Acta Cryst. (1995). D**51**, 1025–1035

Sodium Ions and Water Molecules in the Structure of Poly(dA)·Poly(dT)

BY RENGASWAMI CHANDRASEKARAN,* AKELLA RADHA AND HYE-SHIN PARK

Whistler Center for Carbohydrate Research, 1160 Smith Hall, Purdue University, West Lafayette, IN 47907, USA

(*Received* 3 *October* 1994; *accepted* 8 *February* 1995)

Abstract

Analysis of X-ray diffraction data from a polycrystalline and well oriented fiber of the sodium salt of poly(dA)·poly(dT) shows that this B′-DNA corresponds to a right-handed antiparallel tenfold double-helix of pitch 32.4 Å with C2′-*endo* furanose rings in both strands. The helix contrasts itself from B-DNA in terms of a very narrow minor groove. Difference electron-density maps have revealed that a continuous spine of water molecules, two per base pair, propagates along this groove with the same symmetry as the DNA and establishes new links between the two strands. In addition to this hydrated DNA helix, the monoclinic unit cell (space group $P2_1$) accommodates about 20 sodium ions and 12 water molecules in the vicinity of phosphate groups. These structured guest molecules provide an intricate network of bridges, ranging in size from a single sodium ion to a multiple sodium–water–water–sodium unit, connecting phosphate groups belonging to adjacent DNA helices. The crystallographic R value for this structure is 0.23 for a total of 102 reflections extending out to 3.2 Å resolution.

Introduction

DNA duplexes containing purines in one strand and pyrimidines in the other and interlinked only by two Watson–Crick hydrogen bonds in every base pair always adopt a unique morphology. X-ray fiber diffraction studies have demonstrated that the molecular structures of poly(dA)·poly(dT), poly(dA)·poly(dU), poly(dI)·poly(dC) and poly[d(AI)]·poly[d(CT)] can be described in terms of an antiparallel, right-handed double helix, termed B′-DNA (Arnott & Selsing, 1974; Leslie, Arnott, Chandrasekaran & Ratliff, 1980; Park, Arnott, Chandrasekaran, Millane & Campagnari, 1987; Chandrasekaran, Radha, Park & Arnott, 1989; Chandrasekaran & Radha, 1992). It consists of ten base pairs in a pitch length of about 32.3 Å which is 1.5 Å shorter than that for classical B-DNA (Chandrasekaran & Arnott, 1989). Owing to movement of base pairs into the minor groove by approximately 0.6 Å in every step, the minor groove is narrowed, but the major groove widened, relative to B-DNA. The results further suggest that the spatially equivalent A·T, A·U and I·C base pairs, despite chemical differences, have virtually identical structural behavior. Morphological differences are responsible for DNA bending at junctions between fragments of B- and B′-DNA's. When stretches of about five base pairs alternate between these two conformations, these differences are appropriate to induce significant overall curvature in A,T-rich genomic DNA (Koo, Wu & Crothers, 1986).

Sodium salt of poly(dA)·poly(dT) has so far been trapped in at least three distinct crystalline forms in oriented fibers. Structural details for two of them are already published. The first, a low salt (0.01 *M* NaCl) specimen, is a screw-disordered α-form, obtained at 92% relative humidity (r.h.), and it corresponds to a one-molecule trigonal unit cell (Park, Arnott, Chandrasekaran, Millane & Campagnari, 1987). The second, also a low-salt (0.01 *M* NaCl) specimen, is a polycrystalline β-form at a reduced (77%) r.h. corresponding to a two-molecule monoclinic unit cell (Chandrasekaran & Radha, 1992). The third, obtained with high salt (0.1 *M* NaCl), is also a polycrystalline form, but it has a much smaller, one-molecule monoclinic unit cell. Henceforth, this is referred to as the β_1-form and the previously reported β-form as the β_2-form, the subscript denoting Z, the number of molecules per unit cell. X-ray diffraction analysis of the new β_1-form of poly(dA)·poly(dT) has now been completed. In addition to defining its molecular structure, we have also identified in difference electron-density maps a number of sodium ions and water molecules around the double-helices. Thus, for the first time, we have obtained and are presenting the structural details of a polymeric DNA and its associative properties mediated by cations and water molecules.

Materials and methods

Sample preparation

Poly(dA)·poly(dT) was synthesized enzymically by Dr R. L. Ratliff, Los Alamos National Laboratory, by using *Micrococcus luteus* DNA polymerase as previously described (Leslie, Arnott, Chandrasekaran & Ratliff, 1980). The DNA was re-precipitated in 0.05 *M* Tris–HCl, pH 8.6 buffer containing 0.1 *M* NaCl with ice-cold ethanol. It was then washed successively with two volumes of 80 and 100% ethanol, finally rinsed with acetone, and dried in a vacuum desiccator. A small

* Author for correspondence.

amount of this DNA was dissolved to make an aqueous solution of about 5 mg ml^{-1} concentration. Fibers were then stretched in fiber pullers at 92% r.h., using a few drops of this DNA solution per fiber (Fuller, Hutchinson, Spencer & Wilkins, 1967).

X-ray data

A diffraction pattern was recorded using a flat-film pin-hole camera with a specimen-to-film distance of about 4 cm using nickel-filtered Cu $K\alpha$ radiation (λ = 1.5418 Å) from a microfocus generator. The specimen chamber was continuously flushed with a steady stream of helium gas which, after bubbling through an appropriate salt solution, maintained the fiber at 90% or higher r.h. as desired and also helped to reduce air scattering that could otherwise fog the film. The ring (characteristic spacing 3.035 Å) on a pattern obtained from a DNA fiber dusted with calcite powder was used for internal calibration in measuring unit-cell dimensions.

Diffraction patterns were digitized on an optronics P-1000 rotating-drum microdensitometer with a 100 μm raster step and processed into final Lp-corrected structure amplitudes. The computational details for the various stages from background removal to evaluation of integrated intensities for Bragg reflections are the same as those previously reported (Chandrasekaran, Radha & Ratliff, 1994). The lowest measured intensity of an observed spot was assigned as the threshold value to each unobserved (too weak to be seen) spot in the pattern.

Model building and refinement

The linked-atom least-squares (LALS) refinement technique (Smith & Arnott, 1978) was used to generate molecular models compatible with the observed helical parameters. While bond lengths and bond angles were kept fixed, the conformation angles α to ζ and χ in A and T nucleotides, and the propeller π in the A·T base pair were refined. The nucleotide nomenclature is that of the Cambridge convention (Dickerson, 1989). In the final stages of refinement, endocyclic bond angles, and ν_0 to ν_4, in the furanose ring were also refined subject to ring-closure constraints. Orientation of the helix in the unit cell was a variable packing parameter.

Since the details of the least-squares refinement procedure are the same as those described by Chandrasekaran, Radha & Ratliff (1994), only the key points are given here. The function minimized is of the form $\Omega = \sum w_m \Delta F_m^2 + \sum e_i \Delta \theta_i^2 + \sum k_j \Delta c_j^2 + \sum \lambda_h G_h = X + E + C + L$. The first term on the right-hand side accounts for the sum of the squares of the differences between observed and calculated X-ray structure amplitudes. An unobserved reflection was included in the refinement only when its calculated structure amplitude was equal to or higher than the observed amplitude. The second term minimizes differences between expected/standard values and corresponding conformation and bond angles of the current model. The third term includes both intra- and interchain hydrogen bonds, and differences between acceptable and calculated non-bonded distances for those contacts which are smaller than the acceptable limiting values; this is designed to keep the model free from steric compression. Finally, the fourth term imposes constraints (with Lagrange multipliers λ_h) for helix connectivity and ring closure, and vanishes when the applied constraints are satisfied. During refinement, structure factors were calculated using water-smeared atomic scattering factors (Chandrasekaran & Radha, 1992). Because of the low resolution (~3 Å) of the X-ray data, the thermal parameter for all atoms was fixed at $B = 4$ Å^2 following common practice (Alexeev, Lipanov & Skuratovskii, 1987; Chandrasekaran & Radha, 1992; Chandrasekaran, Radha & Ratliff, 1994). Base-pair related features were computed with Dickerson's *NEWHEL92* program.

Location of guest molecules

Difference Fourier maps were computed using $(2F_o - F_c)$ as coefficients, in which F_c and its phase were derived from the current model corresponding to normal atomic scattering factors (*International Tables for X-ray Crystallography*, 1974). Peaks of positive electron density, not belonging to DNA atoms, were carefully examined for possible sites of sodium ions, and alternately as water molecules, that could associate with at least two functional groups. The choice between sodium and water was ambiguous from electron density alone, but otherwise verified from considerations of its surroundings and relative drop in the crystallographic R value upon its inclusion in the crystal structure.

Results

Unit-cell dimensions and contents

The X-ray diffraction pattern shown in Fig. 1 corresponds to one of the best polycrystalline and well oriented fibers of sodium poly(dA)·poly(dT) at 92% r.h. Bragg reflections are present up to about 3.2 Å resolution which include 75 non-meridional spots. The X-ray data also contain another set of 87 below-threshold spots. For comparison with related polycrystalline structures in fibers, the two respective numbers are 113 and 150 in β_2 poly(dA)·poly(dT) (Chandrasekaran & Radha, 1992) and 74 and 99 in β poly(dA)·poly(dU) (Chandrasekaran, Radha, Park & Arnott, 1989), both of which have similar larger unit cells. The overall intensity distribution is similar to that reported for β_2 B′-DNA (Chandrasekaran & Radha, 1992) and the strong meridional reflection on the tenth layer line indicates tenfold helix symmetry. The observed spots could be indexed with monoclinic unit-cell dimensions a = 18.56(5), b = 22.68(9), c(fiber axis) = 32.40(8) Å, and γ = 99.9(3)° (V = 13 435 Å^3).

Space group was assigned as $P2_1$ on the basis of systematic ($00l$, l odd) absences.

Noteworthy is the novel $a\sin\gamma = b\cos36°$ relationship in contrast to that ($b = a\tan36°$) for B-DNA (Chandrasekaran, Radha & Ratliff, 1994). This suggests that tenfold screw symmetry which relates adjacent base pairs by 36° turn angle corresponds to the best packing arrangement of the DNA helices. Fiber density, measured by the flotation method (using bromobenzene and acetone) was 1.44 g cm^{-3}. There is no room for two, but the unit cell can accommodate one turn of a double helix consisting of ten A·T pairs which accounts for only 53% of the fiber by weight with the remaining 47% of non-DNA density arising from ions and water molecules. If it is only one or the other, the guest molecules in the unit cell may be 240 sodium ions or 306 water molecules. However, an appropriate combination of both is realistic.

Molecular structure and packing arrangement

The starting molecular model was adopted from the similarly high humidity α-form of poly(dA)·poly(dT) (Park, Arnott, Chandrasekaran, Millane & Campagnari, 1987). The molecular asymmetric unit of the tenfold right-handed, Watson–Crick base-paired, antiparallel double helix consists of an A nucleotide in one strand and a T nucleotide in the other so that the two strands are chemically, and may be conformationally, different. This helix was placed at the unit-cell origin with its molecular axis coinciding with the c axis, and the best value for its orientation (μ) about the c axis was selected by examining the variation of intermolecular contacts as a function of μ. The position of minimum steric compression also corresponded to the minimum R value in this search. Using this as the initial set up, the molecular parameters, including those for flexible sugar rings, and the packing parameter were refined along with the X-ray scale factor. At the end of refinement, this model was satisfactory on stereochemical grounds, and its R value (0.28) low and acceptable as in previously reported fiber structures: 0.25 for β_2 poly(dA)·poly(dT) (Chandrasekaran & Radha, 1992), 0.30 for β poly(dA)·poly(dU) (Chandrasekaran Radha, Park & Arnott, 1989), and 0.25 for the B-DNA of poly[d(AATT)] (Chandrasekaran, Radha & Ratliff, 1994). The packing arrangement (Fig. 2) showed interactions between DNA helices that involved peripheral phosphate groups. They were more intimate along the a direction than along b, and there were none along either diagonal in the ab plane.

Aborted models

We also considered two additional possibilities. The first was a 10_1 helix in which the polarity of the DNA was reversed along the z direction. The second was a relaxed 2_1 model consisting of five base pairs per repeating unit and having 60 extra degrees of freedom. Both were refined in a similar fashion as described above. Their final R values were 0.30 and 0.27, respectively. Hence, these alternate models were rejected from further consideration at greater than 99.5% confidence level (Hamilton, 1965).

Inclusion of non-DNA molecules in the crystal structure

The empty columns between helices (Fig. 2) were the search sites for the ordered sodium ions and water molecules. Consistent with space-group requirement, the

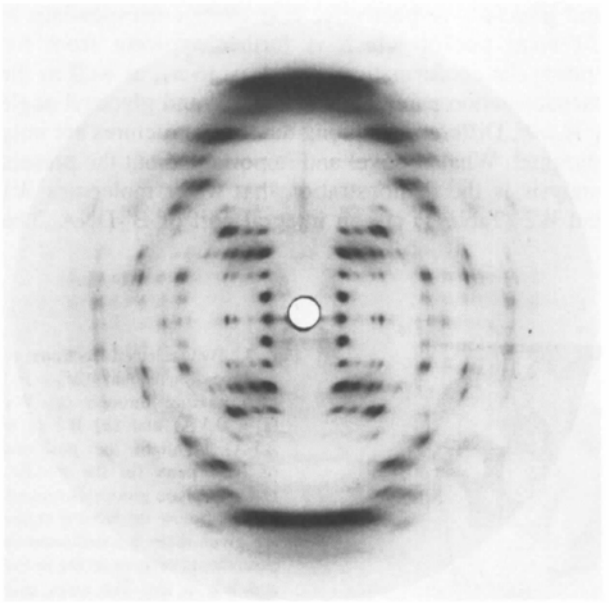

Fig. 1. X-ray diffraction pattern from a well oriented and polycrystalline fiber of the sodium salt of poly(dA)·poly(dT).

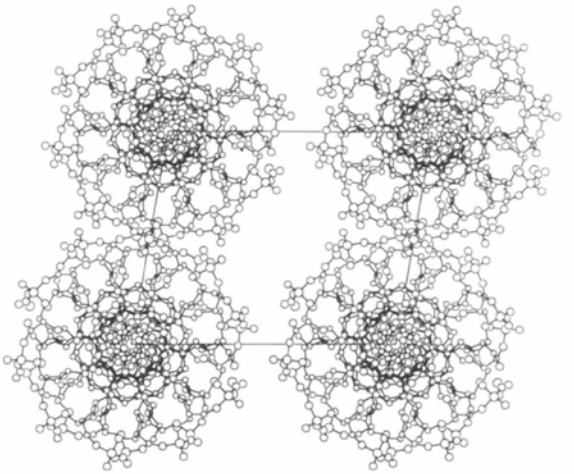

Fig. 2. Packing arrangement of DNA helices in the monoclinic unit cell shows inter digitation only along the a axis.

electron-density maps were computed for one half of the unit cell ($-a/2$ to $a/2$, $-b/2$ to $b/2$ and 0 to $c/2$) in sections normal to c at intervals of $c/33$. The first $(2F_o - F_c)$ map was very informative in revealing five strong peaks, all in the minor groove near functional groups of DNA and they further seemed to adopt the DNA-helix symmetry. Every peak appeared in the proximity of a base pair whose atoms O2 of thymine and N3 of adenine were at the same level as, and a level above, that of an ion/water (Fig. 3) at distances close to 3 Å, respectively. In addition, a furanose ring O atom of the T strand was at a similar distance from each guest peak. One of them was chosen to be the principal peak representing a periodic guest molecule and the DNA symmetry was used to generate the rest. Structure-factor calculation, with normal atomic scattering factors, established the superiority of this peak (Fig. 3*a*) as a water molecule (*W*1) rather than a sodium ion and the R values, before and after inclusion of this, were 0.35 and 0.33, respectively.

The map was helpful to identify four more peaks between the DNA helices. Because of their ability to bridge phosphate groups in their vicinity, they were designated as sodium ions (*S*3 to *S*5 and *S*9) and verified to be the case *vis-à-vis* water molecules. These ions did not follow the DNA-helix symmetry and their inclusion brought down the R value to 0.32.

The second $(2F_o - F_c)$ map, based on the augmented crystal structure, was cleaner than the first. Once again, a series of five strong peaks appeared roughly midway between *W*1 and its helix mates in a regular fashion. This was assigned to be the second periodic water molecule (*W*2) in the minor groove (Fig. 3*b*) which was able to form hydrogen bonds with flanking helical mates of *W*1, as well as with a furanose ring O atom of the A strand shown in Fig. 3(*a*). Another 12 peaks, some near phosphate groups and others between DNA helices were verified to be excellent guest molecules. These peaks, unrelated among themselves, consist of six sodium ions (*S*1, *S*2, *S*6 to *S*8 and *S*10) and six water molecules (*W*3 to *W*8). Their inclusion led to a substantial reduction in R value (0.28). The third $(2F_o - F_c)$ map was nearly featureless and the search for additional guest molecules ended. Specific roles of these molecules in stabilizing the association of DNA helices are given later.

The choice between sodium and water for the 16 guest peaks appearing between DNA helices was based on two criteria. The first was guided by Hamilton's significance test (Hamilton, 1965). This led to the designation of ten of these peaks as ions, five of them at 90% and the remaining five at 75% confidence level, others as water molecules involved in water–ion or water–water links. Secondly, each of these ten alleged sodium ions were tantalisingly located in close proximity of phosphate groups. This arrangement of ions not only confirmed the initial assignment, but also is required on stereochemical grounds for balancing the distribution of negative charge on the periphery of DNA.

The final round of refinement of the crystal structure containing the DNA helix and all the above water molecules and sodium ions, using water-smeared atomic scattering factors, produced very small shifts in the variable parameters, and reduced the R value to 0.23. This covers all the 75 observed and 27 below-threshold spots. The atomic coordinates of this model are given in Table 1. The observed and calculated structure amplitudes are listed in Table 2.

Conformational features

The double helix has a compact shape similar to those of the other two forms, α and β_2. Both strands display the same conformational domain as other B-type structures. The similarities may be verified by comparing respective conformation angles (Table 3) both within and between duplexes. Following the same trend observed in polymers as well as oligomers in this type, in all cases, angles α, β and γ are *gauche*$^-$, *trans* and *gauche*$^+$, respectively; δ is *trans* corresponding to C2′-*endo* pucker which is further apparent from the endocyclic conformation angles, ν_0 to ν_4, as well as the pseudorotation parameters P and τ_m; and glycosyl angle χ is *anti*. Differences among the three structures are only marginal. What is novel and important about the present analysis is the demonstration that water molecules *W*1 and *W*2 (Table 1) are an integral part of B′-DNA. The

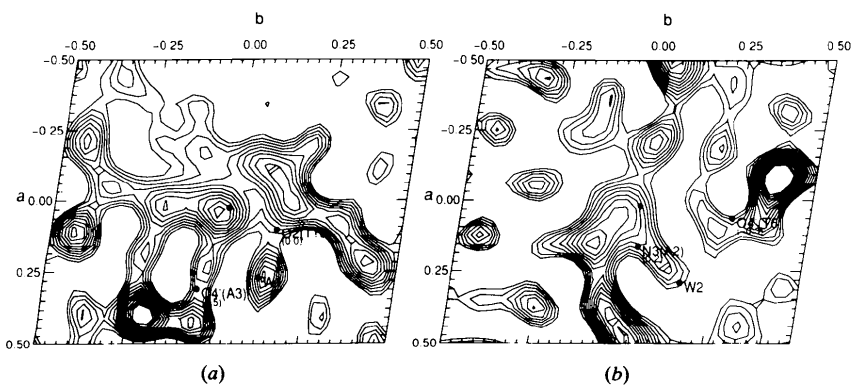

Fig. 3. Two selected sections of a three-dimensional $(2F_o - F_c)$ map passing through (*a*) W1 (z = 0.5 Å) and (*b*) *W*2 (z = 2.5 Å) highlight the positions of base pairs (in the middle) and phosphate groups (strongest peaks). Below the labeled atoms are given their z coordinates in parentheses relative to *W*1 in (*a*) and *W*2 in (*b*). The helix axis (•) passes through the center of the diagram. Negative peaks are suppressed for clarity.

Table 1. *Cartesian and cylindrical polar coordinates of nucleotides in an* A·T *pair, sodium ions and water molecules in the* β_1 *form of poly(d*A)·*poly(d*T)

The a^* axis in the unit cell corresponds to $\varphi = 0$ and $z = 0$. The cylindrical polar coordinates of the nth repeating unit of the base pair and of the two water molecules *W*1 and *W*2 in the helix are given by $[r, \varphi + (n-1)36, z + (n-1)3.24]$.

		x(Å)	y(Å)	z(Å)	r(Å)	φ(°)
Phosphate	P	2.4350	−8.9135	−3.1448	9.2401	−74.72
	O1P	2.8540	−9.7378	−4.3004	10.1474	−73.67
	O2P	1.6752	−9.6163	−2.0869	9.7611	−80.12
	O3′	1.5956	−7.6529	−3.6610	7.8175	−78.22
	O5′	3.7148	−8.1998	−2.5025	9.0020	−65.63
Sugar	C1′	3.2346	−4.6649	−0.9442	5.6766	−55.26
	C2′	3.5040	−5.9370	−0.1475	6.8939	−59.45
	C3′	4.8256	−6.1975	−0.8702	7.8547	−52.09
	C4′	4.5029	−5.9155	−2.3375	7.4343	−52.72
	C5′	4.3573	−7.1343	−3.2275	8.3597	−58.59
	O4′	3.2817	−5.1231	−2.2863	6.0841	−57.36
	H1′	4.0145	−3.9328	−0.7346	5.6199	−44.41
	H12′	2.8010	−6.7383	−0.3754	7.2973	−67.43
	H22′	3.6983	−5.7449	0.9076	6.8324	−57.23
	H3′	5.2314	−7.2046	−0.7744	8.9036	−54.02
	H4′	5.2583	−5.2492	−2.7542	7.4299	−44.95
	H15′	5.3426	−7.4625	−3.5585	9.1778	−54.40
	H25′	3.7589	−6.8769	−4.1014	7.8371	−61.34
Adenine	N1	0.6414	−0.3087	−0.1059	0.7118	−25.70
	C2	1.9297	−0.5585	−0.3346	2.0089	−16.14
	N3	2.5281	−1.7108	−0.5424	3.0526	−34.09
	C4	1.6407	−2.7260	−0.5014	3.1817	−58.96
	C5	0.2977	−2.6268	−0.2789	2.6436	−83.53
	C6	−0.2230	−1.3398	−0.0701	1.3582	−99.45
	N6	−1.5213	−1.0949	0.1596	1.8744	−44.26
	N7	−0.2931	−3.8825	−0.3085	3.8935	−94.32
	C8	0.6954	−4.6884	−0.5442	4.7397	−81.56
	N9	1.9000	−4.0602	−0.6374	4.4828	−64.92
	H2	2.5263	0.2439	−0.3504	2.5380	5.51
	H16	−1.8357	−0.1563	0.3015	1.8423	−75.13
	H26	−2.1751	−1.8510	0.1890	2.8561	−39.60
	H8	0.5850	−5.6790	−0.6239	5.7091	−84.12
Water	*W*1	5.007	1.7894	0.5402	5.3112	19.69
	*W*2	4.8426	3.8224	2.4289	6.1694	38.29
	*W*3	−6.4097	5.2084	1.8098	8.2590	140.90
	*W*4	5.9595	−10.5660	0.0264	12.1221	−60.55
	*W*5	−0.8632	−8.1437	0.9136	8.1893	−96.05
	*W*6	7.0422	−8.5739	8.2870	11.0953	−50.60
	*W*7	−4.3092	−10.5573	10.5205	11.4337	−112.58
	*W*8	−6.5869	−8.5826	9.3625	10.8189	−127.51

		x(Å)	y(Å)	z(Å)	r(Å)	φ(°)
Phosphate	P	−1.3284	9.3450	2.6656	9.4390	98.09
	O1P	−1.3085	10.5053	3.5851	10.5865	97.10
	O2P	−2.3285	9.3963	1.5758	9.6805	103.92
	O3′	−1.5201	7.9977	3.5070	8.1409	100.76
	O5′	0.1302	9.1260	2.047	9.1269	89.18
Sugar	C1′	1.3249	5.6967	0.8164	5.8487	76.91
	C2′	1.1269	6.9223	−0.0692	7.0134	80.75
	C3′	2.1668	7.8044	0.6217	8.0996	74.48
	C4′	1.9284	7.5043	2.1015	7.7481	75.59
	C5′	1.1634	8.5609	2.8745	8.6396	82.26
	O4′	1.2067	6.2393	2.1222	6.3549	79.05
	H1′	2.3132	5.2738	0.6361	5.7588	66.32
	H12′	0.1358	7.3558	0.0647	7.3571	88.94
	H22′	1.4391	6.7281	−1.0953	6.8803	77.93
	H3′	2.0868	8.8779	0.4511	9.1199	76.77
	H4′	2.8830	7.3148	2.5923	7.8624	68.49
	H15′	1.8483	9.3483	3.1891	9.5293	78.82
	H25′	0.7137	8.1098	3.7590	8.1411	84.97
Thymine	N1	0.2901	4.6427	0.6200	4.6518	86.42
	C2	0.7341	3.3516	0.4656	3.4311	77.65
	O2	1.9151	3.0503	0.4839	3.6017	57.88
	N3	−0.2529	2.04025	0.2860	2.4158	96.01
	C4	−1.6131	2.6326	0.2490	3.0875	121.50
	O4	−2.4020	1.6999	0.0816	2.9427	144.71
	C5	−1.9817	4.0182	0.4201	4.4803	116.25
	C6	−1.0422	4.9615	0.5965	5.0698	101.86
	C7	−3.4429	4.3561	0.3927	5.5524	128.32
	H3	0.0475	1.4556	0.1715	1.4564	88.13
	H6	−1.3202	5.9149	0.7130	6.0605	102.58
	H17	−3.5704	5.4295	0.5332	6.4982	123.33
	H27	−3.8659	4.0629	−0.5681	5.6082	133.58
	H37	−3.9546	3.8218	1.1931	5.4995	135.98
Sodium ions	*S*1	−4.8688	2.0065	−2.0217	5.2661	157.60
	*S*2	−4.4114	7.8659	−0.1902	9.0185	119.28
	*S*3	−3.2771	11.2894	16.4784	11.7554	106.19
	*S*4	7.7192	−8.9681	−1.1396	11.8327	−49.28
	*S*5	1.5130	−9.5393	0.7863	9.6585	−80.99
	*S*6	−2.5838	−10.3577	1.7003	10.6751	−104.01
	*S*7	4.5794	−10.3908	9.1268	11.3552	−66.22
	*S*8	7.2149	−4.9422	7.0977	8.7453	−34.41
	*S*9	−5.9433	9.7118	8.9253	11.3860	121.47
	*S*10	−8.2233	−6.0484	8.3338	10.2081	−143.66

spine of water molecules is intimately bound to the DNA helix *via* a series of hydrogen bonds which connect not only adjacent base pairs, but also sugar rings in different strands (Fig. 4). The new interactions repeat regularly for every base pair with the same periodicity as the DNA and thus render unusual stability to the A·T helix.

Morphological features

One of the prominent differences between A- and B-type structures, apart from furanose puckering, is in the positioning of base pairs from the helix axis. This is equally a clear-cut discrimination between B- and B′-DNA. As the base-stacking diagrams in the two cases indicate (Fig. 5), the displacement of base pairs into the minor groove is promiscuous in all three allomorphs (*a* to *c*) of B′-DNA compared to the familiar astride the helix-axis disposition in (*d*) B-DNA, and it is the largest for the β_2-form (Fig. 5*b* and Table 4). Yet, all four panels display the same type of intrastrand overlaps, and no interstrand overlap at all. The forward locations of base pairs are apparently coupled to the reduced pitch, and are also reflected on the surface of the B′-DNA helix; as shown in Fig. 4, the minor groove collapses and the major groove proportionately widens (Table 4). Such peripheral features, unique only to oligo(dA)·oligo(dT) and related stretches, are not seen in any mixed sequence B-DNA.

DNA–DNA interactions and guest molecules

The distribution of all the guest molecules alone in one unit cell as viewed along the *c* axis is shown in Fig. 6. An interesting observation is that the regular string of water molecules, *W*1, *W*2 and their helix mates, appears like a circle at the unit-cell corners. At the rate of two per base pair, these water molecules embedded in the minor groove are responsible for enhancing the rigidity of the A·T helix. In contrast, the ten sodium ions (*S*1 to *S*10) and six water molecules (*W*3 to *W*8), which have

Table 2. *Observed and calculated structure amplitudes for the final model*

In each reflection box, the observed amplitude is given in the first line and the calculated amplitude (italics) in the second line. The curved and rectangular brackets refer to below threshold reflections included in, and rejected from, the least-squares refinement, respectively. M, N and S stand for meridional, systematically absent and saturated reflections, respectively.

h	*k*	*l*=0	1	2	3	4	5	6	7	8	9	10
0	0	M *[1747]*	N *0*	M *[35]*	N *0*	M *[163]*	N *0*	M *[19]*	N *0*	M *[54]*	N *0*	M *[707]*
0	1											
1	0	311	415	276	516	298	[112]	[118]	(126)	146		
-1	1	*348*	*418*	*326*	*483*	*224*	*[92]*	*[81]*	*(178)*	*130*		
		108	(83)	167	382	266	[101]	[108]	[116]	167	307	
1	1	*138*	*(113)*	*61*	*343*	*277*	*[51]*	*[70]*	*[97]*	*163*	*418*	
		(88)	[92]	[96]	347							S
0	2	*(146)*	*[58]*	*[53]*	*200*							*[891]*
		200	[98]	204	371	466	(142)	(149)	(157)			
-1	2	*253*	*[83]*	*53*	*201*	*449*	*(179)*	*(281)*	*(186)*			
2	0	282	(116)									
-2	1	*347*	*(122)*									
		250	[109]	243	345	538	249	298	356	483	[780]	
1	2	*252*	*[62]*	*181*	*217*	*503*	*326*	*320*	*371*	*568*	*[762]*	
2	1	(144)	[142]	[146]	288	[158]	[163]	381	305			
-2	2	*(165)*	*[74]*	*[121]*	*301*	*[124]*	*[104]*	*315*	*272*			
0	3	(138)	[137]	(141)	260	[150]	[160]	420	324	373		
-1	3	*(220)*	[57]	*(193)*	*244*	*[65]*	*[131]*	*393*	*310*	*482*		
2	2	(145)	[145]	(148)	(153)	(158)	[163]					
1	3	*(269)*	*[122]*	*(240)*	*(219)*	*(169)*	*[123]*					
		[144]	[146]	(148)	(152)	(157)	[167]	405	374			
-2	3	*[56]*	*[66]*	*(306)*	*(186)*	*(206)*	*[93]*	*542*	*413*			
-3	1	181	233	(164)	[167]	251	242	310	330	273		
3	0	*231*	*131*	*(264)*	*[119]*	*228*	*222*	*381*	*294*	*356*		
		262	273	434	[165]	302	254	223				
-3	2	*139*	*66*	*286*	*[81]*	*258*	*218*	*164*				
3	1											
-1	4	295	244	361	[182]	315	232	(208)	330	196		
0	4	*304*	*181*	*355*	*[153]*	*298*	*271*	*(221)*	*363*	*294*		
		[169]	[167]	[171]	(177)	273	(189)		[213]	[224]		
2	3	*[137]*	*[59]*	*[120]*	*(241)*	*176*	*(200)*		*[100]*	*[118]*		
-2	4											
-3	3											
1	4	394	(227)	[233]	[237]	[245]	(255)					
3	2	*411*	*(292)*	*[184]*	*[223]*	*[140]*	*(265)*					
-4	1											
4	0											
-1	5											
-3	4											
-4	2											
0	5	(273)	351	668	585		460					
2	4	*(294)*	*356*	*580*	*603*		*352*					
3	3											
4	1	[235]	[233]	491	(240)		[256]					
-2	5	*[113]*	*[129]*	*402*	*(266)*		*[242]*					
-4	3	[242]	[243]	444	[253]		[271]					
1	5	*[130]*	*[66]*	*248*	*[210]*		*[245]*					
4	2											
-3	5	[301]	[302]	308	[315]		[334]					
-4	4	*[238]*	*[125]*	*205*	*[112]*		*[283]*					
3	4											
-1	6	438	258	239	[279]		[298]					
2	5	*373*	*208*	*180*	*[227]*		*[91]*					
0	6	469	287	254	201		403					
-5	1	*267*	*174*	*119*	*161*		*199*					
-2	6											
-5	2											
5	0	564	232	196	[304]		422					
4	3	*431*	*340*	*163*	*[137]*		*234*					
		[263]	[260]	[267]			[290]					
1	6	*[67]*	*[155]*	*[134]*			*[192]*					

Table 3. *Major conformation angles, propeller and pseudorotation parameters in the β_1 form compared with those in the β_2 and α forms of poly(dA)·poly(dT)*

Estimated standard deviations are given in parentheses.

	β_1		β_2		α	
	A	T	A	T	A	T
α	−45 (4)	−41 (5)	−42	−43	−36	−40
β	128 (2)	136 (2)	135	147	127	138
γ	37 (5)	38 (3)	43	40	35	46
δ	139 (2)	141 (3)	136	143	137	133
ε	−119 (3)	−133 (4)	−135	−146	−127	−144
ζ	−170 (4)	−160 (5)	−156	−147	−166	−148
χ	−109 (4)	−115 (4)	−113	−116	−110	−110
π	−12 (1)	−12 (1)	−15	−15	−22	−22
ν_0	−48 (3)	−42 (3)	−49	−38	−45	−45
ν_1	56 (2)	52 (3)	56	49	53	50
ν_2	−43 (3)	−42 (3)	−42	−40	−41	−37
ν_3	19 (4)	19 (4)	16	20	17	13
ν_4	17 (3)	14 (3)	20	11	17	19
P	142	146	140	167	142	138
τ_m	59	54	58	51	52	52

been found in one half of the unit cell around the major groove, and their space-group symmetry (twofold screw along *c* axis) mates constitute 3.2 guest molecules per base pair; they appear to be lined up in a zigzag fashion roughly in the middle and parallel to the *a* direction. Along the *a* or *b* direction, distances between phosphate O atoms belonging to neighboring DNA helices are usually too large to bridge by a single water or sodium ion. Consequently, ions and water molecules join hands in an intricate fashion to close the gaps and render cohesive force around discrete contact regions.

The most conspicuous interactions among the DNA, whose crystallographic repeating unit consisting of five base pairs is denoted as A1A2A3A4A5·T6T7T8T9T10, and the guest molecules are listed in Table 5. Since the positions of the latter are rather vague in low-resolution electron-density maps, in order not to omit any weak interactions, distances up to 3.8 Å are included. It is reasonable to assume 3.2 Å to be the cut-off limit for potential hydrogen bonds and ionic contacts (Schneider, Cohen & Berman, 1992). Examination of Table 5 indicates that some attractive interactions are confined within the double helix, either within a strand or between strands. In addition, there are five examples when (phosphate groups of) two helices that are separated by *a* or *b* in the unit cell are tied together by spacers whose lengths range from a short one- to long four-unit of guest molecules.

As already noted, the periodic water molecules *W*1 and *W*2 (Fig. 3), appearing in the minor groove (Figs. 4*a*, 4*b*), are an integral part of the DNA helix. *W*1 is connected to atom N3 of adenine one step above, atom O2 of thymine almost at the same *z* level, and atom O4′ (furanose ring) one step below in the thymine strand. *W*2, hydrogen bonded to atom O4′ which is at a slightly higher *z* level in the adenine strand, is also hydrogen bonded to flanking *W*1 (below) and its helix mate (above). Together, a string of water molecules, two for every base pair, connects not only the bases but also the main chains in a tandem fashion in phase with the DNA symmetry.

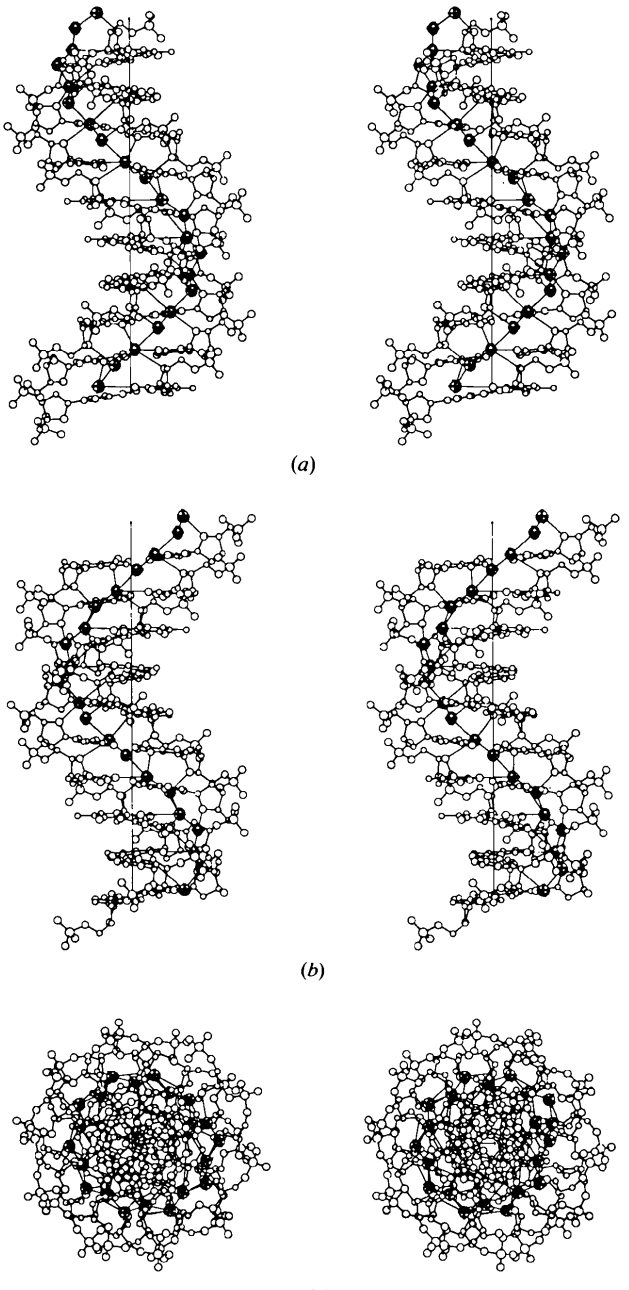

Fig. 4. Three mutually perpendicular stereoviews of one turn of the A·T helix and the periodic water molecules *W*1 and *W*2 (filled circles). The spine of water in the minor groove (*a*) as seen from the same side in contrast with (*b*) that seen from the major groove side. The links from water molecules to DNA atoms are drawn by thin lines. The vertical line is the helix axis. (*c*) The axial projection highlights the circular contour of the spine of water roughly in the middle.

There is no such spine of water molecules in the major groove. However, an isolated short bridge, through a single spacer, occurs in the major groove *via* sodium ion *S*1 connecting atoms N6(A4) and O4(T7) in the fourth base pair. Also, a medium bridge (*i.e.* two spacers) between adjacent phosphate groups occurs in the thymine strand; sodium ion *S*2 and water molecule *W*3 connect atoms O1P(T9) and O1P(T10).

Furthermore, a set of four ion-mediated interhelix interactions between phosphate groups belonging to adjacent helices in the *bc* plane take place in a region halfway across Fig. 7. Three of them are between the adenine strand of the helix on the left and thymine strand of the helix on the right. The first interaction, marked (1) in the first column, and A1–*S*3–T6 in the last column of Table 5, means that the phosphate of A1 is bridged by sodium ion *S*3 to T6. This short bridge is readily possible since the shortest distance between concerned O atoms is only 4.6 Å. The second and third interactions, A2–*S*4–*W*4–*S*3–T6 and A1–*S*5–*W*5–*S*6–T10, are similar in that both utilize medium three-unit sodium–water–sodium spacers to span the gap between the phosphate groups. The fourth is a four-unit sodium–water–water–sodium bridge in T9–*S*9–*W*7–*W*8–*S*10–T7 that connects neighboring thymine strands. This is a unique example of association between two parallel thymine strands. The fifth and final interaction, mediated by a sodium–water–sodium spacer, is A5–*S*7–*W*6–*S*8–T8 in which the link is along the *a* axis and not along *b* as in the rest. This is listed in Table

Table 4. *Morphological parameters in* β_1, β_2 *and* α *forms of poly(dA)·poly(dT) and in* B-DNA

Parameter	β_1	β_2	α	B-DNA
Tip (°)	−0.5	−2.2	7.0	1.7
Inclination (°)	−6.8	−5.6	−5.0	3.1
X displacement (Å)	0.15	−0.82	0.13	0.56
Y displacement (Å)	0.17	−0.08	−0.34	−0.08
Twist (°)	36.0	36.0	36.0	36.0
Rise (Å)	3.24	3.23	3.23	3.38
Slide (Å)	−1.98	−2.28	−1.94	−1.19
D_{xy} (Å)	3.51	3.66	3.44	3.51
	3.61	3.77	3.70	3.51
D_{xyz} (Å)	4.78	4.89	4.72	4.87
	4.85	4.97	4.91	4.87
Minor groove width (Å)	2.73	2.83	2.94	5.98
Major groove width (Å)	14.04	14.05	13.80	11.65

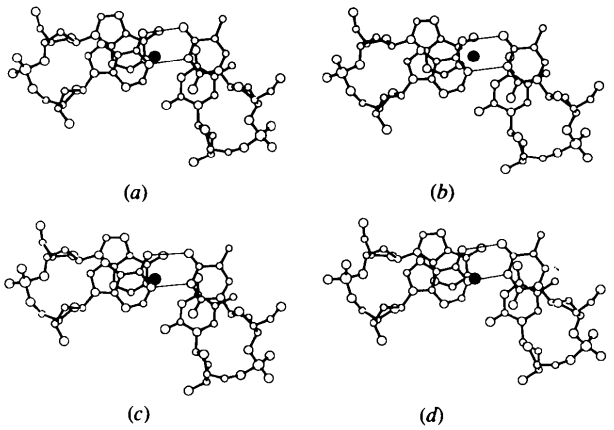

Fig. 5. Intrastrand base-stacking interactions in (*a*) β_1-form compared with those in (*b*) β_2-form and (*c*) α-form of poly(dA)·poly(dT), and contrasted with (*d*) those in B-DNA. The helix axis (•) is astride the base pair in B-DNA, but is in the major groove in others. Hydrogen bonds (thin lines) are shown for the top base pair in all panels.

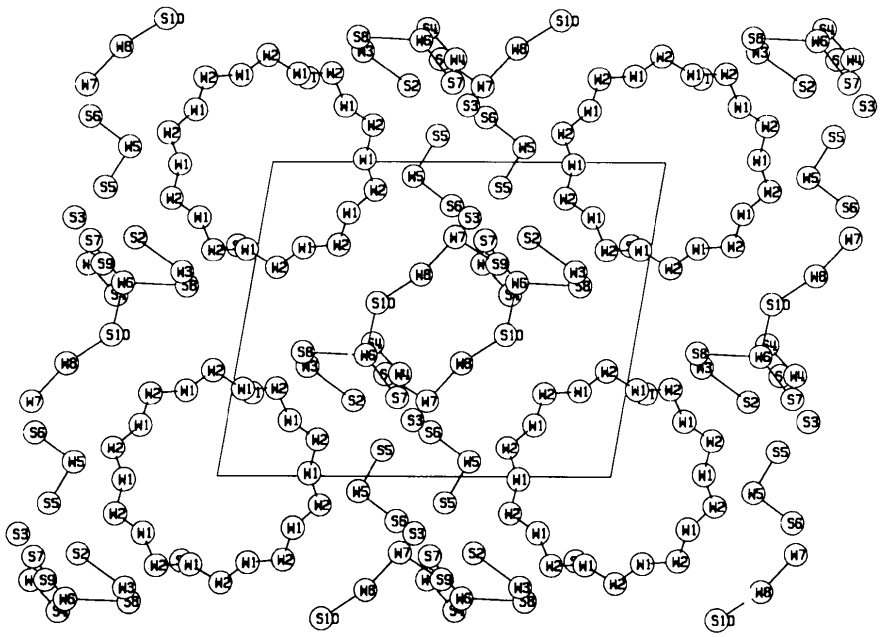

Fig. 6. A *c*-axis projection of the packing arrangement of sodium ions and water molecules exclusively around the unit cell.

Table 5. *Environment of sodium ions and water molecules in the poly(dA)·poly(dT) structure*

Nucleotide position for DNA atom is given in parentheses. *W*1(1) refers the helix mate of *W*1 in the next step. Also, triplet integers in parentheses define unit-cell translations.

Site	*X*	*Y*	*X*···*Y* (Å)	Precursor *P*	*P*—*X*···*Y* (°)	Interaction
Minor groove	N3(A2)	*W*1	3.35	C4	125	Hydration spine
	O2(T10)	*W*1	3.32	C2	170	
	O4′(T6)	*W*1	3.06	C1′	103	
	O4′(A3)	*W*2	3.07	C1′	113	
	*W*1	*W*2	2.79			
	*W*1(1)	*W*2	2.36			
Major groove	N6(A4)	*S*1	3.75	C6	166	
	O4(T7)	*S*1	3.25	C4	120	
Intra strand	O1P(T9)	*W*3	2.92	*P*	152	T9—*W*3—*S*2—T10
	*W*3	*S*2	3.80			
	O1P(T10)	*S*2	3.22	*P*	149	
Inter helix (1)	O1P(A1)	*S*3	3.27	*P*	120	A1—*S*3—T6
	O1P(T6, 010)	*S*3	3.10	*P*	87	
(2)	O2P(T6, 010)	*S*3	2.82	*P*	98	A2—*S*4—*W*4—*S*3—T6
	*W*4	*S*3	2.79			
	*W*4	*S*4	2.65			
	O1P(A2)	*S*4	3.23	*P*	86	
	O2P(A2)	*S*4	2.79	*P*	104	
(3)	O1P(A1)	*S*5	2.89	*P*	137	A1—*S*5—*W*5—*S*6—T10
	*S*5	*W*5	2.76			
	*W*5	*S*6	2.91			
	O1P(T10, 010)	*S*6	2.94	*P*	94	
	O2P(T10, 010)	*S*6	2.92	*P*	95	
(4)	O2P(T9)	*S*9	3.16	*P*	129	T9—*S*9—*W*7—*W*8—*S*10—T7
	*S*9	*W*7	3.28			
	*W*7	*W*8	3.18			
	*W*8	*S*10	3.18			
	O1P(T7, 010)	*S*10	3.10	*P*	154	
(5)	O2P(A5)	*S*7	3.01	*P*	118	A5—*S*7—*W*6—*S*8—T8
	*S*7	*W*6	3.15			
	*W*6	*S*8	3.80			
	O1P(T8, 100)	*S*8	2.98	*P*	150	

5 but not shown. In all the intermolecular bridges, the common feature is the presence of a sodium ion near an appropriate O atom of a phosphate group. The spacer between ions is one or two water molecules.

Discussion

It is well known that DNA morphology is often influenced by surrounding cations and water molecules. Ever since a highly structured water network was discovered in a dinucleoside–drug complex (Neidle, Berman & Shieh, 1980), crystallographic investigations on longer oligomers have uncovered the architecture of ordered water molecules in several examples belonging to the A (Kennard, Cruse, Nachman, Prange, Shakked & Rabinovich, 1986; Bingman, Li, Zon & Sundaralingam, 1992), B (Kopka, Fratini, Drew & Dickerson, 1983) and Z (Chevrier, Dock, Hartmann, Leng, Moras, Thuong & Westhof, 1986) forms of DNA. Heavy atoms such as cobalt have also been located in another example of Z-DNA (Gessner, Quigley, Wang, van der Marel, van Boom & Rich, 1985). The salient features, based on a systematic analysis of available data, have appeared in two elegant review articles (Schneider, Cohen & Berman, 1992; Berman, 1994). Once the spine of water molecules in the minor groove was proposed as a stabilizing force in the A·T region of DNA from a structure analysis of the AATT-embedded Caltech dodecamer (Kopka, Fratini, Drew & Dickerson, 1983), this has been reiterated to varying extent in other dodecamer crystal structures containing longer stretches of 3A's (Edwards, Brown, Spink, Skelly & Neidle, 1992), 5As (DiGabriele, Sanderson & Steitz, 1989) and 6A's (Nelson, Finch, Luisi & Klug, 1987) in one strand and corresponding T's in the complementary strand.

Investigations of water molecules in polymeric DNA have also been forthcoming from parallel fiber studies. For instance, crystalline water in fibers has been reported from neutron diffraction study for general sequence DNA in A (Langan, Forsyth, Mahendrasingam, Pigram, Mason & Fuller, 1992) and for poly[d(AT)] in the D conformation (Fuller, Forsyth, Mahendrasingam, Pigram, Greenall, Langan, Bellamy, Al-Hayalee & Mason, 1989). An earlier X-ray analysis of polycrystalline fiber of poly(C) showed a series of water molecules stabilizing three RNA single helices in a rhombohedral unit cell (Arnott, Chandrasekaran & Leslie, 1976). Patterson methods have also been used to explore the environment of poly(dA)·poly(dT) for ions and water molecules (Alexeev, Lipanov, Skuratovskii, 1992). Save for the first, atomic positions of guest molecules are not given in the other three publications.

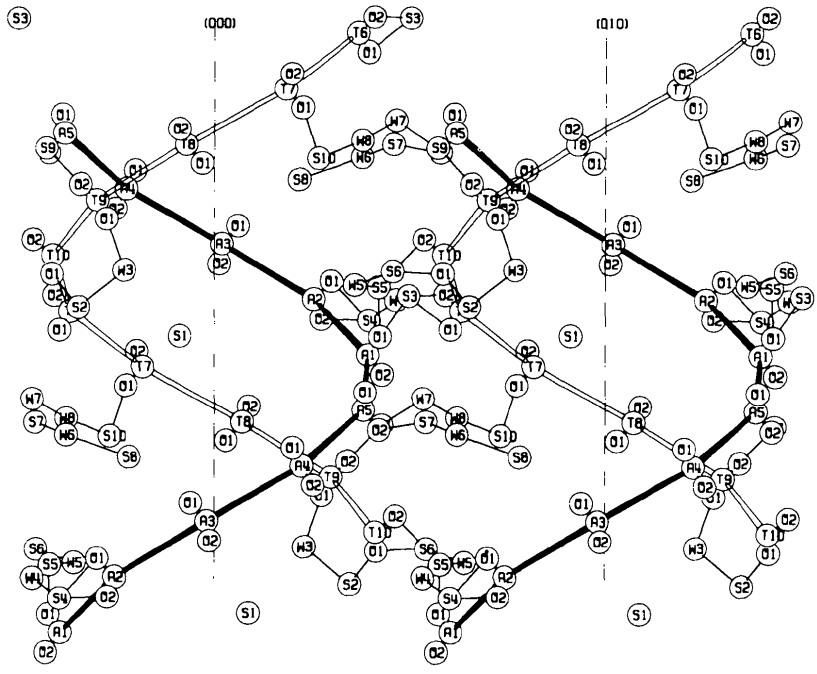

Fig. 7. Schematic representation of the interactions between a pair of A·T helices in the *bc* plane which are mediated and stabilized by sodium ions and water molecules around phosphate groups. The helix axes (vertical lines) are separated by 22.68 Å. The A and T strands are drawn by solid and open bonds, respectively, connecting successive phosphorus atoms labeled A1, A2 . . . and T6, T7 *etc.* Notice the clustering of ions and water molecules (in the middle) connecting phosphate groups of the A strand (left) with those of the T strand (right).

On the other hand, the results of the present work are the first to give structural details on sodium poly(dA)·poly(dT) and the associated guest molecules. The most important revelation is the spine of water in the minor groove. Of the two water molecules *W*1 and *W*2 for every base pair, *W*1 has three ligands which are the same as those reported in the dodecamer structure with 3A's (Edwards, Brown, Spink, Skelly & Neidle, 1992) except that atom O4′ belongs to the T strand in the polymer. What is not observed in the dodecamer, owing to the short AAA stretch, is the equivalent of *W*2 in the polymer which provides a continuous series of water–water hydrogen bonds in a helical fashion and thus extra stability to the helix.

As expected from conventional wisdom, barring *S*1 which is in the major groove, the sodium ions (Table 2) are located outside the helix near phosphate groups. This observation is in agreement with the findings of a preliminary EXAF study on calcium poly(dA)·poly(dT) (Skuratovskii, Hasnain, Alexeev, Diakun & Volkova, 1987). These ions are able to establish contacts between phosphate groups belonging to different helices with the help of neighboring water molecules. As in oligomer structures, this is accomplished by short as well as long bridges, consisting of a single sodium ion in one, as many as two ions and two water molecules in another, and intermediate lengths in three other cases. They add up to a total of five interhelical bridges for every five base pairs and are deemed critical for the stability of the packing of DNA helices. It is not surprising to see several large distances (Table 5) which would normally be interpreted as weak links of the structure. Tight binding might be detrimental, but such loosely bound ions and water molecules are crucial for, and consistent with, transitions among the three allomorphs, α, β_1 and β_2, of poly(dA)·poly(dT) which could be induced in intact fibers by changes in hydration and ionic strength (Arnott & Selsing, 1974; this study).

In single crystals of nucleic acids, about 25% of water is ordered (Schneider, Cohen & Berman, 1992). Likewise, our results show that over 20% of water and sodium ions in the fiber are well organized and their structural roles are lucidly defined. The remaining material might belong to the third and higher shells or form a disordered continuous medium surrounding the DNA and the discrete guest molecules. Description of their structure is beyond the means of X-ray diffraction.*

We dedicate this structure determination to Struther Arnott, FRS, who taught us fiber diffraction as an art, on the year of his sixtieth birthday. This research was supported in part by the National Institutes of Health.

* Atomic coordinates have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1PLY). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0412).

References

ALEXEEV, D. G., LIPANOV, A. A. & SKURATOVSKII, I. YA. (1987). *J. Biomol. Struct. Dynam.* **4**, 989–1012.

ALEXEEV, D. G., LIPANOV, A. A. & SKURATOVSKII, I. YA. (1992). *Int. J. Biol. Macromol.* **14**, 139–144.

ARNOTT, S., CHANDRASEKARAN, R. & LESLIE, A. G. W. (1976). *J. Mol. Biol.* **106**, 735–748.
ARNOTT, S. & SELSING, E. (1974). *J. Mol. Biol.* **88**, 509–521.
BERMAN, H. M. (1994). *Curr. Op. Struct. Biol.* **4**, 345–350.
BINGMAN, C., LI, X., ZON, G. & SUNDARALINGAM, M. (1992). *Biochemistry*, **31**, 12803–12812.
CHANDRASEKARAN, R. & ARNOTT, S. (1989). *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology*, Vol. VII/1b, edited by W. SAENGER, pp. 31–170. Berlin: Springer Verlag.
CHANDRASEKARAN, R. & RADHA, A. (1992). *J. Biomol. Struct. Dynam.* **10**, 153–168.
CHANDRASEKARAN, R., RADHA, A., PARK, H.-S. & ARNOTT, S. (1989). *J. Biomol. Struct. Dynam.* **6**, 1203–1215.
CHANDRASEKARAN, R., RADHA, A. & RATLIFF, R. L. (1994). *J. Biomol. Struct. Dynam.* **11**, 741–766.
CHEVRIER, B., DOCK, A. C., HARTMANN, B., LENG, M., MORAS, D., THUONG, M. T. & WESTHOF, E. (1986). *J. Mol. Biol.* **188**, 707–719.
DICKERSON, R. E. (1989). *J. Biomol. Struct. Dynam.* **6**, 627–634.
DIGABRIELE, A. D., SANDERSON, M. R. & STEITZ, T. A. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 1816–1820.
EDWARDS, K. J., BROWN, D. G., SPINK, N., SKELLY, J. V. & NEIDLE, S. (1992). *J. Mol. Biol.* **226**, 1161–1173.
FULLER, W., HUTCHINSON, F., SPENCER, M. & WILKINS, M. H. F. (1967). *J. Mol. Biol.* **27**, 507–524.
FULLER, W., FORSYTH, V. T., MAHENDRASINGAM, A., PIGRAM, W. J., GREENALL, R. J., LANGAN, P., BELLAMY, K., AL-HAYALEE, Y. & MASON, S. A. (1989). *Physica B*, **156/157**, 468–470.
GESSNER, R. V., QUIGLEY, G. J. WANG, A.-H. J., VAN DER MAREL, G. A., VAN BOOM, J. H. & RICH, A. (1985). *Biochemistry*, **24**, 237–240.
HAMILTON, W. C. (1965). *Acta Cryst.* **18**, 502–510.
International Tables for X-ray Crystallography (1974). Vol IV, pp. 99–101. Birmingham: Kynoch Press. (Present distributor Kluwer Academic Publishers.)
KENNARD, O., CRUSE, W. B. T., NACHMAN, J., PRANGE, T., SHAKKED, Z. & RABINOVICH, D. (1986). *J. Biomol. Struct. Dynam.* **3**, 623–647.
KOO, H.-S., WU, H.-M. & CROTHERS, D. M. (1986). *Nature (London)*, **325**, 821–823.
KOPKA, M. L., FRATINI, A. V., DREW, H. R. & DICKERSON, R. E. (1983). *J. Mol. Biol.* **163**, 129–146.
LANGAN, P., FORSYTH, V. T., MAHENDRASINGAM, A., PIGRAM, W. J., MASON, S. A. & FULLER, W. (1992). *J. Biomol. Struct. Dynam.* **10**, 489–503.
LESLIE, A. G. W., ARNOTT, S., CHANDRASEKARAN, R. & RATLIFF, R. L. (1980). *J. Mol. Biol.* **143**, 49–72.
NEIDLE, S., BERMAN, H. M. & SHIEH, H.-S. (1980). *Nature (London)*, **288**, 129–133.
NELSON, H. C., FINCH, J. T., LUISI, B. F. & KLUG, A. (1987). *Nature (London)*, **330**, 221–226.
PARK, H.-S., ARNOTT, S., CHANDRASEKARAN, R., MILLANE, R.P. & CAMPAGNARI, F. (1987). *J. Mol. Biol.* **197**, 513–523.
SCHNEIDER, B., COHEN, D. & BERMAN, H. M. (1992). *Biopolymers*, **32**, 725–750.
SKURATOVSKII, I. YA., HASNAIN, S. S., ALEXEEV, D. G., DIAKUN, G. P. & VOLKOVA, L. I. (1987). *J. Inorg. Biochem.* **29**, 249–257.
SMITH, P. J. C. & ARNOTT, S. (1978). *Acta Cryst.* A**34**, 3–11.