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°.

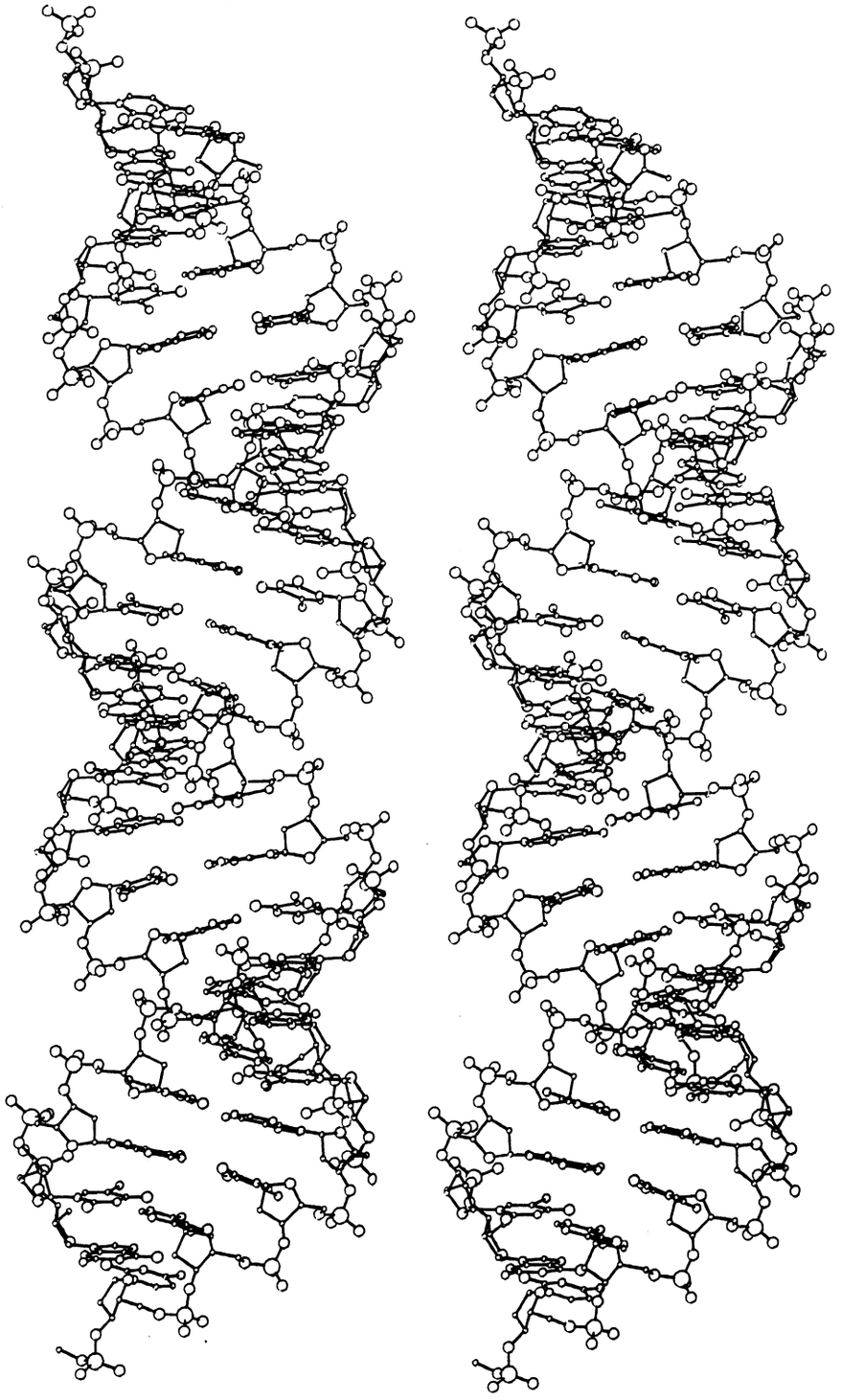

Fig. 23.3.3.1. The A-DNA stereo pair drawing from which Fig. 23.3.1.2 was derived, with repeating sequence -(G-T-A-T-A-C)$_n$-. The impression of the A helix as a ribbon wrapped around an imaginary core is even more strongly developed in this stereo. (From Dickerson, 1983.)

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

that its minor edge approaches the helix surface, making the minor groove very shallow and the major groove cavernously deep. In Z-DNA, it is the major edge of each base pair that is pushed toward the surface, so that the minor groove is deep and the major groove is so shallow as hardly to be characterized as a groove at all. It is sometimes stated that 'Z-DNA has no major groove', but space-filling stereos, such as Fig. 1 of reference Z6 or Fig. 3 of Z23 reveal the shallowest of major grooves running around the helix cylinder, flanked by very slightly higher phosphate backbones.

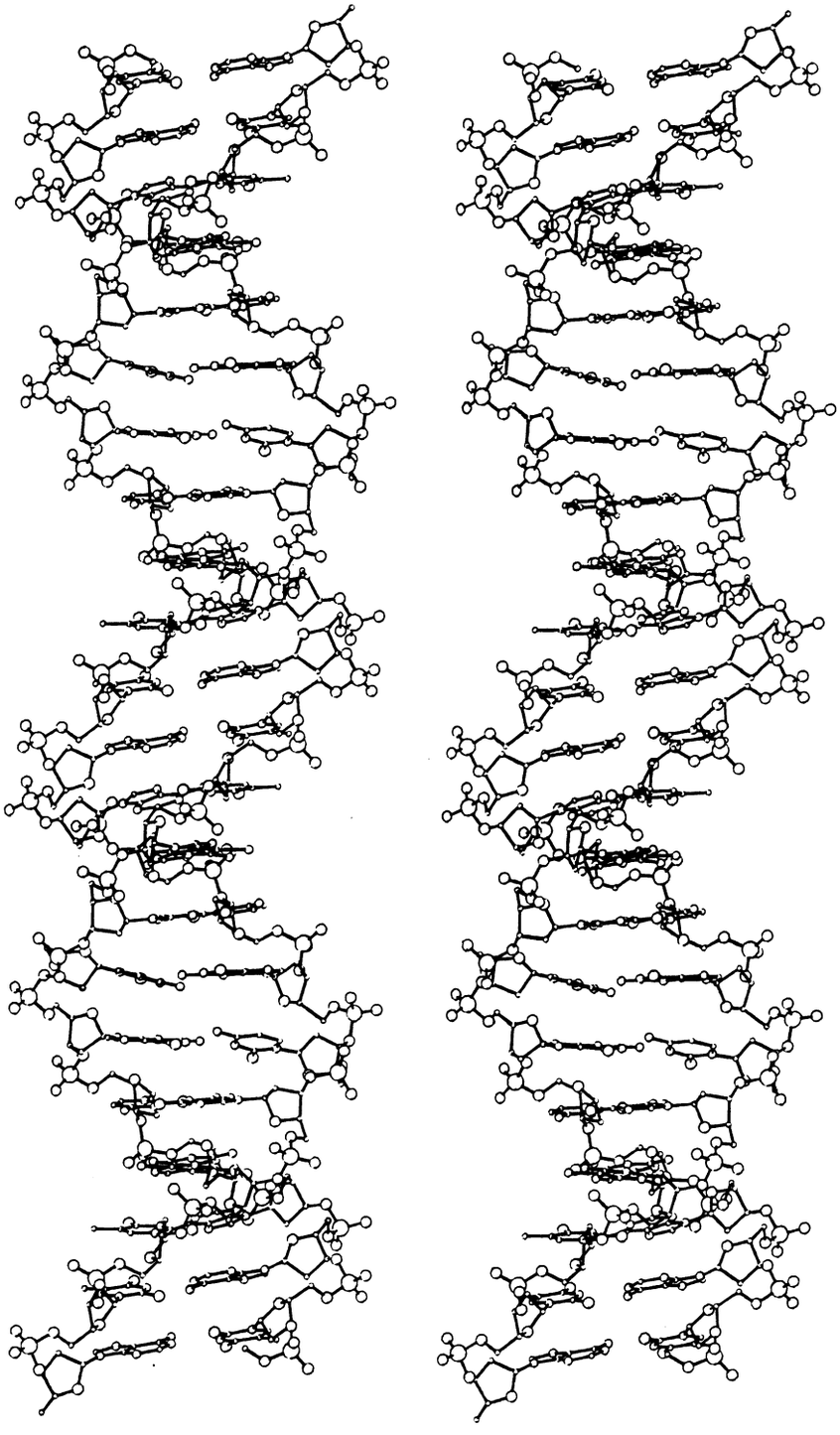

Fig. 23.3.3.2. The B-DNA stereo pair drawing from which Fig. 23.3.1.3 was derived, with repeating sequence -(G-C-G-A-A-T-T-C-G-C)$_n$-. The variation of minor groove widths on the front and back sides of the helix is striking. (From Dickerson, 1983.)

23.3.3.2. *Glycosyl bond geometry*

In both A- and B-DNA, all glycosydic bonds are *anti*, with sugar rings swung to either side away from the minor groove, as in Fig. 23.3.3.4(*a*). As mentioned earlier, when viewed into the minor groove, the backbone chains describe a clockwise rotation, with the chain on the right running downward, and that on the left upward, as in Fig. 23.3.2.1. In Z-DNA, both chains run in the *opposite* direction, leading to a counterclockwise rotation sense viewed into the minor groove. But Z-DNA has yet another striking (and defining) feature. Purines and pyrimidines alternate along each chain. G and C are most strongly favoured by far, but A and T can substitute intermittently at a price in stability. Breaking the strict alternation of purines and pyrimidines is even more unfavourable and is rarely encountered in crystal structures (Table A23.3.1.3). At each purine base, the glycosyl bond is rotated into the minor groove to the *syn* position, as in Fig. 23.3.3.4(*c*). This causes the local backbone directions, defined by sugar ring atoms C4′ and C3′, to be parallel in the two strands. Z-DNA avoids becoming a parallel-chain helix by performing a local chain reversal at each pyrimidine. In Fig. 23.3.3.4(*c*), although the local C4′–C3′ chain direction at the cytosine sugar is downward, the double loop in backbone chain gives it a net upward orientation. In stereo Fig. 23.3.3.3, the ascending backbone chain rises smoothly past each guanine, with a chain path parallel to the helix axis. However, the chain bends abruptly at right angles when passing a cytosine, in a direction tangential to the helix cylinder. Guanine sugar rings point their O4′ oxygen atoms in the backward chain direction (as is also true for all bases in A- and B-DNA), but cytosine sugars point their oxygens in the forward direction. This 'up at G, across at C' pathway and inversion of sugar rings is what produces the zigzag backbone pathway that leads to the name Z-DNA. The O4′ atom of each cytosine sugar is stacked on top of the guanine ring of the subsequent nucleotide, and this stacking of a polar O (or N) on top of a polarizable aromatic ring contributes to the stability of the Z helix, as it does to many other base–base interactions to be discussed later (Bugg *et al.*, 1971; Thomas *et al.*, 1982; B32).

23.3.3.3. *Sugar ring conformations*

Sugar ring conformations in A- and B-DNA have a logical structural basis. The B-DNA backbone is more extended than the A-DNA backbone, with P–P distances of *ca* 6.6 Å along one chain, compared with *ca* 5.5 Å in A-DNA. In turn, C2′-*endo* is a more extended ring conformation than C3′-endo, demonstrable in Fig. 23.3.2.4 by a greater distance between C5′ and O3′ atoms. Hence, it is logical that

the more extended ring conformation should be associated with the more extended backbone chain. In Z-DNA, the extended C2′-*endo* form is adopted at cytosine, where a zigzag double chain reversal must be accommodated, while the more compact C3′-*endo* occurs at the straight backbone segment running past a guanine.

The cramped *syn* glycosyl conformation is strongly disfavoured, although not absolutely forbidden, at pyrimidines, most probably because of steric clash between the pyrimidine O2 and the *syn* ring (Haschmeyer & Rich, 1967; Davies, 1978; Ho & Mooers, 1996; Basham *et al.*, 1998). Hence, the Z-DNA helix is effectively limited to alternating pyrimidine/purine sequences, with a price that must be paid for intermittent substitution of A and T for G and C, and an even higher price paid for breaking the pyrimidine/purine alternation. This is reflected in the X-ray crystal structures listed in Table A23.3.1.3. Only one non-alternating sequence has been completely solved and published: *C-G-G-G-*C-G (Z40), where adoption of the Z form has been forced by 5-methylation of cytosines (*C). A second non-alternating sequence that includes AT base pairs, *C-G-A-T-*C-G (Z13), was solved in 1985, but its coordinates have never been made public. It, too, required methylation of cytosines to induce the Z form. A third sequence, C-C-G-C-G-G (Z42), opens its terminal base pairs to make intermolecular base pairs with crystal neighbours. The 52 remaining Z-DNA structures in Table A23.3.1.3 all have strict alternation of pyrimidines and purines.

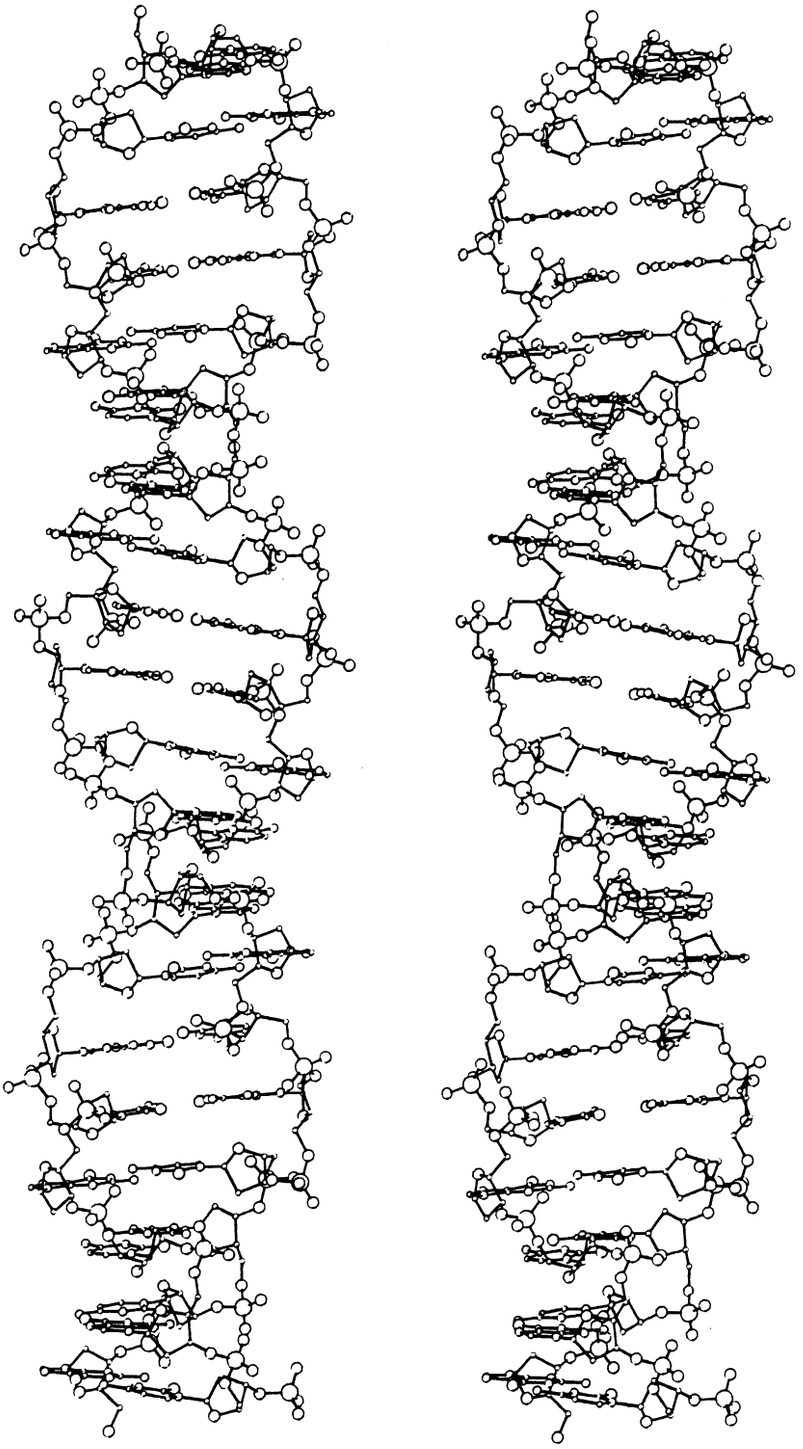

Fig. 23.3.3.3. The Z-DNA stereo pair drawing from which Fig. 23.3.1.4 was derived, with repeating sequence -(G-C-G-C)$_n$-. Note the left-handed zigzag path of the sugar–phosphate backbone, which led to its designation as the Z helix. (From Dickerson, 1983.)

23.3.3.4. *Helical twist and rise, and propeller twist*

The helical repeat unit in Z-DNA is therefore two successive base pairs, rather than the single base pair of A- and B-DNA. Ho & Mooers (1996) propose that the C-G or $^{5'}$pyrimidine-P-purine$^{3'}$ step be considered the fundamental unit of the Z-helical structure, because of the tight overlap between the two base pairs. As can be seen in Fig. 23.3.3.3, in a C-G step the pyrimidine rings from the two base pairs actually stack over one another, whereas the purine rings are packed against neighbouring sugar O4′ atoms. Helix-axis rotation at this step is only −8°, whereas the preceding and following G-C steps have a mammoth −52° twist. Hence, although Z-DNA has 12 base pairs per turn, it technically is not a dodecamer helix, but a hexamer with a two-base-pair repeating unit and a total rotation of −60° per unit.

This virtual restriction to purine/pyrimidine alternation means that Z-DNA cannot be involved in the coding of genetic information. A and B helices have no such restriction; their structures can accommodate a random sequence of bases. Average twist angles are as shown in Table 23.3.3.1, although extreme variation in twist is observed at individual steps in single-crystal structure analyses, from as little as 16° to as much as 55°. Base-sequence preferences for local helix parameters are discussed below.

In both B and Z helices, base pairs are very nearly perpendicular to the helix axis, whereas in the winding double ribbon of A-DNA, the long axis of each base pair is inclined by 10 to 20° away from perpendicularity to the axis. Hence, the rise per base pair for all B-helical steps and for G-C steps of Z-DNA is equal to the thickness of a base pair, 3.4 Å. The rise at a C-G step of

Table 23.3.3.1. *Comparison of structures of A, B and Z helices*

	A	B	Z
Handedness	Right	Right	Left
Helix axis relative to base pairs	Major groove side	Through centre of base pair	Minor groove side
Major groove	Very deep and narrow	Wide, same depth as minor	Very shallow and broad
Minor groove	Shallow and broad	Variable, same depth as major	Very deep and narrow
Glycosydic bonds	*anti*	*anti*	C: *anti* G: *syn*
Minor groove backbone chain sense *	Clockwise	Clockwise	Counterclockwise
Sugar conformation	C3′-*endo* (narrow range)	C1′-*exo*/C2′-*endo* (broad range)	C: C2′-*endo* G: C3′-*endo*
Base pairs per helix repeat	1	1	2
Base sequence limitations	None	None	Alternating $(C\text{-}G)_n$ or close variants
Rise per base pair (average)	2.9 Å	3.4 Å	C-G: 4.1 Å G-C: 3.5 Å
Base pair inclination	10–20°	*ca* 0°	*ca* 0°
Mean twist angle	30–33°	34–36°	C-G: −8° G-C: −52°
Helix repeats per turn	11–12	10–10.5	6 (2 base pairs)
Propeller twist	Often substantial, 0–25°	Often substantial, 0–25°	Usually small
Common biological occurrence	RNA	DNA	None?

* Relative 5′-to-3′ directions of the two backbone chains, when viewed into the minor groove.

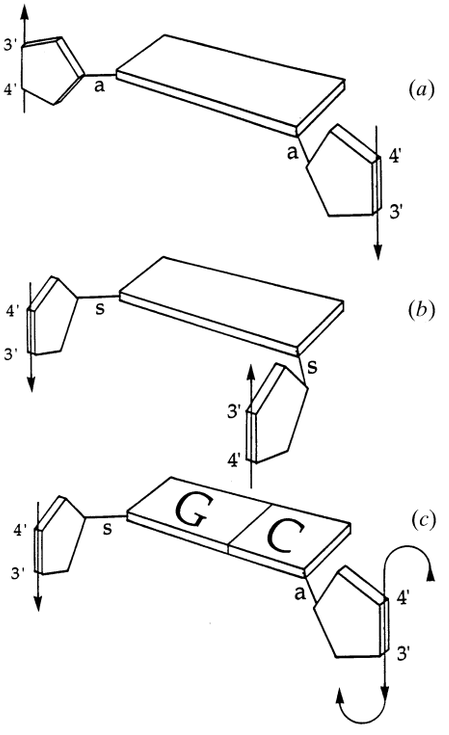

Fig. 23.3.3.4. Glycosyl conformation and chain sense. (*a*) Glycosyl conformations *anti/anti*, backbone chains antiparallel, with clockwise sense when viewed into the minor groove, as here. This is typical for A- and B-DNA. (*b*) Glycosyl conformation *syn/syn*, backbone chains antiparallel, with counterclockwise sense viewed into minor groove. This is not known for any nucleic acid duplex. (*c*) Glycosyl conformation *syn* at G and *anti* at C, with the C4′—C3′ edge of the sugar pointing downward in both strands, which would seem to imply a parallel-stranded helix. However, in Z-DNA, antiparallel strands are achieved by a local reversal of chain direction at each C, as shown here. This produces the zigzag backbone pathway that is characteristic of the Z helix, visible in Fig. 23.3.3.3.

Z-DNA is larger because it involves stacking of a sugar oxygen on each purine ring, not ring stacked on ring. For A-DNA, the rise along the helix axis can actually be less than the thickness of a base pair, because adjacent base pairs are stacked at an incline. The perpendicular distance from one base pair to the next in A-DNA is still 3.4 Å. Both A- and B-DNA exhibit considerable base pair propeller twist, especially at A·T pairs with only two hydrogen bonds rather than three. In contrast, Z-DNA, with predominately G·C pairs, shows only a small propeller twist.

The stacking of base pairs has immediate consequences for crystal growth. For Z-DNA, four base pairs are one-third of a helical turn, and six base pairs are a half turn. Hexamers are the most common crystal form in Table A23.3.1.3 by a large majority. In contrast, octamers and decamers are not simple fractions of a turn, and they stack in a disordered manner. One would predict that dodecamers of Z-DNA might crystallize well if the oligomers were not so long as to fall prey to cylindrical disorder.

By the same principles, B-DNA decamers stack easily and well to build pseudo-infinite helices through the crystal, with ordered cylindrical rods packed in six different space groups. The other common crystallization mode for B-DNA, the dodecamer, has a two-base-pair overlap of ends that both stablizes the crystals and yields a functional ten-base-pair repeat. (See Fig. 2 of Dickerson *et al.*, 1987.) Because the dodecamers are held by their outer two base pairs, the central eight pairs are unobstructed and accessible in the crystal, making dodecamers particularly good subjects for the study of minor-groove binding drugs.

A-RNA duplexes [Table A23.3.1.1, part (*k*)] also stack end-for-end in a manner simulating an infinite A helix, even though the end base pairs are inclined and are not perpendicular to the helix axis. This behaviour has been seen for octamers with roughly two-thirds of a helical turn, for nonamers, and for dodecamers with roughly a full turn.

In contrast, crystals of A-DNA behave quite differently. Regardless of chain length, A-DNA helices crystallize with the outer base pair of one helix packed against one wall of the broad, open and relatively hydrophobic minor groove of another helix.

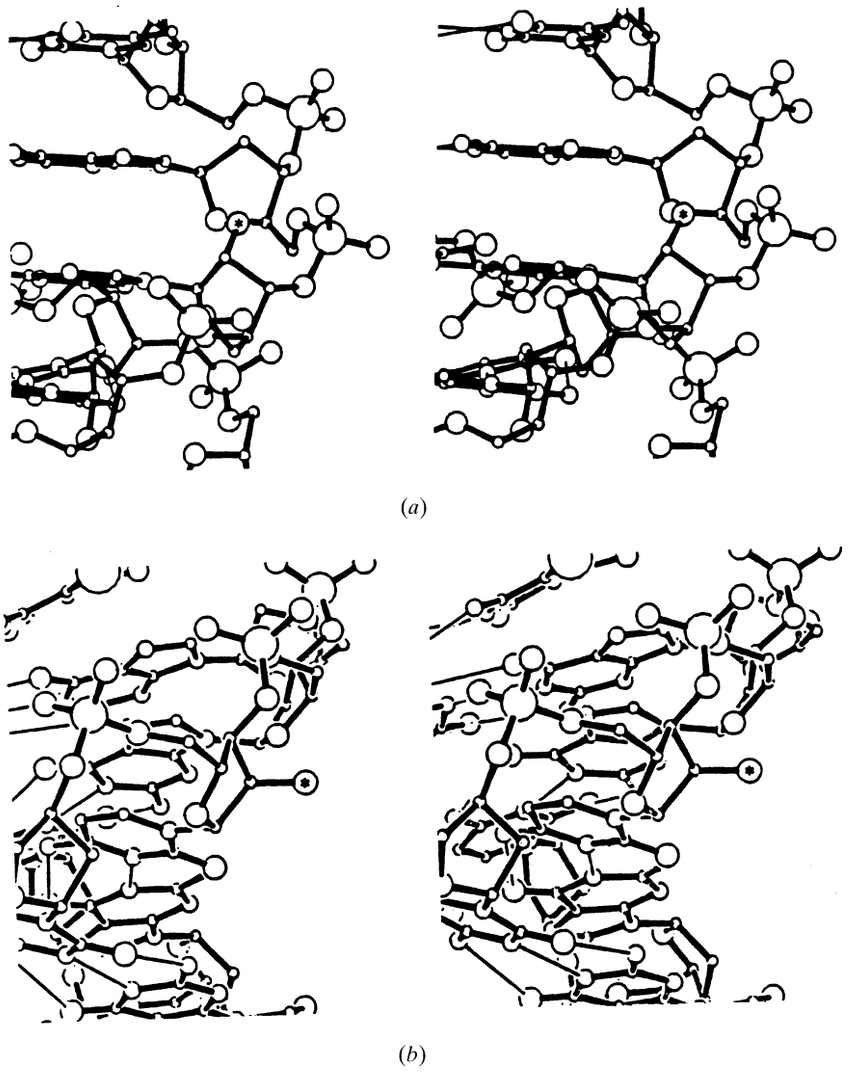

Fig. 23.3.3.5. The role of the C2′-OH in RNA helix geometry. (*a*) Addition of a C2′-OH group (*) to the B-DNA helix leads to close contacts and unallowable steric hindrance with the following O5′ and O4′ atoms and, to a lesser degree, the subsequent base itself. (*b*) A C2′-OH group added to the A-DNA helix extends outward radially from the helix cylinder surface and produces no steric clashes. Hence, A-RNA is quite possible, whereas B-RNA is disallowed.

This packing mode is sufficiently adaptable to accommodate duplexes of lengths four, six, eight, nine, ten and 12 base pairs. Hence, A-DNA does not simulate infinite helices through the crystal lattice, as A-RNA and B- and Z-DNA do.

23.3.3.5. *Allowable RNA helices*

So far this discussion has only been concerned with DNA. Which of the three helix types can be adopted by RNA? Fig. 23.3.3.5 shows that addition of a 2′-OH group to a B-DNA helix [part (*a*)] creates severe steric clash with the phosphate group and sugar ring of the following nucleotide, whereas in an A helix [part (*b*)], the added hydroxyl group extends radially outward from the helix cylinder and causes no steric problems. Hence, the natural helical form for RNA is the A helix, not the B helix. Table A23.3.1.1 shows several single-crystal analyses of A-RNA and RNA/DNA hybrids; Table A23.3.1.2 shows no B-RNA structures. One RNA/DNA hybrid is known as a Z helix: C-G-c-g-C-G (Z24), in which the two central nucleotides are RNA. If one mentally adds an —OH to each C2′ atom in Fig. 23.3.3.3, on the same side of the ring as O3′, it is apparent that the C2′-OH is not inherently incompatible with the Z helix, as it is with the B helix. At guanine sugars, the C2-OH points out and away from the helix, while at cytosine sugars it points away from the base into the spacious minor groove.

23.3.3.6. *Biological applications of A, B and Z helices*

The B helix is the biologically relevant structure for DNA. The A form might logically be adopted at the stage of transient DNA–RNA duplexes during transcription, but elsewhere the B form holds sway. It was once thought that binding of DNA to a protein surface, most particularly nucleosomal winding, might constitute a sufficient dehydration of bound water molecules from the DNA duplex to shift it to the A form. This proved to be false; nucleosomal DNA clearly retains the B conformation. The closest that one comes to biological A-DNA is local deformations upon binding of B-DNA to a few proteins that have been described as ‘A-like distortions’. On the other hand, the A helix has been found repeatedly in RNA duplexes, including tRNA and ribozymes.

The situation is even more restrictive with the Z helix. Although its alternating purine/pyrimidine sequence makes it unusable for genetic *coding*, the suggestion has been made on many occasions that Z-DNA might be an important element in genetic *control* by being involved in negative supercoiling (Herbert & Rich, 1996). It has been shown that a left-handed DNA conformation can be induced by negative superhelical stress, but it is not absolutely clear that this induced, left-handed conformation is the same as the Z helix seen in crystal structures of small oligomers. As noted by Herbert & Rich (1996), after nearly twenty years of enquiry, it is still far from certain that Z-DNA itself has any demonstrable biological role.*

A major stumbling block is the cumbersome mechanism that must be invoked to explain a B-to-Z interconversion. As mentioned previously, a simple twisting of the helix from right to left is not sufficient, because the backbone chains run in opposite directions in the two forms. Fig. 23.3.3.6 demonstrates the steps that must still be undertaken after both B and Z helices have been unwound so as to remove all of their helical character. Note the opposite sense of the backbone strands in B [part (*a*)] and Z [part (*e*)]. In order to accomplish the interconversion, base pairs of B-DNA must be

* Rich and co-workers (Schwartz *et al.*, 1999) have recently solved the crystal structure of the Zα domain of the human editing enzyme ADAR1 in a complex with a six-base-pair Z-DNA helix of sequence CGCGCG. This left-handed hexamer may suffer from the same length *versus* conformation uncertainty mentioned later in this chapter in connection with oligonucleotide crystals, especially since protein–DNA contacts in the Zα complex occur only with the zigzag phosphate backbone, which is not that dissimilar in Z-DNA (Fig. 23.3.3.3) and Z(WC)-DNA (Fig. 23.3.3.7). Nevertheless, it is encouraging to see a short segment of Z actually making contact with its protein in a presumably biologically relevant context.

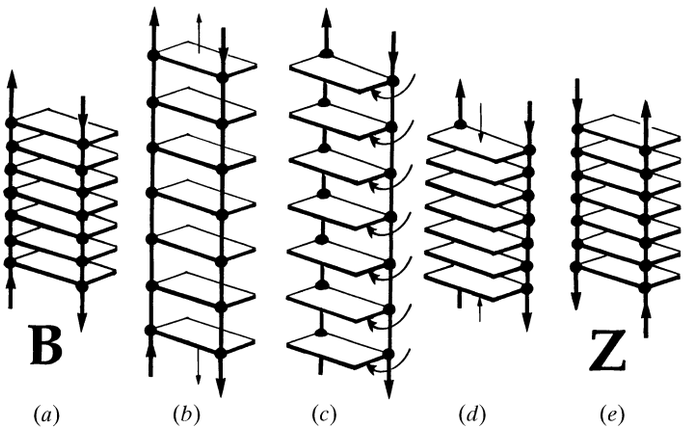

Fig. 23.3.3.6. Interconversion of a B to a Z helix. Because the strands have opposite directions in B (*a*) and Z (*e*), interconversion must involve opening up the helix (*b*), flipping each base pair to the other side (*c*), and re-stacking base pairs (*d*). (*d*) and (*e*) are identical upon rotation about a vertical axis.

pulled apart, as in part (*b*), and each base pair swung around to the opposite side of the backbone 'ladder' [part (*c*)]. This would automatically lead to *syn* conformations at both ends of the base pair, as drawn in Fig. 23.3.3.4(*b*). Returning pyrimidines to an *anti* conformation would create the zigzag backbone chain (Fig. 23.3.3.4*c*). Base pairs can then be re-stacked, as in parts (*d*) and (*e*) in Fig. 23.3.3.6 (which differ only by rotation of the entire helix about the vertical), to yield the backbone geometry of a Z helix. This is the simplest interconversion and one which was recognized and proposed in the very first Z-DNA structure paper (Z1). Other alternatives have been suggested, involving breaking individual base pairs, swinging the bases independently around their backbone chains, and re-forming the pairs. But one kind of special mechanism or another must be invoked if a B-to-Z interconversion is to be achieved.

23.3.3.7. *'Watson–Crick' Z-DNA*

Ansevin & Wang (1990) have proposed an alternative left-handed double helix, with many of the properties of Z-DNA, but possessing the same backbone chain orientations as A- and B-DNA. With such a helix, a B-to-Z conversion would require only a twisting of the duplex about its axis – no separation of bases or unpairing, and no pulling apart of the stack. Ansevin & Wang did not challenge the X-ray crystal structure analyses of short Z-DNA oligomers. Instead, they suggested that Z-DNA was globally the most stable form, adopted in short oligomers where chain unravelling and rearrangement is easy, but that their 'Watson–Crick' Z-DNA or Z(WC)-DNA was the structure that was actually produced by *in vitro* or *in vivo* manipulations of long DNA duplexes. They noted that most solution measurements focus on only two characteristics of the DNA: left-handedness and a dinucleotide repeat, both shared by Z-DNA and Z(WC)-DNA.

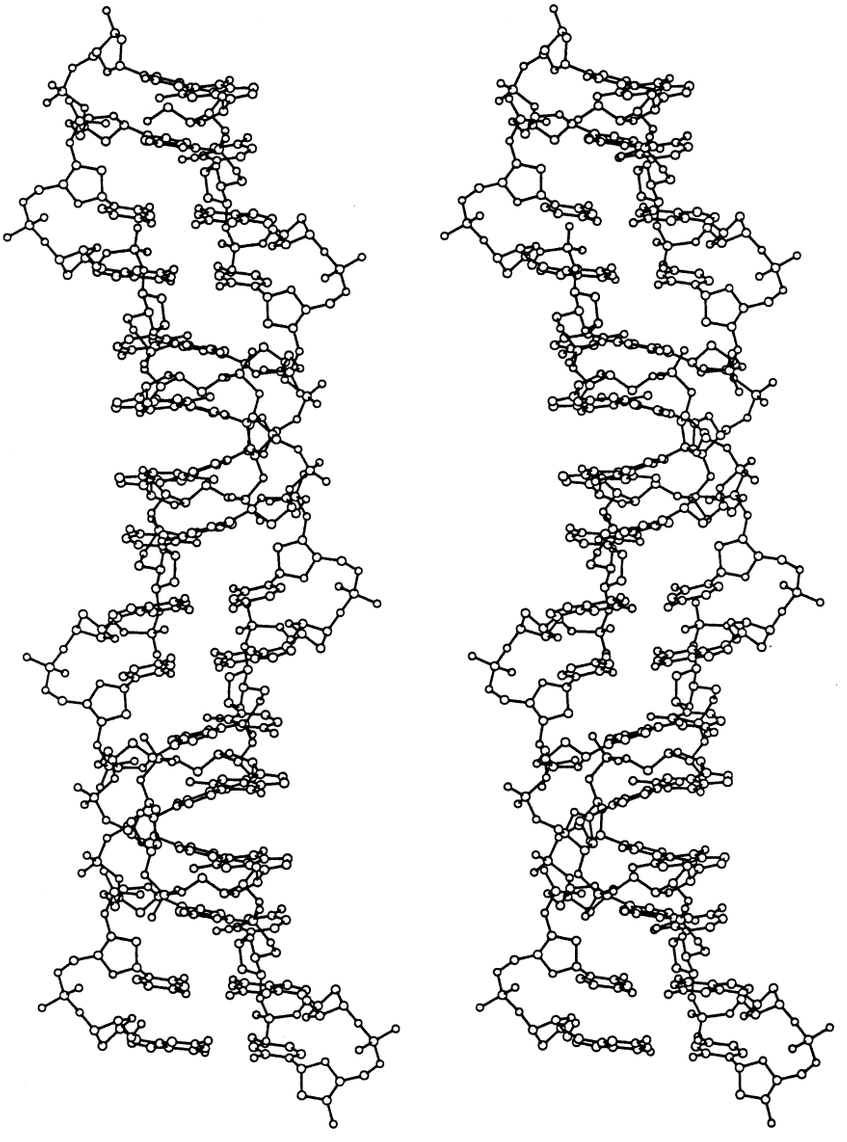

Fig. 23.3.3.7. Z(WC)-DNA, or 'Watson–Crick Z-DNA', a proposed left-handed, zigzag, alternating purine/pyrimidine helix with many of the properties of Z-DNA, but with the backbone chain sense found in A- and B-DNA (Ansevin & Wang, 1990). Coordinates courtesy of Allen T. Ansevin.

The Z(WC) helix is shown in Fig. 23.3.3.7, and a different stereo view appears as Fig. 7 of Dickerson (1992). Like Z-DNA, it is left-handed, with a deep minor groove and shallow major groove. Cytosines with *anti* glycosyl bonds and guanines with *syn* bonds alternate along each backbone strand. However, sugar puckering is reversed: cytosines are C3′-*endo*, while guanines are C2′-*endo*. In Z-DNA, the backbone chain runs parallel to the helix axis past G, and at right angles to the axis past C. In Z(WC)-DNA, this is reversed: parallel to the helix past C, and at right angles past G. Because of efficient stacking of base pairs, the logical two-base-pair structural unit in Z-DNA is $^{5'}$C–G$^{3'}$; in Z(WC)-DNA it is $^{5'}$G–C$^{3'}$. One such unit is clearly visible in the centre of Fig. 23.3.3.7. This behaviour is reflected in local twist angles:

Helix	C–G	G–C	Sum
Z-DNA	−8°	−52°	−60°
Z(WC)-DNA	−70°	+10°	−60°

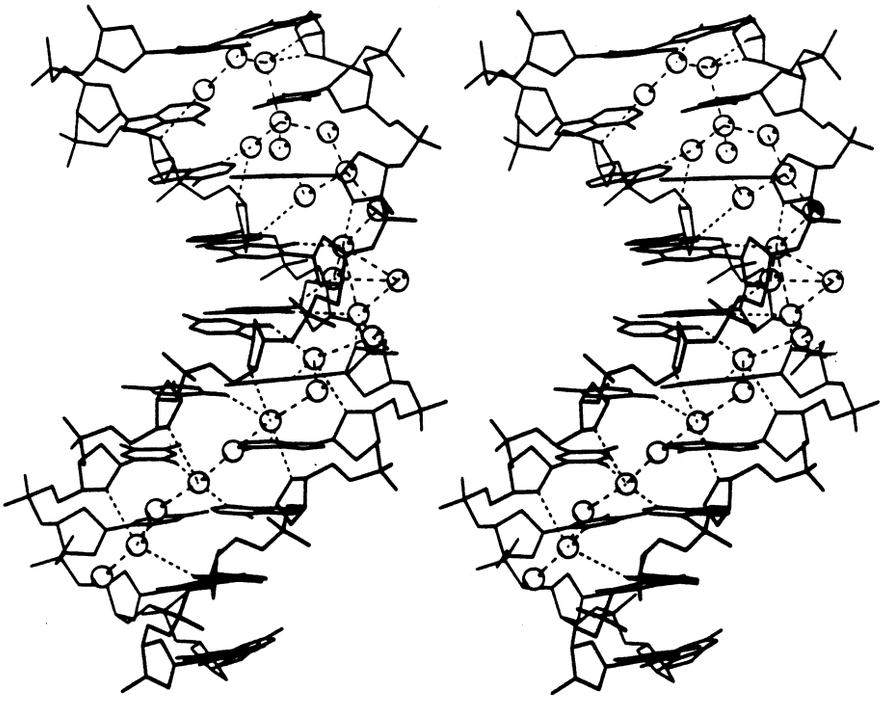

Fig. 23.3.4.1. Structure of C-A-A-A-G-A-A-A-A-G (B107). The lower half of the helix, with -A-A-A-A-G, exhibits the narrow minor groove commonly associated with the AT region of the helix and a single zigzag spine of hydration, as was first seen in C-G-C-G-A-A-T-T-C-G-C-G (B1–B6). The upper half, with C-A-A-A-G-, has the wider minor groove of general-sequence B-DNA and two separate rows of hydrating water molecules along the two walls of the wider groove.

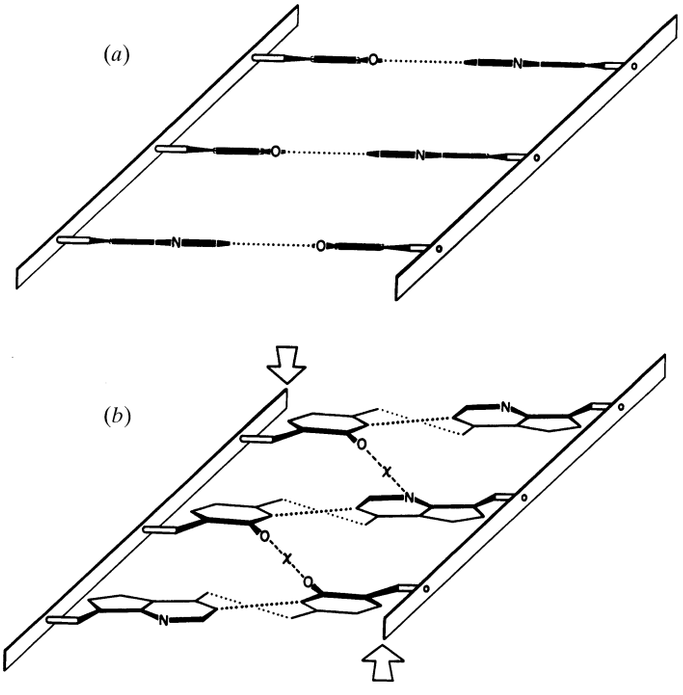

Fig. 23.3.4.2. Relationship between minor groove width and propeller twist. (*a*) View into the minor groove of B-DNA, with base pairs seen on edge and with the sugar–phosphate backbones shown schematically as inclined ladder uprights. (*b*) Consequences of propeller twisting the base pairs. Glycosyl bonds connected to sugar C1′ atoms are all displaced upward in the right strand and downward in the left strand. This shifts the backbone chains as indicated by the arrows. Hence, the gap between the chains is decreased, and the minor groove is narrowed.

The Ansevin–Wang helix has been sedulously ignored since its publication in 1990, especially by crystallographers. The Science Citation Index lists an average of *one* citation of their paper per year since publication, most commonly by spectroscopists. Ho & Mooers (1996) are almost alone among crystallographers in coupling the B-to-Z interconversion dilemma to the possible existence of a different kind of left-handed structure in long polynucleotides. Of course the Z(WC)-DNA structure, as presented here, is only a model; it could be far from the true structure in many respects. But its interest lies in the fact that a left-handed alternating helix with 'standard' backbone dirctions *can* be built with reasonable bond geometries and with properties that fit the various physical measurements as well as Z-DNA. It calls into question not the correctness of the Z-DNA structure obtained from short oligomers with free helix ends, but the relevance of that structure to the production of left-handed regions in longer duplexes with constrained ends.

23.3.4. Sequence–structure relationships in B-DNA

Two channels of information exist in B-DNA by which base sequence is expressed to the outside world. One of these is the Watson–Crick base pairing of A with T and G with C that is used in the storage of genetic information and in replication and transcription. The other channel, used in control and regulation of the expression of this genetic information, involves the hydrogen-bonding patterns of base-pair edges along the floors of the grooves and any systematic deformations of local helix structure that result explicitly from the base sequence.

The simplest and most direct expression of this second channel is the passive reading of hydrogen-bonding patterns along the floor of the major and minor grooves. This readout mechanism was first

references