International Tables for Crystallography (2006). Vol. F, Section 23.3.2, pp. 588–596.

23.3. Nucleic acids

BY R. E. DICKERSON

23.3.1. Introduction

In 1953, James Watson and Francis Crick solved the structure of double-helical DNA (Watson & Crick, 1953; Crick & Watson, 1954). So what has a dedicated cadre of X-ray crystallographers been doing for the subsequent 45 years? That is the subject of this chapter: the advance of our knowledge of nucleic acid duplexes, primarily from single-crystal X-ray diffraction, and the biological implications of this new knowledge. The focus will be primarily on DNA because much more is known about it, but DNA/RNA hybrids and duplex RNA will also be considered. Because the emphasis is on the geometry of the nucleic acid double helix, exotic structures, such as quadruplexes, hammerhead ribozymes and aptamers, will be omitted, as will larger-scale structures such as tRNA.

Fibre diffraction showed that there were two basic forms of DNA duplex: the common B form and a more highly crystalline A form (Fig. 23.3.1.1) that, in some but not all sequences, could be produced by dehydrating the fibre (Franklin & Gosling, 1953; Langridge *et al.*, 1960; Arnott, 1970; Leslie *et al.*, 1980). A- and B-DNA are contrasted in Figs. 23.3.1.2 and 23.3.1.3. The high-humidity B form has base pairs sitting squarely on the helix axis and roughly perpendicular to that axis. In contrast, in the low-humidity A form, the base pairs are displaced off the helix axis by *ca* 4 Å and are inclined 10–20° away from perpendicularity to that axis. The two grooves in B-DNA are of comparable depth because base pairs sit *on* the helix axis, but the major groove is wider than the minor because of asymmetry of attachment of base pairs to the backbone chains. In A-DNA, the minor groove is broad and shallow, whereas the major groove is cavernously deep (all the way from the surface of the helix, to the helix axis, and beyond) but can be quite narrow.

Pohl and co-workers had shown in the 1970s that alternating poly(dC-dG) is special in that it undergoes a reversible salt- or alcohol-induced conformation change (Pohl & Jovin, 1972; Pohl, 1976). Hence, it was not surprising that when DNA synthesis methods advanced to the stage where oligonucleotide crystallization became feasible, two separate research groups – those of Alexander Rich at MIT and Richard Dickerson at Caltech – elected to synthesize, crystallize and solve a short, alternating C-G oligomer. The result was a third family of DNA duplexes, Z-DNA (Fig. 23.3.1.4), first as the hexamer C-G-C-G-C-G (Z1) and then the tetramer C-G-C-G (Z3). (References to A-, B- and Z-DNA structures are listed at the end of Tables A23.3.1.1, A23.3.1.2 and A23.3.1.3 in the Appendix, respectively. They are cited by numbers beginning with A, B or Z.) Single-crystal analyses of the traditional helix types soon followed: B-DNA as C-G-C-G-A-A-T-T-C-G-C-G (B1), and A-DNA as both C-C-G-G (A1) and G-G-T-A-T-A-C-C (A2).

23.3.2. Helix parameters

23.3.2.1. *Backbone geometry*

Before making detailed comparisons of the three helix types, one must define the parameters by which the helices are characterized. The fundamental feature of all varieties of nucleic acid double helices is two antiparallel sugar–phosphate backbone chains, bridged by paired bases like rungs in a ladder (Fig. 23.3.2.1). Using the convention that the positive direction of a backbone chain is from 5′ to 3′ within a nucleotide, the right-hand chain in Fig. 23.3.2.1 runs downward, while the left-hand chain runs upward. A- or B-DNA is then obtained by twisting the ladder into a right-handed helix. But Z-DNA cannot be obtained from Fig. 23.3.2.1 simply by giving it a left-handed twist; both backbone chains run in the wrong direction for Z-DNA. A more complex adjustment is required, and this will be addressed again later.

The conformation of the backbone chain along each nucleotide is described by six torsion angles, labelled α through ζ, as shown in Fig. 23.3.2.2. An earlier convention termed these same six angles as ω, φ, ψ, ψ', φ', ω' (Sundaralingam, 1975), but the alphabetical nomenclature is now generally employed. Torsion angles are defined in Fig. 23.3.2.3, which also shows three common configurations: *gauche*$^-$ (−60°), *trans* (180°) and *gauche*$^+$ (+60°). These three configurations are especially favoured with sp^3 hybridization or tetrahedral ligand geometry at the two ends of the bond in question, because their 'staggered' arrangement minimizes ligand–ligand interactions across the bond. An 'eclipsed' arrangement with ligands at −120°, 0° (*cis*), and 120° is unfavourable because it brings substituents at the two ends of the bond into opposition. Table 23.3.2.1 lists the mean values and standard deviations of all six main-chain torsion angles for A-, B- and Z-DNA, as recently observed in 96 oligonucleotide crystal structures (Schneider *et al.*, 1997).

23.3.2.2. *Sugar ring conformations*

The type of ligand–ligand clash just mentioned is an important element in ensuring that five-membered rings, such as ribose and deoxyribose, are not ordinarily planar, even though the internal bond angle of a regular pentagon, 108°, is close to the 109.5° of tetrahedral geometry. A stable compromise is for one of the four ring atoms to lie out of the plane defined by the other four, as in Fig. 23.3.2.4. This is termed an 'envelope' or E conformation, by analogy with a four-cornered envelope having a flap at an angle. Intermediate 'twist' or T forms are also possible, in which two adjacent atoms sit on either side of the plane defined by the other three, but this discussion will focus on the simple envelope conformations. In most cases, the accuracy of a nucleic acid crystal structure determination is such that it would be difficult to distinguish clearly between a given E form and its flanking T forms. For this reason, most structure reports consider only the E alternatives.

A convenient and intuitive nomenclature is to name the conformation after the out-of-plane atom and then specify whether it is out of plane on the same side as the C5′ atom (*endo*) or the opposite side (*exo*). Ten such conformations exist: five *endo* and

This chapter is dedicated to Irving Geis, who died on 22 July 1997 at the age of 88, just as the chapter was begun. Irv was a pioneer in the representation of protein and DNA structures, beginning with illustrations for *Scientific American* articles on myoglobin (Kendrew, 1961), lysozyme (Phillips, 1966), cytochrome *c* (Dickerson, 1972) and DNA (Dickerson, 1983). He was coauthor with the present writer of *Structure and Action of Proteins* (Dickerson & Geis, 1969) and two later textbooks (Dickerson & Geis, 1976, 1983) and contributed drawings and paintings to a great number of other books and articles, most notably Voet & Voet's *Biochemistry* (Voet & Voet, 1990, 1995), which is a veritable gallery of Irv's art. His meticulous and carefully thought-out diagrams and drawings of myoglobin and haemoglobins have never been matched. More information about his life, work and art may be found in three articles by the present author (Dickerson, 1997*a*,*b*,*c*). Irv saw his role as one of bringing an understanding of protein structure to life scientists and sometimes referred to himself half-humorously as 'the Andreas Vesalius of molecular anatomy'. In view of the formative influence that his art exerted on the first generation of protein crystallographers and molecular biologists, it is more appropriate to remember Irv as the Leonardo da Vinci of macromolecules. As of late 2000, nearly all of Irving Geis' work – paintings, drawings, illustrations and correspondence – is being preserved for study as the Geis Archives at the Howard Hughes Medical Institute, Washington DC.

five *exo*. In Fig. 23.3.2.4 (top), pushing the C3′ atom of the C3′-*endo* conformation into the plane of the ring would tend to push C2′ below the ring, passing through a T state and creating a C2′-*exo* conformation. C2′ can, in turn, be returned to the ring plane if C1′ is pushed above the ring, forming C1′-*endo*, and so on, around the ring. In this way, a contiguous series of alternating *endo/exo* conformations is produced, as listed in Table 23.3.2.2.

This ten-conformation *endo/exo* cycle can be generalized to a continuous distribution of intermediate conformations, characterized by a pseudorotation angle, *P* (Altona *et al.*, 1968; Altona & Sundaralingam, 1972), with the ten *endo/exo* conformations spaced 36° apart (Table 23.3.2.2). Fig. 23.3.2.5 shows the calculated potential energy of conformations around the pseudorotation cycle (Levitt & Warshel, 1978). Note that C2′-*endo* and C3′-*endo* are most stable, that the pathway between them along the right half of the circle remains one of low energy, but that a large 6 kcal mol^{-1} potential energy barrier (1 kcal mol^{-1} = 4.184 kJ mol^{-1}) effectively forbids conformations around the left half of the circle.

As Fig. 23.3.2.4 indicates, the main-chain torsion angle, δ, is sensitive to ring conformation, because the C5′—C4′ and C3′—O3′ bonds that define the angle shift as ring puckering changes. The idealized relationship between torsion angle, δ, and pseudorotation angle, *P* (Saenger, 1984), is

$$\delta = 40^\circ \cos(P + 144^\circ) + 120^\circ.$$

Fig. 23.3.2.6 shows the observed torsion angles, δ, and pseudorotation angles, *P*, from X-ray crystal structure analyses of synthetic DNA oligonucleotides: 296 examples from A-DNA and 280 from B-DNA. The most striking aspect of this plot is the radically different behaviour of A- and B-DNA. The prototypical sugar conformation for A-DNA obtained from fibre diffraction modelling, C3′-*endo*, is, in fact, adhered to quite closely in A-DNA crystal structures.

However, B-DNA shows a quite different behaviour. Although earlier fibre diffraction led one to expect C2′-*endo* sugars, the actual experimental distribution is quite broad, extending up the right-hand side of the pseudorotation circle of Fig. 23.3.2.5, through C1′-*exo*, O1′-*endo* and C4′-*exo*, in some cases all the way to C3′-*endo* itself. Indeed, the mean value of δ observed in B-DNA oligomer crystal structures is 128° rather than 144° (Table 23.3.2.1), making C1′-*exo* a better description of sugar conformation in B-DNA than C2′-*endo*. Old habits die hard, however, and the B-DNA sugar conformation is still colloquially termed C2′-*endo*, a designation of historical significance but of little practical value. The apparent greater malleability of the B helix compared to A may indeed be one feature that makes B-DNA particularly suitable for expressing its base sequence to drugs and control proteins *via* local helix structure changes.

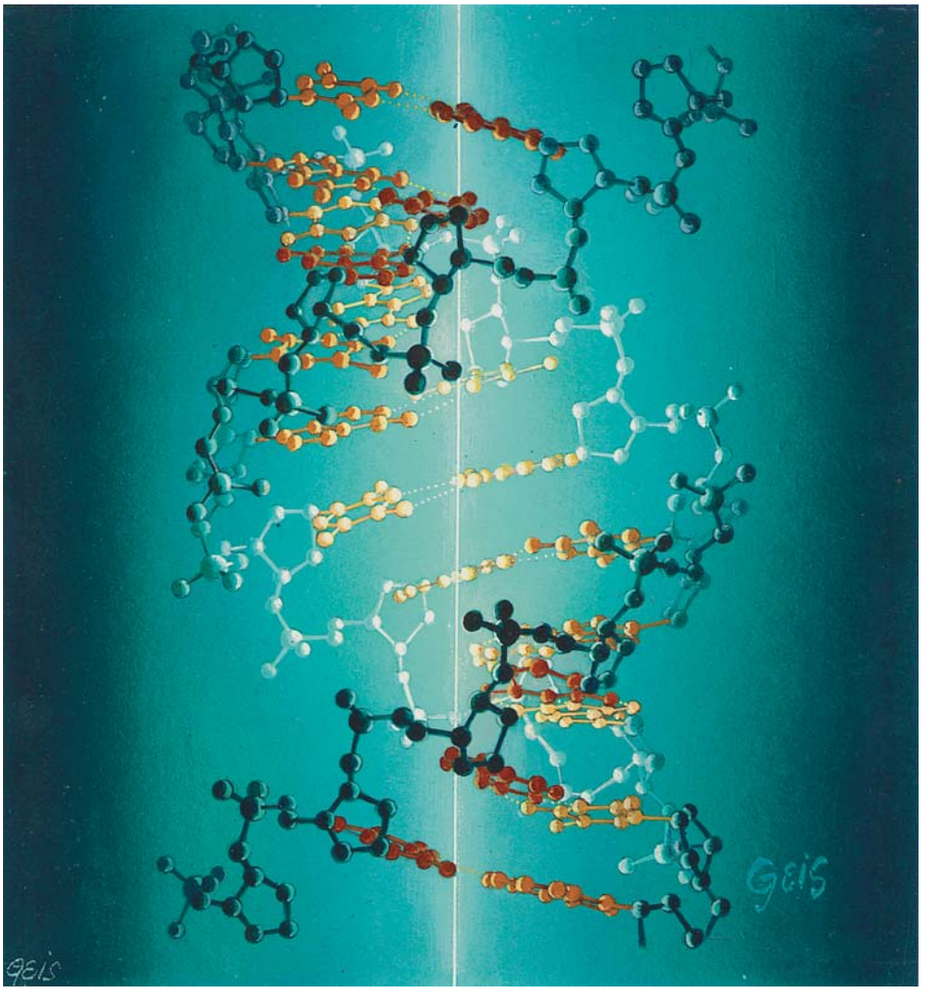

Fig. 23.3.1.1. 'Hot wire' painting of A-DNA by Irving Geis. Geis produced two dramatic paintings of horse-heart cytochrome *c*, in which the sole light source was the central iron atom within the haem, producing a glowing 'molecular lantern' effect. One painting showed this central luminous haem surrounded by hydrophobic side chains; the other featured the polar side chains extending out from the surface. These are to be seen today on the front and back covers of Voet & Voet's *Biochemistry* (Voet & Voet, 1990, 1995). In the present A-DNA painting, Geis chose the imaginary central axis of the helix as a monofilament light source, thereby reversing the conventional illumination: atoms lining the deep major groove glow brightly, whereas the outer surface of the helix is in dark silhouette. Geis struggled with the B helix as an artistic subject, but was never satisfied with the results. Hence, this glowing A-DNA helix represents his nucleic acid artistic legacy. Reprinted courtesy of the estate of Irving Geis. Rights owned by Howard Hughes Medical Institute.

23.3.2.3. *Base pairing*

The key to the biological role of DNA is that one of the two purines can pair with only one of the pyrimidines: A with T, and G with C. Hence, genetic information present in one strand is passed on to the complementary strand. The standard two-base pairs are shown in Fig. 23.3.2.7 along with the conventional numbering of the atoms. Backbone sugar and phosphate atoms are primed while base atoms are unprimed, as, for example, C1′ and N9 at opposite ends of a purine glycosydic bond. The G·C base pair is held together by three hydrogen bonds, whereas an A·T pair has only two. This means that A·T pairs show less resistance to propeller twisting (counter-rotation of the two bases about their common long axis), and this will have an effect on minor groove width, as seen later. The patterns of hydrogen-bond acceptors (*A*) and donors (*D*) on the major and minor groove edges of base pairs are important elements in recognition of base sequence by drugs and control proteins.

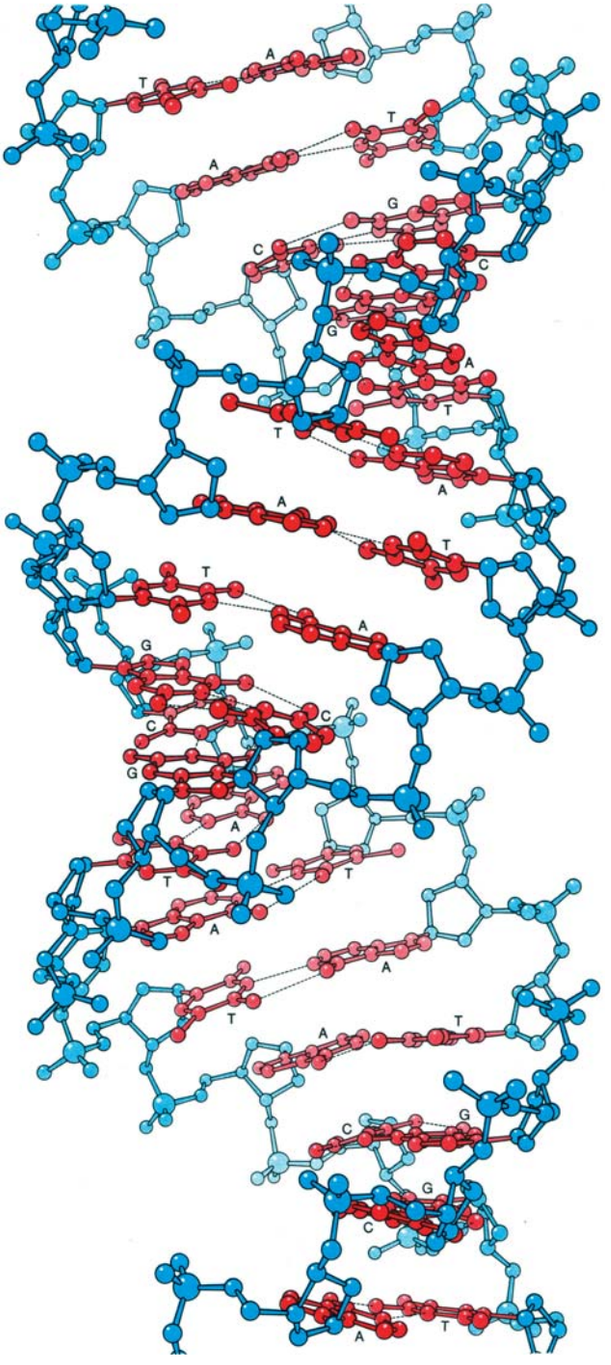

Fig. 23.3.1.2. Infinite A-DNA helix, generated from the X-ray crystal structure of the hexamer G-G-T-A-T-A-C-C (references A2 and A7 in Table A23.3.1.1) by deleting the outer base pair from each end and stacking images of the resulting truncated hexamer so their outer phosphate groups overlapped. This generates an endless helix that exhibits the local structural features of the X-ray crystal structure. Note the degree to which the A helix resembles an antiparallel double-stranded ribbon wound around an invisible helical core (the 'hot wire' axis of Fig. 23.3.1.1). (From Dickerson, 1983.) Reprinted courtesy of the estate of Irving Geis. Rights owned by Howard Hughes Medical Institute.

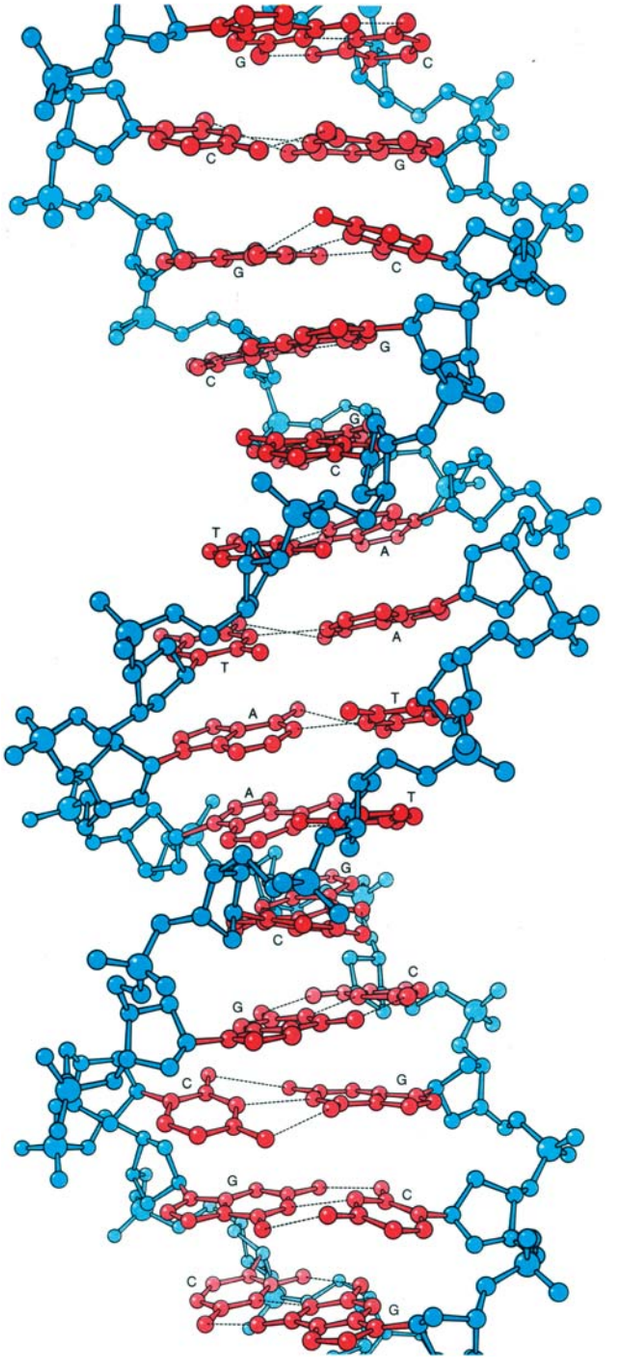

Fig. 23.3.1.3. Infinite B-DNA helix, generated in a similar manner to Fig. 23.3.1.2 from the central ten base pairs of the dodecamer C-G-C-G-A-A-T-T-C-G-C-G (B1–B5). Note that the minor groove is narrow in the AT region facing the viewer at the centre, but appreciably wider in the GC regions on the back side of the helix at top and bottom. Propeller twisting, or deviations of bases from coplanarity within one pair, is one sequence-dependent aspect of DNA that was not suspected from the averaged structures obtained from fibres. (From Dickerson, 1983.) Reprinted courtesy of the estate of Irving Geis. Rights owned by Howard Hughes Medical Institute.

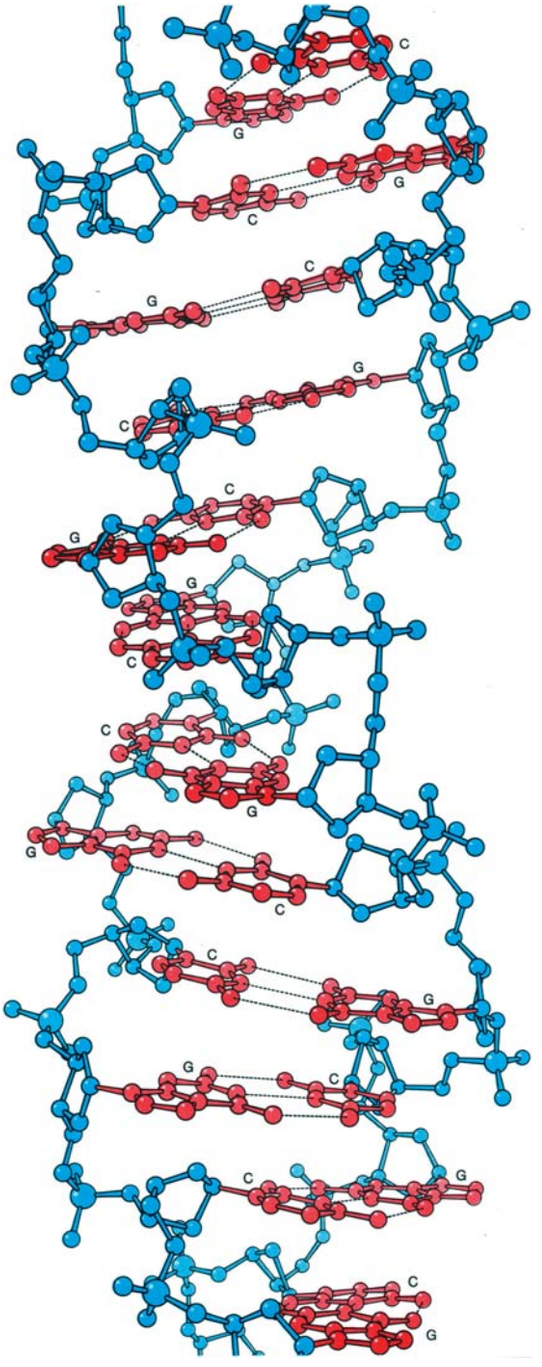

Fig. 23.3.1.4. Infinite Z-DNA helix, generated as before from the central four base pairs of the hexamer C-G-C-G-C-G (Z1). G and C bases alternate along each chain. The sugar–phosphate backbone adopts a pronounced zigzag pathway, rising vertically past each guanine, but travelling horizontally across the helix at cytosines. Hence, the formal helix repeat is two base pairs, G followed by C, rather than a single base pair, as in the A and B helices. Note that the structures of Z-DNA and A-DNA are in many ways the inverse of one another. The Z helix is left-handed, tall and slim, with a deep minor groove, a flattened major groove and small propeller twist. The A helix is right handed, short and broad, with a deep major groove, a shallow minor groove and large propeller twist. (From Dickerson, 1983.) Reprinted courtesy of the estate of Irving Geis. Rights owned by Howard Hughes Medical Institute.

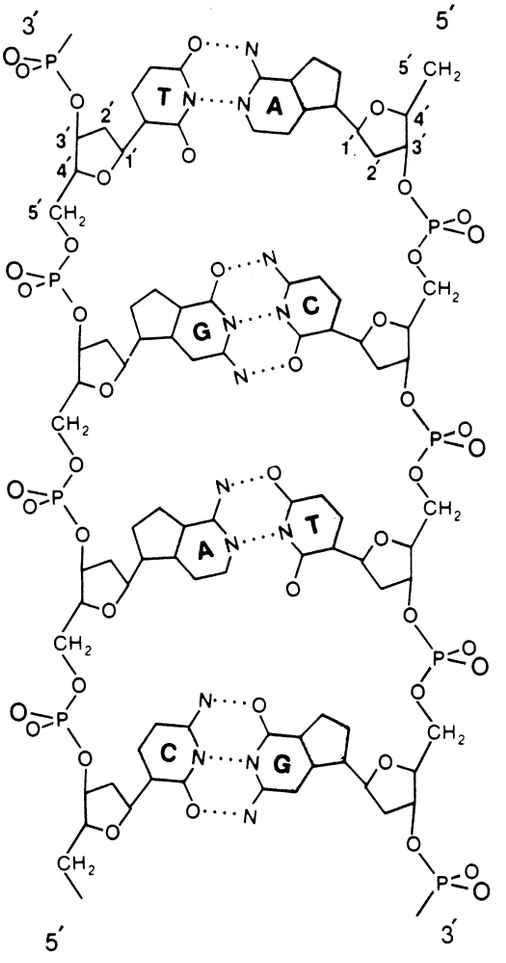

Fig. 23.3.2.1. Unrolled schematic of A- or B-DNA, viewed into the minor groove. Paired bases are attached to backbone chains that run in opposite directions: downward on the right and upward on the left. Z-DNA differs from A- and B-DNA in that the two backbone chains run in opposite directions from those shown here. Hence, Z-DNA cannot be obtained from A- or B-DNA by simple twisting around the helix axis.

Other related but nonstandard base pairs are compared in Fig. 23.3.2.8. Inosine (I) is useful in studying properties of DNA in that, when paired with cytosine (C), it creates a G·C-family base pair having overall similarity to A·T. Similarly, diaminopurine (DAP) [also known as 2-aminoadenine (2aA)], when paired with thymine (T), creates a G·C-like pair from A·T-family bases. Hence, in a given experimental situation, one can unscramble the relative significance of number of hydrogen bonds *versus* identity and location of exocyclic groups.

The conventional Watson–Crick base pairing of Fig. 23.3.2.7 uses the hexamer 'end' of the purine base. A different type of base pairing was proposed many years ago by Hoogsteen (1963), in which the upper edge of the purine was used: N7 and N6/O6. Hoogsteen base pairing is shown between the left-hand two bases in each part of Fig. 23.3.2.9. Note that in Hoogsteen base pairing of A and T, each ring provides both a hydrogen-bond donor and an acceptor. Guanine cannot do this, since both its N7 and O6 positions are acceptors. As a consequence, in a G·C pair, C must supply both of the hydrogen-bond donors. It can only form a Hoogsteen base pair with G when the cytosine ring is protonated. This would lead one to expect triplex formation only at low pH. However, the stability of a triplex can, to a certain extent, alter the pK_a of the N—H proton itself. (Recall the shift in pK_a of buried Asp and His

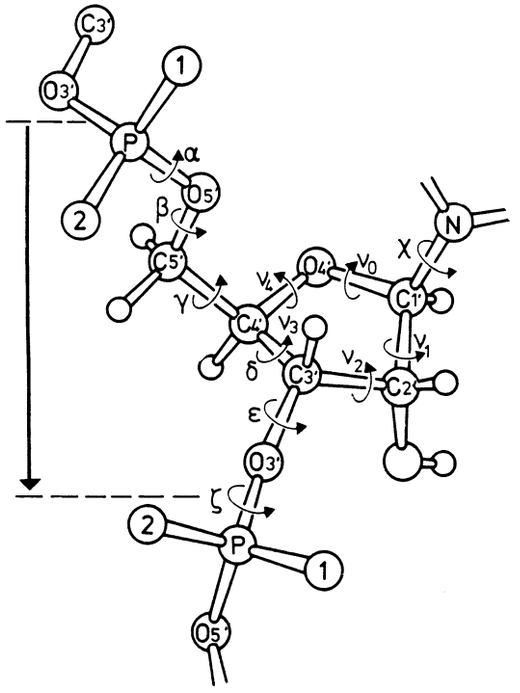

Fig. 23.3.2.2. Sugar–phosphate backbone of RNA and DNA polynucleotides. One nucleotide begins at a phosphorus atom and extends just short of the phosphorus atom of the following nucleotide, with the conventional positive direction being P→O5′—C5′—C4′—C3′—O3′→P, as indicated by the arrows. Main-chain torsion angles are designated α through ζ, and torsion angles about the five bonds of the ribose or deoxyribose ring are ν_0 through ν_4, as shown. If one imagines atoms O3′—P—O5′ as a hump-backed bridge, as one crosses the bridge in a positive chain direction, oxygen atom O1 is to the left and O2 is to the right. These oxygens, accordingly, are sometimes designated O_L and O_R. The —OH group attached to the C2′ atom of the ribose ring in RNA shown here is replaced by —H in the deoxyribose ring of DNA. Atom N to the right is part of the base attached to the sugar ring: N1 in pyrimidines and N9 in purines. Torsion angle χ is defined by O4′—C1′—N1—C2 in pyrimidines and O4′—C1′—N9—C4 in purines.

groups in the active sites of enzymes.) Hence, with a single-chain DNA, G-A-G-A-G-A-A-C-C-C-C-T-T-C-T-C-T-T-C-T-T-T-C-T-C-T-C-T-T, that folds back upon itself twice to build a triplex, NMR experiments indicate a significant amount of triplex remaining even at pH 8.0 (Sklenár & Feigon, 1990; Feigon, 1996).

23.3.2.4. *Helix parameters*

An important advantage of single-crystal oligonucleotide structures over fibre-based models is that one can actually observe local sequence-based departures from ideal helix geometry. B-DNA fibre models indicated a mean twist of *ca* 36° per step, or ten base pairs per turn, whereas A-DNA fibre patterns indicated less winding: *ca* 33° per step or 11 base pairs per turn. Twist, rise per base pair along the helix axis, horizontal displacement of base pairs off that axis, and inclination of base pairs away from perpendicularity to the axis are all intuitively obvious parameters. But when single-crystal structures began appearing in great numbers in the mid-1980s, it became imperative that uniform names and definitions be used for these and for less obvious, but increasingly significant, local helix parameters.

An EMBO workshop on DNA curvature and bending, held at Churchill College, Cambridge, in September 1988, led to an agreement on definitions and conventions that was published

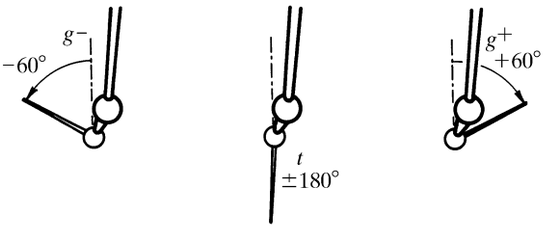

Fig. 23.3.2.3. Definition of torsion angles. A positive angle results from clockwise rotation of the farther bond, holding the nearer bond fixed. Torsion angle +60° is designated as *gauche*$^+$ or g^+, angle 180° is *trans* or *t* and angle −60° is *gauche*$^-$ or g^-.

simultaneously in four journals (Dickerson *et al.*, 1989). Fig. 23.3.2.10 shows the reference frames for two successive base pairs, and Figs. 23.3.2.11 and 23.3.2.12 illustrate local helix parameters involving rotation and translation, respectively. Subsequent experience has shown the most useful parameters to be inclination, propeller, twist and roll among the rotations, and *x* displacement, rise and slide among the translations. As mentioned at the beginning of this chapter, inclination and *x* displacement are the two properties that best differentiate A- from B-DNA. The four most widely used computer programs for calculation of local helix parameters are

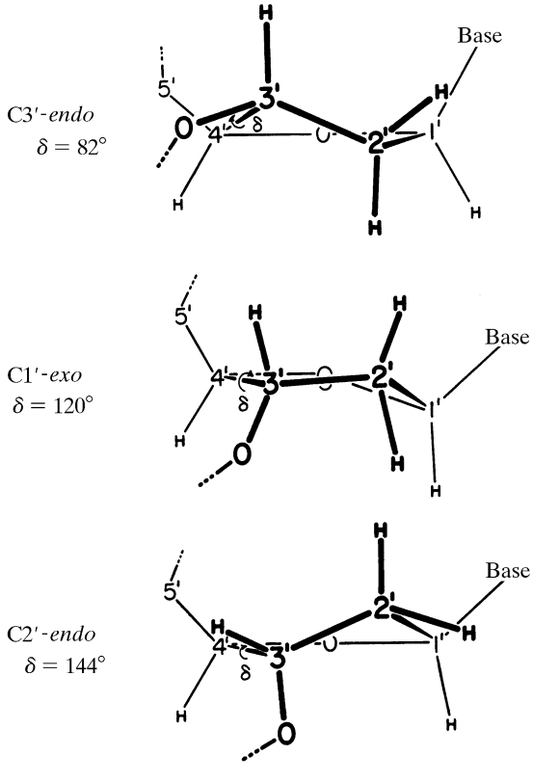

Fig. 23.3.2.4. The three most common furanose ring geometries. The planar form of the five-membered ribose or deoxyribose ring is unstable because of steric hindrance from side groups; one of the five atoms prefers to pucker out-of-plane on one side of the ring or the other. Puckering toward the same side of the ring as the C5′ atom is termed *endo*, and puckering toward the opposite 'outside' surface is termed *exo*. The main-chain torsion angle δ is related to sugar ring conformation because of the motion undergone by the C3′—O3′ bond during changes in puckering.

Table 23.3.2.1. *Average torsion-angle properties of A-, B- and Z-DNA (°)*

Values listed are mean torsion angles, with standard deviations in parentheses. Conformations are only approximate; — indicates a non-*gauche/trans* conformation. B_{II} and Z_{II} are less common variants. For δ, the sugar ring geometry is quoted in place of *gauche/trans*. χ for B-DNA combines pyrimidines and purines. Values were obtained from a sample of 30 A-DNAs, 34 B-DNAs, 22 Z-DNAs and ten nonstandard DNAs in the Nucleic Acid Database. From Schneider *et al.* (1997).

	α	β	γ	δ	ε	ζ	χ
A-DNA	293 (17)	174 (14)	56 (14)	81 (7)	203 (12)	289 (12)	199 (8)
Conformation	g^-	*t*	g^+	C3′-*endo*	*t*	g^-	*t*
B-DNA	298 (15)	176 (9)	48 (11)	128 (13)	184 (11)	265 (10)	249 (16)
Conformation	g^-	*t*	g^+	C1′-*exo*	*t*	g^-	g^-
B_{II}-DNA		146 (8)		144 (7)	246 (15)	174 (14)	271 (8)
Conformation		—		C2′-*endo*	g^-	*t*	g^-
Z_I-DNA – purines	71 (13)	183 (9)	179 (9)	95 (8)	95 (8)	301 (16)	63 (5)
Conformation	g^+	*t*	*t*	O4′-*endo*	g^+	g^-	g^+
Z_{II}-DNA – purines					189 (12)	52 (14)	58 (5)
Conformation					*t*	g^+	g^+
Z_I-DNA – pyrimidines	201 (20)	225 (16)	54 (13)	141 (8)	267 (9)	75 (9)	204 (98)
Conformation	*t*	—	g^+	C2′-*endo*	g^-	g^+	*t*
Z_{II}-DNA – pyrimidines	168 (16)	166 (14)					
Conformation	*t*	*t*					

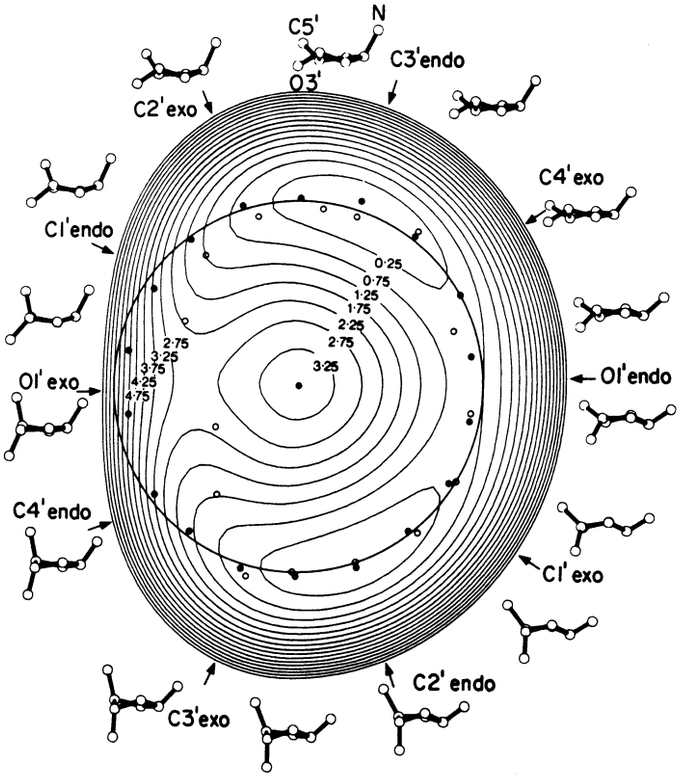

Fig. 23.3.2.5. Potential plot of all furanose ring conformations. Energies are in kcal mol^{-1}. The distance from the central point gives the maximum displacement of the out-of-plane atom from the plane of the other four. The circle is a constant-displacement trajectory chosen to pass through the potential minima on the right three-quarters of the plot. C2′-*endo* and C3′-*endo* are especially favoured, whereas O1′-*exo* on the left is highly disfavoured. The path from C2′-*endo* through C1′-*exo*, O1′-*endo* and C4′-*exo* to C3′-*endo* is a low-energy path, and many examples all along this path are known in B-DNA helices. Reprinted with permission from Levitt & Warshel (1978). Copyright (1978) American Chemical Society.

Table 23.3.2.2. *Sugar ring conformations, pseudorotation angles and torsion angle δ*

Ring conformation	Pseudorotation angle (°)	Torsion angle δ (°)
C3′-*endo*	18	82
C4′-*exo*	54	82
O4′-*endo*	90	96
C1′-*exo*	126	120
C2′-*endo*	162	144
C3′-*exo*	198	158
C4′-*endo*	234	158
O4′-*exo*	270	144
C1′-*endo*	306	120
C2′-*exo*	342	96

NEWHELIX by Dickerson (B7, B46), *CURVES* by Lavery & Sklenar (1988, 1989), *BABCOCK* by Babcock & Olson (Babcock *et al.*, 1993, 1994; Babcock & Olson, 1994) and *FREEHELIX* (Dickerson, 1998*c*). *NEWHELIX* was the earliest of these, but it performs all calculations relative to a best overall helix axis. This is satisfactory for single-crystal DNA structures, but makes the program unusable for the 180° bending observed in some protein–DNA complexes. *CURVES* is especially convenient for mapping the axis of a bent or curved helix. *FREEHELIX*, which evolved from *NEWHELIX*, calculates all parameters relative to local base-pair geometry, without assuming an overall axis, and permits display of normal vector plots that are especially useful in analysing bending in DNA–protein complexes (Dickerson & Chiu, 1997).

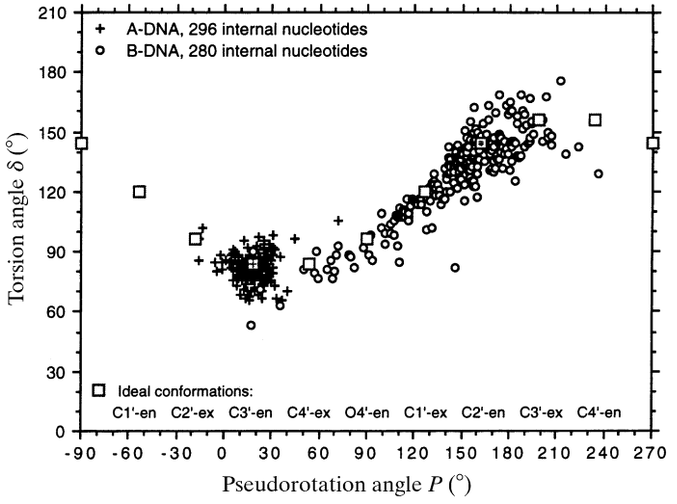

Fig. 23.3.2.6. Plot of observed sugar conformations in 296 nucleotides of A-DNA (crosses) and 280 of B-DNA (open circles). Open squares mark ideal relationships between torsion angle δ (vertical axis) and pseudorotation angle P (horizontal axis) from the expression $\delta = 40° \cos(P + 144°) + 120°$. Deviations from this ideal curve for real helices arise, because the amplitude of pseudorotation (or displacement of one atom from the mean plane of the others) varies from one ring to another. Note the tight clustering of A-DNA points around C3′-*endo* and the broader distribution of B-DNA conformations.

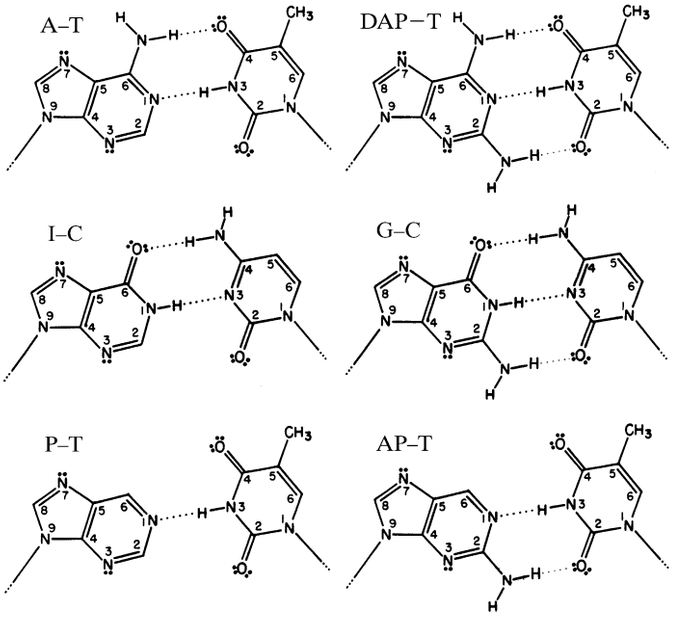

Fig. 23.3.2.8. Alternative purines and pyrimidines, and possible base pairings. Purines: P = purine; AP = 2-aminopurine; A = adenine or 6-aminopurine; DAP = 2,6-diaminopurine (also known as 2aA = 2-aminoadenine); G = guanine; I = inosine. Pyrimidines: T = thymine (uracil if methyl group is absent); C = cytosine. DAP–T is a nonstandard AT-family analogue of G–C, and I–C is a nonstandard GC-family analogue of A–T.

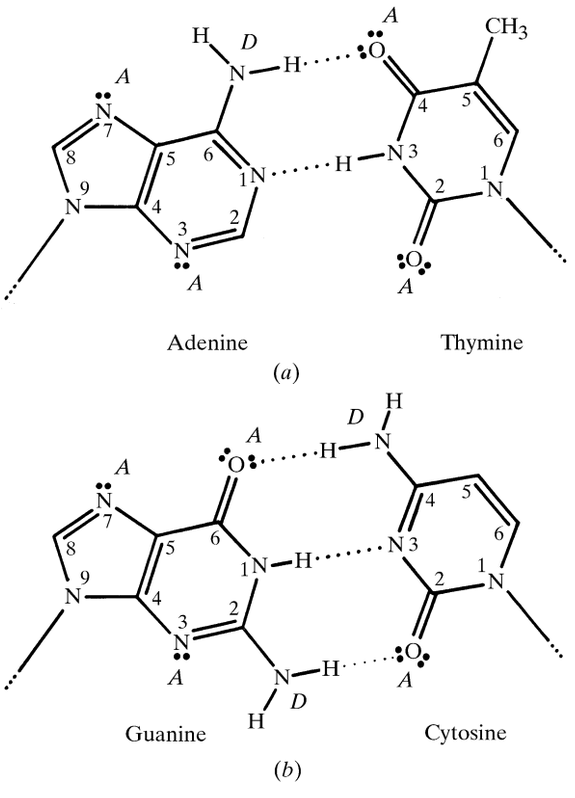

Fig. 23.3.2.7. A·T and G·C base pairs with minor groove edge below and major groove edge above. *A* is a hydrogen-bond acceptor, *D* is as hydrogen-bond donor.

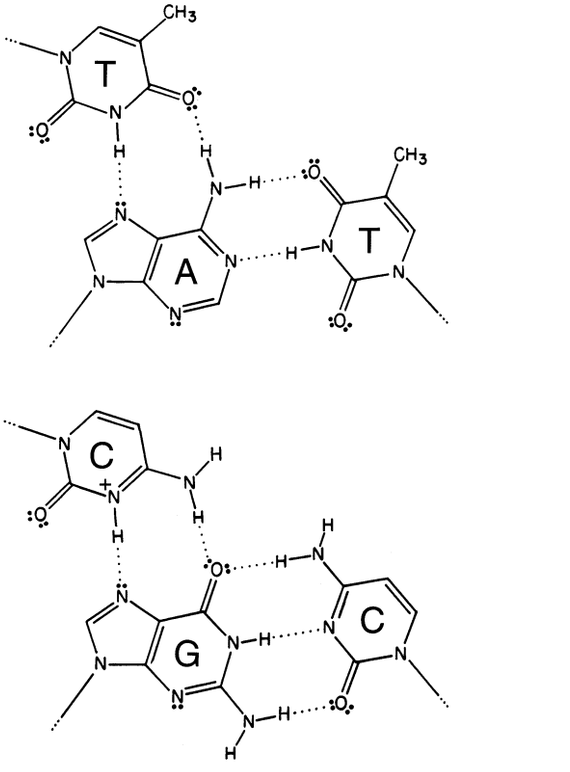

Fig. 23.3.2.9. Watson–Crick pairing of a purine (A or G) with a pyrimidine to its right (T or C), and Hoogsteen pairing of the same purine with a pyrimidine above it. This combination of Watson–Crick and Hoogsteen pairing is found in triple helices or triplexes. Note that Hoogsteen pairing of G and C can only occur at a pH at which C is protonated, because the extra proton is essential for the second hydrogen bond.

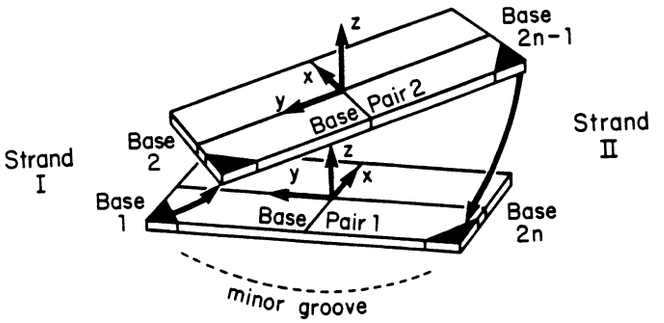

Fig. 23.3.2.10. Definitions of local reference axes (x, y, z) at the first two base pairs of an n-base-pair double helix. Base 1 is paired with base $2n$, base 2 with base $2n-1$ *etc*. Shaded corners represent attachment points to sugar rings. Curved arrows denote 5′-to-3′ 'positive' directions of each backbone chain. Note that when looking into the minor groove, as here, the two strands illustrate a clockwise rotation, upwards on the left and downwards on the right. This is true for A- and B-DNA, but for Z-DNA, the sense of the two backbone strands is reversed.

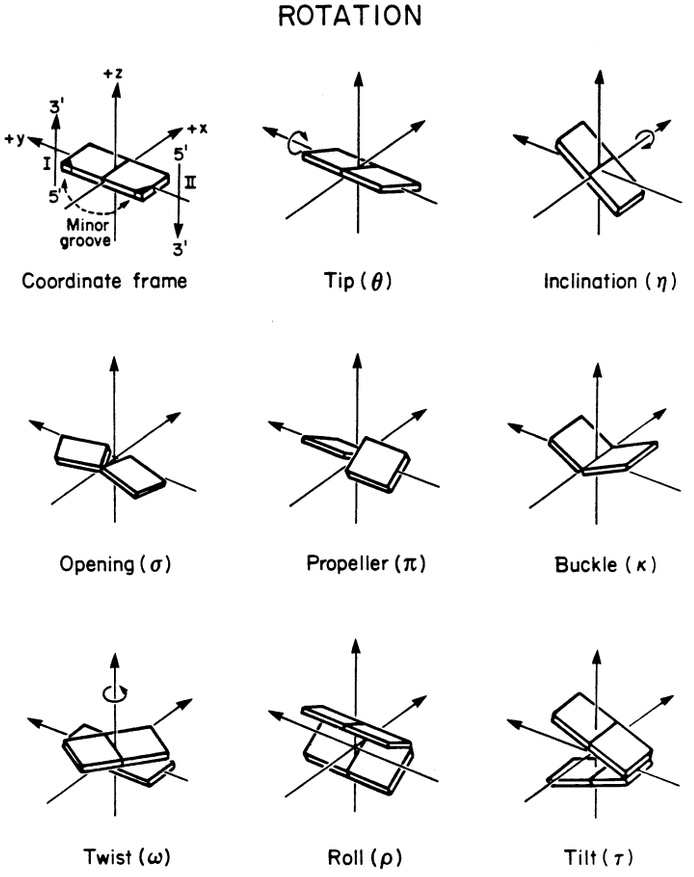

Fig. 23.3.2.11. Local helix parameters involving rotations. Tip and inclination describe the orientation of a base pair relative to the helix axis, produced by rotation about the base-pair long axis or short axis, respectively. Opening, propeller and buckle describe rotations of the two bases of a pair relative to one another. Twist, roll and tilt describe changes of orientation from one base pair to the next, *via* rotations about the z, y and x axes, respectively.

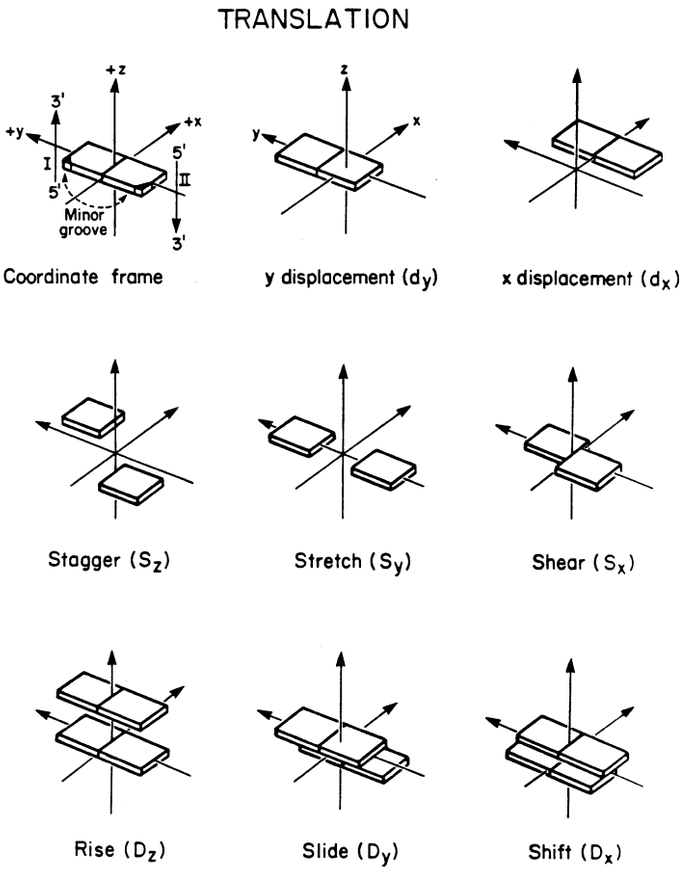

Fig. 23.3.2.12. Local helix parameters involving translations. y and x displacements describe shifts of a lone base pair along its long or short axis, respectively. Stagger, stretch and shear describe displacements of the two bases of a pair relative to one another. Rise, slide and shift describe displacements from one base pair to the next, *via* translations along the z, y and x axes, respectively.

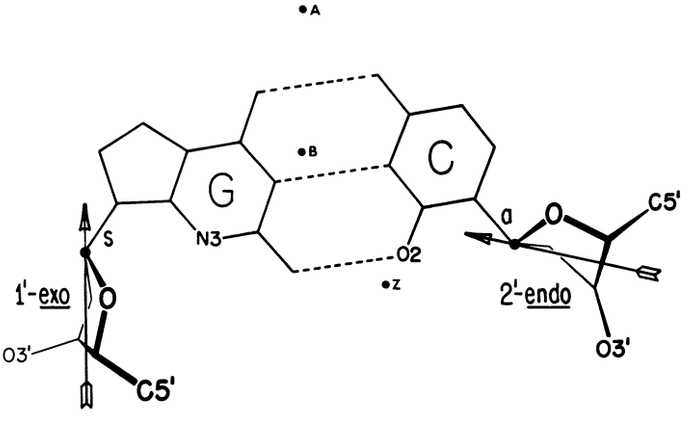

Fig. 23.3.2.13. *Syn versus anti* orientation about the glycosyl bond connecting sugar and base. Right: *anti* conformation, with χ *ca* 210°. Left: *syn* conformation, with χ around 60°. Both A- and B-DNA only employ the *anti* geometry; Z-DNA uses *anti* for pyrimidines and *syn* for purines, as shown here. Note that the 5′-to-3′ direction in both rings is down into the paper. Hence, antiparallel backbone chains can be achieved only by a zigzag chain geometry with local chain reversals, as shown later in Fig. 23.3.3.4. Black dots labelled A, B and Z indicate the position of the helix axis relative to the base pairs in A-, B- and Z-DNA.

23.3.2.5. *Syn/anti glycosyl bond geometry*

The glycosyl bond angle, χ, about the bond connecting a sugar ring to a base is a special case of torsion angle, and is defined by O4′—C1′—N1—C2 for pyrimidines and O4′—C1′—N9—C4 for purines. In A- and B-DNA, the normal range of χ is 160 to 300°. This is known as the *anti* conformation (right-hand side of Fig. 23.3.2.13) and swings the sugar ring out away from the minor groove edge of the base pair. In Z-DNA, pyrimidines also exhibit the *anti* glycosyl bond conformation, but purines adopt the *syn* geometry shown on the left-hand side of Fig. 23.3.2.13. Now the sugar ring is rotated so that it intrudes into the minor groove, and χ lies in the range 50 to 90°.

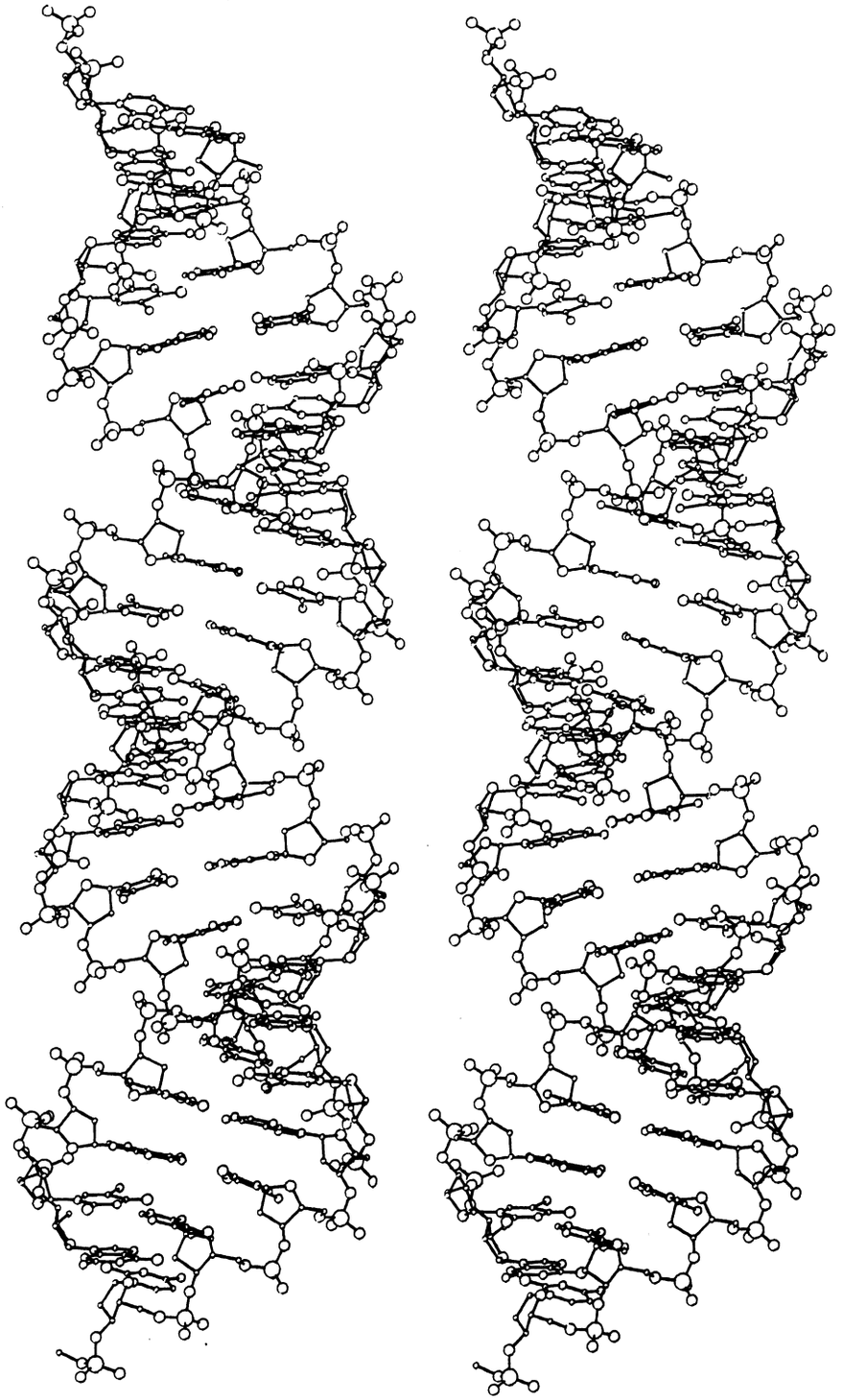

Fig. 23.3.3.1. The A-DNA stereo pair drawing from which Fig. 23.3.1.2 was derived, with repeating sequence -(G-T-A-T-A-C)$_n$-. The impression of the A helix as a ribbon wrapped around an imaginary core is even more strongly developed in this stereo. (From Dickerson, 1983.)

23.3.3. Comparison of A, B and Z helices

Figs. 23.3.3.1–23.3.3.3 show the original stereo pairs that were re-drawn by Irving Geis in preparing Figs. 23.3.1.2–23.3.1.4. These stereo pairs were constructed from X-ray structures of A-, B- and Z-DNA oligomers by deleting the outermost base pair from each end, eliminating the backbone as far as the first phosphate group, and then stacking these trimmed-down helices on top of one another, with phosphate groups overlapping, to create an infinite helix. They are improvements over the idealized infinite helices generated from fibre diffraction in that they display local variation in helix parameters that only single-crystal analyses can reveal. In the present context, they are good subjects for discussion of the differences between the three helix types.

23.3.3.1. *x displacement and groove depth*

A-DNA (Wahl & Sundaralingam, 1996, 1998), B-DNA (Berman, 1996; Dickerson, 1998*b*) and Z-DNA (Ho & Mooers, 1996; Basham *et al.*, 1998) have each been the subject of recent reviews, to which the reader is referred for details that cannot be covered here. The distinctive properties of the three helices are listed in Table 23.3.3.1. The most obvious distinction is handedness: A and B are right-handed helices, whereas Z is left-handed. Moreover, the position of each base pair relative to the helix axis is quite different. As noted in Fig. 23.3.2.13, the helix axis passes through base pairs in B-DNA, lies on the minor groove side of base pairs in Z-DNA, and on the major groove side in A-DNA. In terms of the helix parameters of Fig. 23.3.2.12, A-DNA has a typical *x* displacement of $d_x = +3$ to $+5$ Å, B-DNA has $d_x = -1$ to 0 Å, and Z-DNA has $d_x = -3$ to -4 Å. There is virtually no overlap between these three ranges; *x* displacement, d_x, in fact, is a better criterion for differentiating the three classes of helix than is sugar ring conformation.

A direct consequence of these *x* displacement values is great differences in depths of major and minor grooves. Both grooves are of equivalent depth in B-DNA because base pairs sit on the helix axis. In A-DNA, a base pair is pushed off-axis so

references