Helix Geometry, Hydration, and G·A Mismatch in a B-DNA Decamer

GILBERT G. PRIVÉ, UDO HEINEMANN, SRINIVASAN CHANDRASEGARAN, LOU-SING KAN, MARY L. KOPKA, RICHARD E. DICKERSON*

The DNA double helix is not a regular, featureless barber-pole molecule. Different base sequences have their own special signature, in the way that they influence groove width, helical twist, bending, and mechanical rigidity or resistance to bending. These special features probably help other molecules such as repressors to read and recognize one base sequence in preference to another. Single crystal x-ray structure analysis is beginning to show us the various structures possible in the B-DNA family. The DNA decamer C-C-A-A-G-A-T-T-G-G appears to be a better model for mixed-sequence B-DNA than was the earlier C-G-C-G-A-A-T-T-C-G-C-G, which is more akin to regions of poly(dA)·poly(dT). The G·A mismatch base pairs at the center of the decamer are in the *anti-anti* conformation about their bonds from base to sugar, in agreement with nuclear magnetic resonance evidence on this and other sequences, and in contrast to the *anti-syn* geometry reported for G·A pairs in C-G-C-G-A-A-T-T-A-G-C-G. The ordered spine of hydration seen earlier in the narrow-grooved dodecamer has its counterpart, in this wide-grooved decamer, in two strings of water molecules lining the walls of the minor groove, bridging from purine N3 or pyrimidine O2, to the following sugar O4′. The same strings of hydration are present in the phosphorothioate analog of G-C-G-C-G-C. Unlike the spine, which is broken up by the intrusion of amine groups at guanines, these water strings are found in general, mixed-sequence DNA because they can pass by unimpeded to either side of a guanine N2 amine. The spine and strings are perceived as two extremes of a general pattern of hydration of the minor groove, which probably is the dominant factor in making B-DNA the preferred form at high hydration.

IN THE 8 YEARS SINCE ADVANCES IN SYNTHETIC METHODS made single crystal structure analysis of DNA oligonucleotides possible, several examples of A and Z helices have become known, but fewer of the biologically significant B form. The dodecamer: C-G-C-G-A-A-T-T-C-G-C-G and its variants (*1, 2*) stood alone as examples of B helices until the recent structure analysis of the phosphorothioate analog of G-C-G-C-G-C (*3*). The dodecamer became the paradigm for high-humidity B-DNA, and in combination with several A-helix structures, led to a set of rules for predicting sequence-dependent local variations in helix structure (*4, 5*). The dodecamer also gave rise to the concept of a "spine of hydration": a zig-zag string of water molecules running down the floor of the especially narrow minor groove of AT-rich regions of the helix (*6–9*). This spine has been implicated in stabilizing the B form relative both to unwinding and to conversion to the A or Z state. The width of the minor groove has also been shown to be an important structural feature that is recognized by both proteins and drugs (*10–14*).

We now describe the crystal structure analysis of a B-DNA decamer of sequence: C-C-A-A-G-A-T-T-G-G, which forms a double helix with normal Watson-Crick base-paired ends, but with two G·A mispairings in the center. The structure is of interest both as another of the all-too-rare examples of crystal structures of B-DNA, and as an example of non–Watson-Crick pairing. Such mispairings arise in DNA through recombination and errors of replication, and if uncorrected, may lead to point mutations in subsequent rounds of replication. In *Escherichia coli,* misincorporated bases that escape the proofreading activity of DNA polymerase are recognized and repaired by a mismatch repair system. The efficiency of repair depends both on the nature of the mispair and on the flanking DNA sequences (*15, 16*). Little is understood yet of the structural features used in discriminating mispairs from normal Watson-Crick base pairs. The G·A pairing also is of interest because it occurs naturally near the inner end of the anticodon stem in many transfer RNA's (tRNA's) (*17, 18*).

Crystallization and structure analysis. The DNA decamer: C-C-A-A-G-A-T-T-G-G was synthesized by liquid phase phosphotriester methods (*19*). Monoclinic crystals were grown at 4°C by vapor diffusion, with solutions 3 m*M* in decamer and 0.7*M* in $MgCl_2$ being equilibrated against a reservoir of 45 percent 2,4-methylpentanediol. The space group is *C*2, with the following cell dimensions: $a = 32.52$ Å; $b = 26.17$ Å; $c = 34.30$ Å; and $\beta = 118.9°$. Data were collected on a Nicolet P1-bar diffractometer with 5072 independent reflections observed above the one-sigma (1σ) level in the resolution range from 8.0 to 1.3 Å. A cluster of strong reflections corresponding to 3.4 Å spacing suggested that the molecule was a B helix aligned along the *c* axis, and such proved to be the case. Since the asymmetric unit contained a single decamer strand, or alternatively five base pairs of the double helix, the helix was constrained to lie on a crystallographic twofold axis.

The structure was solved by the molecular replacement method (*20, 21*), using as starting model a self-complementary decamer C-C-A-A-G-C-T-T-G-G with central G·C base pairs rather than G·A. Refinement, first with CORELS (*22*) and then with NUCLSQ (*23, 24*), reduced the residual error or *R* factor from an initial $R = 0.514$

G. G. Privé, M. L. Kopka, and R. E. Dickerson are in the Molecular Biology Institute of the University of California at Los Angeles, Los Angeles, CA 90024. U. Heinemann is at the Institut für Kristallographie, Freie Universität, Berlin, and was a postdoctoral colleague at UCLA. S. Chandrasegaran and L.-S. Kan are in the Division of Biophysics, School of Hygiene and Public Health, Johns Hopkins University, Baltimore, MD 21205.

*To whom correspondence should be addressed.

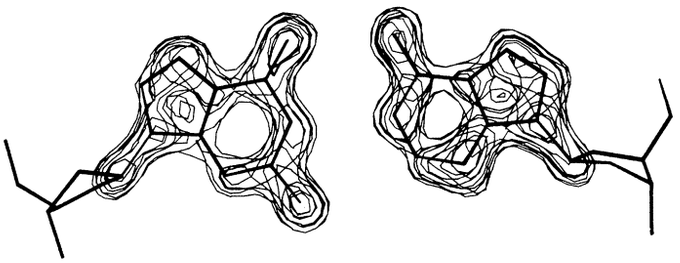


Fig. 1. Difference electron density map, at 1.3 Å resolution, of the mismatch decamer in the region of a G·A base pair. The two bases and their attached sugar C1′ atoms were deleted from the final refined structure in order to calculate this map. Compare figure 3 of (*25*).

at 3.8 Å resolution, to $R = 0.338$ at 1.3 Å. No restraints were placed on torsion angles, including those of the sugar rings.

At this point we removed the two central G·C base pairs from the phasing model, and calculated a difference map to examine the region of the missing pairs. The map showed that the cytosine of each pair had been badly fitted. The difference electron density clearly was that of an adenine in the *anti* conformation about its glycosyl bond, not the *syn* that has been reported in a different B-DNA setting (*25*). We then built *anti-anti* G·A base pairs into the center of the decamer helix, and further refinement by NUCLSQ converged at $R = 0.276$ without solvent. Gradual addition of solvent atoms as they appeared in successive difference maps led to a final R factor for 8.0 to 1.3 Å data of $R = 0.212$ for the 5072 1σ reflections, or $R = 0.185$ for the 3405 2σ data. The asymmetric unit of the final model contains 204 DNA atoms (half of the decamer helix), 45 water molecules treated as oxygen atoms, two $Mg(H_2O)_6$ complexes, and one $Mg(H_2O)_5$ complex bound directly to a DNA phosphate oxygen by a Mg–O bond. The appearance of the G·A base mispair after refinement is shown in Fig. 1. Both the original x-ray data and the final refined coordinates have been placed in the public domain by deposition with the Brookhaven Protein Data Bank.

Overall helix structure. The structure of the decamer helix is compared in Fig. 2 with that of the other available full turn of B-DNA, the dodecamer: C-G-C-G-A-A-T-T-^{br}C-G-C-G (*2*) (^{br}C, 5-bromocytosine). For ease of reference, these two helices are designated in what follows as the decamer and the dodecamer, respectively. The most striking difference is the width of the minor groove in the decamer, contrasting with the very narrow minor groove in the central -A-A-T-T- region of the dodecamer. A useful measure of the width of the minor groove opening is the shortest P-P distance across the groove, less 5.8 Å to represent two phosphate group van der Waals radii. For the decamer, this opening width averages 7.2 Å, whereas for the central region of the dodecamer it is only 3.7 Å.

The greater width for the decamer arises in part because of the presence of the two central G·A mispairs, which in the *anti-anti* conformation are 2.0 Å wider than a conventional Watson-Crick base pair; that is, 12.5 Å between C1′ atoms instead of 10.5 Å. Part of the greater width also may reflect the mixed sequence character of the center of the decamer helix, which lacks the homogeneous run of A·T pairs that has been proposed as causing a special narrowing of the groove of the dodecamer (*2*).

Each G·A base pair is held together by two hydrogen bonds (Figs. 2 and 3), from guanine N1 to adenine N1 (2.91 Å), and from guanine O6 to adenine N6 (3.01 Å). In addition, each G·A base pair has a strong 24.8° propeller twist (Table 1), which makes possible an unprecedented hydrogen bond of normal length (3.03 Å) between guanine N2 and the thymine O2 of the adjacent A·T base pair. This H bond gives that A·T base pair a high 18.3° propeller twist, and pushes its thymine O4 into close contact (3.0 Å) with the N6 atom of adenine A3. This contact may be responsible for flattening the propeller twist of base pair A3·T17 to only 11.1°.

The four central base pairs seem to constitute a structural block or unit (Fig. 2A), well-stacked internally, and separated from the outer three base pairs to either end by large roll angles, θ_R, and large vertical spacings, h (Table 2). Roll angles for each individual strand at steps 4 to 6 within this block are virtually zero, in contrast to a 16° roll at the ends of this block. This may be a special feature arising from the central mismatch base pairs. Control experiments—examination of the structures of self-complementary sequences analogous to the decamer, but with -T-A- and -G-C- instead of -G-A- pairing are needed to sort out the special attributes of mispairings.

As has been noted earlier (*2, 26*), a particular hallmark of the B helix by comparison with its A and Z cousins is its flexibility and its ability to make minor adjustments in local helix structure to accommodate special sequences. Although the C1′ to C1′ separation across a base pair increases suddenly by 2.0 Å at the central *anti-anti* G·A base pairs, the corresponding horizontal P-P distances adjust more gradually, spreading the widening over the four central base pairs of the "block". By contrast, only a minor adjustment in backbone geometry is needed for a G·A pair in an *anti-syn* geometry,

Table 1. Torsion angles and propeller twist for base pairs.

Base	Main chain torsion angles*							ϵ–ζ	Helix type	Sugar pucker	Propeller twist
	α	β	γ	δ	ϵ	ζ	χ				
C1			48	128	−169	−81	−120	−88	B_I	C1′-exo	12.6
C2	−78	184	47	143	−96	−195	−86	+99	B_{II}	C2′-endo	9.6
A3	−68	147	47	140	−175	−89	−85	−86	B_I	C2′-endo	11.1
A4	−71	175	47	131	−92	−181	−89	+90	B_{II}	C2′-endo	18.3
G5	−89	143	49	139	−144	−172	−90	+29	B_{II}	C2′-endo	24.8
A6	−61	153	60	140	−183	−77	−103	−104	B_I	C2′-endo	24.8
T7	−66	177	56	100	−162	−87	−129	−74	B_I	O4′-endo	18.3
T8	−68	175	58	135	−97	−192	−91	+95	B_{II}	C2′-endo	11.1
G9	−76	148	55	140	−179	−92	−86	−87	B_I	C2′-endo	9.6
G10	−62	169	46	106			−110			O4′-endo	12.6
Mean:	−71	164	51	130	−144	−130	−99				15.3
			B_I (avg.)		−173	−85					
			B_{II} (avg.)		−107	−185					

*Main chain torsion angles are defined by: P–α–O5′–β–C5′–γ–C4′–δ–C3′–ϵ–O3′–ζ–P. For definitions of glycosyl bond angle χ, base pair propeller twist, correlation between sugar pucker, and torsion angle δ, and diagrams of B_I compared to B_{II} conformations, see (*2*) and Fig. 4.

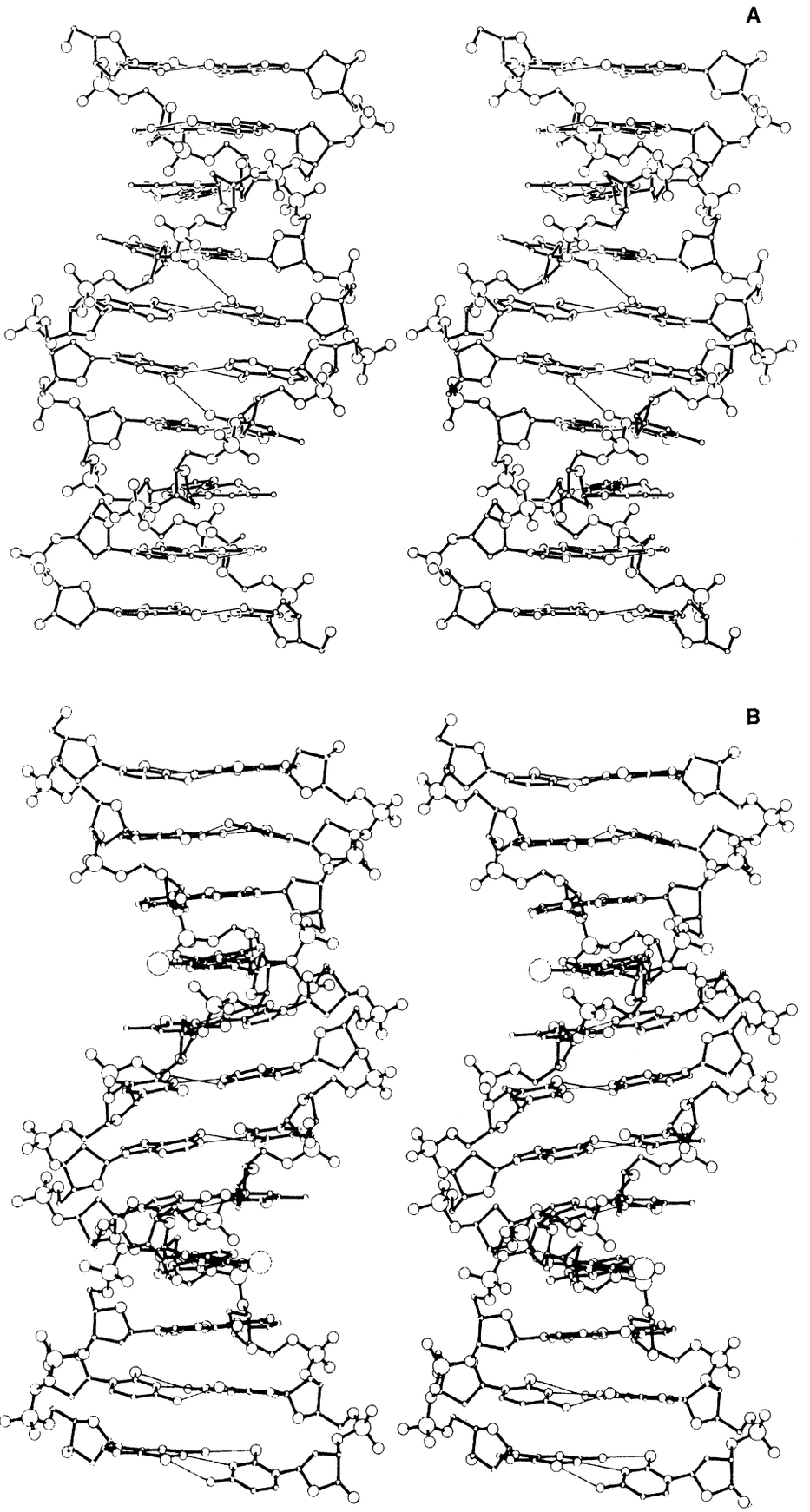


Fig. 2. Comparative stereo drawings of the B-DNA decamer and dodecamer. (**A**) Decamer with sequence C-C-A-A-G-A-T-T-G-G, with two central G·A mispairs. Each strand is numbered sequentially from the 5′ end: C1, C2, A3, A4, G5, and so on for strand 1, and C11, C12, A13, and so on for strand 2. As viewed directly into the minor groove at center, strand 1 runs from C1 at upper left to G10 as lower left, and strand 2 runs from C11 at lower right to G20 at upper right. The two strands are identical by space group symmetry; a crystallographic twofold axis is perpendicular to the page at the center of this drawing. For reference, steps from one base pair to the next are numbered 1 through 9 from the top of the helix to the bottom. (**B**) Dodecamer of sequence C-G-C-G-A-A-T-T-^{br}C-G-C-G (*2*). Minor groove view as before. Strand 1 runs from C1 at upper left to G24 at lower left, and strand 2 from C13 at lower right to G24 at upper right. The entire dodecamer is the asymmetric unit; it contains no crystallographic twofold axis. Two large dotted spheres are bromine atoms on 5-bromocytosines at positions C9 and C21.

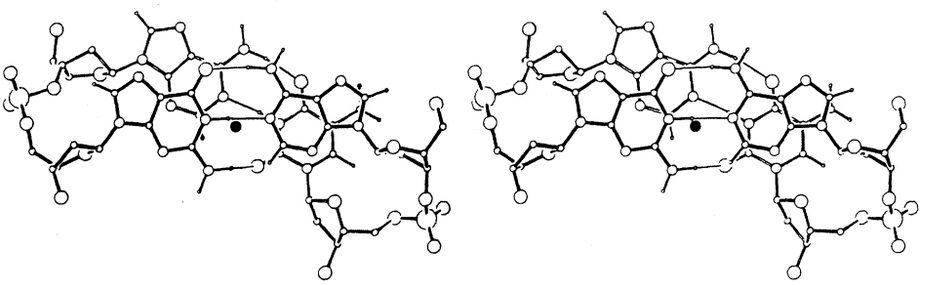


Fig. 3. Stereo diagram of base pairs G15·A6 (thick bonds) over A14·T7 (thin bonds), viewed down the helix axis (black dot). Phosphate P7 at right is in the B_I conformation (see Fig. 4), and P15 at left is in B_{II}. One effect of the B_{II} conformation at left is to de-stack the two purine bases.

as has been reported in a different setting (*25*). Even with the significant *anti-anti* helix enlargement observed in this structure, adaptation is made via small changes in several individual backbone torsion angles, each of which in isolation lies within the range observed with normal Watson-Crick base pairs.

B_I and B_{II} phosphate backbone conformations. Two conformations about the C3′–O3′–P segment of the backbone chain have been observed in B-DNA oligomers (Fig. 4): the more common B_I form in which torsion angles ϵ(C3′–O3′) and ζ(O3′–P) have the conformation (t, g^-), and the rarer B_{II} form with torsion angles of conformation (g^-, t). The B_I form is the one that has been considered "canonical" for years from fiber diffraction model-building (*27*), and B_{II} was more recently proposed as an alternative general structure for B-DNA (*28*). Of the 22 phosphates in the dodecamer, 20 were found to have the B_I conformation, and only the penultimate phosphate in each strand was B_{II} (*2*). In contrast, four of the nine phosphates in the decamer are B_{II} (Table 1), as are four of the ten phosphates in the two independent strands of the B-helical phosphorothioate analog of G-C-G-C-G-C (*3*), hereafter referred to as the hexamer.

A useful practical measure of these conformational alternates is the torsion angle difference ϵ–ζ, which is about −90° for B_I and about +90° for B_{II} (Table 1). A B_{II} phosphate conformation tends to lock the preceding sugar ring into a C2′-endo sugar pucker (Fig. 4), whereas B_I allows a greater degree of flexibility. The greater proportion of B_{II} phosphates in the decamer and hexamer than observed in the earlier dodecamer probably accounts for the fact that sugar conformations tend to be more clustered around the traditional B-DNA C2′-endo conformation in the decamer and hexamer (see sugar puckers in Table 1), in contrast to the broader range of conformations observed in the dodecamer (*2*). There is a clear correlation between B_I or B_{II} as measured by ϵ–ζ, and β for the subsequent nucleotide (*29*). A B_I phosphate is followed by a β at about 180°, and a B_{II} phosphate by a β at about 150°. This effect arises because of the "crankshaft" connection through the α or P–O5′ bond.

The effect of a B_{II} phosphate conformation is to allow the bases preceding and following that phosphate to slip away from one another and destack (*3*), as can be seen at left in Fig. 3. Bold face values of individual twist angles in Table 2 are for B_{II} steps, and these values tend to be larger than average. While it is difficult to unscramble cause and effect, the occurrence of unusually large overall helix twist angles at steps 2 and 5, and the presence of B_{II} phosphates at both sides of these steps, are logically correlated.

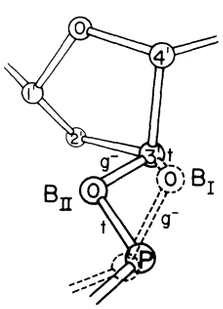


Fig. 4. Definition of B_I and B_{II} phosphate backbone conformations in B-DNA. In B_I, torsion angle ϵ is *trans* (about −180°) and ζ is *gauche*$^-$ (about −60°), whereas in B_{II}, ϵ is *gauche*$^-$ and ζ is *trans*. In B_I the C3′–O3′–P "elbow" angle points roughly parallel to the helix axis in the direction of the minor groove, whereas in B_{II} the elbow is twisted to point inward toward the helix axis. This swinging in of bond O3′–C3′ also moves the C3′–C2′ bond in a direction that forces a C2′-endo sugar pucker. The sugar ring has more conformational freedom in the B_I conformation [from (*2*)].

The B_{II} phosphate backbone conformation introduces an added element of variability in structure that causes the simple "Calladine's Rules" for predicting local helix structure variation to fail (*4, 5*). These rules were an attempt to account for local variations in helical twist and base pair roll angles, by hypothesizing that twist and roll underwent minor adjustments to relieve the steric clash that occurred when two purines occurred on opposite strands at two adjacent base pairs. They predict that helix step 2, from C2 to A3, should have a smaller than average helix twist and should open its roll angle toward the minor groove (positive θ_R). Exactly the opposite is observed—namely 51° twist and negative θ_R. Step 2 avoids purine-purine clash by damping down the propeller twists of the two base pairs to about 10° (a good Calladine strategy), and by destacking the bases via phosphate geometry (beyond the Calladine model). Moreover, the perturbation expected by the Calladine model between G19 and A8 is minor, by comparison with the major deformations rippling down toward the ends of the helix from the central block of four with its two G-A mispairs. The Calladine model is therefore incomplete. It considers only steric clash in accounting for local helix perturbations, and ignores such important factors as electrostatic interactions, hydrogen bonding, and hydration. Furthermore, it considers only one type of clash: cross-chain contacts between purines. The construction of a good general model for predicting local helix variation in B-DNA from base sequence requires examples of many more B oligomer structures than are now available.

Hydration and helix stabilization. DNA conventionally is spoken of as having three components: deoxyribose sugars, phosphates, and bases. But physical chemical measurements and x-ray structure studies in recent years have brought home the fact that DNA really has four essential structural elements: sugars, phosphates, bases, and bound water molecules. Hydration is a requirement for preservation of the integrity of the B helix, and it is a matter of no small interest to locate the water molecules responsible for this stabilization effect.

The dodecamer was observed from the x-ray structure analysis to have a well-defined spine of hydration, a zig-zag string of water molecules running down the floor of the minor groove in its -A-A-T-T- center. Such a spine was proposed as a major stabilizing influence in AT-rich DNA (*6, 7*). An equivalent spine was not observed in the hexamer (*3*), nor would it be expected since the spine is disrupted by the N2 amine groups of guanines. It therefore became of considerable interest, during the decamer analysis, to see whether anything resembling a spine would be found in the adenine-rich center of the helix.

The disposition of water molecules, and of two symmetry-related hydrated magnesium ions, down the minor groove of the decamer is shown in Fig. 5. (Four other hydrated magnesiums, not shown, are found in the major groove.) Both the presence of guanine N2 amine groups and the greater width of the minor goove in the decamer

Table 2. Helical parameters between steps of the decamer. Base-pair steps are as defined in Fig. 2A, from top to bottom of the helix. Step n is from base pair n to base pair $(n + 1)$. Steps 7 to 9 are not tabulated, as they are symmetrically related to steps 3 to 1 (compare steps 4 and 6). Individual parameter values for strands 1 and 2 are identified by subscripts 1 and 2, and the third value is obtained by considering the entire base pair. t is the helical twist or rotation from one base pair to the next, and h is the vertical rise along the helix axis. Boldface values of t indicate B_{II} phosphate conformations, as defined in Fig. 4. Roll angle θ_R refers to the angle opened between two base pairs by virtue of rotation about their long axes, and is positive if the angle opens toward the minor groove. Tilt, θ_T, is the angle produced by rotation about the short axis of the base pairs, and is positive if the angle opens toward the outside of the helix, or toward strand 1 for the combined tilt (*2*). Idealized fiber B-DNA (*27*) values for twist, rise, roll, and tilt are 36.0°, 3.38 Å, −3.6°, and 0°, respectively, and the propeller twist is negative: −4.4°.

Step	Helical twist (°)			Vertical rise (Å)			Roll (°)			Tilt (°)		
	t_1	t_2	t	h_1	h_2	h	θ_{R1}	θ_{R2}	θ_R	θ_{T1}	θ_{T2}	θ_T
1	34.5°	26.4°	29.1°	3.95	3.48	3.71	+5.4	+4.0	+4.3	8.8	−4.1	+1.3
2	**47.5°**	**47.6°**	50.6°	2.98	3.26	3.12	−7.1	−8.0	−6.9	6.4	−2.8	+1.7
3	20.7°	36.0°	26.9°	3.93	4.70	4.32	+14.1	+18.7	+16.5	4.6	−5.1	−3.6
4	**34.2°**	32.3°	32.6°	3.37	3.01	3.19	+0.3	−0.3	−7.9	6.9	−5.9	+0.7
5	**41.2°**	**41.2°**	41.4°	2.60	2.60	2.60	−2.4	−2.4	+2.1	1.8	−1.8	0.0
6	32.3°	**34.2°**	32.6°	3.01	3.37	3.19	−0.3	+0.3	−7.9	5.9	−6.9	−0.7

Table 3. Distances from minor groove waters in the decamer, to base N3, O2, and sugar O4′ atoms

Water No.	Nucleotide	Atom	Distance (Å)
41	C1	O4′	2.7
6	C2	O2	2.6
6	A3	O4′	3.0
14	A3	N3	2.6
14	A4	O4′	3.3
30	A4	N3	2.8
30	G5	O4′	3.5
3	G5	N3	3.0
3	A6	O4′	2.6
8	A6	N3	3.0
21	T17	O2	2.8
21	T18	O4′	2.6
21	T18	O2	2.4
13	T18	O2	2.9
13	G19	O4′	2.8
33	G19	N3	2.7
44	G20	O4′	3.3
44	G20	N3	3.2
44	Water No. 33		3.1

would argue against a true central spine of hydration, and indeed none is present. Instead, along each wall of the minor groove lies a string of water bridges, from a purine N3 or pyrimidine O2 of base n to the O4′ atom of deoxyribose $(n + 1)$. Minor irregularities are found here and there, but the general pattern is clear (Fig. 6A). The strings of hydration are present in both A·T and G·C regions of the helix; they are not disrupted by guanines because they pass by on either side of the N2 amines.

The very same ribbons of hydration line the two walls of the minor groove in the phosphorothioate hexamer (Fig. 6B), although they were not specifically noted by the original authors (*3*). Water molecules are not observed at 2 of the 14 sites, but the essential pattern of base-to-sugar bridging is plain.

The minor groove hydration in these three B helices—dodecamer, decamer, and hexamer—leads to conclusions that may be of general applicability in the stabilization of B-DNA by hydration. The two strings of hydration down the walls of a wide mixed-sequence minor groove and the single spine down the middle of a narrow A·T groove can be regarded as, conceptually, the two extremes of minor groove hydration. If in an imaginary experiment, we were to remove all of the guanine N2 amines from the floor of the groove, and then push the walls of the groove together, pairs of water molecules from the two strings would approach and coalesce at the center, as water molecules 8 and 8, 21 and 3, 13 and 30, 33 and 14 in Fig. 6A. Weakening the bonds to sugar O4′ would leave the merged waters with a binding pattern identical to that of the first hydration shell of the dodecamer spine of hydration (Fig. 6C). Bridging these sites with waters in a second hydration shell then would produce the complete spine. Both the double string and the spine have two water molecules per base pair, although differently arranged.

In AT-rich DNA, the spine of hydration may play a significant role in stabilizing the narrow minor groove of the B helix, relative to interconversion to the A form in which that groove is opened up on the surface of the helix (*6, 7*). The apparent conservatism of ordered minor groove hydration in B-DNA, whether in the form of a single spine or double string, suggests that water structure is as important in stabilizing the B form in mixed G·C and A·T sequences as it is in runs of consecutive A·T pairs.

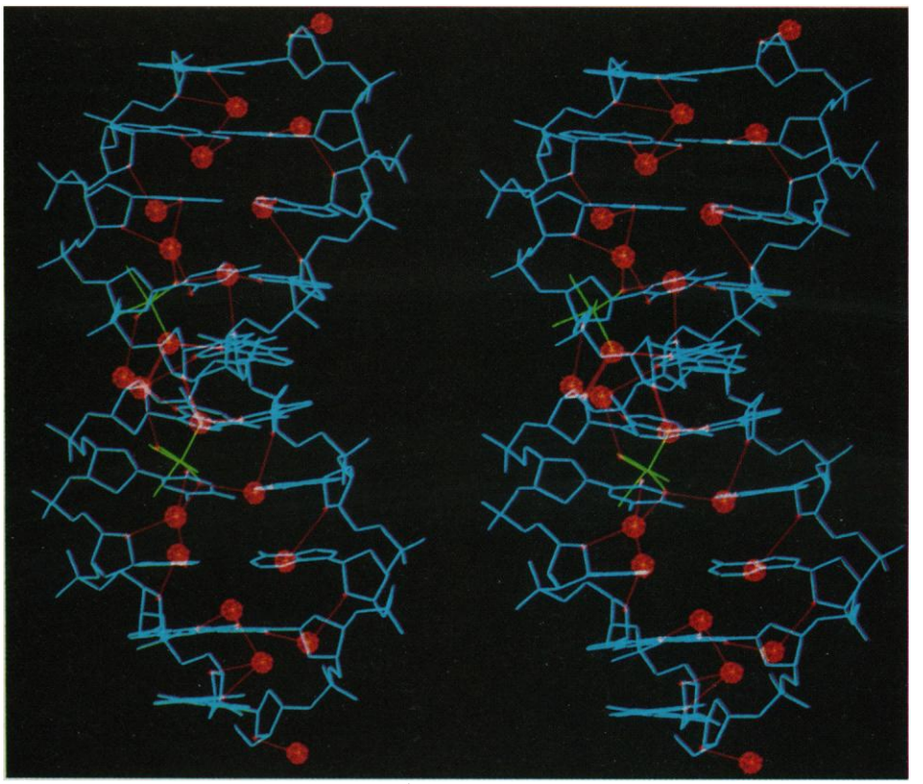
Varieties of B-DNA structure. The total number of distinct B-DNA oligomer crystal structures now stands at three: the dodecamer C-G-C-G-A-A-T-T-C-G-C-G, the mismatch decamer C-C-A-A-G-A-T-T-G-G, and the phosphorothioate hexamer analog of G-C-G-C-G-C. These are enough to show the variety inherent in a B helix, but not enough to allow one to generalize with confidence from sequence to detailed local helix structure. Minor groove width varies with A·T compared to G·C content, but the case of the decamer is clouded by the additional factor of the two G·A mispairings. Further analysis of C-C-A-A-G-C-T-T-G-G and C-C-A-A-T-A-T-T-G-G should help to separate the two effects.

This decamer now has been studied both in the crystal by x-ray diffraction and in solution by NMR (*30*) with the same conclusion. The G·A base pairs are in the *anti-anti* conformation about their glycosyl bonds (Figs. 1 and 3). The same geometry has been observed by nuclear magnetic resonance (NMR) for G·A mispairs in C-G-A-G-A-A-T-T-C-G-C-G (*31*), and by x-rays in the anticodon stem of transfer RNA (*17, 18*). The only reported example of G·A with *anti-syn* geometry has been the x-ray study of C-G-C-G-A-A-T-T-A-G-C-G (*25*), and this has not yet been examined by NMR methods. Theoretical calculations have suggested that the energy difference between *anti-anti* and *anti-syn* is small (*32*), and it may be that one or the other conformation might be assumed, depending upon local conditions.

The dodecamer spine of hydration, with two water molecules per base pair, now has been complemented by the hexamer and decamer double strings of water molecules, again with two waters per base pair. The spine may be regarded as a coalescence of the strings in narrow minor groove regions lacking N2 amines along the floor of the groove. The spine may be typical of general DNA sequences containing only A·T and T·A base pairs. Or, if Chuprina's calculations (*9, 33*) are correct in indicating that TpA disrupts the spine, it may only be found in poly(dA)·poly(dT) and in short AT-containing regions lacking a TpA step. The observed differences between dodecamer and decamer structures could well model the often-remarked but structurally unknown differences between poly(dA)·poly(dT) and poly(dA-dT)·poly(dA-dT) helices. In any event, the decamer or hexamer with their strings of hydration down the walls of a wider minor groove probably are better models for general, mixed-sequence B-DNA.

Whether as spine or strings, the regular hydration structure down the minor groove of B-DNA, with two waters per base pair,

Fig. 5. Side view of the decamer, showing those water molecules (red spheres) and hydrated magnesium ions (green octahedral crosses) that are found within the minor groove. Hydrogen bonds to base edge N and O, and to sugar O4′ atoms, are drawn as thin lines. The minor groove opens to the left and is seen in the upper half of the helix. The crystallographic twofold axis lies horizontally in the plane of the paper through the center of the helix.



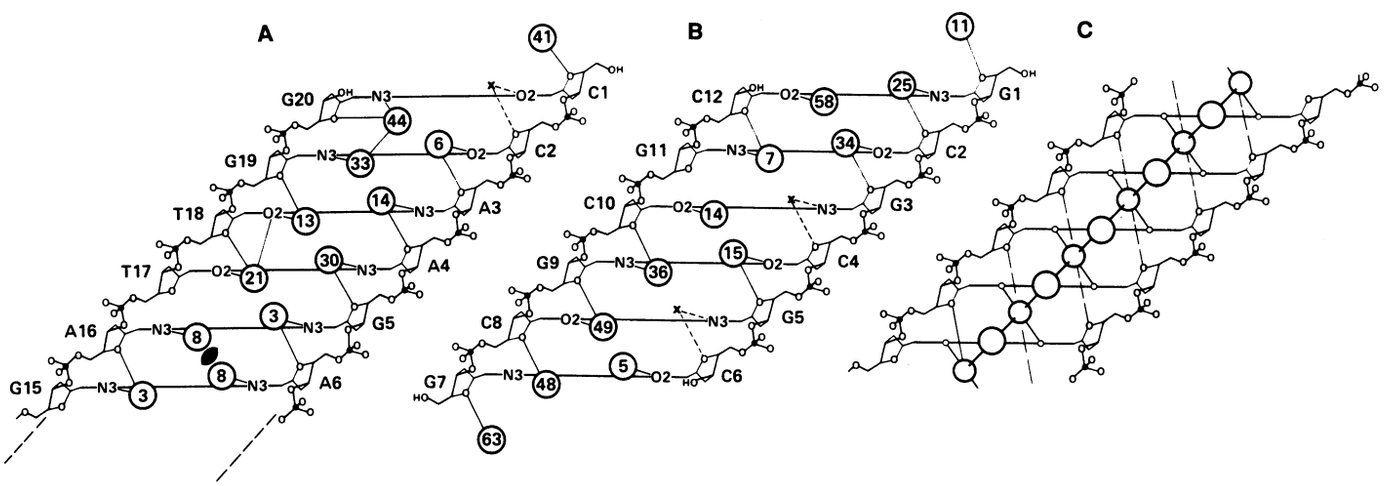


Fig. 6. Comparison of minor groove water structures in three B-helical oligomers. (**A**) Strings of water molecules (numbered circles) down the two walls of the minor groove in the decamer: C-C-A-A-G-A-T-T-G-G. Thin lines indicate hydrogen bonds from waters to adenine N3, thymine O2, or sugar O4′ atoms. Distances are given in Table 3. X marks an unfilled hydration site. The entire helix is repeated about the twofold axis marked between the two symmetry-equivalent waters No. 8 (**B**) Strings of water molecules down the walls of the minor groove in the phosphorothioate hexamer analog of G-C-G-C-G-C (*3*). Same conventions as before, except that the entire double helix is drawn. Hydrogen bond distances to waters are listed in (*3*). (**C**) Idealized diagram of the spine of hydration, as found in the central AT region of C-G-C-G-A-A-T-T-C-G-C-G (*6, 7*). The smaller heavy circles are first shell water oxygens and the larger circles are second shell water oxygens, nearer to the viewer. Hydrogen bonds to O2 and N3 (here marked only by small circles) are in light line and distances to sugar O4′ atoms are dashed. The coordination from a base edge N3 or O2, to a first shell water, to a sugar O4′, is the same as the coordination of one water molecule of a string in (A) or (B).

probably makes a major contribution toward stabilizing that helical form. The principle of economy of hydration of phosphate groups (*34*) may also be a corollary factor. As with dodecamer and hexamer, the phosphates of the decamer are seen to be individually hydrated. Only two phosphate-water-water-phosphate long bridges are observed, and no short bridges involving a single water molecule. Economy of hydration, or the ability of A-DNA to hydrate its phosphates fully with a smaller number of water molecules, may explain why the A form is favored at low water activity, and ordered and systematic hydration of the minor groove in B-DNA may be critical in stabilizing that form under conditions of high water activity, both factors being operative in general, mixed A·T and G·C sequences.

REFERENCES AND NOTES

1. H. R. Drew *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 2179 (1981).
2. A. V. Fratini *et al.*, *J. Biol. Chem.* **257**, 14686 (1982).
3. W. B. T. Cruse *et al.*, *J. Mol. Biol.* **192**, 891 (1986).