The Anatomy of A-, B-, and Z-DNA

Richard E. Dickerson, Horace R. Drew
Benjamin N. Conner, Richard M. Wing
Albert V. Fratini, Mary L. Kopka

Summary. Recent advances in DNA synthesis methods have made it possible to carry out single-crystal x-ray analyses of double-stranded DNA molecules of predetermined sequence, with 4 to 12 base pairs. At least one example has been examined from each of the three known families of DNA helix: A, B, and Z. Each family has its own intrinsic restrictions on chain folding and structure. The observed solvent positions in these crystal structures have confirmed earlier fiber and solution measurements, and have led to proposals explaining the transitions from B to A and from B to Z helices. Prospects are improving for an understanding of the mode of bending of DNA in chromatin, and the way in which specific DNA sequences are recognized by drug molecules and repressor proteins.

When the first single-crystal x-ray analyses of protein molecules began emerging in the early 1960's (*1, 2*), there was no reason why corresponding analyses of DNA of varied sequence would not have been equally absorbing—except that pure, homogeneous DNA for crystallization was not available. Instead, attention turned to diffraction patterns of drawn fibers of DNA under varying salt, alcohol, and humidity conditions, circular dichroism, infrared spectroscopy, and other solution techniques. From this work came the realization that double-helical DNA structures could be categorized into two families, A and B, with fundamental differences in the way in which the base pairs were stacked and the sugar-phosphate backbone strands were wrapped around the helix axis (*3–5*).

Development of triester methods of DNA synthesis (*6*) made it possible for the first time to prepare tens of milligrams of pure DNA of predetermined sequence, and to grow crystals suitable for x-ray analysis. The first DNA tetramer examined, of sequence pApTpApT (*7, 8*), proved not to form a regular double helix, but subsequent structure analyses have led to examples of both the A and B families (*9–16*) and to a new and unexpected left-handed Z helix (*17–21*). From a comparison of the three helical types at the molecular level we can, in the first place, begin to understand the principles of DNA helix formation and the factors affecting transformations between helix types. A second highly desirable goal is an understanding of the molecular basis for the recognition of particular base sequences by proteins such as repressors and restriction enzymes. A third goal, under active pursuit by several groups of investigators, is the unraveling of the way in which specific intercalating and nonintercalating drugs bind to DNA, and why they sometimes show preferences for particular DNA sequences. Fiber diffraction by its very nature can only provide an averaged or idealized type structure, and cannot establish what is happening at one base pair or type of base. This fine detail is what single-crystal structure analysis of DNA is beginning to provide.

A-, B-, and Z-DNA at the Molecular Level

The newest addition to the roster of DNA helix types studied by single-crystal x-ray analysis is A-DNA, in the form of the double-helical tetramer d(iodoCpCpGpG) or ^{I}CCGG (*9, 10*), and more recently the octamer GGTATACC (*11*). (The tetramer was synthesized with 5-iodocytosine at the first position to assist in structure analysis.) The B-DNA helix is represented by the dodecamer d(CpGpCpGpApApTpTpCpGpCpG) or CGCGAATTCGCG (*9, 12–16*) and by the related CGCGAATT^{Br}CGCG with 5-bromocytosine, a structural analog of 5-methylcytosine, at the ninth position along each chain of the double helix (*22*). The parent dodecamer has a 19° overall bend in helix axis, arising at least in part from intermolecular contacts in the crystal. But the simultaneous presence of bromocytosine at the ninth position and a high alcohol concentration [60 percent methylpentanediol (MPD)] causes the helix to straighten out in a manner that, under certain conditions, is reversible within the crystal. The straight-helix, high-MPD crystal form also has a more well-defined hydration of the phosphate backbone than does the native structure, and this will be important later in the discussion of the transition of the B to A helix. One other contribution to B helix geometry presently is known: daunomycin intercalates between the CpG steps of the hexamer CGTACG (*23*), and, although the geometry around the intercalation sites is distorted from that in a free B helix, the central TpA step is sufficiently unperturbed to yield useful information about the B helix. Three left-handed Z-DNA structures are known: the CGCG tetramer in its low-salt (*17*) and high-salt (*18, 19*) forms, and the hexamer CGCGCG with several cation variants (*20, 21*).

The structures of A-, B-, and Z-DNA as deduced from these single-crystal analyses are compared in Fig. 1, a to f. The first impression from A and B helices is how well these structures fit the model helices that have been derived over the past 25 years from fiber diffraction data. A-DNA is a wide, stubby helix with bases canted sharply to the helix axis, a cavernous major groove, and a minor groove that is almost too shallow to be termed a groove at all. B-DNA is slimmer and more elongated, with base planes essentially perpendicular to the helix axis, and with a narrow minor groove and wide major groove of comparable depth. The Z helix continues this logical progression: it is even thinner and more elongated for the same number of base pairs. With Z, the minor groove is cavernously deep, and the major groove is completely flattened out on the surface of the molecule. The Z helix is almost the complete inverse of A; it is long, thin, and left-handed rather than short, fat, and right-handed and shows major and

This work was carried out by Dr. Dickerson and his colleagues while at the California Institute of Technology. Richard E. Dickerson, Horace R. Drew, Benjamin N. Conner, and Mary L. Kopka have recently moved to the Molecular Biology Institute, University of California at Los Angeles, Los Angeles 90024. Richard M. Wing is a professor of chemistry at the University of California, Riverside, Riverside 92521. Albert V. Fratini is a professor of chemistry at the University of Dayton, Dayton, Ohio 45469.

0036-8075/82/0430-0475$01.00/0

minor grooves that have exchanged character. The two features characteristic of Z-DNA are the reversal of helix direction and the use of two consecutive nonequivalent base pairs as the helical repeating unit, the latter conferring on the sugar-phosphate backbone a zigzag conformation that is responsible for the Z designation.

Table 1 lists helix parameters as deduced from crystal structure analyses of examples of the three helical families. There are few surprises, except perhaps for the range of variability of individual helix parameters around mean values similar to those found by fiber diffraction. These crystal structures bear roughly the same relation to the fiber diffraction work as the structures of myoglobin, hemoglobin, and lysozyme did in 1965 to the fiber diffraction analyses that originally suggested the α helix, β sheet, and collagen triple helix (*24*). What we see in these discrete DNA molecules are variations on a fiber-derived theme; as with proteins, it is these variations that eventually will provide the most information about the functioning of the macromolecules.

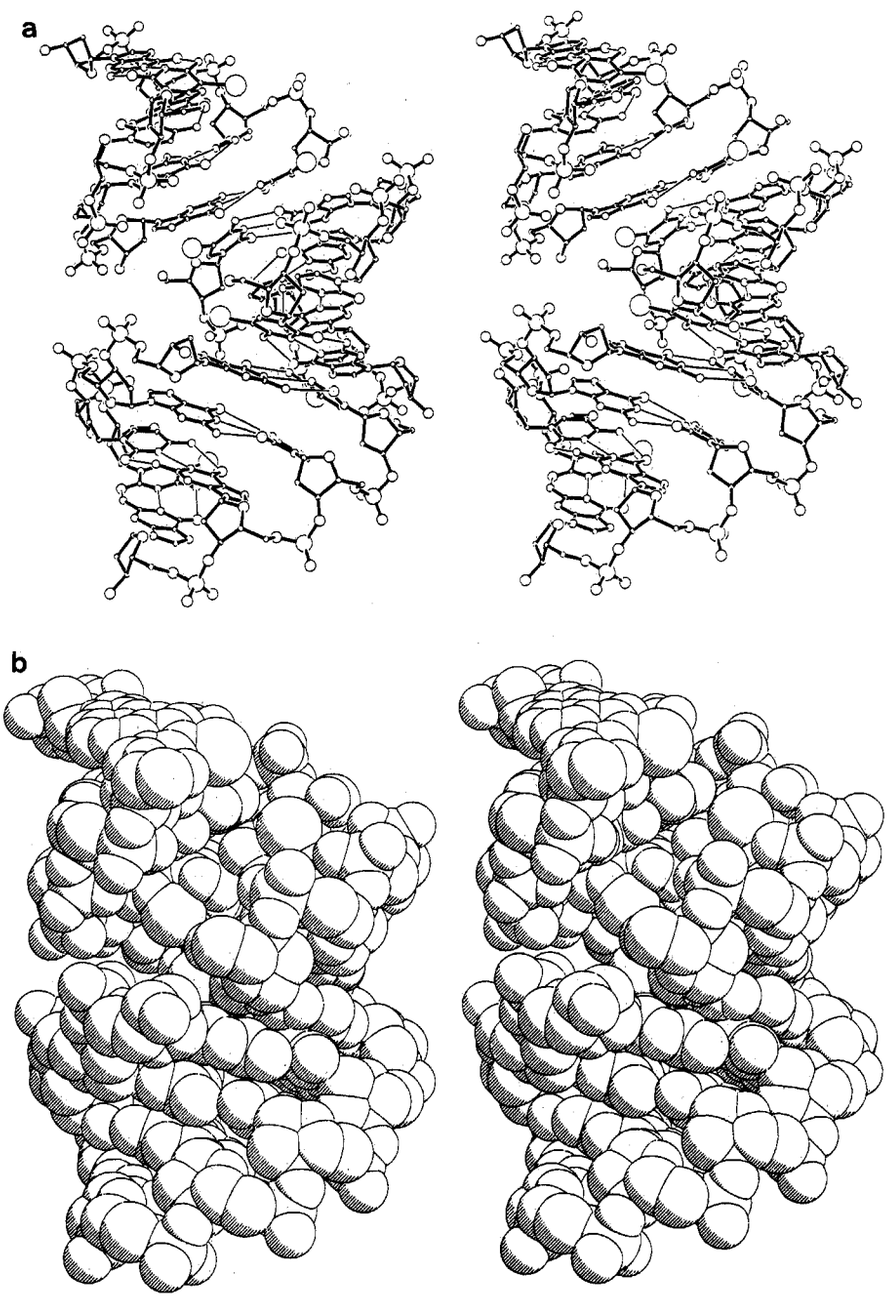


Fig. 1. Uniform-scale skeletal and space-filling drawings of A-, B-, and Z-DNA as determined from single-crystal x-ray structure analyses of the oligomers at the California Institute of Technology. In each case 12 base pairs are depicted: a complete molecule for the B-DNA dodecamer, and three molecules stacked along a common helix axis for the tetramers. The double helix, tetramer or dodecamer, is the crystallographically unique unit in each structure. (a and b) The A-helical tetramer d(iodo-CpCpGpG) or ^{I}CCGG. Bases are numbered sequentially from the 5′ to 3′ end for one strand and then the other: C1 to G4 and C5 to G8, with base pair G4 · C5 at top and C1 · G8 at bottom. Atoms in order of descending radius are I, P, O, N, C. The distortion in C5 ring plane orientation is a consequence of intermolecular packing in the crystal.

As an example, the mean global rotation per base pair in the B-DNA dodecamer is 35.9°, in precise agreement with the 10.0 base pairs per turn derived from fiber diffraction. The standard deviation of 4.2° indicates a broad range of variation in individual twist values, and these in fact range from as low as 27.4° (that is, 13.1 base pairs per turn) to as high as 41.9° (8.6 base pairs per turn). But the fact that the same variations occur at corresponding positions in three independently refined structures of sequence CGCGAATTCGCG and the related CGCGAATT^{Br}CGCG suggests that the individual values of helical twist, as well as the overall average, are properties of this particular sequence. Furthermore, Lomonossoff *et al.* (*25*) have found that deoxyribonuclease I cleaves each phosphate bond of CGCGAATTCGCG in solution with a rate constant that can be correlated with the individual helical twist angle at that phosphate in the crystal structure (*14*). This is a strong argument for the essential identity of helix structure in solution and crystal. A sequence dependence of helical twist values has been noted in other solution experiments as well (*26, 27*), with 9.9 base pairs per turn reported for double-stranded poly(dA) · poly(dT), and 10.5 for both the alternating polymer poly(dA–dT) and random-sequence DNA. More such relationships should emerge from both crystal and solution studies.

Several sequence-related generalizations about local variations in helix parameters have begun to emerge from x-ray analysis of the dodecamer (*14, 16*) and these may be the beginning of a sequence and structure lexicon of DNA helix behavior. As with helical twist angles, a concordance will constantly be sought between crystal results and solution studies done with various techniques (*28*). Only for B-DNA have such sequence or structure relationships begun to appear, or has much been learned yet about the kinematics of helix bending. The CCGG A-DNA helix is too short, and the alternating $(CG)_n$ sequence of Z-DNA is too monotonous, for useful information on sequence and bending to result. In this article we consider the intrinsic constraints on backbone conformation in A-, B-, and Z-DNA, and the observed hydration structures and their implications for the B-to-A helix transition.

Boundary Conditions on Chain Conformation

The observed chain geometry in the three helix families indicates that they are subject to quite different structural constraints. For proteins, the planarity of the amide bond means that only two angles per amino acid residue are necessary to specify the main chain conformation, and the conformation of a folded protein chain is often plotted as a series of points on a (ϕ, ψ) torsion angle or Ramachandran plot (*29, 30*). Six main chain torsion angles must be specified for nucleic acids and, although torsion angle nomenclature has been a bone of contention among investigators in the past, a consensus seems to have been reached on the definitions in Fig. 2 (*31*). One of these main chain torsion angles, δ, is particularly informative because it identifies the sugar puckering or conformation of the deoxyribose ring, a parameter that has played an important part in the earlier studies of DNA structure. One atom of the five-membered furanose ring will be out of plane to one side or the other (Fig. 3), and a conformation is designated as Xn′-*endo* or Xn′-*exo*, where Xn′ identifies the atom and *endo* and *exo* describe whether the out-of-planeness is on the same side of the ring as the C5′ atom or the opposite side.

While sugar conformation is the natural structural feature to consider when building a model of double-helical DNA, torsion angle δ actually is the more accurate parameter to measure from electron density maps produced by crystal structure analysis. Even though the resolution of the map may be insufficient to resolve all five ring atoms clearly, or to establish which one is most out of plane, the unambiguous pathway of the main chain past the sugar ring can establish the value of torsion angle δ to within 5° or 10°. Levitt and Warshel (*32*) have calculated the relationships between sugar pucker and torsion angle δ, and approximate measures of the relative energies of different conformations. Three common sugar conformations are depicted in Fig. 3. The C3′-*endo* or C2′-*exo* conformation has traditionally been assumed for A-DNA, and C2′-*endo* or C3′-*exo* for B-DNA, but as will be seen, this is an oversimplification.

The other especially informative torsion angle is glycosyl angle χ, between the C1′ atom of the sugar ring and the nitrogen of its attached base (Fig. 2a). When the ring is turned so that it is bent toward the minor groove as at left in Fig. 4, the conformation is described as *syn*, and χ has a value around +70°. For purines, the *syn* conformation is somewhat restricted because of close contacts between the sugar ring and purine atom N3, and for pyrimidines the *syn* conformation is disfavored—the geometry of attachment to a six-membered ring brings the sugar too close to atom O2. In contrast, in the less restrictive *anti* conformation (Fig. 4, right) the sugar ring is bent away from the minor groove and from serious steric hindrance; χ then is free to adopt a broad range of values, from −80° to −180°. The characteristic

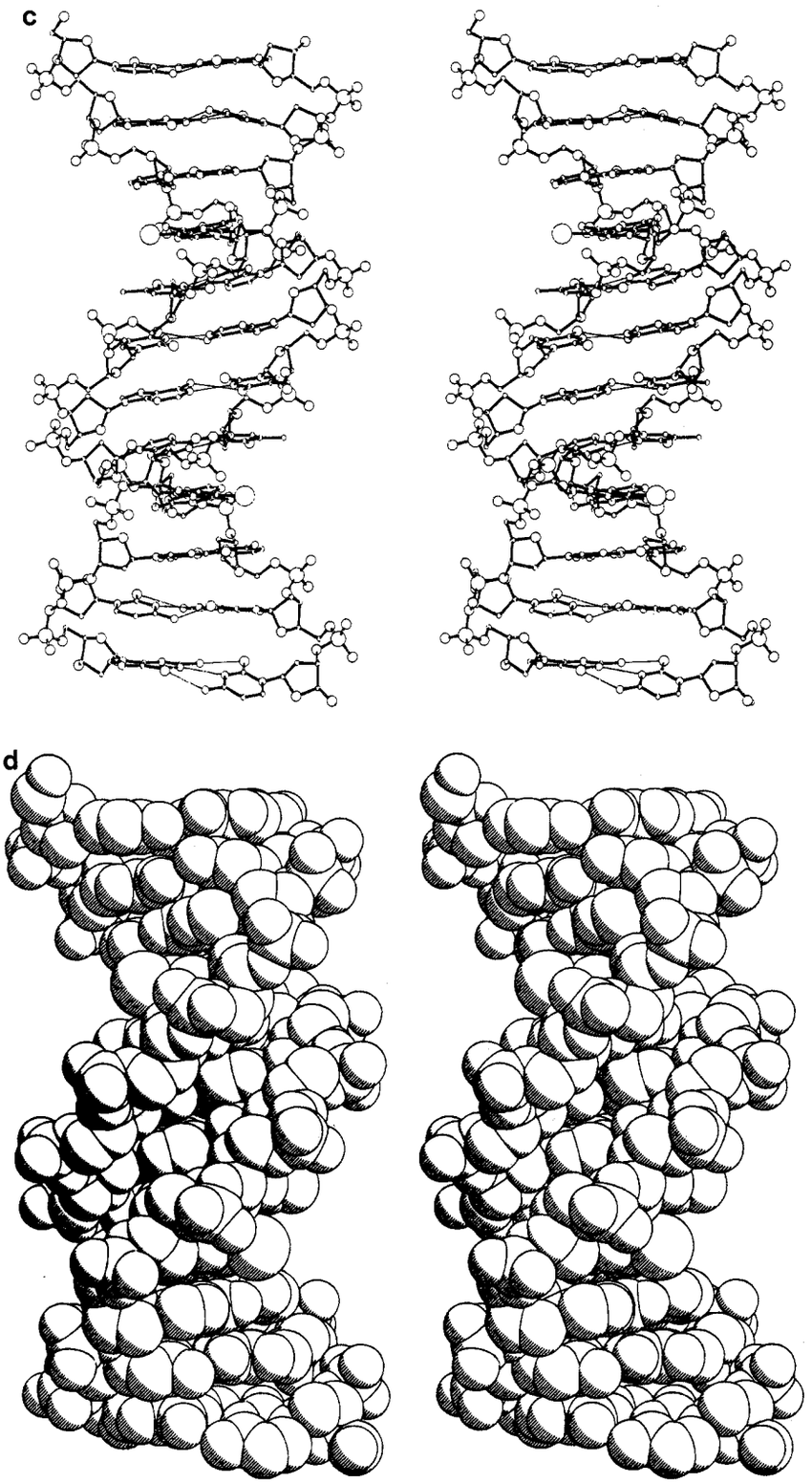


Fig. 1. (c and d) The B-helical dodecamer CGCGAATT^{Br}CGCG. Strands numbered C1 to G12 and C13 to G24 from 5′ ends, with base pair C1 · G24 at top and G12 · C13 at bottom. Atoms in order of descending radius are Br, P, O, N, C. The large propeller twist of C13 at bottom is produced by interaction of its N4 amino group with a backbone phosphate on a neighboring molecule.

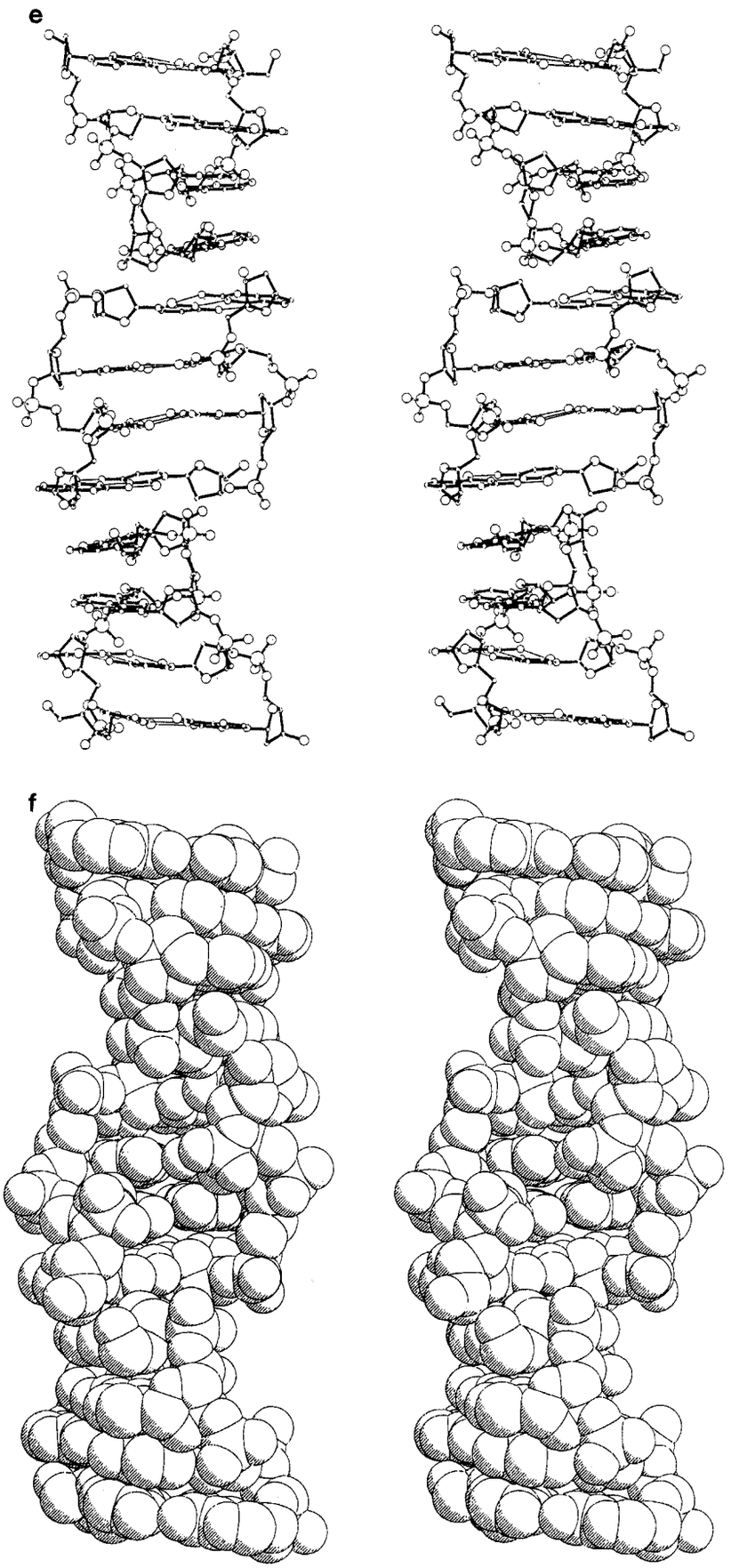


Fig. 1. (e and f) The Z-helical tetramer CGCG in its high-salt form. The base numbering is as for CCGG, with G4 · C5 at top and C1 · G8 at bottom. Each individual molecule in these six stereo drawings is depicted exactly as it appears in its crystal structure, although the stacking of tetramers in the crystals is different from that shown here. No attempt has been made to idealize or regularize the helices or to patch in the missing phosphates that would connect tetramers in a continuous helix.

zigzag chain path in Z-DNA arises because of the alternation of *syn* and *anti* conformations at guanines and cytosines, respectively. In all other DNA helices, only the *anti* conformation is found.

The other five main chain torsion angles also vary to some extent among the three helix families, but are not as characteristic or as diagnostic as δ and χ. The plots of χ as a function of δ for A-, B-, and Z-DNA, analogous to Ramachandran plots for proteins (Figs. 5 to 7) illustrate the different nature of the constraints on chain conformation in the three families (see *9, 13, 16, 19*). In A-DNA the conformations cluster around the classical fiber-derived value for an A helix, whereas in B-DNA the conformations are distributed over a broad range, but χ and δ are linearly correlated. Both χ and δ are uncorrelated in Z-DNA, but each is subject to individual constraints: at guanines, χ is constrained by steric hindrance to a narrow *syn* range around +70° whereas δ is relatively free, and for cytosines, δ is constrained around the C2′-*endo* position whereas χ has greater freedom.

The strong clustering of A-DNA conformation points around the classical fiber value (A_F in Fig. 5) is remarkable only in the light of the lack of such a well-defined clustering for B-DNA. We do not yet have a feeling for the basis for this relative uniformity of the A helix. The correlation between χ and δ for B-DNA can be understood from Fig. 8. At δ = 120° (Fig. 8c) the C3′–C2′ and C4′–O1′ bonds are eclipsed and the four atoms lie in a plane, with C1′ in the *exo* configuration or the *endo* (not shown). Closing down torsion angle δ moves C2′ to the left and O1′ to the right, resulting in O1′-*endo* (Fig. 8b) or the even more extreme C3′-*endo* (Fig. 8a). Conversely, opening up δ as in Fig. 8d moves C2′ to the right and O1′ to the left, yielding C2′-*endo*. Glycosyl angle χ is correlated with these motions because rotating O1′ to the right tends to make χ more negative, and rotating it to the left increases χ. The ideal χ_0 values shown for each conformation in Fig. 8 are perturbed by rotation of the base plane away from perpendicularity to the helix axis or by tilting of the C4′–C3′ bond of the furanose ring away from the helix axis. [For a quantitative analysis of this effect, see (*22*).] But the correlation between χ and δ (and hence sugar conformation) is sufficiently strong to be immediately apparent in Fig. 6.

The reasons for the absence of χ and δ correlation in Z-DNA, and the imposition of separate constraints on χ and δ for guanines and cytosines, are to be found in the special nature of the zigzag backbone (Fig. 1, e and f). A detailed

discussion is given in (*9*) and (*19*), but the arguments can be summarized here. As has already been mentioned, the *syn* conformation for guanines restricts χ to a narrow range to either side of +70°, whereas the *anti* cytosines are not thus restricted and can adopt a broader range of values. The differences in δ restrictions between cytosine and guanine in Z-DNA arise from the different geometry of the backbone chain past the two types of base. As Fig. 4 indicates, decreasing δ at a *syn* guanine with the main chain pathway held more or less fixed pushes the end of the base pair in the direction shown by the arrow at left. This represents only a slipping of one end of the base pair out of the base stack in the direction of the major groove (which is not really a groove at all), a relatively unfettered motion. In contrast, closing down δ at an *anti* cytosine requires that the base pair slip along its own long axis within the base stack (arrow at right in Fig. 4), transmitting the deformation to the opposite chain on the other side of the helix. This is intrinsically a more difficult motion; accordingly, for Z-DNA, δ is more constrained at cytosines than at guanines.

In summary, each of the three families of helix has its own individual kinematic constraints. At present, the small range of variation in A-DNA is more surprising than are the individual restraints on freedom of motion found in B- and Z-DNA. There seems to be no justification for characterizing Z′, Z_I, and Z_{II} as separate helical subtypes, especially since all actual Z-DNA molecules: CGCGCG and high- and low-salt CGCG, always consist of mixtures of Z_I, Z_{II}, and intermediate conformations (*19, 21*). It is an interesting exercise to combine Figs. 5 to 7 into one grand conformation plot for double-helical DNA, and to try to establish boundaries of allowed zones for all three helices—to create the DNA equivalent of a protein Ramachandran plot. There is one important difference: these DNA conformation plots are derived strictly

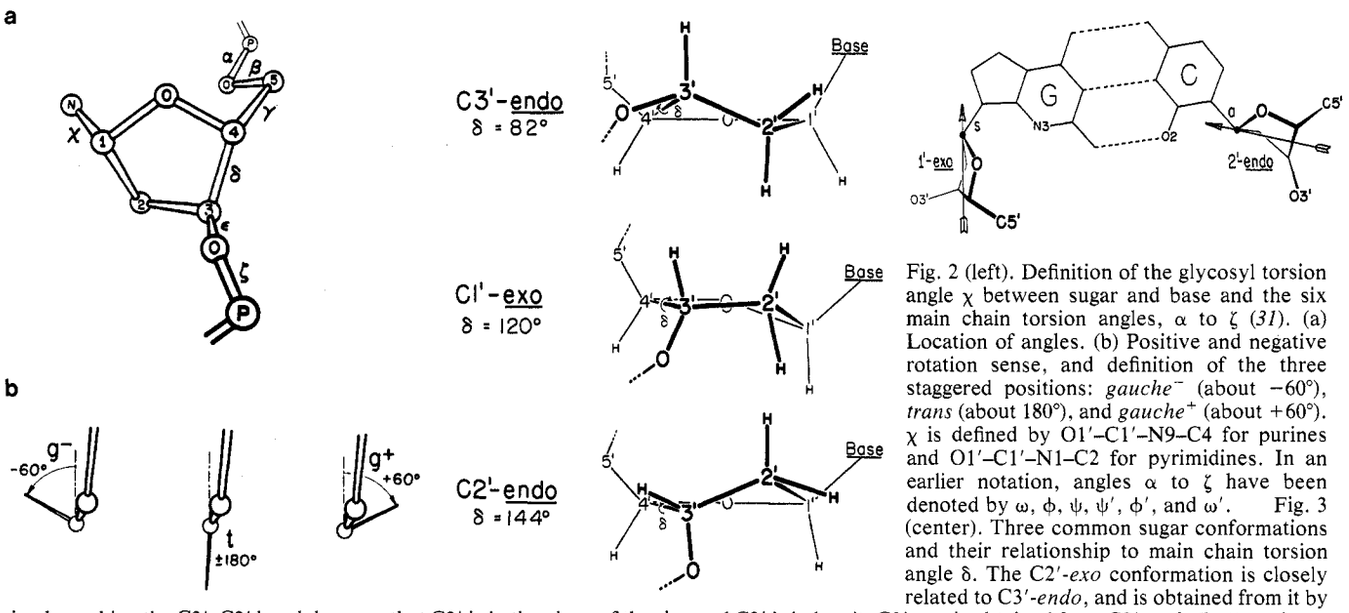


Fig. 2 (left). Definition of the glycosyl torsion angle χ between sugar and base and the six main chain torsion angles, α to ζ (*31*). (a) Location of angles. (b) Positive and negative rotation sense, and definition of the three staggered positions: *gauche*$^-$ (about −60°), *trans* (about 180°), and *gauche*$^+$ (about +60°). χ is defined by O1′–C1′–N9–C4 for purines and O1′–C1′–N1–C2 for pyrimidines. In an earlier notation, angles α to ζ have been denoted by ω, ϕ, ψ, ψ', ϕ', and ω'. Fig. 3 (center). Three common sugar conformations and their relationship to main chain torsion angle δ. The C2′-*exo* conformation is closely related to C3′-*endo*, and is obtained from it by simply pushing the C3′–C2′ bond down so that C3′ is in the plane of the ring and C2′ is below it. C3′-*exo* is obtained from C2′-*endo* (bottom drawing) in exactly the same way. Fig. 4 (right). *Syn* and *anti* conformations of the C1′–N glycosyl bond between sugar and base. The *s* at left marks a *syn* conformation, with the sugar ring turned toward the minor groove, and the *a* at right labels an *anti* conformation, with the sugar rotated away from the minor groove. Arrows indicate the direction of motion when the main chain torsion angle δ (O3′–C3′–C4′–C5′) is decreased, with the main chain pathway held more or less fixed. For *anti* the motion is a translation along the long axis of the base pairs, and for *syn* it is a more easily accomplished swinging out of the guanine end of the base pair toward the major groove. The *syn* conformation is found only in Z-DNA, and even there only at guanine positions.

Table 1. Double-helical DNA parameters from single-crystal analyses.

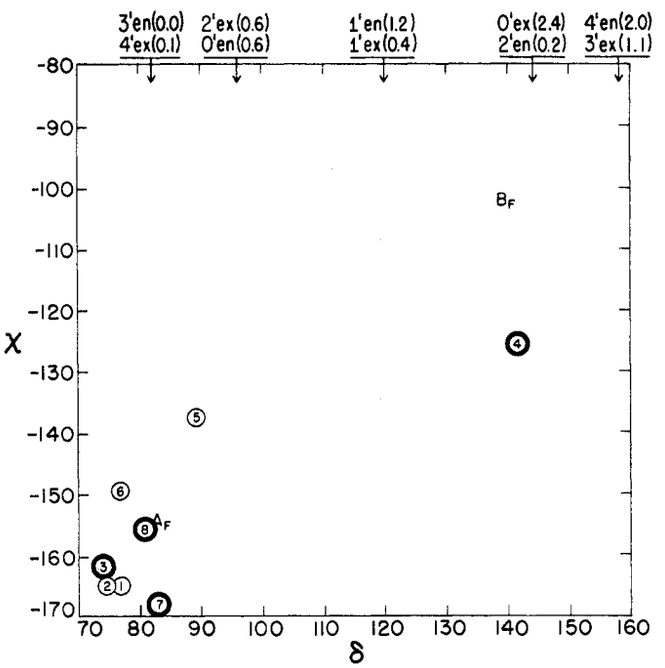
Property	Helix		
	A	B	Z
Example:	^{I}CCGG	CGCGAATTCGCG, CGCGAATT^{Br}CGCG	CGCG
Helix sense	Right-handed	Right-handed	Left-handed
Repeating helix unit	One base pair	One base pair	Two base pairs
Rotation per base pair	33.6°	t_t = 38.0° (4.4°)* t_g = 35.9° (4.2°)*	−60°/2
Mean base pairs per turn	10.7	10.0 (1.2)*	12
Inclination of base normals to helix axis	+19°	−1.2° (4.1°)	−9°
Rise per base pair along helix axis	2.3 Å	3.32 Å (0.19 Å)*	3.8 Å
Pitch per turn of helix	24.6 Å	33.2 Å	45.6 Å
Mean propeller twist	+18°	+16° (7°)*	~ 0°
Glycosyl angle conformation	*anti*	*anti*	*anti* at C, *syn* at G
Sugar pucker conformation (Figs. 5 to 7)	C3′-*endo*	O1′-*endo* to C2′-*endo*	C2′-*endo* at C, C2′-*exo* to C1′-*exo* at G
References	(*9, 10*)	(*9, 12–16, 22*)	(*18, 19*)

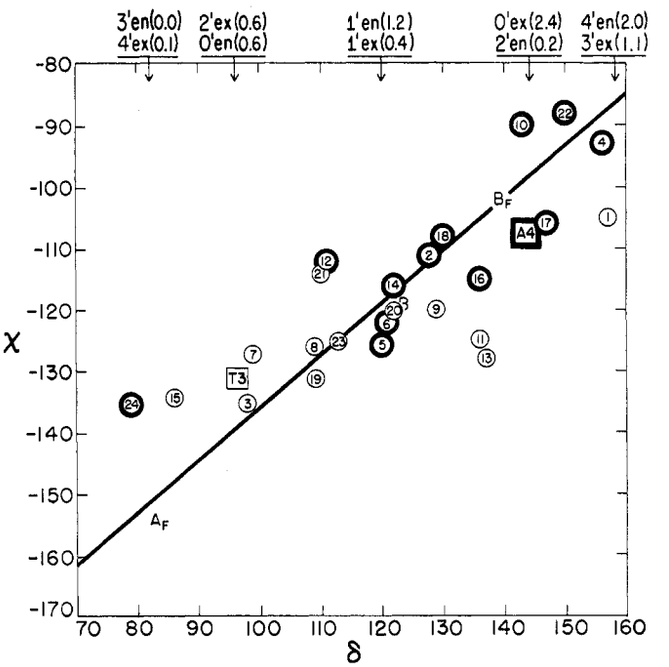
*Mean and standard deviation over 36 bases or 33 base steps in three independently refined dodecamers: CGCGAATTCGCG with bent helix axis, and CGCGAATT^{Br}CGCG under conditions in which its axis is straight and bent. The quantity t_g is the global twist angle as measured from outside the helix, whereas t_t is the local value considering the two base pairs in isolation. They differ because the local helix axis frequently deviates from the best overall axis [see figure 4 in (*14*)].

from observed experimental data, not theory, and this is so because the added five torsion angle degrees of freedom in DNA make it too complicated to carry out the simple steric clash analysis that was used originally for proteins. But as more DNA structures are added to these experimental conformation plots, they will gradually define boundaries of allowable zones that will lead us to principles of DNA helix structure.

DNA Hydration and the Transition of the B to the A Helix

From the very first fiber diffraction studies of DNA, it was apparent that water is an important factor in helix structure. Two characteristic x-ray fiber patterns are observed when stretched fibers are dried: the B pattern around 92 percent relative humidity (RH) and the A pattern when the fiber is dried to 75 percent RH in the absence of salt (*33*). Further drying below 55 percent RH leads to increasing disorder and deterioration of the quality of the diffraction pattern. The B form is stabilized by salt: a 10 percent salt content by weight is sufficient to prevent appearance of the A pattern no matter how low the RH, but with a salt content of 0.4 percent or less the A form can be observed to persist all the way up to 98 percent RH, although





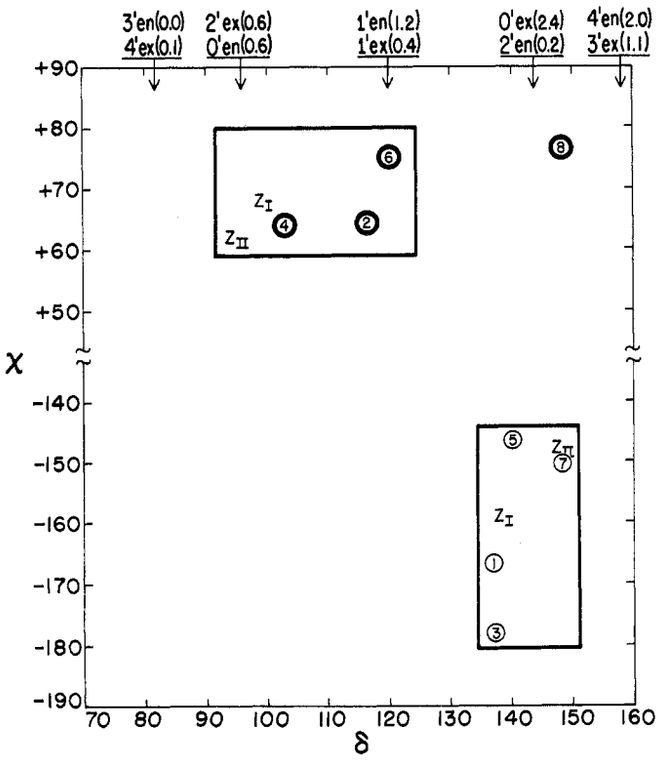


Fig. 5 (upper left). Conformation plot of χ as a function of δ for the A-DNA tetramer ^{1}CCGG (*9*, *10*). Bases are identified by number, and purines are distinguished by a dark ring. A_F and B_F mark the classical helix conformations as established by fiber diffraction. Sugar conformations associated with particular δ values are indicated along the top margin, along with estimates of their energy in kilocalorie per mole relative to C3′-*endo* (*32*). All of the points cluster around the classical A helix region except for guanine G4 at the right. But G4 is a 3′-terminal base, and hence its sugar is not subject to the steric constraints that would be experienced by a sugar ring in the middle of a continuous helix. (It also is in a region of tight nonhelical intermolecular packing within the crystal.) Fig. 6 (upper right). Conformation plot for χ as a function of δ for the B-DNA dodecamer CGCGAATTCGCG (*9*, *12–16*). Purines again are distinguished from pyrimidines by darker rings. This plot differs from that for A-DNA in four respects: (i) The distribution of conformations is broad, rather than being confined to the vicinity of point B_F, (ii) χ and δ are linearly correlated, (iii) purines have systematically higher χ and δ values than pyrimidines, and (iv) the two members of one base pair tend to adopt configurations that are equally far to either side of the midpoint of the plot. This latter feature has been termed the principle of anticorrelation (*13*). Squares marked T3 and A4 represent the unperturbed central base step of the intercalation complex of daunomycin with CGTACG (*23*). That B-helical structure, where not disturbed by intercalation, exhibits all four of the above features. Sugar conformations in B-DNA can range all the way from O1′-*endo* through C1′-*exo* to the classical C2′-*endo*. (Base G24 again is an anomalous configuration at a 3′ terminus.) Fig. 7 (lower left). Conformation plot of χ as a function of δ for high-salt Z-DNA CGCG (*18*, *19*) and for idealized model Z_I and Z_{II} helices based on the CGCGCG and low-salt CGCG structures (*21*). Stereo drawings of superpositions of the Z, Z_I, and Z_{II} helices are to be found in figures 1 and 2 of (*19*). Guanines are confined to the upper *syn* region, while cytosines are clustered to the lower right in the C2′-*endo* region. In Z-DNA, χ and δ are not correlated but are subject to separate constraints. Guanine G8 again is a lone 3′-terminal anomaly.

this probably is a kinetic or hysteresis effect rather than true equilibrium (*34*). Ethanol, isopropanol (the crystallizing medium for CCGG), and other alcohols favor the B-to-A transition (*35–37*). The transition is sharp and cooperative, occurring over a narrow range of humidity or alcohol concentration.

In a classic series of studies, Falk *et al.* (*38–40*) used gravimetric analysis and infrared and ultraviolet spectroscopy with fibers and films of sodium DNA and lithium DNA to completely reconstruct the course of events during the hydration of DNA—a scenario that our x-ray crystal structure analyses have generally confirmed. In this schema, the free phosphate oxygen atoms along the DNA backbone are the most tightly hydrated, with an average of two water molecules per phosphate, the P–O–C and C–O–C oxygens of phosphates and sugars are next, and the N and O atoms on edges of base pairs in major and minor grooves are the least strongly hydrated. From 0 percent to around 55 percent RH, the double helix initially is highly disordered, with imperfect base stacking, but an increasingly regular helix structure develops as the phosphate groups become hydrated. The phosphates are completely hydrated above 60 percent RH, and the now well-ordered A helix is stabilized by the acquisition of four or five water molecules per base pair, hydrating N and O atoms in the major and minor grooves. Above 75 percent RH, the grooves are fully hydrated with a monolayer of water, and further water accumulates as a less ordered or "liquid" zone filling the grooves. The A-to-B helix conversion takes place in a narrow humidity range, with midpoint between 75 and 83 percent RH (*37*). Beyond about 80 percent RH, the grooves are completely filled with liquid water so that further hydration causes the fibers to swell.

Falk's observations of fiber behavior have been confirmed in solution by buoyant density measurements and ultracentrifugation. Hearst and Vinograd (*41*) showed that 12 water molecules are associated with each base pair at 70 percent RH. (More precisely, at a water activity of $a_w = 0.70$, but the two are the same if deviations from ideality in water behavior can be neglected.) From comparisons of DNA of varying base composition, Tunis and Hearst (*42, 43*) established that A · T base pairs bind two more water molecules than G · C pairs do at $a_w = 0.70$, a conclusion that agrees with SCF (self-consistent field) ab initio theoretical calculations on hydration states of A · T and G · C base pairs (*44*).

Our single-crystal results on the hydration states of A-DNA (CCGG) and B-DNA (CGCGAATTCGCG at room temperature and 16 K, and its 9-bromo derivative) seem in retrospect to be simply pictorial realizations of the above-mentioned work, with the crucial difference that we now can see exactly where the water molecules are. The phosphate backbone of A-helical CCGG in 80 percent isopropanol is thoroughly hydrated (*10*). Each of the phosphate groups has associated with it two or three well-defined solvent peaks, and many of these first hydration shell ligands themselves are bridged by other solvent molecules. In particular, because the major groove opening is so narrow, a network of solvent molecules arches across the opening of this groove, connecting phosphates on opposite sides (Fig. 9). Base-edge N and O groups at the bottom of the groove are coated with a monolayer of ordered solvent, and an irregular network of solvent molecules fills the space between this and the opening of the

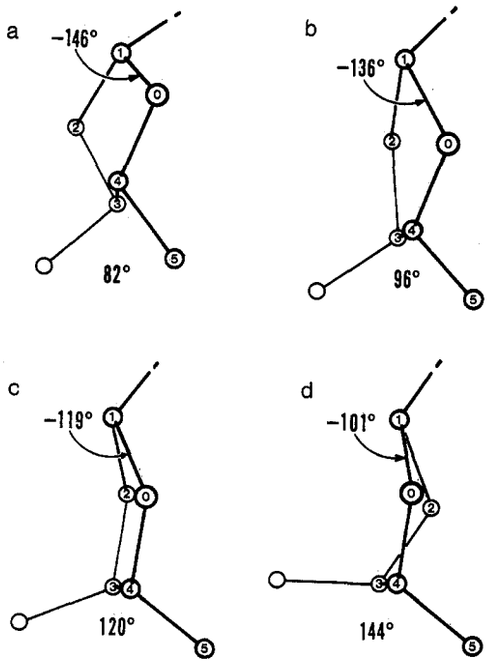


Fig. 8. The four most common sugar pucker conformations as seen approximately down the helix axis, perpendicular to the base planes. (a) C3′-*endo*, (b) O1′-*endo*, (c) C1′-*exo*, (d) C2′-*endo*. Each conformation is associated with a characteristic value of C5′–C4′–C3′–)3′ main chain torsion angle δ, shown below each diagram. Each conformation also is associated with a characteristic ideal glycosyl angle χ_o, the negative angle at the top left of each drawing. Atoms O1′ and C1′ through C5′ are numbered 0 and 1 through 5. [From (*22*)].

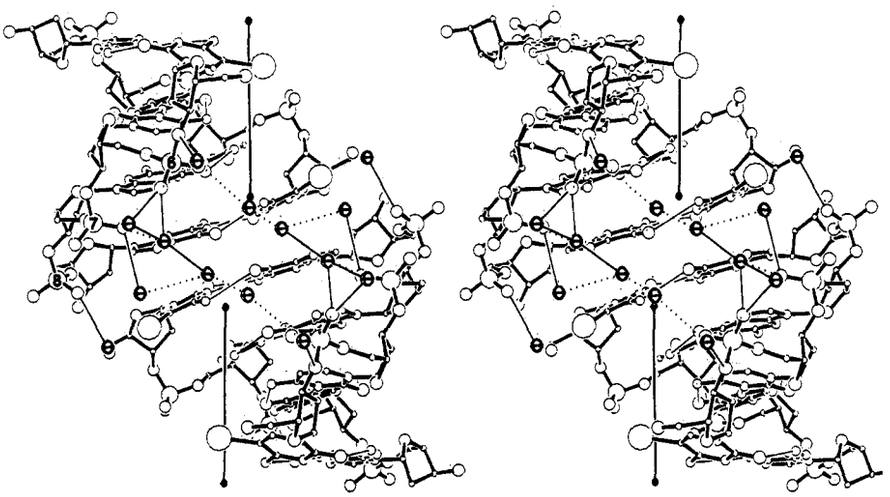


Fig. 9. Two double-stranded A-helical CCGG tetramers as they actually appear in the crystal. Their respective helix axes (vertical lines) are not quite coincident, being displaced from one another horizontally by 1.5 Å, and inclined at an angle of 8°. Even so, the two tetramers effectively form an A-helical octamer, which is barely long enough to begin defining the two edges of the major groove at phosphates P6 and P7 on the two molecules. (P6 through P8 are marked on the upper tetramer.) One consequence of the inclination of local axes between tetramers is an opening up of the major groove from its normal 7.8-Å phosphate separation as in Fig. 1, to 10.9 Å. In spite of this, the opposing phosphates P6 and P7 are laced together by a network of water molecules (crossed spheres). Thin lines are hydrogen bonds of 3.4 Å or less, and distances of 3.4 to 3.8 Å are dotted. [From (*10*); courtesy of *Nature (London)*]

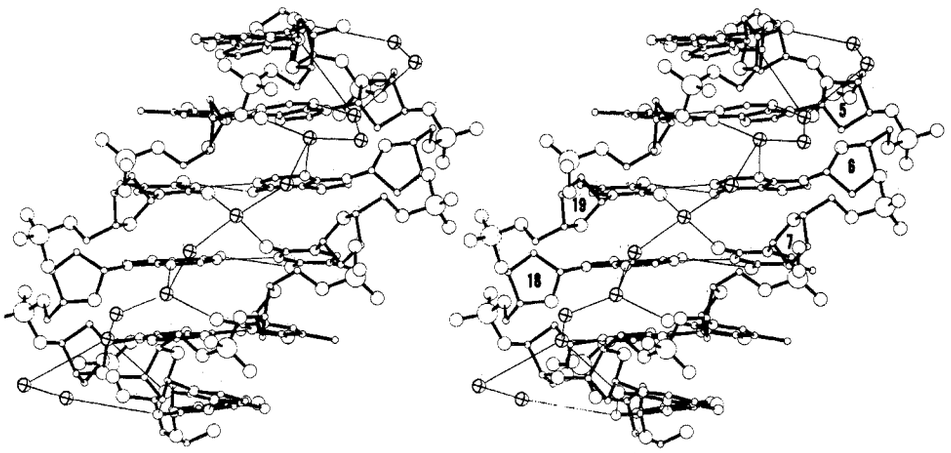


Fig. 10. The spine of hydration (crossed circles) down the minor groove of B-helical CGCGAATTCGCG. Water molecules bridge thymine O2 and adenine N3 atoms of bases on adjacent base steps that are brought closer together by the helical rotation. A second layer of solvation connects these first-shell waters and gives them a local tetrahedral coordination. This zigzag spine of hydration is disrupted by the N2 amino groups of guanines. [From (*15*); courtesy of *Journal of Molecular Biology*]

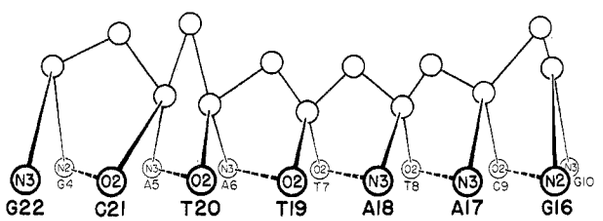


Fig. 11. The spine of hydration (open circles) as it would appear if the minor groove were unrolled from the helix and flattened. Circles with N3, N2, and O2 are atoms on the bases that are identified beneath. Dashed lines connect bases that are hydrogen-bonded in each base step. Water molecules bridge bases on adjacent steps. In the center, the hydration can reasonably be described as hydrogen bonding with tetrahedral water coordination, but at the two ends the spine pulls away from the floor of the minor groove, ending in two plumes of disordered water molecules (not shown). [From (*15*); courtesy of *Journal of Molecular Biology*].

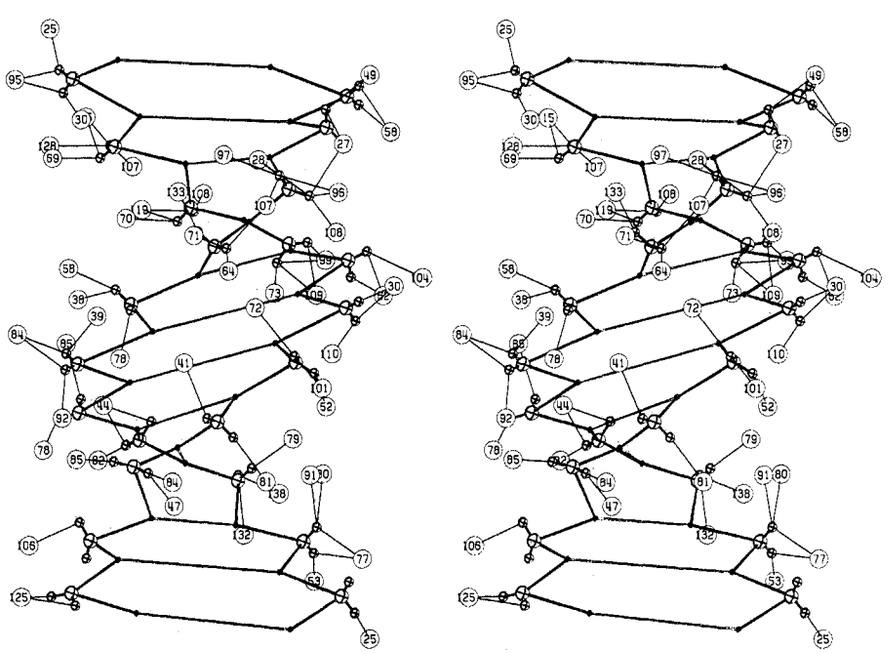


Fig. 12. Skeletal representation of the CGCGAATT^{Br}CGCG molecule in 60 percent MPD, showing hydration around the phosphate backbone. The DNA helix is represented only by PO_2 groups and C1′ sugar atoms, with straight lines connecting C1′ atoms to symbolize base pairs. Numbered circles are solvent peaks, and only those peaks are shown that lie within 4.25 Å of a free (unesterified) phosphate oxygen atom. Other solvent peaks not shown are associated with N and O atoms on base pair edges in major and minor grooves. Some of the peaks depicted here may well be magnesium ions rather than water molecules, as judged by their short bond distances: Nos. 96 (1.93 Å), 99 (2.16 Å), 28 (2.27 Å), 82 (2.30 Å), 73 (2.34 Å), and 64 (2.39 Å) are likely candidates. There is frequent recurrence of a pattern in which each free oxygen of a phosphate is hydrated, with another hydration peak shared between them to form a "W" of hydration, as with solvent peaks 30–95–25 at phosphate P2 at upper left, or peaks 39–84–92 at phosphate P19 at center left.

groove. In contrast, the minor groove is very dry, with only two of the accessible hydration sites along base edges occupied.

The B-DNA dodecamer structures show a different pattern of hydration. In CGCGAATTCGCG at room temperature (*15*), the N and O atoms on base edges within the major groove again have a uniform monolayer of hydration, with a looser network of second and higher hydration shell peaks, becoming disordered bulk water near the wide opening of the groove. But the minor groove has a distinctive and very regular two-layer hydration structure that we believe to be largely responsible for the stability of the B helix. The N3 atoms of adenine and the O2 of thymine on adjacent base pairs are bridged by first-shell solvent molecules, and these in turn are given a local tetrahedral environment by a bridging second hydration shell (Fig. 10). Together these make up a zigzag "spine" of hydration that lines the floor of the minor groove (Fig. 11). The groove is too narrow to continue an ice-like tetrahedral coordination of the second shell peaks, so the rest of the groove is filled with a less regular but well defined network, frequently connecting to phosphate oxygens along the rims of the groove. Nothing at all like this is seen in A-DNA, where the groove is so shallow as to be scarcely more than a helical indentation around the molecular surface.

The degree of localized, ordered hydration along the phosphate backbone of the B-DNA dodecamer depends on the particular crystal conditions. Native CGCGAATTCGCG in 35 percent MPD at room temperature has a high overall temperature factor of B equal to 36, with higher values for the phosphate backbone than for the base pairs. Discrete solvent peaks along the phosphate backbone generally are absent, except at three places where a water molecule is immobilized between a phosphate oxygen and a thymine methyl group. The hydrating molecules are invisible to the x-ray analysis, which can see only the

average structure over all molecules, and hence only the ordered solvent positions. But cooling the crystals to 16 K, or transferring CGCGAATT^{Br}CGCG crystals to 60 percent MPD, lowers the temperature factors and produces localized hydration along the backbone. In the high-MPD crystals, with an overall B = 16, 48 well-defined solvent peaks are associated with the 22 phosphates (Fig. 12), in addition to the minor groove spine and the major groove hydration layers already encountered. A typical pattern is for each of the two unesterified oxygen atoms on a phosphate to be associated with one hydrating water molecule, with a third molecule shared between them to form a "W" of hydration: water-oxygen-water-oxygen-water. Comparable behavior is observed in the low-MPD crystals at 16 K (*45*). The effect of high alcohol content in both A- and B-DNA, and ultra-low temperature in B-DNA, seems to be to freeze out domains of ordered water molecules that otherwise would be disordered or liquid.

The minor groove hydration spine present in all of the B dodecamers is defined best in the central AATT region, and begins to fray away from the bottom of the groove as it enters the CGCG regions at either end of the helix. At the very ends of the helix the minor groove is blocked by an interlocking overlap of molecules along a 2_1 screw symmetry axis (*12–16*), but the integrity of the spine begins to deteriorate before the blocked region of the groove is reached. The reason is not hard to see from the structure; the N2 amino groups of guanines (Fig. 13) intrude both physically and chemically: they get in the way of the regular pattern of bonding involving purine N3 and pyrimidine O2 atoms, and they introduce hydrogen atom donors where only hydrogen acceptors were before. The regular spine of hydration is a special attribute of AT-rich regions and is disrupted by guanines.

AT-rich DNA polymers have been known for years to behave differently from GC-rich polymers. Neither poly(dA–dT) nor poly(dI–dC), both of which lack the minor groove N2 amino group, undergo the salt-induced transformation from right-handed B helix to left-handed Z that is observed in poly(dG–dC) (*46, 47*). Similarly, poly(dI) · poly(dC), poly(dA) · poly(dT), and the light satellite crab DNA that is principally (dA) · (dT), all fail to adopt the A-helical structure at any point during dehydration. They remain in the B helix structure to unusually low relative hydrations, and then convert to a disordered semihelical structure

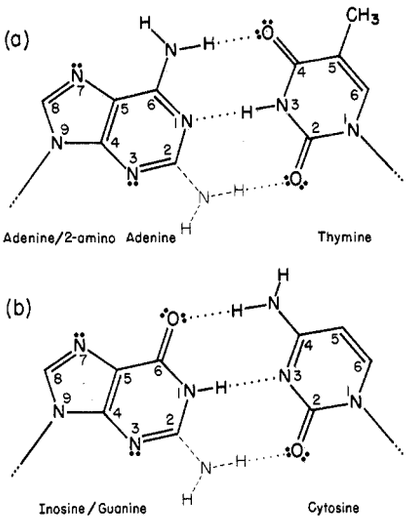


Fig. 13. Base pairs with and without the purine N2 amino group. In each case the base named to the left of the slash (adenine and inosine or hypoxanthine) lacks the N2 amino group, and that to the right of the slash (2-amino adenine and guanine) has it. B-DNA sequences with N2 amino groups are consistently easier to convert to the A structure than those without, and we attribute this to the disruptive effect of the amino group on the spine of hydration, which stabilizes the B structure. 2-Amino adenine, which has not yet been tested for its effects on helix structure, should behave like guanine.

when enough solvent molecules have been removed (*48*). Poly(dA–dT) only adopts the A conformation over a short range of hydration, 75 to 65 percent RH, and becomes disordered below this point. In systematic fiber diffraction studies of a great many synthetic DNA polymers of varied sequences, the data of Leslie and co-workers reveal that an almost 1:1 relation exists between the absence of N2 amino groups on purines (only A or I, not G), and the inability of the fiber to be driven from the B to the A conformation by dehydration [(*49*); see table 3 in (*15*) or table 1 in (*9*)].

On the basis of this evidence and our observations of hydration structure in the A and B helices, we propose that the minor groove spine is a major structural element in stabilizing the B form of DNA, and that the weakening and disruption of this spine is a necessary first step in the B-to-A helix transition. In this model, guanines favor the transition because they disrupt the spine. Part of the cooperativity of the B-to-A transition may occur because a break in the spine at one point tends to propagate in both directions when the unsecured ends of the spine fray away from the bottom of the groove more or less, as is observed at the borders of the AATT center of the dodecamer. During the transition from B-DNA to A-DNA, the minor groove loses both its deep profile and its ordered spine of water molecules, whereas the major groove retains most of its first hydration shell. The phosphates along the backbone, which are highly hydrated in B-DNA, remain so in A-DNA. But the proximity of the two backbone chains across the opening of the major groove in A-DNA leads to a lacing of water molecules across the opening, from phosphates on one chain to those on the other, as depicted in Fig. 9. When even these phosphate-hydrating water molecules are removed, then the order inherent in the helix structure is gradually destroyed.

A comparable analysis of the effect of hydration state on the B-to-Z transition cannot, at present, be made. Details of hydration for the low-salt Z helix structures CGCG and CGCGCG have not been published, and in the high-salt CGCG structure most of the potential hydration sites are co-opted by chloride ions, or by Mg^{2+} ions bridging phosphates on adjacent molecules (*18, 19*). Phosphate groups on different strands of the same helix come quite close to one another across the minor groove in Z-DNA (Fig. 1, e and f), and it has been suggested that high salt might favor the B-to-Z conversion because of cationic shielding of phosphate-phosphate repulsions in the Z form (*20*). But it was argued [in (*18*)] that the principal factor involved in the transition should be a lowered water activity rather than ionic shielding per se, since the same effect is brought about in poly(dG–dC) by an increase in salt concentration (which would lead to screening of phosphate repulsions) or alcohol concentration (which would exacerbate the phosphate-phosphate repulsions because of the change in dielectric constant of the medium). Of course, a high alcohol concentration might also favor formation of site-specifically bound cation complexes with phosphates, so in a sense both of these environmental factors, salt and alcohol, are interrelated. The recent discovery that a trivalent cation, cobalt hexammine, can selectively induce formation of Z-DNA (*50*) probably is an example of allosteric conversion by site-specific binding of a particular cation. In any case, the fact that high salt favors the transition of B to Z but not of B to A, although phosphate separations are drastically reduced in both transitions, is something that will have to wait for explanation until more data are available on specific cation locations in A, B, and Z helices.

Since the same AT-rich and IC-rich sequences that resist conversion from the B to the A helix also fail to transform to the Z structure even when a sequence alternation of purines and pyrimidines might seem to permit it, one is tempted to seek a common factor in the absence of minor groove N2 amino groups, and the presence of a stable hydration spine. It may be that this spine of water provides just enough free energy to keep the B helix from slipping over into either the A or Z helix under circumstances where, with a different sequence of bases, it would do so.

Prospects

Two subjects that have barely been mentioned in this discussion are the influence of base sequence on helix structure and the mechanics of bending of the DNA double helix. Several sequence-specific effects have been noted in the B-DNA dodecamer (*13, 14*). Comparable sequence effects on the A helix may be found when structures longer than the CCGG tetramer are examined. An even more interesting possibility, however, is that sequence-dependent structural variation might *not* be found in A-DNA. The structural uniformity of the A helix observed in the plot of χ as a function of δ (Fig. 5) may be a clue that the mechanical properties of this type of helix are not such as to be influenced by local variations in the base sequences. If these variations do play any role in recognition by proteins, then it may be that DNA evolved as the primary storage medium for genetic information rather than RNA, precisely because its structure was more malleable and more expressive of base sequence.

The bending properties of double-helical DNA are especially interesting because its winding around the nucleosome core in chromatin requires that it be able to wrap in an arc of radius 45 Å (*51, 52*). Little is known about the detailed bending behavior of B-DNA; proposals have been made that this chromatin bending is smooth and continuous (*53, 54*), or that it involves helix breaks or kinks every 20 (*55*), 10 (*56*), or 5 base pairs (*57*). The 19° curvature in the helix backbone of the CGCGAATTCGCG dodecamer corresponds to a radius of curvature of 112 Å, roughly twice that of the nucleosome core, and an energy of bending of approximately 0.25 to 0.50 kcal per mole of dodecamer (*12*). The fact that a combination of bromination at the ninth position and 60 percent MPD induces the helix to straighten out means that we can examine in detail what happens at each base pair when a helix bends and straightens. This will be reported later (*22*), but the best way of describing what we believe we see is to term the bending "annealed kinking," or kinking of the helix backbone without unstacking of base pairs. The requirement for 5-bromocytosine at the ninth base position for unbending to occur in this dodecamer is interesting in view of the suggestions that its close analog, 5-methylcytosine, may be involved in several eukaryotic gene control functions (*58–61*). Can any of the recognition properties of 5-methylcytosine be ascribed to its effect on the mechanical properties of the double-helical DNA in which it is found?

These analyses of the structure and properties of double-helical DNA are only one of the areas opened up by improvements in oligonucleotide synthesis. The binding to DNA of antitumor drugs such as daunomycin (*23*) and cisplatin [*cis*-dichlorodiamminoplatinum-(II)] and its *trans*- and tetraammino analogs (*62*) illustrates an approach that has great medical as well as theoretical possibilities. Although structures now are known for three proteins that recognize and bind specifically to particular DNA sequences—catabolite activator protein (*63*) and the *cro* (*64*) and lambda (*65*) repressors, no structure has yet been determined of the complex of such a recognition protein with its DNA. It is hoped that this situation can be changed soon for this catabolite activator protein, and for other examples also. Triester synthesis methods have opened up an entirely new field of research, and the coming two decades should be as informative for nucleic acid structure analysis as the previous two have been for proteins.

References and Notes

1. J. C. Kendrew, R. E. Dickerson, B. E. Strandbery, R. G. Hart, D. R. Davies, D. C. Phillips, V. C. Shore, *Nature (London)* **185**, 422 (1960); M. F. Perutz, M. G. Rossmann, A. F. Cullis, H. Muirhead, G. Will, A. C. T. North, *ibid.*, p. 416.
2. C. C. F. Blake, D. R. Koenig, G. A. Mair, A. C. T. North, D. C. Phillips, V. R. Sarma, *ibid.* **206**, 757 (1965).
3. J. D. Watson and F. H. C. Crick, *ibid.* **171**, 737 (1953).
4. R. Langridge, D. A. Marvin, W. E. Seeds, H. R. Wilson, C. W. Hooper, M. H. F. Wilkins, L. D. Hamilton, *J. Mol. Biol.* **2**, 38 (1960).
5. S. Arnott, in *Organization and Expression of Chromosomes*, V. G. Allfrey, E. K. F. Bautz, B. J. McCarthy, R. T. Schimke, A. Tissieres, Eds. (Dahlem Conference, Berlin, 1976), pp. 209–222.
6. K. Itakura, N. Katagiri, C. P. Bahl, R. H. Wightman, S. A. Narang, *J. Am. Chem. Soc.* **97**, 7327 (1975).
7. M. A. Viswamitra, O. Kennard, P. G. Jones, G. M. Sheldrick, S. Salisbury, L. Favello, Z. Shakked, *Nature (London)* **273**, 687 (1978).
8. Abbreviations for the bases are A, adenine; C, cytosine; G, guanine; T, thymine; and U, uracil. The corresponding nucleosides are adenosine, cytidine, guanosine, thymidine, uridine. dA, Deoxyadenosine; dT, deoxythymidine; poly(A), polyadenosine; poly(dA), polydeoxyadenosine; poly(dI), polydeoxyinosine; poly(dA) · poly(dT), a double-stranded polymer; poly(dA–dT), an alternating polymer [see *Eur. J. Biochem.* **15**, 206 (1982)]. The symbols C1, N9, O2, and the like specify the position of the atom in the base molecule; the primes indicate the numbering in the sugar.
9. R. E. Dickerson, H. R. Drew, B. N. Conner, in *Biomolecular Stereodynamics*, R. H. Sarma, Ed. (Adenine, New York, 1981), vol. 1, pp. 1–34.
10. B. N. Conner, T. Takano, S. Tanaka, K. Itakura, R. E. Dickerson, *Nature (London)* **295**, 294 (1982).
11. Z. Shakked, D. Rabinovich, W. B. T. Cruse, E. Egert, O. Kennard, G. Sala, S. A. Salisbury, M. A. Viswamitra, *Proc. R. Soc. London Ser. B* **213**, 479 (1981).
12. R. M. Wing, H. R. Drew, T. Takano, C. Broka, S. Tanake, K. Itakura, R. E. Dickerson, *Nature (London)* **287**, 755 (1980).
13. H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, R. E. Dickerson, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 2179 (1981).
14. R. E. Dickerson and H. R. Drew, *J. Mol. Biol.* **149**, 751 (1981).
15. H. R. Drew and R. E. Dickerson, *ibid.* **151**, 535 (1981).
16. R. E. Dickerson and H. R. Drew, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7318 (1981).
17. J. L. Crawford *et al.*, *ibid.* **77**, 4016 (1980).
18. H. R. Drew, T. Takano, S. Tanaka, K. Itakura, R. E. Dickerson, *Nature (London)* **286**, 567 (1980).
19. H. R. Drew and R. E. Dickerson, *J. Mol. Biol.* **152**, 723 (1981).
20. A. H.-J. Wang, G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel, A. Rich, *Nature (London)* **282**, 680 (1979).
21. A. H.-J. Wang, G. J. Quigley, F. J. Kolpak, G. van der Marel, J. H. van Boom, A. Rich, *Science* **211**, 171 (1981).
22. A. V. Fratini, M. L. Kopka, H. R. Drew, R. E. Dickerson, in preparation.
23. G. J. Quigley, A. H.-J. Wang, G. Ughetto, G. van der Marel, J. H. van Boom, A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 7204 (1980).
24. R. E. Dickerson, in *The Proteins*, H. Neurath, Ed. (Academic Press, New York, ed. 2, 1964), vol. 2, pp. 603–678.
25. G. P. Lomonossoff, P. J. G. Butler, A. Klug, *J. Mol. Biol.* **249**, 745 (1981).
26. L. J. Peck and J. C. Wang, *Nature (London)* **292**, 375 (1981).
27. D. Rhodes and A. Klug, *ibid.*, p. 378.
28. D. J. Patel *et al.*, *Biochemistry* **21**, 428 (1981); D. J. Patel, A. Pardi, K. Itakura, *Science*, in press.
29. C. Ramakrishnan and G. N. Ramachandran, *Biophys. J.* **5**, 909 (1965).
30. R. E. Dickerson and I. Geis, *The Structure and Action of Proteins* (Benjamin, Menlo Park, Calif., 1968).
31. C. Altona *et al.*, "Abbreviations and Symbols for the Description of Conformations of Polynucleotide Chains," IUB-IUPAC Joint Commission on Biochemical Nomenclature, Discussion Draft, February 1981.
32. M. Levitt and A. Warshel, *J. Am. Chem. Soc.* **100**, 2607 (1978).
33. M. H. F. Wilkins, *Science* **140**, 941 (1963).
34. J. Texter, *Prog. Biophys. Mol. Biol.* **33**, 83 (1978).
35. V. I. Ivanov, L. E. Minchenkova, A. K. Schyolkina, A. I. Poletayev, *Biopolymers* **12**, 89 (1973).
36. V. I. Ivanov, L. E. Minchenkova, E. E. Minyat, M. D. Frank-Kamenetskii, A. K. Schyolkina, *J. Mol. Biol.* **87**, 817 (1974).
37. G. Malenkov, L. Minchenkova, E. Minyat, A. Schyolkina, V. Ivanov, *FEBS Lett.* **51**, 38 (1975).
38. M. Falk, K. A. Hartman, Jr., R. C. Lord, *J. Am. Chem. Soc.* **84**, 3843 (1962).
39. ———, *ibid.* **85**, 387 (1963).
40. ———, *ibid.*, p. 391.
41. J. E. Hearst and J. Vinograd, *Proc. Natl. Acad. Sci. U.S.A.* **47**, 1005 (1961).
42. M.-J. B. Tunis and J. E. Hearst, *Biopolymers* **6**, 1325 (1968).
43. ———, *ibid.*, p. 1345.
44. A. Goldblum, D. Perahia, A. Pullman, *FEBS Lett.* **91**, 213 (1978).
45. H. R. Drew, S. Samson, R. E. Dickerson, *Proc. Natl. Acad. Sci., U.S.A.*, in press.
46. F. M. Pohl and T. M. Jovin, *J. Mol. Biol.* **67**, 375 (1972).
47. D. J. Patel, L. L. Canuel, R. M. Pohl, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 2508 (1979).

48. J. Pilet, J. Blicharski, J. Brahms, *Biochemistry* **14**, 1869 (1975).
49. A. G. W. Leslie, S. Arnott, R. Chandrasekaran, R. L. Ratliff, *J. Mol. Biol.* **143**, 49 (1980).
50. M. Behe and G. Felsenfeld, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 1619 (1981).
51. J. T. Finch, L. C. Lutter, D. Rhodes, R. S. Brown, B. Rushton, M. Levitt, A. Klug, *Nature (London)* **269**, 29 (1977).
52. R. D. Kornberg and A. Klug, *Sci. Am.* **244**, 52 (February 1981).
53. M. Levitt, *Proc. Natl. Acad. Sci. U.S.A.* **75**, 640 (1978).
54. J. L. Sussman and E. N. Trifonov, *ibid.*, p. 103.
55. F. H. C. Crick and A. Klug, *Nature (London)* **255**, 530 (1975).
56. H. M. Sobell, C. Tsai, S. G. Gilbert, S. C. Jain, T. D. Sakore, *Proc. Natl. Acad. Sci. U.S.A.* **73**, 3068 (1976).
57. V. B. Zhurkin, Y. P. Lysov, V. I. Ivanov, *Nucleic Acids Res.* **6**, 1081 (1979).
58. A. Razin and A. D. Riggs, *Science* **210**, 604 (1980).
59. M. Ehrlich and R. Y.-H. Wang, *ibid.* **212**, 1350 (1981).
60. M. Groudine, R. Eisenman, H. Weintraub, *Nature (London)* **292**, 311 (1981).
61. T. Mohandas, R. S. Sparkes, L. J. Shapiro, *Science* **211**, 393 (1981).
62. R. M. Wing, P. Pjura, H. R. Drew, R. E. Dickerson, in preparation.
63. D. B. McKay and T. A. Steitz, *Nature (London)* **290**, 744 (1981).
64. W. F. Anderson, D. H. Ohlendorf, Y. Takeda, B. W. Matthews, *ibid.*, p. 754.
65. C. O. Pabo, personal communication.
66. We thank Keiichi Itakura of the City of Hope National Medical Center, who synthesized DNA oligomers and trained members of his group and ours in triester synthesis methods (especially Peter Dembek, Shoji Tanaka, and Chris Broka). This project evolved from an earlier one whose goal was to solve the structure of the *lac* repressor-operator complex (which we may yet manage). We thank S. Narang of the National Research Council in Ottawa and R. Wu of Cornell University for help in getting the original project going; A. D. Riggs of the City of Hope Research Laboratory, J. M. Rosenberg, now of the University of Pittsburgh, and O. B. Kallai for constant encouragement, ideas, and hard work during the earlier repressor-operator phase of the project; T. Takano for help with refinement methodology, and L. Casler for the line drawings. Supported by NIH grants GM-12121 and GM-24393, NSF grant PCM79-13959, a grant from the Upjohn Company, two NIH predoctoral traineeships (H.D. and B.C.), and a fellowship (DRG-507) from the Damon Runyon–Walter Winchell Cancer Fund (H.D.). This is contribution No. 6506 from the Division of Chemistry and Chemical Engineering, California Institute of Technology.

Steel Recycling and Energy Conservation

Bruce Hannon and James R. Brodrick

Summary. The potential for energy conservation through increased use of steel scrap by the U.S. steel industry is examined. It is concluded that increased use of scrap would reduce energy use, but it is not economical, due mainly to volatile scrap prices. Other energy-saving technologies exist, but it is likely that energy will be conserved through reduced use of steel as rising energy costs are passed through to consumers.

In this article we discuss the potential for energy conservation through increased recycling in the U.S. steel industry. The industry, which plays a fundamental role in the U.S. economy, is regarded as very energy intensive. As international competition for metallurgical grade coal and public concern about waste have grown, the increased use of scrap steel to save energy in the steel industry has become an important topic. To our knowledge, this subject has not previously been studied in a comprehensive way; no research results are available in the open literature that quantify the total energy impact on the economy of shifts between the various steelmaking technologies or changes in the amount of recycling.

We calculate the historic energy intensity of the U.S. steel industry and the intensity under a variety of technological mixes needed to handle significant increases in recycling. The energy intensity is the direct and indirect energy required per ton of steel produced by the industry. Direct energy is defined as all the energy used on site by the steel manufacturer per ton of finished steel produced. Indirect energy is the energy used elsewhere in the United States and the world to furnish the manufactured goods, energy, and services used by the steelmaker per ton of finished steel. The total energy intensity is the sum of the direct and indirect energy intensities. After calculating the decreases in total energy intensity that result from increased steel scrap recycling, we determine the dollar cost of saving this energy and compare this dollar cost with the marginal costs of energy from new sources.

Our results show that most investigators have greatly understated the total energy cost of finished steel, and that two different types of furnaces used in steelmaking—the basic oxygen and open hearth furnaces—require essentially the same amount of energy per ton of finished steel. We find that the principal way to reduce energy use in steelmaking is to use more scrap, but even if all the steel were made from scrap the total energy intensity of steel would be reduced by only about 6 percent, saving less than 1 percent of total annual U.S. energy use.

An economic analysis of the cost of saving this energy indicates that it is very unlikely that rising energy prices will encourage significant energy savings in the steel industry. However, there may be other reasons for the industry to increase its use of scrap, such as relative rises in labor costs in the iron and iron ore industries and relatively high ore taxes. Several energy-conserving alternatives to scrap recycling exist, as we discuss in the concluding section of this article. However, extensive new capital investment will be required in the steel industry to reduce energy use and to keep the rise in the cost of production at or below the inflation rate.

Scrap and the Steelmaking Processes

Molten iron and iron and steel scrap are mixed with special additives in the steelmaking processes, which are of three basic types: the open hearth furnace, the basic oxygen furnace, and the electric arc furnace. The steel mills of the 1960's employed a mixture of all three technologies, but in recent years the open hearth process has been largely

Bruce Hannon is associate professor in geography and director of the Energy Research Group, Office of Vice Chancellor for Research, University of Illinois at Urbana-Champaign, Urbana 61801. James R. Brodrick is an engineer with Carrier Corporation, Research Division, Syracuse, New York 13221.

0036-8075/82/0430-0485$01.00/0