structure was refined by using steepest descent and conjugate gradient algorithm until the rms energy gradient relative to atomic coordinate changes was less than 0.1 kcal mol^{-1} Å^{-1}.

The initial conformation was built by generating a parallel regular B and A helix α-d(TCTAAAC)–β-d(AGATTTG) using coordinates derived from diffraction measurements on classical antiparallel DNA fibers (Arnott & Hukins, 1972).

The chirality of the C1′ atoms of the first strand (α-strand) was modified to set the base in an α-position with respect to the furanose ring. Most of the steric contacts were removed on the Evans and Sutherland PS 390 graphic display by using the FRODO program (Jones, 1978; Pflugrath et al., 1984) (Figure 1).

The resulting structures were then energy refined by using AMBER, including distance constraints to ensure the Watson–Crick hydrogen-bond base pairing. These additional constraints were removed for the refinement of the structure in solution.

These structures provide starting conformations for the next energy minimization in which the constraints determined from the NMR measurements were incorporated. The refined structures labeled A and B are derived from the A and B initial conformations, respectively.

Table I: rms Difference[a] of Interproton Distances (Å) for the Constrained and Relaxed Structures

	all[a]	intraresidue[a]	interresidue[a]	relative[b] rms difference
constrained structure	0.18	0.17	0.18	4.3
relaxed structure	0.47	0.49	0.45	15.5

[a] rms = $(\sum(R_{cal} - R_{exp})^2/N^{1/2}$. [b] Relative rms difference = $100[(\sum(R_{cal} - R_{exp})^2/N)/R_{exp}]^{1/2}$.

An effective harmonic potential representing the interproton constraints was added to the total energy function of the system: $E_{NOE} = K_r(r_{ij} - r_{ij}^{\circ})^2$, where r_{ij} is derived from NMR data. The force constant K_r was chosen to be 25 kcal mol^{-1} Å^{-1}, when $r_{ij} < 3$ Å and 15 kcal mol^{-1} Å^{-1} when $r_{ij} > 3$ Å. These values are rather small compared to bond-stretching force constants. Larger force constants lead to a better agreement between calculated and experimental interproton distances but can induce a distortion of the geometry of the base.

The NMR measurements provided us information on the sugar conformation in terms of the pseudorotation angle P; thus, sugar pucker constraints were applied by forcing the torsion angles ν_0(C4′–O1′–C1′–Cα) and ν_4(C3′–C4′–O1′–C1′) and the sugar interproton distances to their ideal values derived from the pseudorotation relationship $\nu_j = \nu_{max}\cos(P + (4\pi/5j) + (2\pi/5))$ (Altona & Sundaralingam, 1972), where ν_{max} is the maximal amplitude of puckering ($\pi/5$ in our calculation), P is the pseudorotation angle, and j enumerates the bond.

Generation of the Refined Structure

The comparison of experimental and theoretical NOE values has been made from the complete relaxation matrix, which requires the knowledge of the geometry of the overall molecule.

The ambiguity of this procedure has been removed by the introduction of an iterative process in which the target distance between each pair of protons was continuously readjusted.

Using the H5–H6 distance as internal reference (2.46 Å), the 60 shortest proton–proton distances were estimated from $r = r(\text{ref})(\eta_{ref}/\eta)^{1/6}$. These data and the pseudorotation angle P of the sugar previously determined (Lancelot et al., 1987) were used for a first constrained energy minimization in which only the shortest interproton distances (mainly intraresidue base–sugar distances H8/H6–H2′/H2″) and the puckering constraints were included. The coordinates of the resulting structure were then used for a further determination of the interproton distances from the 2D NOE data.

The new target distances r(exp) were calculated by a least-squares procedure in which the data were fitted by

$$r(\text{exp}) = r(\text{A})(\eta_A/\eta)^{1/6}$$

where r(A) is the interproton distance for the refined structure, η the experimental NOE and η_A the NOE calculated by minimizing the sum of the deviations for the four values τ_i of the mixing times $\sum|\eta(\tau_i) - \eta_A(\tau_i)|$. At this step, the geometry of the molecule, refined at the previous step, was used to calculate the complete relaxation matrix.

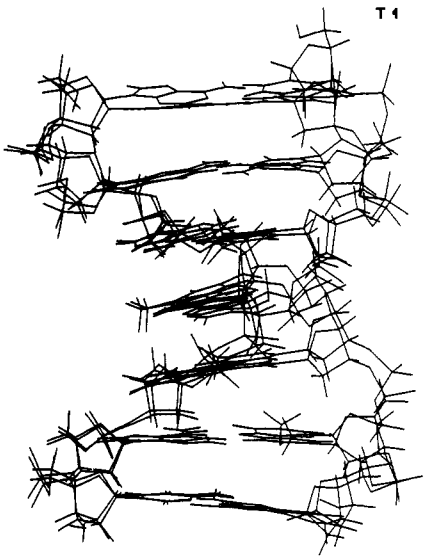

FIGURE 2: View of the superposition of the two A and B refined structures of the unnatural duplex α-d(TCTAAAC)–β-d-(AGATTTG).

This allows an improvement of the distance data set. The initial values were corrected, and the interresidue base–base and base–sugar interproton distances constraints were included in the next refinement. After three iterations, reasonable agreement was found between the calculated and all the experimental interproton distances (87 values) constituting the constraint data base. During the last step of the refinement, interproton constraints including protons of the methyl group were added. At the end of the refinement, the final structure was reminimized without any constraint, leading to a "relaxed" structure that would be compared to the "constrained" structure.

Three-Dimensional Structure

The rms differences between the two A and B refined structures were only 0.51 Å. The A and B conformations were very similar, and the best fit superposition is shown in Figure 2. In fact, the main differences occur at the T1 terminal residue; this position is weakly defined with respect to the adjacent C2 residue (Table III). If this residue is omitted in the calculation of the atomic rms, its value drops to 0.33 Å. Since the B conformation is 34 kcal mol^{-1} more stable than the A conformation, we will only describe its conformational parameters. The differences between the A and B forms are examined in the discussion.

Stereoviews of the final structure and the superposition of this constrained nucleotide with the relaxed one are shown in Figures 3 and 4. The rms differences between the calculated and experimental interproton distances r(exp) are given in Table I and the individual energy terms in Table II. It is clear for Table I that even using small force constants for our constraint energy terms, we obtain a good agreement between the calculated and experimental interproton distances, the agreement being the same for the base–sugar intraresidue distances and for the base–sugar interresidue distances. For more details, we listed in Table III the calculated distances

Table II: Individual Energy Term for the Constrained and Relaxed Structures

	energy (kcal/mol)								
	total	potential[a]	bond	angle	torsion	elec	VDW	H-bond	constraints
constrained structure	−356	−415	9.0	68	175	−554	−102	−1.0	58.6
relaxed structure without constraints	−469	−469	5.3	49	161	−545	−138	−1.5	0.0

[a] The total energy includes the constraints energy, whereas the potential energy does not.

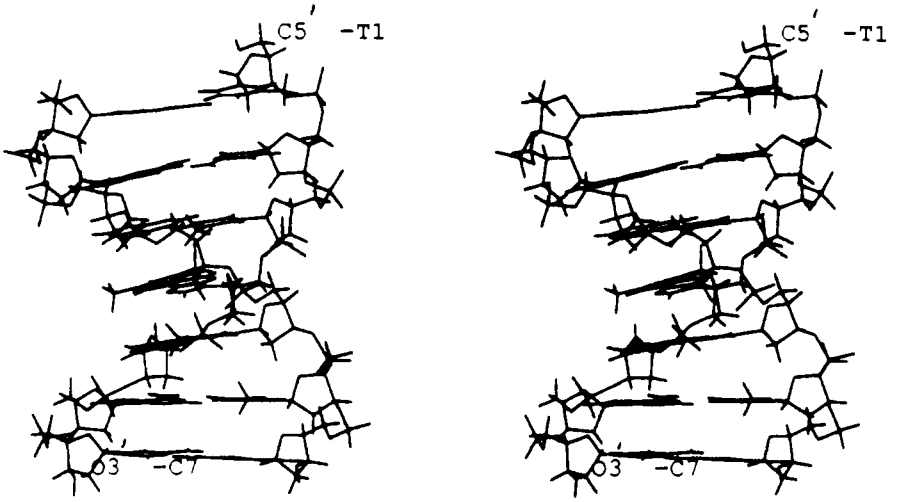

FIGURE 3: Stereoview of the constrained structure of the unnatural duplex α-d(TCTAAAC)-β-d(AGATTTG).

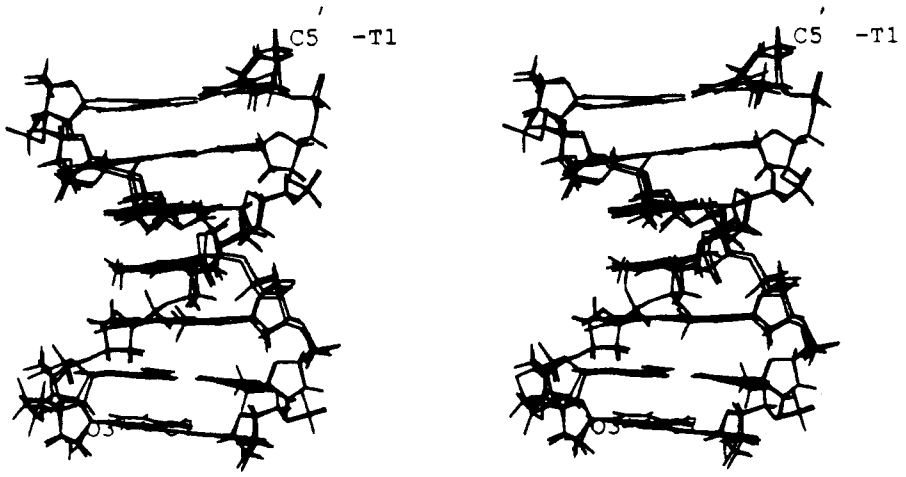

FIGURE 4: Stereoview of the superposition of the constrained and relaxed structures of the unnatural duplex α-d(TCTAAAC)-β-d(AGATTTG).

and the differences between the experimental and calculated interproton distances for each residue, and in Table IV, we compared the sugar puckerings of the refined structure with the pseudorotation angle *P* deduced from experimental data. Except for the adenine A3 and the thymine T5 of the β-strand, all the *P* values calculated for the constrained structure fall in the range of the experimental data. Generally the agreement is better for the short distances for which the experimental data are the most precise, and this agreement is about the same for each residue of both strands (Table III). The larger disagreement occurs for the internucleotide distances and fixes the position of adenine 5 with respect to adenines 4 and 6 in the α-strand. Any attempt to improve these results by increasing the constraints of the interproton distances leads to an unrealistic distortion of the bases.

The introduction of the NMR constraints in the refinement leads to a substantial improvement of the fit between the experimental and calculated distances (Table I). However, the constrained and relaxed structures are globally the same (Figure 4). The rms difference between the Cartesian coordinates of both structures is only 0.33 Å. The removal of the constraints on the pseudorotation angle *P* does not change drastically the final puckering of the sugars of both strands except for the T4 residue (Table IV), whereas the release of all constraints leads to notable modifications of the sugar puckering. These results show that the intraresidue constraints were necessary and sufficient to provide sugar conformations.

The average values of the torsion angles and of various helical parameters are given in Tables V and VI. The torsion

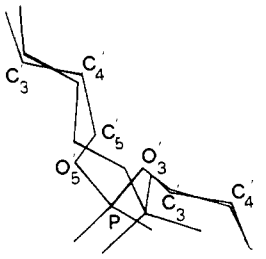

FIGURE 5: View of the two limit conformations of the backbone of the α-strand in the unnatural duplex α-d(TCTAAAC)-β-d(AGATTTG).

angles on the β-strand show little variation from one residue to another; they are typical of a B-DNA conformation: g^-, t, g^+, t, and g^-, respectively, for α, β, γ, ε, and ζ (Fratini et al., 1982). On the α-strand, α, β, and γ, in contrast, tend to occur in two distinct states: g^-, g^+, and g^+ in the majority of the cases but g^+, g^-, and t at positions C2 and A4. The differences between the two structures are illustrated in Figure 5 and correspond globally to a rotation of P, O5′, and C5′ atoms; O3′, C3′, and C4′ remain in the same positions. The angle δ in each case reflects the sugar conformation. The glycosyl C1′–N torsion angles χ correspond to an anti conformation of the β-strand (χ = −100°). On the α-strand, the χ value of 135° corresponds to a syn position of the base with respect to the furanose ring; however, the inversion of the chirality of the C1′ atom leads to a 120° rotation of the base in the α-position compared to the β-form. More generally,

Table III: Calculated Distances and Difference between Experimental and Calculated Distances

	(A) Intranucleotide (H8/H6) with Sugar Atoms						
	H1′	H2′	H2″	H4′	H3′	H5′	H5″
	Calculated/Difference						
α T1	3.75	4.14	3.08	2.98			
	0.09	−0.25	−0.10	−0.08			
α C2		4.11	3.10	2.89			
		−0.25	−0.22	−0.18			
α T3	3.49	3.83	2.72	2.98	3.79		
	−0.23	−0.16	−0.13	−0.08	−0.07		
α A4	3.88		2.99	3.11	3.94		
	0.15		−0.20	−0.18	−0.23		
α A5	3.86	4.19	2.99	3.09	4.11		
	0.10	−0.43	−0.24	−0.22	0.26		
α A6	3.66	4.13	3.05	2.95			
	−0.20	−0.27	−0.22	−0.08			
α C7			3.19	2.76			
			−0.13	−0.11			
	Calculated/Difference						
β A1		2.58	4.03			4.36	3.53
		−0.07	0.09			0.04	−0.02
β G2	3.75	2.11		4.74			
	−0.16	0.03		−0.17			
β A3	3.80	2.19	3.76				3.96
	−0.22	0.00	0.33				0.02
β T4	3.69	2.10	3.41				
	−0.01	0.13	−0.02				
β T5	3.77	2.09		4.82			
	0.09	−0.01		−0.08			
β T6	3.68	2.23	3.57				
	−0.17	0.10	0.13				
β G7	3.92	2.30					
	0.02	0.06					

	(B) Internucleotide[a]						
	H1′	H2′	H2″	H6/H8	H5/CH3	H5′	H5″
	Calculated/Difference						
α T1	3.40					2.86	3.74
	0.03					−0.02	−0.19
α C2					3.03[b]		
					−0.10		
α T3	3.22				3.66[b]		
	−0.05				0.15		
α A4	2.66			5.59[b]	3.39[b]	4.82	3.25
	−0.08			0.19	−0.35	−0.91	−0.04
α A5	2.40	4.25	4.87			4.49	
	0.25	−0.24	−0.52			−0.95	
α A6	2.88	4.21		5.45[b]			
	0.22	−0.36		0.03			
	Calculated/Difference						
β G2	4.40	4.50	2.88				
	0.07	−0.10	0.24				
β A3	3.28			4.70[c]	3.10[c]		
	−0.03			−0.08	0.02		
β T4	3.45	3.41	2.40		3.18[c]		
	−0.03	−0.03	0.32		−0.21		
β T5	3.58				3.45[c]		
	0.07				−0.33		
β T6	3.70	4.19	2.55				
	0.10	−0.01	0.11				
β G7		4.05					
		0.01					

[a] α-strand (H6/H8) of residue i with atom of residue $i + 1$. β-strand (H6/H8) of residue i with atom of residue $i - 1$. [b] Distances with atom of residue $i - 1$. [c] Distances with atom of residue $i + 1$.

it has been shown that for the same position of a sugar relative to the base in the α- and β-anomer $\chi_\alpha = \chi_\beta + 200°$. Thus, the distance between the H6/H8 proton of the base and the H1′ atom of the sugar remains in the range of the value observed for a natural oligonucleotide with a base in an anti position (Figure 6).

No general feature emerges from Table VI for the base roll and tilt. Conversely, the values of the helical twist are alternately small and large. The largest one (45°) differs notably from the idealized value for a B-DNA structure (36°) (Arnott et al., 1975); however, values as large as 50° have been observed for the X-ray structure of a B decamer (Privé et al., 1987).

The propeller twist angle for each base pair (Table VI) is relatively small, except for the T1–A1 base pair at the 5′ end of the helix.

It is clear from Figures 3 and 4 that the constrained and the relaxed structures appear to be bent. We tried to access the degree of bending by calculating the angle between local helical axes of consecutive steps. The largest value, 13°, occurs

Table IV: Calculated (P_{theo}, P_{rel}) and Experimental (P_{exp}) Pseudorotation Angle for the Sugars[a]

α	P_{theo}	P_{rel}	P_{exp}	β	P_{theo}	P_{rel}	P_{exp}
T1	207	196	205–215	A1	187	179	185–195
C2	210	210	200–210	G2	192	194	
T3	219	214	215–225	A3	188	194	175–185
A4	210	200	210–220	T4	180	158	180–190
A5	212	199	205–215	T5	173	181	180–190
A6	201	206	200–205	T6	183	188	180–190
C7	196	179		G7	190	189	185–195

[a] The P_{theo} and P_{rel} values were calculated by running the iterative procedure with and without constraints on the puckering angles P, respectively.

Table V: Torsion Angles[a]

α	T1	C2	T3	A4	A5	A6	C7		
α		62	−86	69	−92	−86	−91		
β		−141	133	−133	130	122	140		
γ	61	−175	54	171	49	47	38		
δ	154	157	149	153	147	153	147		
ε	−73	−87	−73	−81	−85	−78			
ζ	178	−165	−172	−161	−171	−170			
χ	129	145	129	138	135	137	133		
β	A1	G2	A3	T4	T5	T6	G7	12-mer[b]	10-mer[c]
α		−74	−70	−66	−70	−70	−66	−62	−71
β		178	−179	−177	178	179	178	169	164
γ	52	67	61	47	62	66	56	55	51
δ	151	150	144	138	148	150	152	121	142
ε	−161	−179	177	−177	−176	−174		−173	−173
ζ	−125	−119	−87	−113	−118	−127		−105	−107
χ	−118	−92	−97	−96	−96	−93	−97	−117	−99

[a] Main-chain torsion angles are defined by $P\text{—}O5'_{\beta}\text{—}C5'_{\gamma}\text{—}C4'_{\delta}\text{—}C3'_{\epsilon}\text{—}O3'_{\zeta}\text{—}P$; χ is the glycosyl bond torsion angle. The values for ε and ζ are the average values for a B_I conformation. [b] Dodecamer d(CGCGAATTCGCG) (Fratini et al., 1982). [c] Decamer d(CCAAGATTGG) (Privé et al., 1987).

Table VI

(A) Helical Parameters for Each Base Step

	twist (deg)	roll (deg)	tilt (deg)
T1–A1	38.2	8.4	−10.8
C2–G2	43.0	0.9	1.5
T3–A3	35.6	−7.6	−4.8
A4–T4	39.0	2.0	−0.8
A5–T5	38.0	−3.8	6.0
A6–T6	45.0	4.2	5.3
C7–G7			
DMP7[a]	35.9 ± 4.5	−2.3 ± 6.0	0 ± 1.5

(B) Propeller Twist for Each Pair of Bases

pair	propeller twist (deg)	pair	propeller twist (deg)
T1 A1	9	A5T5	3
C2G2	1	A6T6	3
T3A3	6	C7G7	5
A4T4	5	*a*	18.7 ± 7

[a] Fratini et al. (1982), average values.

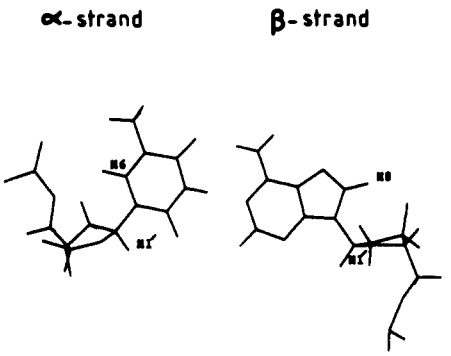

FIGURE 6: Relative position of the base with respect to the sugar in the α- and β-strands. Although the average χ value for both strands differs by 200°, it appears already that the H8/H6–H1′ distances are about the same.

between steps 1 and 2; for steps 2–3, 3–4, and 4–5 the local axes are almost in line, while the angle for steps 5–6 and 6–7 is intermediate, 6°.

All the bends do not occur in the same direction. Thus, the total angle between the local helical vectors of steps 1 and 7 is 18°. This bending is probably a consequence of the base sequence for this parallel duplex; however, no correlation can be found between the bending and the nature of the step (pyr–pur or pur–pyr). Similarly, there is a rather poor stacking of the bases, especially for step 6–7. The stacking is purely an intrastrand stacking typical of B-DNA and is almost nonexistent between consecutive thymines (Figure 7).

To analyze the origin of this feature, we ran a molecular mechanics calculation for the natural antiparallel β-d(CAAATCT)–β-d(AGATTTG). The resulting structure shows a similar bend and poor intrastrand stacking of the bases. The stacking of the consecutive thymines remains very bad. The stacking is almost nonexistent for step 3–4, and there is some overlap of the bases between steps 6 and 7. Thus, the absence of stacking is not a characteristic feature of a parallel duplex and depends more probably on the base sequence; however, it is difficult to correlate the stacking of the base with the base sequence. In any case, if correlation exists, it should be different for parallel and antiparallel duplexes. The quasi absence of stacking is balanced by good hydrogen-bond interactions between the Watson–Crick base pairs. The hydrogen bonds do not deviate from linearity by more than 10°, the X···H (X = O, N) distances ranging from 1.80 to 1.86 Å.

Conclusion

In this work, we presented the first determination of the three-dimensional structure of an unnatural oligonucleotide for which no crystallographic data are yet available.

Owing to the errors involved in the determination of a DNA structure using NMR data, the refined conformation satisfied reasonably the experimental constraints.

The errors mentioned above can arise from several sources:

(i) The accuracy of the spectral parameters measurements depends on several experimental factors (signal/noise ratio, digital resolution, phasing, etc.), the errors being obviously much larger on small NOE than on stronger ones.

(ii) Error may occur during the conversion of the NOESY data into interproton distances. However, these errors have been reduced by the introduction of an iterative process that allows us to take into account the spin diffusion by using the structure refined in a previous step to determine the interproton distances from the NOESY maps.

(iii) The dynamics of the molecule may cause errors. In solution, DNA experiences large amplitude motions (Behling & Kearns, 1986; Holbrook & Kim, 1984; Levy et al., 1981); and the distances derived from NMR measurements may be considered as an average value of the fluctuating interproton distances. Our energy-minimized structure gives only a static picture of the conformation of the molecule. It is tempting to run a molecular dynamics simulation to obtain a time-averaged structure; however, the simulation of these motions by molecular dynamics is presently restricted by the short time range ratio of the simulations (10 ps) compared to the very low rate of the motions (1 ns).

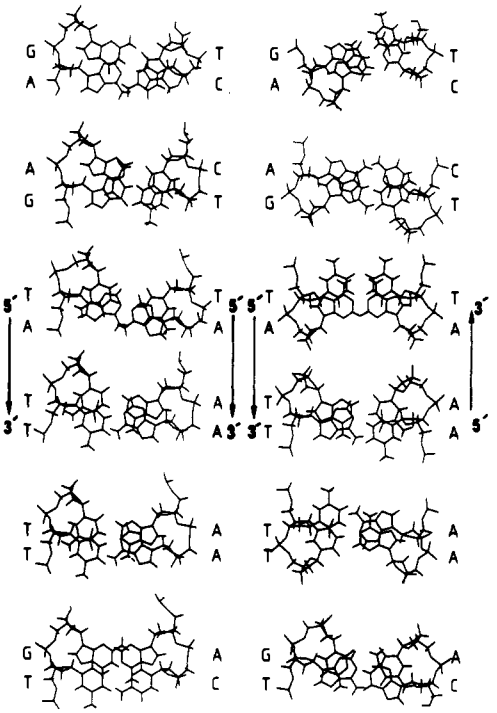

FIGURE 7: Scheme of the six individual base-pair steps of the constrained structure viewed down the global helix axis. (Left) Unnatural duplex between the parallel strands α-d(TCTAAAC) and β-d(AGATTTG). (Right) Natural duplex between the antiparallel β-d(CAAATCT) and β-d(AGATTTG).

Apart from the intrinsic limitations of the procedure, we note that compared to other molecular mechanics calculations incorporating distance constraints derived from NMR measurements (Behling et al., 1987), we have used quite a large data set involving not only intraresidue interactions between the base and the sugar protons but also interresidue interactions that define the positions of the bases relative to each other. The distance data set was sufficient to define the conformation of sugar and the relative position of the base with respect to its own sugar and to the sugar of the neighboring residue. Moreover, the sugar conformation obtained by this method was in agreement with the conformation deduced from the coupling constants (Table IV). Two phosphodiester backbone conformations (Figure 5) corresponding to two sets of torsion angles (Table V) allowed us to link the C3′ and C5′ atoms of two neighboring sugars; these spatial positions were determined by NMR constraints. Although the H8 or H6–H5′/H5″ correlation time was difficult to determine, it is worth noting that the calculated H8 or H6–H5′/H5″ distances were in agreement with the values to the experimental data.

We conclude that we can be confident in the relation of positions of residues, but the phosphodiester backbone conformation may be questionable.

Although it has been argued that inclusion of the counterions and hydrogen-bonded water molecules improves the solution structure model (Suzuki et al., 1986; Zhou et al., 1987), we

did not try to include solvent molecules in our molecular mechanics calculations. As already noted (Nilges et al., 1987), the energy-minimized structure including NOE constraints should reflect the solution structure since the interproton distances introduced as constraints are measured for the solute molecule.

The use of a correlation time of 2.25 ns allows a good fit of almost all the base–base or base–sugar inter- or intraresidues connectivities, but the simulated values of the NOE between H2′ and H2″ are rather different from the experimental data. It has been recently reported (Nerdal et al., 1988) that the spin diffusion rather than local motion was responsible for the apparent H2′–H2″ cross-relaxation rate being slower than expected when scaled to the fixed H5–H6 cytosine rate for a 12 base pair natural DNA duplex. For our 7 base pair unnatural hybrid DNA duplex where the NOE-simulated values have been calculated by taking into account all spin diffusion effects, this disagreement is probably due to the internal motion of the sugar in the whole oligonucleotide. In fact, the dynamical behavior of a short 7 base pair unnatural DNA at 22 °C whose T_m value is 33 °C should be rather different from that of a 12 base pair at 42 °C. We have tentatively taken into account this effect by using the model of internal motion developed by Woessner (1962). A good fit of the data was obtained with an internal correlation time of 0.65 ns. The same simulation was made for the H1′–H2″ connectivities, which vary only slightly with the conformation of the sugar, and a good fit of the data was obtained for an internal correlation time of 0.75 ns. The discrepancy between the two internal correlation times is probably due to the fact that the internal motions of the H2′–H2″ and H1′–H2″ vectors are different and not isotropic. This difficulty in simulating the NOE in the sugar moiety has led us to determine the sugar conformation by using only the H8/H6–sugar proton connectivities.

It is worth noting that the final A and B structures were very similar as shown by their atomic rms difference (0.51 Å) and Figure 2, while the atomic rms difference between the initial and final structures was 2.75 Å. The rms differences between the initial and final coordinates were 3.36 and 1.72 Å for A and B, respectively.

The main differences between the A and B structures concern the geometry of the phosphodiester backbone. In the B structure, two groups of torsional angles were alternatively found in the oligonucleotide sequence (Table V), while only one group is obtained for the A structure: $\alpha = 85 \pm 5°$, $\beta = 120 \pm 5°$, $\gamma = 42 \pm 7°$, $\epsilon = -85 \pm 4°$, $\zeta = -165 \pm 5°$, and $\chi = 135 \pm 5°$. As already reported by Pardi et al. (1988), we found that the aromatic bases and sugars are more precisely defined than the sugar–phosphate backbone. Indeed, only proton–proton distance NMR data were introduced in the constraint data base.

Running the AMBER program from the starting structures without NMR constraints led to a structure significantly different from the constrained structure. Moreover, the potential energy obtained for these constraint-free structures was higher than the corresponding value obtained for the constrained structure.

Although we have used molecular mechanics calculation, which is well-known to converge to an energy minimum close to the starting structure, it appears clearly here that the introduction of NOE constraints increases the ability of this method to allow us to reach a conformation different from the starting one.

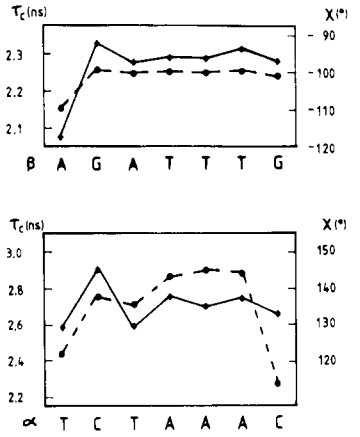

FIGURE 8: Representation of the χ glycosyl angle (◆) and the best effective correlation time (●) for each base of both strands. (Top) β-strand. (Bottom) α-strand. Some correlation between χ and τ_c appears for both strands (see text).

Table VII: Calculated Distances and Difference between Experimental and Calculated Distances between the Methyl Group (H7) and Other Protons

distance	calculated	difference
α H6(C2)–H7(T1)	3.03	−0.10
α H8(A4)–H7(T3)	3.39	−0.35
β H8(A3)–H7(T4)	3.10	0.02
β H2′(A3)–H7(T4)	2.73	−0.02
β H2″(A3)–H7(T4)	3.36	0.15
β H6(T4)–H7(T5)	3.18	−0.21
β H6(T5)–H7(T6)	3.45	−0.33

We have tried to obtain some information on the motion of each deoxyribose moiety relative to its base by determining for each nucleotide the effective correlation time able to give simultaneously the best fit for the intraresidue NOE H8/H6–sugar protons. Figure 8 shows some relative correlation between this τ_c effective and the χ glycosyl angle for each nucleotide. We tentatively explain this correlation by remarking that the values of these NOE are a function of the internal motion of the sugar; the amplitude and spectrum in frequency may be a function of the χ glycosyl angle. Data on only one oligonucleotide cannot allow us to decide upon the existence of a correlation, and comparisons with several other molecules are desirable.

Until now the proton–methyl group distances were not used (Broido et al., 1985; Jamin et al., 1985; Suzuki et al., 1986; Zhou et al., 1987, 1988) or cautiously used (Tropp, 1980; Tropp & Redfield, 1981; Nerdal et al., 1988) for the structure refinement of oligonucleotides. To assess the reliability of proton–methyl group distance determination, our data base has been enlarged by the inclusion of methyl protons–H6/H8 distances. The conversion of the NOE intensities into distances as well as the refinement of the structure on the basis of these experimental distances is made by assuming a free rotation of the methyl hydrogens around the C6–C7 bond. Examination of the refined structure on the graphic system shows that this assumption of free rotation may not hold for each methyl group. More specifically, in the β-strand, during the rotation, the T5 and T6 methyl protons may come respectively into closer contact with the H2′ and H2″ protons of T4 and T5 [r(H···H) < 1.90 Å], the steric hindrance being much larger for T6 than for T5. This may explain the discrepancy observed between the experimental and the refined distances for this residue (Table VII). However, except for this specific residue, the agreement is quite satisfactory and the inclusion of these distances in our data set allows us to confirm the

positions of the thymines with respect to their neighboring residues.

Registry No. α-d(TCTAAAC)-β-d(AGATTTG), 121810-01-3.

REFERENCES

Abragam, A. (1978) in *The Principles of Nuclear Magnetism*, Claredon Press, Oxford, U.K.

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

Arnott, S., & Hukins, D. W. L. (1972) *Biochem. Biophys. Res. Commun. 47*, 1506–1509.

Arnott, S., Campbell Smith, P. J., & Chandrasekharan, R. (1975) *CRC Handbook of Biochemistry and Molecular Biology. Nucleic Acids*, Vol. 2, pp 411–422, CRC Press, Cleveland, OH.

Asseline, U., Delarue, M., Lancelot, G., Toulmé, F., Thuong, N. T., Montenay-Garestier, T., & Hélène, C. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 3297–3301.

Behling, R. W., & Kearns, D. R. (1986) *Biochemistry 25*, 3335–3346.

Behling, R. W., Rao, S. N., Kollman, P., & Kearns, D. R. (1987) *Biochemistry 26*, 4674–4681.

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

Broido, M. S., James, T. L., Zon, G., & Keepers, J. W. (1985) *Eur. J. Biochem. 150*, 117–128.

Drobny, G., Pines, A., Sinton, S., Witekamp, D. P., & Wemmer, D. (1979) *Faraday Div. Chem. Soc. Symp. 13*, 49–55.

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

Holbrook, S. R., & Kim, S.-H. (1984) *J. Mol. Biol. 173*, 361–388.

Jamin, N., James, T. L., & Zon, G. (1985) *Eur. J. Biochem. 152*, 157–166.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Lancelot, G., Guesnet, J. L., Roig, V., & Thuong, N. T. (1987) *Nucleic Acids Res. 15*, 7531–7547.

Levy, G. C., Hilliard, P. R., Levy, L. F., & Rill, R. L. (1981) *J. Biol. Chem. 256*, 9986–9989.

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

Marion, D., & Lancelot, G. (1984) *Biochem. Biophys. Res. Commun. 124*, 774–783.

Morvan, F., Rayner, B., Imbach, J. L., Lee, M., Hartley, J. A., Chang, D. K., & Lown, J. W. (1987) *Nucleic Acids Res. 15*, 7027–7044.

Nerdal, W., Hare, D. R., & Reid, B. R. (1988) *J. Mol. Biol. 201*, 717–739.

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

Noggle, S. H., & Schirmer, R. E. (1971) in *The Nuclear Overhauser Effect: Chemical Applications*, Academic Press, New York.

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

Pardi, A., Hare, D. R., & Wang, C. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8785–8789.

Pflugrath, J. W., Saper, M. A., & Quiocho, F. A. (1984) *Methods and Applications in Crystallographic Computing* (Halland, S., & Ashirka, T., Eds.) p 407, Clarendon Press, Oxford, U.K.

Privé, G. G., Heineman, V., Chandrasegaran, S., Kan, L. S., Kopka, M., & Dickerson, R. (1987) *Science 238*, 498–504.

Scalfi Happ, C., Happ, E., Nilges, M., Gronenborn, A. M., & Clore, G. M. (1988) *Biochemistry 27*, 1735–1743.

Suzuki, E., Pattabiraman, N., Zon, G., & James, T. L. (1986) *Biochemistry 25*, 6854–6865.

Thuong, N. T., Asseline, U., Roig, V., Takasugi, M., & Hélène, C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5129–5133.

Tropp, J. (1980) *J. Chem. Phys. 72*, 6035–6043.

Tropp, J., & Redfield, A. G. (1981) *Biochemistry 20*, 2133–2140.

Weiner, P., & Kollman, P. (1981) *J. Comput. Chem. 2*, 287–303.

Weiner, S. J., Kollman, P. A., Case, D. A., Singh, V. C., Ghio, C., Alagona, G., Profeta, S., & Weiner, P. (1984) *J. Am. Chem. Soc. 106*, 765–784.

Woessner, D. E. (1962) *J. Chem. Phys. 37*, 647–654.

Zhou, N., Bianucci, A. M., Pattabiraman, N., & James, T. L. (1987) *Biochemistry 26*, 7905–7913.

Zhou, N., Manogaran, S., Zon, G., & James, T. L. (1988) *Biochemistry 27*, 6013–6020.