Crystal Structure of the B-DNA Hexamer d(CTCGAG): Model for an A-to-B Transition

Markus C. Wahl,* Sambhorao T. Rao,‡ and Muttaiya Sundaralingam*‡
Laboratory of Biological Macromolecular Structure, Departments of ‡Chemistry and *Biochemistry, The Ohio State University, Columbus, Ohio 43210-1002 USA

ABSTRACT The crystal structure of the B-DNA hexamer d(CTCGAG) has been solved at 1.9 Å resolution by iterative single isomorphous replacement, using the brominated derivative d(CG^{5Br}CGAG), and refined to an *R*-factor of 18.6% for 120 nonhydrogen nucleic acid atoms and 32 water molecules. Although the central four base pairs form a typical B-form helix, several parameters suggest a transition to an A-like conformation at the termini. Based on this observation, a B-to-A transition was modeled, maintaining efficient base stacking across the junction. The wide minor groove (~6.9 Å) is reminiscent of that in the side-by-side double drug·DNA complexes and hosts a double spine of hydration. The global helix axes of the pseudo-continuous helices are at an acute angle of 60°. The pseudocontinuous stacking is reinforced by the minor groove water structure extending between the two duplexes. The crossover point of two pairs of stacked duplexes is at the stacking junction, unlike that observed in the B-DNA decamers and dodecamers. This arrangement may have implications for the structure of a four-way DNA junction. The duplexes are arranged around a large (~20 Å diameter) channel centered on a 6_2 screw axis.

INTRODUCTION

B-DNA single crystals are obtained primarily from two classes of molecules: orthorhombic dodecamers and monoclinic, trigonal, hexagonal, or orthorhombic decamers of the general sequences d(CCNNNNNNGG) and d(CGANNNNTCG) (Joshua-Tor and Sussman, 1993). Whereas some members of the decamer class yield crystals that diffract to the best resolutions observed for right-handed DNA structures, the dodecamers diffract notoriously to only low resolution, usually around 2.5 Å. Although much insight, pertaining to sequence and packing effects, hydration and metal ion coordination, as well as drug binding in the minor groove, was gained from the analysis of the three-dimensional structures of these molecules and their complexes, our picture of DNA helical fine structure is still incomplete.

In our continuous effort to compare A-DNA and A-RNA molecules of the same sequence, the DNA hexamer d(CTCGAG) was investigated by x-ray crystallography. The sequence was chosen because the double-helical stems of some tRNAs contain the sequence element CUC·GAG. Unexpectedly, the hexamer crystallized as a right-handed double helix with overall B-form conformation. Z-form (Wang et al., 1979) and A-form (Mooers et al., 1995) hexamer DNAs and a modified B-DNA hexamer with phosphorothioate groups in the backbone (Cruse et al., 1986) have been determined crystallographically. Very recently, an unmodified non-self-complementary B-DNA hexamer with an unusual A·T base pair has been described at relatively low resolution (Tari and Secco, 1995). We report here the crystal structure of an unmodified self-complementary B-DNA hexamer, d(CTCGAG), at 1.9 Å resolution. The duplex displays a wide minor groove that is occupied by a double spine of hydration, A-type structural features at the termini, and novel crystal packing interactions.

Received for publication 20 December 1995 and in final form 29 February 1996.

Address reprint requests to Dr. Muttaiya Sundaralingam, Laboratory of Biological Macromolecular Structure, Departments of Chemistry and Biochemistry, The Ohio State University, 1060 Carmack Road, Columbus, OH 43210-1002. Tel.: 614-292-2925/2999; Fax: 614-292-2524; E-mail: sunda@biot.mps.ohio-state.edu.

0006-3495/96/06/2857/10 $2.00

MATERIALS AND METHODS

Oligonucleotide synthesis and crystallization

The chemical syntheses of d(CTCGAG) and d(CT^{5Br}CGAG) were performed on an Applied Biosystems model 381 nucleic acid synthesizer (Foster City, CA) employing the phosphoramidite chemistry and a standard 64-step coupling cycle for 1-μmol samples. After deprotection and purification by ion-exchange fast-performance liquid chromatography and ethanol precipitations, the lyophilized sample was dissolved in water to a concentration of 2 mM (duplex). Annealing was not necessary. Crystals were grown by hanging drop vapor diffusion. The unmodified hexamer crystallized after 4 weeks at 4°C as chunky hexagonal rods from a solution containing 1 mM oligonucleotide duplex, 5 mM $MgCl_2$, 1 mM spermine·$(HCl)_4$, 30 mM sodium cacodylate buffer (pH 7.0), and 5% 2-methyl-2,4-pentanediol, equilibrated against 40% 2-methyl-2,4-pentanediol. The crystallization was very dependent upon low concentrations of spermine; higher levels of spermine yielded thin needles. The brominated hexamer crystallized only after several months in the 4°C cold room without the presence of spermine and with the $MgCl_2$ concentration raised to 100 mM.

Data collection, structure solution, and refinement

Diffraction data were collected on a Siemens four-circle goniostat equipped with a multiwire proportional counter and a MacScience rotating Cu anode operated at 40 kV/100 mA, generating graphite-monochromated CuKα radiation (λ = 1.5418 Å). The crystal-to-detector distance was 12.0

cm. Two ϕ scans (150°) with $\omega/2\theta$ settings of 340°/340° and 320°/320°, respectively, and a single 60° ω scan at 2θ of 320° yielded data up to 1.9 Å resolution for the unmodified sequence and up to 2.0 Å for the brominated hexamer. The data, collected with a 0.25° frame width, could be integrated in space groups $P6_222$ and $P6_422$ with R_{sym} values of 4.99% (unmodified) and 3.80% (brominated), and unit cell dimensions of $a = b = 40.14$ Å, $c = 44.57$ Å (unmodified), and $a = b = 40.32$ Å, $c = 45.23$ Å (brominated). There is one hexamer single strand in an asymmetric unit.

The derivative data were scaled to the native data between the 8.0 and 2.0 Å shells and above the 3s intensity level. The space group $P6_222$ was initially chosen and later proved to be correct. An automated interpretation of the isomorphous difference Patterson map, using locally developed routines, indicated a single peak consistent with the features in the Patterson map, which was chosen as the most likely position of the bromine atom. The native and derivative structure factor amplitudes and the assumed bromine position were used in the program package PHASES (W. Furey, University of Pittsburgh) to calculate SIR maps. After three cycles of iterative single isomorphous replacement (ISIR) (Wang, 1985), assuming a solvent content of 35%, a readily interpretable ISIR map was obtained (Fig. 1). Projections of the map onto the *a/b* plane clearly showed features of planes of base pairs. A B-DNA model was superimposed on the ISIR Fourier map with the graphics package CHAIN (Sack and Quiocho, 1992) (Fig. 1). The direction of the helix axis of the model was consistent with the meridional reflection observed at ~3.4 Å along the ⟨110⟩ direction.

An unmodified fiber model positioned in the ISIR density was refined against the corresponding structure factor amplitudes in X-PLOR (Brünger, 1992). Positional and B-factor refinement and a single round of simulated annealing (Brünger, 1988) produced an *R*-factor of 23.4% for data up to 1.9 Å above the 2σ level (1547 reflections, 83%). Thirty-two spherical peaks above 3σ in the F_o-F_c difference map with distances of <3.4 Å to polar nucleic acid atoms were interpreted as water molecules and manually positioned with CHAIN. After further positional refinement, water positions were confirmed with F_o-F_c difference omit maps, omitting 10 water molecules at a time. The model was also checked against the ISIR density and manually adjusted at several positions. After positional and B-factor refinement the *R*-factor converged at 18.6%. The sequence numbering scheme is shown in Fig. 2. Statistics are summarized in Table 1, and the coordinates will be deposited in the Nucleic Acid Data Bank (Berman et al., 1994) (NDB ID: BDF068). Residues of the duplex are numbered 1 through 6 in the first strand and 7 through 12 when referring to the Watson-Crick paired, twofold related strand.

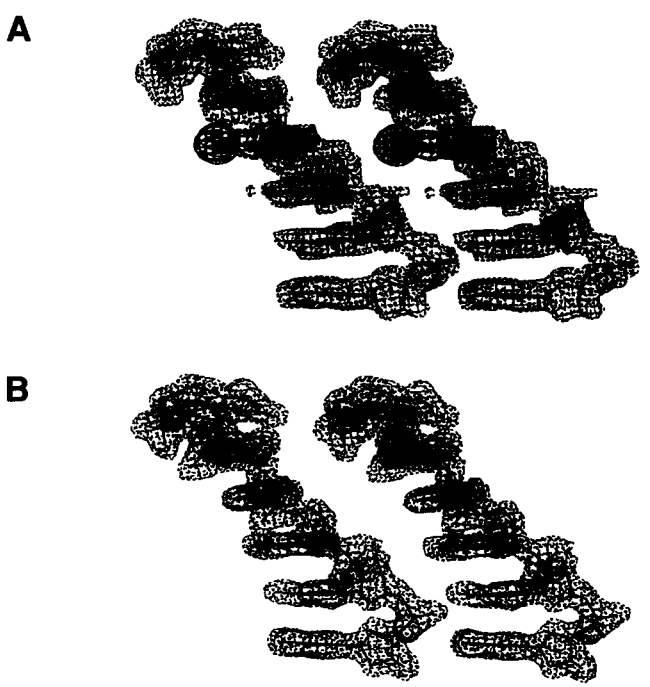

FIGURE 1 (*A*) ISIR density after three rounds of single isomorphous replacement phasing and solvent flattening. The spherical density for the bromine atom at residue 3 is seen in the difference map. A fiber B-DNA model of the hexamer fits well into the density. (*B*) Final $2F_o$-F_c Fourier map.

RESULTS

Molecular structure: transition to the A-form at the termini

The hexamer d(CTCGAG) crystallized as a right-handed double helix with overall B-form conformation. The crystallographic twofold axis passing through the center of the duplex renders the two molecular halves structurally equivalent. Although the global structure can be described as B-DNA, the hexamer exhibits interesting local structural variations when examined in more detail. Whereas the central four nucleotides adopt standard B-DNA geometry, the conformations of the end base pairs suggest a transition to the A-form. Table 2 shows helical and backbone parameters for the duplex and for B- and A-DNA fiber models. The parameters were calculated with the program NUPARM (Bhattacharyya and Bansal, 1990) based on a local helix axis. The structural differences between A-form and B-form helices are generally represented most clearly in the twist angle (A: 33°, B: 36°), the rise (A: 2.7 Å, B: 3.4 Å), the inclination (A: ~20°, B: ~0°), the *x*-displacement (A: ~-4 Å, B: ~0 Å), the glycosidic torsion angle (A: ~-160°, B: ~-100°), and, most decisively, the sugar pucker (A: C3′-*endo*, B: C2′-*endo*) (values in parentheses giving fiber model values). The C2′-*endo* puckering in B-form helices leads to a longer 5′-phosphate-to-3′-phosphate distance (~7 Å), whereas the C3′-*endo* puckering yields a shorter separation (~5.8 Å). In the present duplex the rise and twist angles are B-like throughout, but for both the 5′- and 3′-terminal residues the inclination, the *x*-displacement, the glycosidic torsion angles, and the sugar puckering adopt A-like values (Fig. 3). A shortening of the P-P distance concomitant with the N-type sugar puckering cannot be observed because of the missing phosphate groups at the

TABLE 1 Data collection and refinement statistics

	d(CTCGAG)	d(CT^{5Br}CGAG)
Resolution	8.0–1.9 Å	8.0–2.0 Å
R_{sym}	4.99%	3.80%
No. of reflections ($F > 2\sigma$)	1547	1143
Data completeness	83%	71%
Space group	$P6_222$	$P6_222$
Cell constants	$a = b = 40.14$ Å; $c = 44.57$ Å	$a = b = 40.32$ Å; $c = 45.23$ Å
Final *R*-factor	18.6%	
Final model	120 nucleic acid atoms 32 water molecules	
Deviations from ideal geometry	0.02 Å (bonds) 2.9° (angles)	

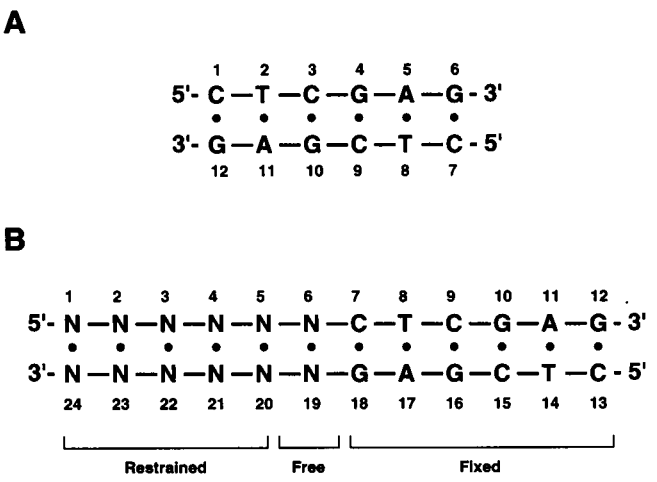

FIGURE 2 Numbering of the hexamer crystal structure (*A*) and the dodecamer models of the B-to-A transition (*B*). The treatment of the three molecular parts during the energy minimization is indicated. The A-form model distal to the junction is restrained, the central portion is without constraints, and the observed crystal structure is fixed.

ends. However, the consecutive O5′-O5′ distances also display a dependence on the sugar pucker, with fiber A-DNA showing lower values (~5.6 Å) than fiber B-DNA (~5.9 Å). Clearly the terminal steps in the present hexamer exhibit shortened O5′-O5′ distances (5.8 and 6.0 Å) compared to the central steps (average 6.3 Å), although the distances are longer than for the fiber models in both the terminal and the central parts (Table 2). The roll angle and the propeller twist also switch toward larger values at the termini. The larger propeller twist is achieved by only tipping the terminal cytosines. The guanines remain coplanar with the preceding base pair to maintain efficient base stacking.

Model for an A-B junction

The unusual conformation of all four terminal nucleotides of the duplex motivated us to attempt the modeling of an A-B junction. Four heptamer sequences with average fiber A-form geometry terminating in a C·G base pair were built, each having one of the four possible base pairs, A·T, T·A, G·C, and C·G, following the terminal C·G. The terminal C·G's of the model heptamers were superimposed on that of the hexamer duplex, leaving 12-base pair model structures (Fig. 2). The resulting structures were subjected to energy minimization in X-PLOR using the PARAM11.DNA parameter file. During the energy minimization all atoms of the hexamer structure (residues 7–18, with the exception of O5′ of d(C)1; Fig. 2) were kept in fixed positions while the remainder of the structure was allowed to vary. Only the first base pair of the A-form model (residues 6 and 19; Fig. 2), the one bordering the junction, was allowed to freely adjust its geometry. To maintain the fiber A-form geometry for residues 1–5 and 20–24 (Fig. 2), distances between three pairs of atoms in adjacent base pairs and the Watson-Crick hydrogen bonding distances were restrained.

All resulting models were quite similar, independent of which base step was chosen across the junction (Fig. 4). In the A-form part of the structure all sugars remained in the C3′-*endo* puckering mode during the energy minimization. The helical axes of the A-form and B-form parts, determined with CURVES (Lavery and Sklenar, 1988), are at an (acute) angle of ~26° with respect to each other (Fig. 4, *A*, *B*). The bending seems to be mainly due to the large roll angle adopted by the junction base pair step and the following step in the A-form half. These two base steps also exhibit a reduced twist angle (~30°). Base stacking across the junction is efficiently maintained, but the stacking geometry is typical of neither the A-form nor the B-form (Fig. 5). The junction is basically made up of two base pairs, the terminal C·G of the hexamer and the first base pair (residues 6 and 19) of the fiber model, with residues outside the junction in standard A- or B-form conformation. The conversion from an A- to a B-form helical environment (or vice versa) can thus be accomplished over a very short stretch of sequence.

In an analogous fashion, B-DNA fiber models were added to the termini of the current structure to assess the effect of the N-type sugar pucker on bending. Bends of ~14° appeared at the junction (Fig. 4, *C*, *D*). The structure changed only insignificantly when energy minimized, mainly adopting some degree of propeller twisting.

Hydration: a double spine in the minor groove

A total of 32 water molecules were found in the structure, 26 in the first hydration shell and six in the second shell. Nine water molecules bind only to backbone atoms, whereas ten and seven fall into the major and minor grooves, respectively. The hydration of backbone and major groove atoms is presented in Table 3. Because of the long P-P distance (~7 Å) associated with the C2′-*endo* sugar pucker, no single water is found to bridge consecutive anionic phosphate oxygens, as is common in A-DNA. However, bridges of two water molecules are found (Table 3). All anionic phosphate oxygens except the O_2P atoms of residues G4 and A5 are hydrated. The water molecule that hydrates O1P of G4 has no other conventional hydrogen bonding partners. It has some short contacts, which may constitute CH···O hydrogen bonds, with (C8)-H (C-O distance: 3.22 Å) and possibly (C2′)-H2 (C-O distance: 3.34 Å) of G4.

Water molecules in the major groove bind to the exocyclic O4, N4, and O6 atoms of thymine, cytosine, and guanine, respectively, and all purines are hydrated at N7 (Table 3). The C7·G10 base pair carries a short string of three waters in the major groove, with the waters contacting the base hetero atoms, N4, O6, and N7, whereas in other pairs this string is interrupted. The waters that hydrate the guanine N7 and O6 atoms are in both cases within hydrogen bonding distance of each other.

TABLE 2 Helical and backbone parameters

Residue	Twist (°)	Roll (°)	Rise (Å)	Inclination (°)	x-Displacement (Å)	Propeller twist (°)	χ angle (°)	O5′-O5′ distance (Å)	Pseudorotation phase angle (*P*, °) and sugar pucker
C					−1.2	−21	−168		20.2 C3′-*endo*
	37	6	3.4	10				5.8	
T					0.5	−6	−106		162.0 C2′-*endo*
	35	1	3.4	2				6.1	
C					0.7	−7	−84		154.4 C2′-*endo*
	38	5	3.3	8				6.4	
G					0.5	−8	−93		155.4 C2′-*endo*
	35	1	3.4	2				6.3	
A					0.5	−5	−110		141.5 C1′-*exo*
	37	6	3.4	10				6.0	
G					−1.2	−21	−130		70.5 C4′-*exo*
B-DNA	36		3.4	0	0		−100	5.9	C2′-*endo*
A-DNA	33		2.6	20	−4.4		−160	5.6	C3′-*endo*

The minor groove of the present hexamer is relatively wide (~6.9 Å). As a consequence there is space to accept a double spine of water molecules (Fig. 6). A network of three- and four-membered water rings is seen in the hydrogen bonding. The four-membered rings are composed entirely of water molecules, whereas the three-membered rings involve two waters and an O4′ sugar atom. The central eight water molecules, constituting an uninterrupted double spine (Fig. 6), are related by the twofold axis passing through the center of the duplex. This water "sheet" exhibits a twist which matches that of the minor groove. The water molecules are hydrogen bonded to the sugar O4′ atoms, endocyclic N3 atoms of the purines, and exocyclic N2 (guanine) and O_2 (cytosine and thymine) atoms, thereby becoming deeply buried in the groove (Fig. 6, edge-on views). The groove hydration and width suggest that the duplex could possibly accommodate ligands mimicking the water structure, e.g., a side-by-side arrangement of a groove-binding drug. This possibility is currently being investigated. Interestingly, the longitudinal connections of the double spine (along the direction of the minor groove) are disrupted (distance ~4 Å) at the helix termini, coincident with the switch to the A-form conformation. These water molecules could still be interacting by coulombic attraction.

A different network of water molecules, with stretches of a single spine and pentagonal water rings involving the O3′ atoms from two symmetry related terminal guanosines, extends between the minor grooves of two head-to-tail stacked duplexes (Fig. 6), i.e., over the region where the structure is more A-like. The terminal O3′ atoms serve as latching points for the two waters that fall between the terminal base pair planes, and participate in the formation of the pentagonal rings. Other anchoring points for these waters again involve O4′ sugar atoms and the N3 atoms of the terminal guanines, maintaining close contacts with the walls and floor of the grooves.

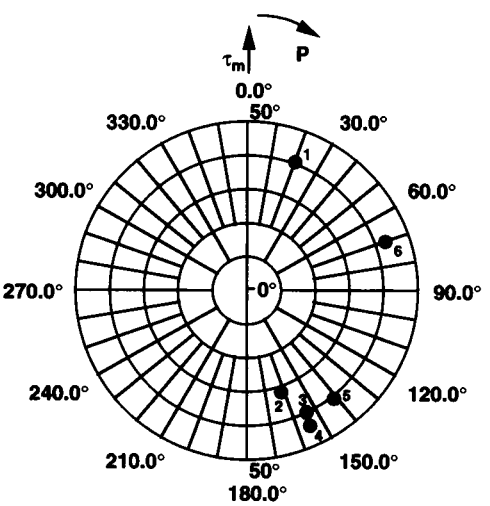

FIGURE 3 Pseudorotation cycle representing the sugar puckering amplitude (radius, τ_m) and phase (azimuth, *P*) for the six crystallographically independent nucleotides (●). The terminal residues 1 and 6 adopt an N-type (C3′-*endo*) pucker, whereas the central four nucleotides have an S-type (C2′-*endo*) sugar pucker.

Crystal packing: model for a four-way junction

In the crystal lattice the molecules stack pseudo-continuously end to end, forming a 60° (acute) angle between successive global helix axes (Fig. 7 *A*). The pseudo-infinite helices are therefore zigzagging through the crystal rather than forming straight lines. The rise and tilt across the gap are 2.7 Å and 0.4°, respectively. The helical axes are offset to allow partial stacking of the cytosine rings on top of each other and of the cytosine O4′ atoms on top of the six-membered guanine rings (Fig. 7 *B*). A large channel (~20 Å diameter) is formed around the 6_2 screw axis (Fig. 7 *C*). Such channels are usually seen with A-DNA structures but not in B-DNA crystals. Stacked rows of helices cross in a fashion that positions the gaps between the helices at the stacking junction (Fig. 8). Because the termini of adjacent

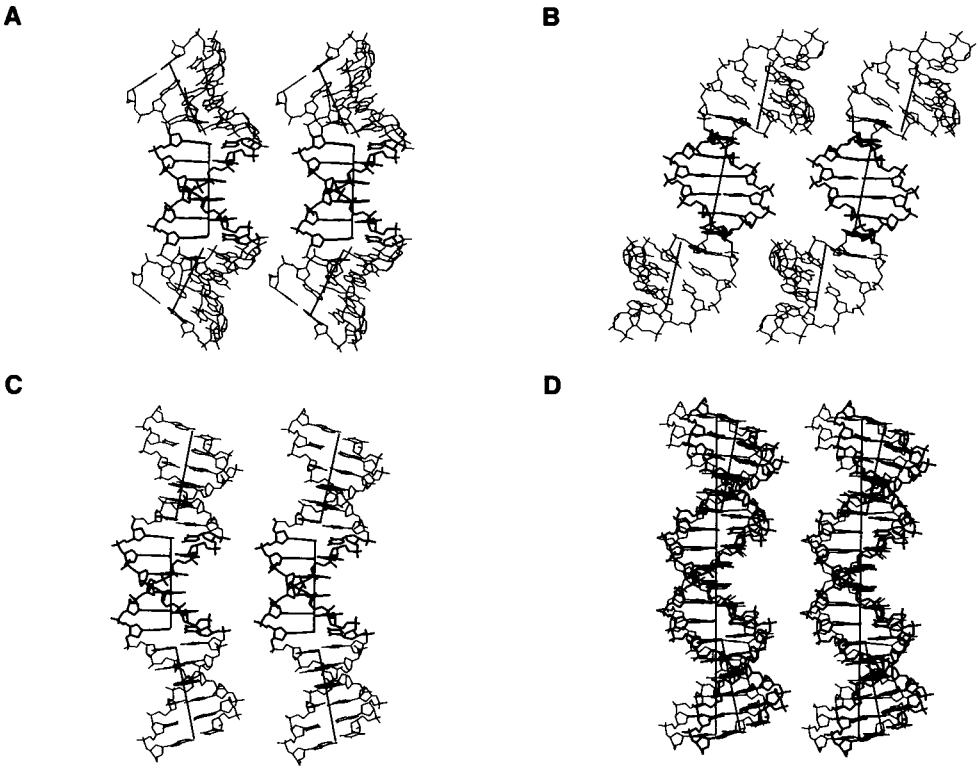

FIGURE 4 (*A*, *B*) Two stereo views showing the B-to-A transition using the hexamer crystal structure (central portion; B-form in thick lines and A-form in thin lines). For brevity, only hexamer A-form models have been used. However, they could be of any length. The upper two-thirds of the structure were energy minimized, and the lower portion was added by applying the crystallographic twofold symmetry. The best linear global helix axes of the three parts are shown. The choice of the base step (ApC, CpC, GpC, TpC) at the junction does not influence the energy-minimized structures (not shown). (*C*, *D*) Two stereoviews with fiber B-DNA models at the termini. In *C* the thin lines represent fiber B-models at either end. Helix axes of the parts are shown, emphasizing the bends. (*D*) Superposition of the entire model in *C* (shown now in *thick lines*) on a fiber B-DNA model of the same sequence (*thin lines*). The superposition was achieved through minimization of the RMS differences between atomic positions of the central six base pairs of the crystal structure. The bending of the model structure is clearly visible.

helices in the pseudo-continuous columns are efficiently stacked, no indication of a strand exchange is observed, as is known to occur in Holliday junctions. Nevertheless, it is tempting to consider the arrangement of four duplexes as a model for a four-way junction. Fig. 8 depicts four duplexes, indicating the possible sites for strand exchange, and mainly serves to illustrate the closeness of the backbone termini. The organization of duplexes is different from the packing observed in B-DNA decamers and dodecamers, where a reciprocal fit of the backbone of one molecule into the groove of a neighboring molecule is observed. This crossed geometry has been discussed as a model for the Holliday junction (e.g., Timsit and Moras, 1991; Goodsell et al., 1995). In the present example, the backbones of the crossed helices are abutting, resembling possibly an early stage of helix association on the pathway to the Holliday junction. Whereas the Holliday junction probably exhibits the backbone-into-groove type of packing, the more open, crossed geometry may be an intermediate during the formation of the junction. There are no metal ions discernible in the present structure, whereas the more intimate contacts in the longer duplexes are usually mediated by Mg^{2+} or Ca^{2+} ions. Interestingly, in a recent report on the nuclear magnetic resonance structure of the RuvC-Holliday junction complex, an opened junction has been observed, which may be modeled more closely by the current crystal packing.

DISCUSSION

A-B junctions and bent B-DNA

Among the most interesting features seen in the crystal structure of the hexamer d(CTCGAG) are the B-to-A transition at the termini, the double spine of hydration in the wide minor groove, and the novel packing interactions. Whereas DNA in solution preferentially adopts the B-form, RNA and chimerical helices are usually found in the A-form (Saenger, 1984). A distortion of the DNA double helix into an A-like conformation may be observed, e.g., upon binding of certain proteins (L. J. Kim et al., 1993; Y. Kim et al., 1993). Because of the interaction of DNA with ligands and because of its joint occurrence with RNA during replication, transcription, and reverse transcription, transitions between the two forms of right-handed helices (A-B junctions) can be expected in the genome. Considering the decisive differences between A- and B-form helices, these junctions may

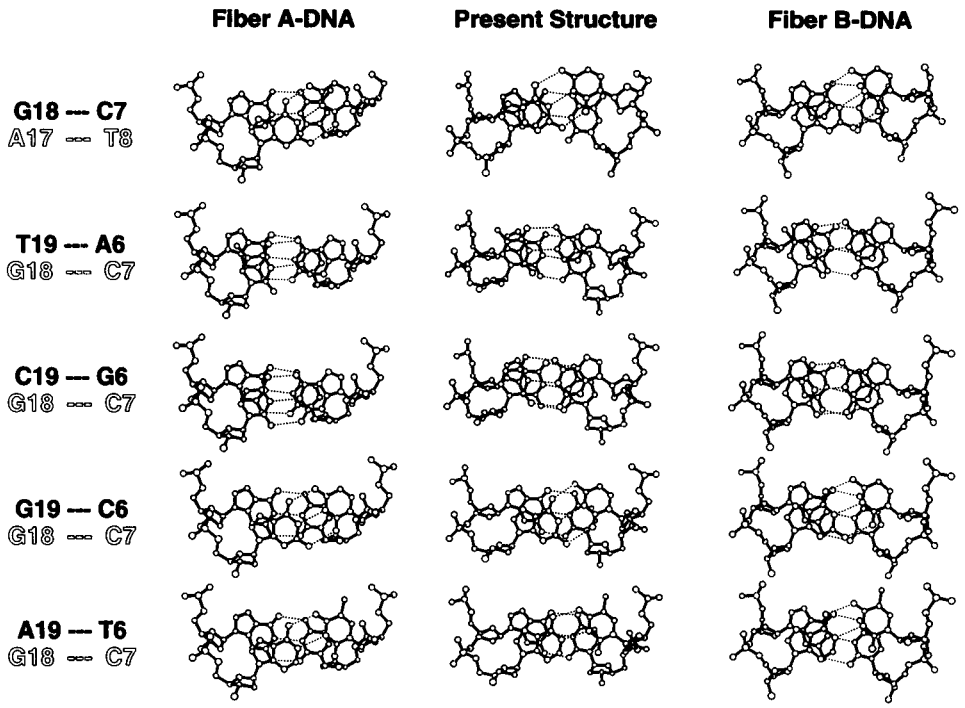

FIGURE 5 Base pair stacking of residues 7–8/17–18 (*first row of structures*; numbering as in Fig. 2) and residues 6–7/18–19 (*second through fifth row of structures*), comprising the A-B junction (*central column of structures*). The energy-minimized stacking geometry for all four possible base steps (ApC, CpC, GpC, TpC) that were tried for the A-form models is shown in structural rows two through five. The central column of structures shows the stacking observed in the crystal structure (*first row*) and the A-form model parts (*rows 2–5*). For comparison, the first and third columns of structures show the stacking of these base steps as observed for fiber A-form and fiber B-form DNA, respectively. The junction base steps exhibit neither pure A- nor pure B-type stacking.

comprise interesting structural characteristics. Although nuclear magnetic resonance experiments detected the presence of A-B junctions in chimerical molecules (Selsing and Wells, 1978; Selsing et al., 1978a), no direct observation in atomic detail of the transition points has been made so far. The deviations at the termini of the present hexamer from the B-form conformation are likely to be due to the crystal packing (end effect) rather than the particular sequence of the molecule. In general, DNA molecules are more susceptible to structural perturbations at the termini compared to the interior of the duplexes. For example, in the recently determined B-DNA hexamer d(CGGTGG)·d(CCACCG), one of the four independent terminal residues adopts a C2′-*exo* sugar pucker. However, the transition is particularly clear in the present structure because all terminal nucleotides have an N-type sugar pucker. Whatever the source of the distortion, it makes an interesting observation reflecting an energetically feasible conformation of the base pairs. B-to-A transitions occurring in the cell are probably also forced under the influence of some DNA-external factor, be it the puckering preference of an associated RNA strand or the induced fit upon binding of a protein or a drug.

TABLE 3 Hydration of the major groove and the backbone phosphates*

Residue	Atom	Water	Distance (Å)
C1	O5′	W13	2.97
T2	O1P	W13	2.77
		W7	2.73
	O2P	W27	3.21
		W9	2.64
	O4	W10	2.72
C3	O1P	W17	3.43
		W24	3.26
		W35	3.02
	O2P	W35	3.48
	N4	W18	3.19
G4	O1P	W30	2.71
	O6	W33	2.49
	N7	W32	2.88
A5	O1P	W34	2.62
	N7	W31	2.83
G6	O1P	W29	3.33
		W35	2.71
	O2P	W36	3.10
	O6	W16	2.77
	N7	W37	2.90

*The minor groove hydration is shown in Fig. 6.

The modeling studies, which are based on structures composed to 50% of an experimentally determined DNA helix, suggest that the conformation of the terminal C·G pairs may be an effective way to initiate an A-B junction. Although they are the physical termini of the present structure, the C·G pairs can be readily accommodated in a continuous helix, serving as terminators of the B-DNA conformation. In the models, the conversion from the B-form to the A-form is completed within two base pairs. The junction base pairs maintain efficient stacking interactions, and there is a pronounced bend between the two helical conformations. A bending angle of ~26° has also been found in the junction modeled by Selsing et al. (1978b), but details of the stacking interactions were different. Verification of the bending angle may come, e.g., from gel retardation experiments. DNA molecules therefore seem flexible enough to undergo the distortions required for such a transition under relatively mild external forces. Molecular dy-

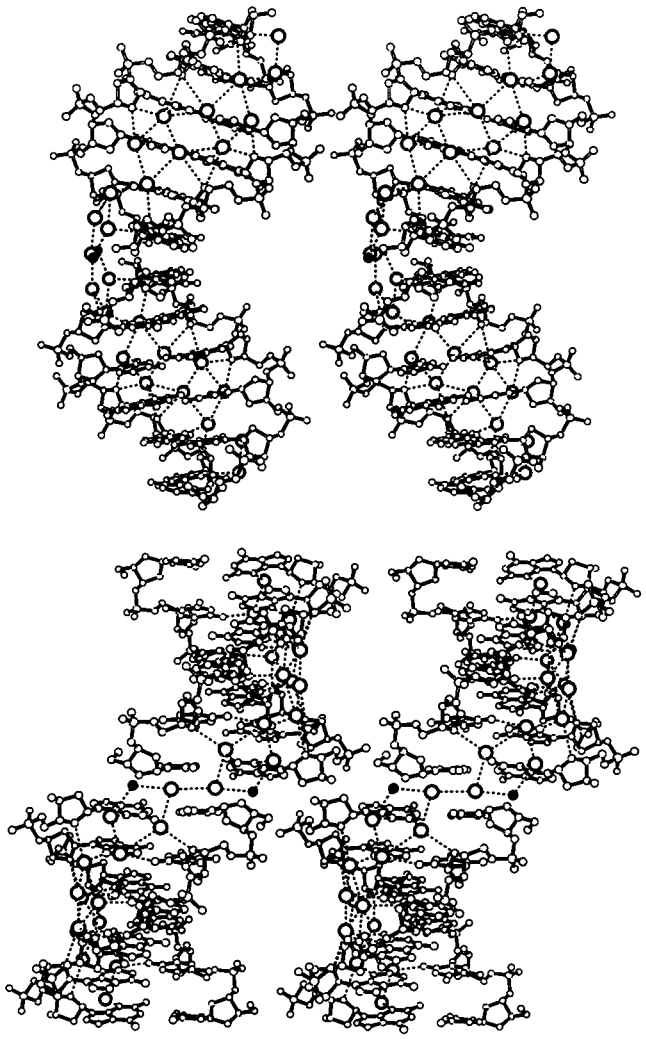

FIGURE 6 Two orthogonal stereo views of two pseudo-continuously stacked duplexes showing the water molecules and hydrogen bonding patterns in the minor grooves. The double spine of hydration (W19, W20, W21, W22, and symmetry equivalents) is intact for the central two-thirds (B-form) of each hexamer duplex but seems to be disrupted at the termini (stacking junction). In the central parts the hydrogen bonding results in a network of four-membered rings in the middle and three-membered rings at the sugar-phosphate backbone. The hydration network spanning the minor grooves of two head-to-tail stacked duplexes (W8, W26, W28, W38, and symmetry equivalents) consists of five-membered rings involving backbone O3′ atoms of symmetry-related duplexes (●) and a single spine at the interface of the helices.

namics simulations using long trajectories have suggested a considerable flexibility of the DNA molecule that also extends to the sugar puckering (Sekharudu et al., 1993).

Alternatively, the unusual conformation of the termini (C3′-*endo* sugar pucker, larger *x*-displacement and inclination) may represent a local structural perturbation of a B-form helix. When B-DNA fiber models were added to the termini, it became clear that the N-type puckering of one base pair would lead to a bend in the helix axis of ~14°. The bending occurred as a result of the increased inclination of the terminal base pairs, which was matched by the base pairs in the abutted fiber models. A different kind of bending is observed when only one sugar changes its pucker in a continuous helix, because different intrastrand phosphate-phosphate distances are associated with C3′-*endo* (~5.8 Å) and C2′-*endo* (~7 Å) puckers, leading to backbone contraction and extension. In the octamer d(GTGTACAC), a pronounced bend at the ends of the A-DNA occurred because the penultimate sugar rings adopted a C2′-*endo* pucker (Thota et al., 1993).

Comparison to related sequences

Under acidic conditions (~pH 4.5), the tetramer d(TCGA) forms a parallel-stranded duplex with C^+·C, G·G, and A·A base pairs (Wang and Patel, 1994). The 5′-TCGA-3′ motif is contained in the center of the current hexamer. Being only two bases longer, the present hexamer prefers a right-handed antiparallel conformation when crystallized at neutral pH. Similarly, long sequences containing the central 5′-TCGA-3′, like the decamer d(CTCTCGAGAG), are found as standard B-DNA duplexes (Goodsell et al., 1995). The d(TCGA) tetramer and other homo base-paired parallel-stranded motifs (Robinson et al., 1992; Robinson and Wang, 1993) were found to be converted to a B-form helix when the pH was raised (Robinson et al., 1992). It seems, therefore, that the parallel-stranded version is unlikely to be encountered under physiological conditions.

Hydration

Unlike in A-tract B-DNA, where only a single spine of hydration or a single groove-binding drug molecule can be accommodated in the narrow (~4.5 Å) minor groove, the expanded minor groove in the current structure (~6.9 Å) hosts a double water spine, which can be described as a sheet of waters following the curvature of the minor groove. The waters associate, among themselves and with nucleic acid atoms, into networks of three- and four-membered hydrogen bonded rings (Lipscomb et al., 1994). This hydrogen bonding pattern is different from the pentagonal water rings found in some A-DNA crystal structures in the deep (major) groove (Kennard et al., 1986). It is not clear whether such hydration networks actively influence the groove width or are encountered as a consequence of the wide groove. However, there are some striking analogies between the groove hydration and the way in which the grooves accommodate drugs of the lexitropsin class. Narrow minor grooves, formed at A/T-rich centers, can bind a single minor-groove-binding drug molecule (Kopka et al., 1985), whereas wide minor grooves are known to bind two drug molecules simultaneously side by side (Chen et al., 1994, 1995). The water structure of uncomplexed DNA with a narrow minor groove (single spine) (Kopka et al., 1983) is reminiscent of the mode of drug binding (single

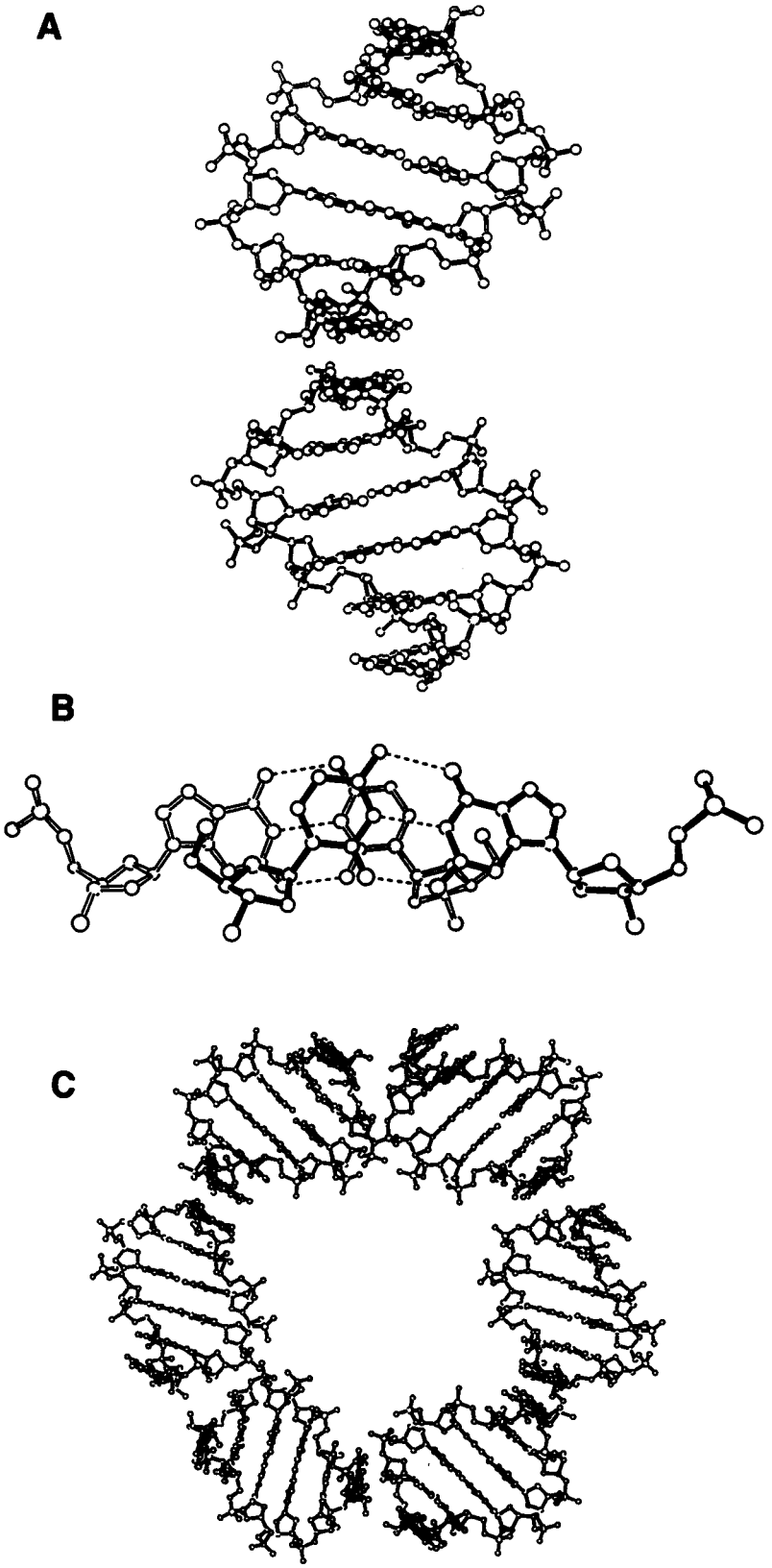

FIGURE 7 Features of the crystal packing of d(CTCGAG). (*A*) Two pseudo-continuously head-to-tail stacked duplexes are inclined at an (acute) angle of 60°. (*B*) Stacking at the junction of the duplexes (*A*, *middle*), showing the partial interstrand stacking. (*C*) Arrangement of six duplexes around the 6_2-screw axis, showing the large channel (diameter ~20 Å) in the center.

drug) (Kopka et al., 1985) and suggests that a similar analogy (double water spine, double drug) may hold in the case of wide minor grooves (Chen et al., 1994, 1995). The presence of guanine bases will oppose binding of drugs in the minor groove of d(CTCGAG). However, in 2:1 (drug: DNA) complexes the exocyclic guanine amino group gen-

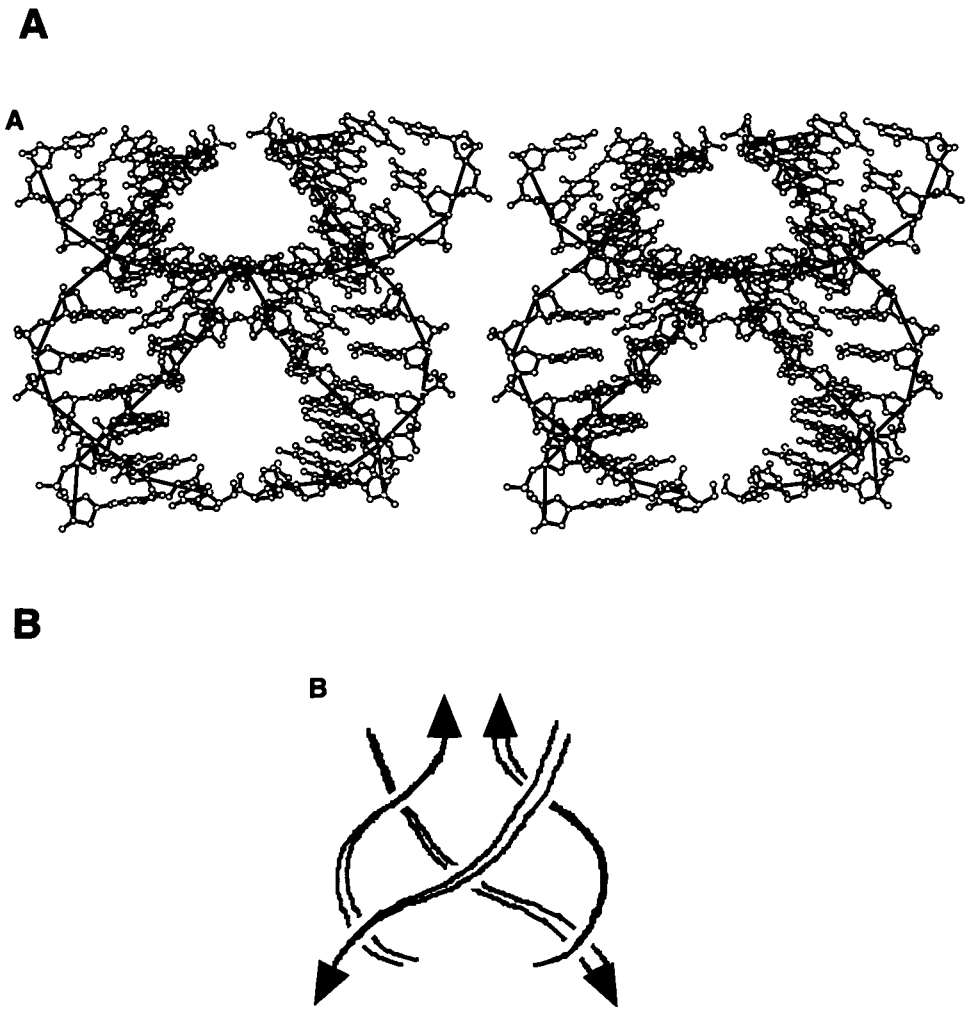

FIGURE 8 (*A*) Stereoview showing the possible strand crossover between four nearest-neighbor duplexes (two pairs of stacked duplexes). Thick lines connect consecutive C3′ atoms. The crossover point of the stacked columns of the duplexes is at the stacking junction. (*B*) Schematic representation of the topology of the junction in *A*.

erally poses less of a steric hindrance (Geierstanger and Wemmer, 1995).

A separate network of water molecules extends between the minor grooves of two pseudo-continuously stacked duplexes. Because the involved nucleotides are for the most part in an A-like conformation, this hydration pattern is quite different from the double spine in the center of the duplex. It is clear that such efficient interduplex water networks should stabilize the observed stacking pattern of the duplexes. It is possible that the tendency of the waters to associate among themselves and with the nucleic acid molecules in the observed fashion may play an active role in the stacking of the duplexes, e.g., dictating the twist and rise across the stacking junction. In this regard they may be reminiscent of the helix-unwinding intercalative drugs.

Crystal packing and its relation to the Holliday junction

The large channels formed around the 6_2-screw axis are an unusual feature for B-DNA crystal structures. Because a large portion of the unit cell volume is occupied by the channels, the volume per base pair (1727 Å^3) falls at the high end of the B-DNA range. A similar situation, where tight packing of the molecules leads to channels in the lattice, is usually seen in A-DNA structures. However, the local packing motif of the hexamer resembles that of B-DNAs, with molecules stacked on top of each other, whereas A-DNA duplexes tend to stack onto the shallow grooves of symmetry-related molecules. The present crystal packing seems to be an intermediary situation between A- and B-DNA packing schemes.

It is possible to create a Holliday junction from the packing motif of four nearest-neighbor duplexes (two pseudo-continuously stacked pairs). At the interface of the four duplexes, the termini of stacked duplexes approach each other closely, and strand exchange may occur (Fig. 8). Without changing the conformation, the backbones are considerably kinked at the crossover points. It is possible that these kinks are smoothed by local conformational changes in a Holliday junction. The current packing may simply model a double-helical approach before strand exchange, which subsequently may be accomplished by shifting from

the abutting backbone to the interdigitated backbone packing, as seen in other B-DNA structures.

This work was supported by National Institutes of Health grant GM-17378, an OSU Presidential Fellowship to MCW, and an Ohio Regents Eminent Scholar Chair to MS.

REFERENCES

Berman, H. M., W. K. Olson, J. Westbrook, A. Gelbin, T. Demeny, and D. Beveridge. 1994. The Nucleic Acid Data Base: A Guide to the Use of a Relational Database of Nucleic Acid Crystal Structures. Center for Computational Chemistry, Rutgers University, New Brunswick, NJ, and Wesleyan University, Middletown, CT.

Bhattacharyya, D., and M. Bansal. 1990. Local variability and base sequence effects in DNA crystal structures. *J. Biomol. Struct. Dyn.* 8:539–572.

Brünger, A. T. 1988. Crystallographic Refinement by Simulated Annealing. Clarendon Press, Oxford.

Brünger, A. T. 1992. X-PLOR—A System for X-ray Crystallography and NMR. Yale University Press, New Haven, CT.

Chen, X., B. Ramakrishnan, S. T. Rao, and M. Sundaralingam. 1994. Side by side binding of two distamycin A drugs in the minor groove of an alternating B-DNA duplex. *Nature Struct. Biol.* 1:169–174.

Chen, X., B. Ramakrishnan, and M. Sundaralingam. 1995. Crystal structures of the B-form DNA-RNA chimers complexed with distamycin. *Nature Struct. Biol.* 2:2–4.

Cruse, W. B. T., S. A. Salisbury, T. Brown, R. Cosstick, F. Eckstein, and O. Kennard. 1986. Chiral phosphorothioate analogues of B-DNA. The crystal structure of Rp-d[Gp(S)CpGp(S)CpGp(S)C]. *J. Mol. Biol.* 192: 891–905.

Geierstanger, B. H., and D. E. Wemmer. 1995. Complexes of the minor groove of DNA. *Annu. Rev. Biophys. Biomol. Struct.* 24:463–493.

Goodsell, D. S., K. Grzeskowiak, and R. E. Dickerson. 1995. Crystal structure of C-T-C-T-C-G-A-G-A-G. Implications for the structure of the Holliday junction. *Biochemistry.* 34:1022–1029.

Joshua-Tor, L., and J. L. Sussman. 1993. The coming age of DNA crystallography. *Curr. Opin. Struct. Biol.* 3:323–335.

Kennard, O., W. B. T. Cruse, J. Nachman, T. Prange, Z. Shakked, and D. Rabinovich. 1986. Ordered water structure in an A-DNA octamer at 1.7 Å resolution. *J. Biomol. Struct. Dyn.* 3:623–647.

Kim, L. J., D. B. Nikolov, and S. K. Burley. 1993. Co-crytal structure of TBP recognizing the minor groove of a TATA element. *Nature.* 365: 520–527.

Kim, Y., J. H. Geiger, S. Hahn, and P. B. Sigler. 1993. Crystal structure of a yeast TBP/TATA-box complex. *Nature.* 365:512–520.

Kopka, M. L., A. V. Fratini, H. R. Drew, and R. E. Dickerson. 1983. Ordered water structure around a B-DNA dodecamer. A quantitative study. *J. Mol. Biol.* 163:129–146.

Kopka, M. L., C. Yoon, D. Goodsell, P. Pjura, and R. E. Dickerson. 1985. The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proc. Natl. Acad. Sci. USA.* 82:1376–1380.

Lavery, R., and H. Sklenar. 1988. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* 6:63–91.

Lipscomb, L. A., M. E. Peek, F. X. Zhou, J. A. Bertrand, D. Vanderveer, and L. D. Williams. 1994. Diversity of water ring size at DNA interfaces: hydration and dynamics of DNA-anthracycline complexes. *Biochemistry.* 33:3649–3659.

Mooers, B. H., G. P. Schroth, W. W. Baxter, and P. S. Ho. 1995. Alternating and non-alternating dG-dC hexanucleotides crystallize as canonical A-DNA. *J. Mol. Biol.* 249:772–784.

Robinson, H., G. A. van der Marel, J. H. van Boom, and A. H. Wang. 1992. Unusual conformation at low pH revealed by NMR: parallel-stranded DNA duplex with homo base pairs. *Biochemistry.* 31:10510–10517.

Robinson, H., and A. H. Wang. 1993. 5′-CGA sequence is a strong motif for homo base-paired parallel-stranded DNA duplex as revealed by NMR analysis. *Proc. Natl. Acad. Sci. USA.* 90:5224–5228.

Sack, J., and F. A. Quiocho. 1992. CHAIN—Crystallographic Modeling Program. Baylor College of Medicine, Houston, TX.

Saenger, W. 1984. Principles of Nucleic Acid Structure. Springer Verlag, New York.

Sekharudu, C. Y., N. Yathindra, and M. Sundaralingam. 1993. Molecular dynamics investigations of DNA triple helical models: unique features of the Watson-Crick duplex. *J. Biomol. Struct. Dyn.* 11:225–244.

Selsing, E., and R. D. Wells. 1978. Polynucleotide block polymers consisting of a DNA/RNA hybrid joined to a DNA/DNA duplex. *J. Biol. Chem.* 254:5410–5416.

Selsing, E., R. D. Wells, C. J. Alden, and A. Struther. 1978a. Bent DNA: visualization of a base-paired and stacked A-B conformational junction. *J. Biol. Chem.* 254:5417–5422.

Selsing, E., R. D. Wells, T. A. Early, and D. R. Kearns. 1978b. Two contiguous conformations in a nucleic acid duplex. *Nature.* 275: 249–250.

Tari, L. W., and A. S. Secco. 1995. Base-pair opening and spermine binding—B-DNA features displayed in the crystal structure of a *gal* operon fragment: implications for protein-DNA recognition. *Nucleic Acids Res.* 23:2065–2073.

Thota, N., X. H. Li, C. Bingman, and M. Sundaralingam. 1993. High-resolution refinement of the hexagonal A-DNA octamer d(GTGTACAC) at 1.4 Å. *Acta Crystallogr. D.* 49:282–291.

Timsit, Y., and D. Moras. 1991. Groove-backbone interaction in B-DNA—implication for DNA condensation and recombination. *J. Mol. Biol.* 221:919–940.

Wang, A. H. J., G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel, and A. Rich. 1979. Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature.* 282:680–686.

Wang, B. C. 1985. Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol.* 115:90–112.

Wang, Y., and D. J. Patel. 1994. Solution structure of the d(TCGA) duplex at acidic pH. A parallel-stranded helix containing C^+·C, G·G and A·A pairs. *J. Mol. Biol.* 242:508–526.