

DNA structure: Revisiting the Watson–Crick double helix

Manju Bansal

Institute of Bioinformatics and Applied Biotechnology, ITPL, Bangalore 560 066, India and
Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

Watson and Crick's postulation in 1953, exactly 50 years ago, of a double helical structure for DNA, heralded a revolution in our understanding of biology at the molecular level. The fact that it immediately suggested a possible copying mechanism for the genetic material aroused the maximum interest, but the structure itself (often referred to as the B-DNA structure, by association with the corresponding X-ray fibre pattern) has also attained an almost iconic status. It was for a long time regarded as being the only biologically relevant structure, even though Watson and Crick had themselves pointed out that the structure could readily undergo changes, depending on the environment. Subsequent studies on synthetic poly-nucleotides as well as naturally occurring DNA sequences with certain repeat patterns, have established that the DNA molecule could have intrinsic, as well as environment-induced, structural polymorphism. Some of the structures show only minor differences, from the canonical structure, while a few are completely different, even in their essential features, such as handedness, base-pairing scheme or number of strands. The various DNA structures have

been characterized as A, B, C, etc. and there is a DNA structure associated with 18 other letters of the English alphabet. Only the letters F, Q, U, V and Y are left to choose from, to describe any new forms of DNA structure that may appear in future! Apart from these structures, with a one-letter 'name', there are several other generic descriptions of DNA structure and many such structures have been characterized in recent years. It is reasonable to expect that the 3 billion base pairs in the human genome (leading to a 2 m long molecule which is packaged in a microscopic size nucleus) will exhibit a variety of polymorphic forms, which could be important for the biological packaging as well as function of DNA. It is also clear that the ideal and canonical Watson–Crick DNA structure is indeed just that – an ideal and aesthetic representation of the DNA molecule, which actually possesses a chameleon-like property of adapting itself to its environment and function by twisting, turning and stretching. Some of these structures are discussed here as also the lacunae in our current understanding of the dynamics of DNA structure.

THE right-handed double helical structure (Figure 1) proposed in 1953 by Watson–Crick for deoxyribose nucleic acid (commonly known as DNA) is the most well-recognized structure for this polymeric molecule¹. While Watson–Crick were undoubtedly the first to propose an essentially correct model for DNA structure, a wide variety of available data was used by them to arrive at this 'canonical' model for DNA, in particular the nucleotide base composition data of Chargaff (Table 1) and information from the X-ray fibre diffraction pattern (Figure 2) of B-form DNA, as recorded by Rosalind Franklin². It was commonly believed for several decades, that this B-form is the only structure of DNA that has biological relevance, even though Rosalind Franklin's fibre diffraction data^{2,3} for A and B forms (Figure 2) had clearly shown that the DNA molecule could readily undergo structural transitions depending on the environment, viz. variation in relative humidity in this case. Fibre diffraction studies in the sixties and seventies also revealed several other forms of DNA structure for synthetic oligo- and poly-nucleotides, depending on the base sequence and environment. Subsequent biochemical and structural studies showed that regions of genomic DNA, under various

physiological conditions, can assume different structures, particularly when some well-defined sequence motifs or repeats occur. It was probably the characterization of such sequence repeats as 'junk DNA' by Francis Crick⁴ that put a damper on the study of such sequences till recently.

The most characteristic feature of the Watson and Crick structure for DNA (shown schematically in Figure 1 *a*), is the presence of two polynucleotide strands coiling around a common axis and being linked together by a specific hydrogen bond scheme¹ between the purine and pyrimidine bases (Figure 1 *b*), viz. adenine (A) with thymine (T) and guanine (G) with cytosine (C). This specific base pair formation (often referred to as the Watson–Crick pairing), readily accounted for Chargaff's base composition data (Table 1) and more importantly, suggested a possible copying mechanism for the genetic material. The fact that this structure could explain how the DNA molecule could code for and transmit the genetic information, aroused considerable interest among structural biologists. However, as pointed out recently by Olby⁵, it was only after the role of DNA in protein synthesis became apparent, that the biochemists began to take serious interest in the structure. Once accepted, the structure itself (often referred to as the B-DNA structure, by association with the corresponding X-ray fibre pattern) has become one of the most recognized symbols or icons of 20th century science

*For correspondence. (e-mail: mbansal@ibab.ac.in; mb@mbu.iisc.ernet.in)

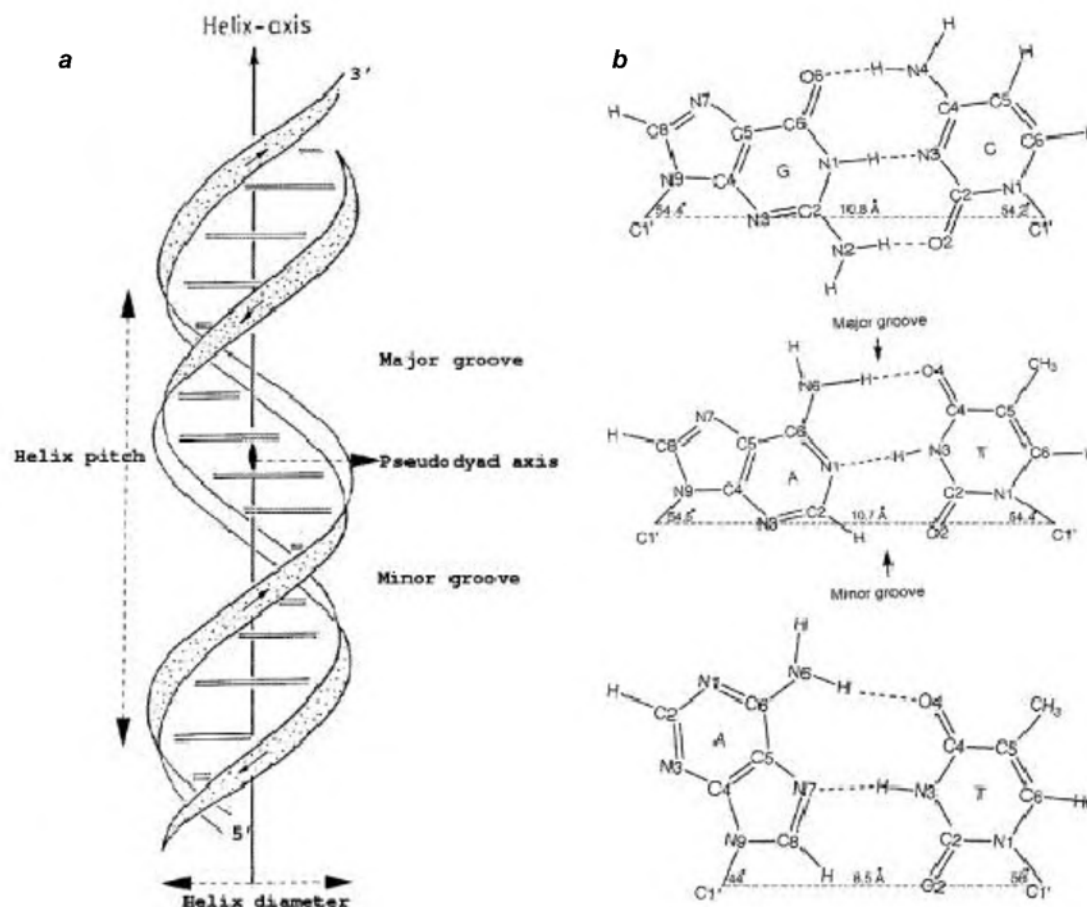


Figure 1. *a*, A schematic diagram of the Watson-Crick double helix. The base pairs are represented by horizontal bars and the sugar phosphate backbone of the two chains, related by a two-fold rotation axis perpendicular to the helix, are represented by ribbons running in opposite directions. The 5' and 3' ends are labelled for the ascending strand. The helix axis, as well as the pitch, diameter and the major and the minor grooves of the double helix, have also been indicated. *b*, Base pairs G : C, and A : T with Watson-Crick and A : T with Hoogsteen hydrogen bond scheme are shown as line diagrams. The base atoms are numbered according to the standard nomenclature and the hydrogen bonds between them are represented by dashed lines. The C1'-C1' distance and the \angle C1'-C1'-N1/N9 angles are also indicated in each case. The minor and major groove face of the base pair is indicated in the case of Watson-Crick A : T base pair. (Reproduced with copyright permission from IUCr, from paper of Ghosh and Bansal¹⁷).

and the golden jubilee of its discovery is being celebrated worldwide.

The DNA molecule is essentially a polynucleotide or a polymer chain formed by phosphate diester groups joining β -D-deoxyribose sugars through their 3' and 5' hydroxyl groups (Figure 3). The backbone of the DNA molecule thus consists of six single bonds about which rotations can take place (also indicated in Figure 3) giving rise to various possible conformations/structures for the polymeric chain. As mentioned above, the canonical Watson-Crick DNA model is a two-stranded helical structure, in which the two chains are held together by hydrogen bonds between the purine (A,G) and pyrimidine (T,C) bases. There are 10 nucleotides per turn, separated by $+36^\circ$ rotation and 3.4 \AA translation along the helix axis, in each of the two chains and the two chains are aligned

Table 1. A representative sample of Erwin Chargaff's 1952 data, listing the base composition of DNA from various organisms.

An interesting feature of this data was that the amount of adenine is nearly equal to that of thymine and the amount of guanine nearly equals that of cytosine. This feature which is referred to as Chargaff's rule, was crucial for Watson and Crick to arrive at their double helical model for DNA and is found to be nearly true in most cases. Any significant deviation from this rule (as in ϕ X174 below) implies that the DNA is single stranded

Source	Mol% of bases				Ratios		%GC
	A	G	C	T	A/T	G/C	
ϕ X174	24.0	23.3	21.5	31.2	0.77	1.08	44.8
Maize	26.8	22.8	23.2	27.2	0.99	0.98	46.1
Octopus	33.2	17.6	17.6	31.6	1.05	1.00	35.2
Chicken	28.0	22.0	21.6	28.4	0.99	1.02	43.7
Rat	28.6	21.4	20.5	28.4	1.01	1.00	42.9
Human	29.3	20.7	20.0	30.0	0.98	1.04	40.7

in mutually anti-parallel orientations (Figures 1 *a* and 4). In 1953 itself, Rosalind Franklin had shown from fibre diffraction studies^{2,3} that DNA can inter-convert between two well-defined forms, viz. A and B (Figure 2). The molecular structures corresponding to these two forms were later shown to be essentially similar in their handedness, chain orientation and hydrogen bonding scheme. Subsequently it has become clear that the DNA molecule has an enormous ability to undergo structural changes depending on its environment by twisting, turning and stretching, leading to a pantheon of DNA structures⁶. Several of these structural polymorphs of DNA have now been experimentally characterized using X-ray diffraction, NMR or other spectroscopic studies and are found to vary considerably from the Watson–Crick type structure. Some of the structures, such as the A, B', C and D forms of DNA, differ only slightly in their local geometries, from the Watson–Crick base-paired duplex structure. Other structures are completely different, even in their essential features, such as handedness, base-pairing scheme or number of strands. This wide range of structural variability is possible due to the inherent conformational flexibility of the polynucleotide backbone^{6–8}, arising due to rotations about the various single bonds (Figure 3). Also, in addition to the three hydrogen bonded GC and two hydrogen

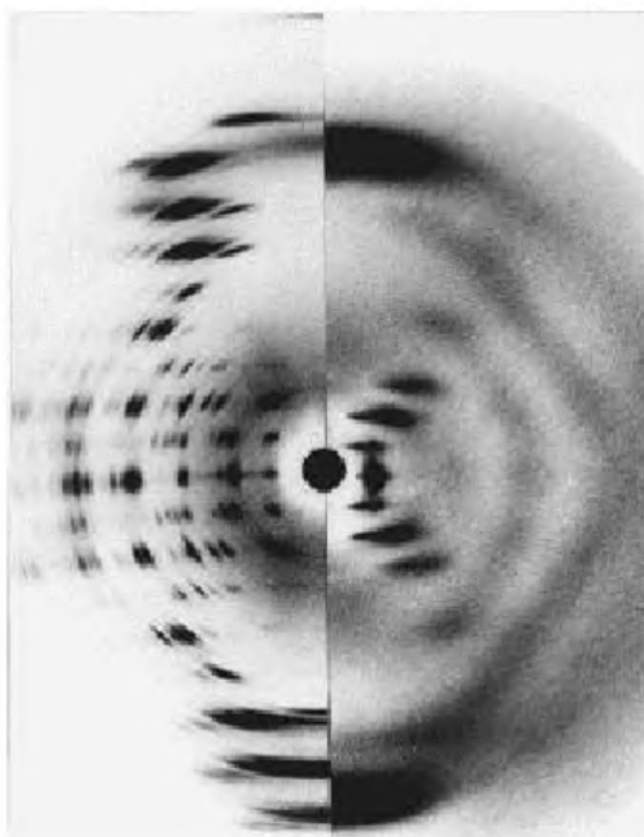


Figure 2. A composite picture showing the X-ray fibre diffraction patterns for the A-form (left half) and B-form (right half) of DNA (taken from www.mpinf-heidelberg.mpg.de).

bonded AT base pairs, as proposed by Watson and Crick (Figure 1 *b*), there are 27 other distinct possibilities of forming at least two hydrogen bonds between any pair of bases⁷. One of the earliest reported crystal structures of a base pair between adenine and thymine in fact showed a non-Watson–Crick hydrogen bond pattern and is referred to as the Hoogsteen pair⁹ – wherein the N7 atom of adenine is involved in pairing with the thymine base (as shown in Figure 1 *b*) rather than the N1 atom, as is the case in Watson–Crick AT base pair.

One of the best characterized, non-B form, DNA structure is the A-form, corresponding to the X-ray pattern³ recorded for fibres of the sodium salt of calf thymus DNA under conditions of low relative humidity (shown on the left in Figure 2). The structure which gives rise to this pattern^{3,10} is also a right-handed double helix, with 11 residues per turn and a translation of 2.56 Å along the helix axis. The Watson–Crick hydrogen bonded base pairs in this structure are considerably displaced from the helix axis, as well as being inclined to it, both in fibres of polynucleotides¹⁰, as well as oligonucleotide crystals¹¹ (Figure 4). The B-form on the other hand is observed for DNA fibres, under conditions of high relative humidity^{2,12} and is generally characterized in fibres and single crystals, by about 10 residues per turn, as in the standard Watson–Crick model described above. In this structure the base pairs are translated by about 3.4 Å and are located nearly astride the helix axis and normal to it^{12,13} as seen in Figure 4. Two other forms of double-stranded helices with right-handed twist, but with slightly different structural parameters are the C and D-DNA structures. C-form was first observed for the lithium salt of calf-thymus DNA and has about 9.3 residues per turn of helix¹⁴, while D-DNA has 8 residues in one turn and is observed for sodium salts of poly[d(A-T)].poly[d(A-T)] as well as poly[d(A-A-T)].poly[d(A-A-T)] sequences¹⁵. All the four forms (A–D) are stable within a range of ionic and humidity conditions. It is also interesting to

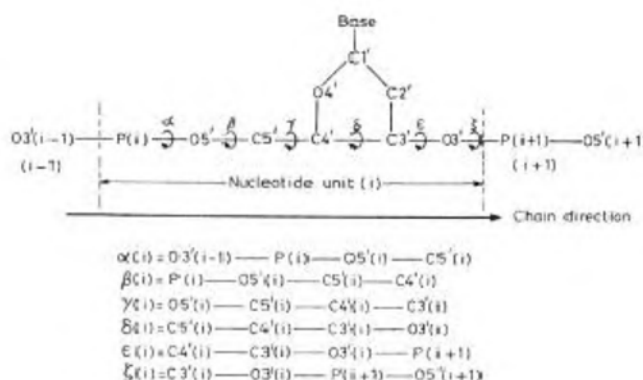


Figure 3. A schematic drawing showing the repeating unit of a polynucleotide chain and the nomenclature for the six backbone torsion angles – α , β , γ , δ , ϵ , ξ . The 5' to 3' direction of chain progression is also indicated. The purine and pyrimidine bases are attached to the C1' atom of the furanose sugar ring in each nucleotide.

note that a random sequence DNA, in solution, has a helical repeat intermediate between the A and B forms, viz. ~ 10.5 units per turn¹⁶. In addition, most single crystal studies of B-DNA molecules, with a variety of base sequences show a sequence-dependent variability along their length leading to the overall molecule very often being slightly bent or curved. A detailed analysis of all available nucleic acid/oligonucleotide structural data also suggests that, the A, B or C type dinucleotide steps may occur at contiguous locations in a single molecule, depending on the base sequence.

It has now been established that there are about 3 billion base pairs in the human genome, leading to a 2 m long thread of DNA, which is compacted by a factor of about 10^{-6} when embedded in a human cell nucleus. The free double helical DNA molecule, or even the X-shaped chromosome are seen only for short periods during the entire cell cycle. Hence it is clear that the inherent flexibility, as well as protein-induced structural changes of the DNA molecule, must play an important role in enabling DNA to pack itself and also to perform its various functions. A few of the rather unusual and interesting variants of DNA structure, which have been discussed in the literature, are described here, while a complete list or glossary of currently known DNA structures has been recently published elsewhere¹⁷.

It is interesting that the first crystal structures of dinucleotides with ribose sugars, viz. ApU and GpC showed that these molecules form mini double helices with Watson–Crick type base pairs^{18,19}. However, some of the fundamental tenets of the Watson–Crick structure, viz. the uniform nature and right-handedness of the DNA double helix, were overthrown when the first few crystal structures of deoxy-oligonucleotides were reported in the late seventies. One of the first oligonucleotide structures reported was of the tetra-nucleotide d(pApTpApT), which showed a non-uniform structure, with the central d(TpA) fragment being unstacked and taking up a non-helical conformation²⁰. The next base-paired DNA double helix reported from single-crystal X-ray diffraction analysis of d(CGCGCG), also differed from the Watson–Crick structure in several important features, particularly the handedness of the helix²¹. Subsequently dubbed the Z-form DNA, to describe the characteristic zigzag path traced by the backbone, it is a left-handed duplex structure with a dinucleotide repeat unit (Figure 4). Generally confined to alternating purine (G) and pyrimidine (C) sequences, it has six dinucleotides per turn and exhibits the characteristic zigzag backbone due to distinctly different conformations for the two residues in the dinucleotide repeat^{21,22}.

New structures with non-Watson–Crick base pairs, consisting of purine–pyrimidine and other mismatch base combinations, are also being reported on a regular basis. While this was an accepted fact for multi-stranded triplex or quadruplex structures and isolated base pairs in duplexes, a hexamer with the sequence d(ATATAT) has been

recently reported²³ in which all six AT base pairs in the duplex show the Hoogsteen hydrogen bond scheme (Figure 1 b). This indicates that non-Watson–Crick base pairs can be readily formed and are quite stable, except perhaps for the G.C. base pair, which is unique in that it is stabilized by three hydrogen bonds in the standard Watson–Crick arrangement. The occurrence of non-Watson–Crick base pairs does introduce some irregularity in the double helical structure, which may in fact play a role in sequence-dependent recognition of DNA.

As mentioned above, several model structures were reported from fibre diffraction studies, which could only be described by helical arrangements with more than two chains, giving rise to triple and tetraplex or quadruplex structures. A DNA double helix of poly d(A).poly d(T) was under certain conditions shown to accommodate a third chain, of poly d(T) in its major groove and stabilized by the formation of hydrogen bonds with the purine (or A) strand²⁴. This type of structure first proposed from X-ray diffraction studies of polynucleotide fibres formed with a stoichiometric ratio of 1 : 2 of poly(purine) and poly(pyrimidine) chains, has subsequently been shown from NMR studies, to also form for other types of sequences^{25,26} containing stretches of purines or pyrimidines, with the third strand being either pyrimidine or purine rich (Figure 5). Similarly an intra-molecular triple helical structure can be formed at low pH conditions, by DNA sequences containing long stretches of poly(purine).poly(pyrimidine) and is believed to play a role in transcriptional control of gene expression²⁷. In this structure, termed as H-DNA, the pyrimidine rich chain partly dissociates from its complementary strand and folds back on itself, to lie in the major groove of the Watson–Crick duplex and hydrogen bonds to the purine rich strand.

A family of four-stranded quadruplex DNA structures, with Hoogsteen type hydrogen bonds between four guanines forming stable G-tetrads, has also been characterized for a variety of G-rich DNA sequences which are present in naturally occurring DNA, particularly in telomeric regions of chromosomes^{28,29}. These structures can be formed with a parallel arrangement of four strands, as observed for poly(G) and some G-tract containing oligonucleotides³⁰ as well as by folded back chains of oligonucleotides (Figure 6) with sequences which have runs of G's with A/T interruptions^{31,32}. Another unusual, four-stranded DNA structure results from inter digitation of two parallel-stranded duplexes of oligo (C), stabilized by hemi-protonated C–C⁺ base pairing, arranged in an anti-parallel orientation and has been dubbed as I-DNA^{33,34}.

These are only a few of the well studied DNA structures, but a whole variety of other structures have also been reported during the last 50 years. A survey of all reported DNA structures¹⁷ characterized by various physico-chemical or biochemical experiments is indeed revealing. Following the nomenclature adopted by the original fibre diffraction workers, of assigning the letters A, B, C,

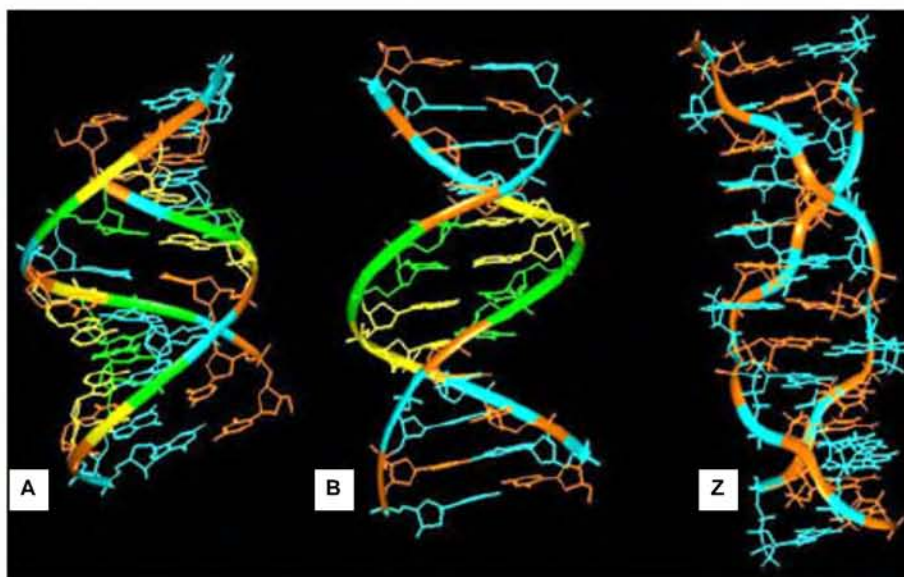


Figure 4. Representative crystal structures for A-, B- and Z-DNA are shown here (coordinates correspond to structures described in refs 11, 13 and 22 respectively). The nucleotides are colour-coded (cytosine in orange, guanine in cyan, thymine in green, and adenine in yellow) and a ribbon is superposed on the backbones, connecting the phosphorus atoms. A and B-DNA are both right handed, nearly uniform double helical structures, while Z-DNA is a left handed double helix with a dinucleotide repeat and the backbone follows a zig-zag path.

etc. to identify the various polymorphic forms of DNA, there has been a tendency to characterize each structure by a one-letter code and currently there are DNA structures associated with 21 of the 26 letters of the English alphabet! Only the letters F, Q, U, V and Y have not been used and are still available, to describe any new forms of DNA structure that may appear in future.

There are also several other generic descriptions of DNA. For example: Form V DNA is used to describe supercoiled DNA, which may contain regions of right-handed and left-handed DNA as in pBR322 plasmid³⁵. During genetic recombination, one of the strands from each of two duplex DNA molecules exchange, to form a four-way junction structure, known as a Holliday junction³⁶. Its occurrence has recently been confirmed by crystal structure analysis of a protein–DNA complex³⁷, as well as in a free oligonucleotide (shown in Figure 7) with an inverted repeat sequence³⁸ d(CCGGTACCGG). It is interesting to note that some related decamer sequences, for example d(CCGGCGCCGG) show a normal duplex structure, with two neighbouring molecules packed in an X-arrangement, but no strand exchange taking place³⁹, indicating the sequence-specific nature of this structure.

Thus the helical structure of DNA is seen to be highly adaptable and can assume various forms. What has been captured by the X-ray diffraction, as well as other structural studies and described above are essentially snapshots of static DNA structures. They do not however provide any insight into how the various structures interconvert between the various forms, to perform the diverse functions. In addition, the DNA in eukaryotic systems is wrapped

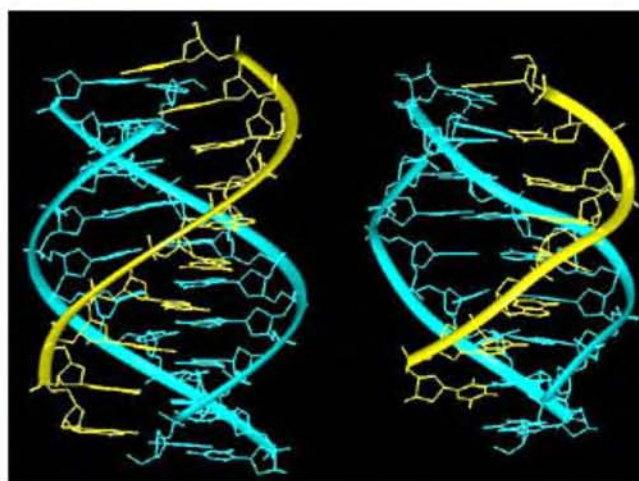


Figure 5. A model structure for a DNA triple helix with C.G.G. triplets²⁵ is shown (on left), along with a triple helix structure reported from NMR (on right), which contains a mixture of T.A.T. and C.G.G. triplets²⁶. The Watson–Crick duplex is shown in cyan, while the third strand is shown in yellow, with ribbons tracing the backbone.

around a histone protein core to form a bead-like structure called nucleosome. A crystal structure of this complex has been solved recently, but the higher order organization of a string of nucleosomes, finally leading to the X-shaped chromosome, is still not firmly established. Like the iconic representation of the DNA molecule as a uniform double helix, it is now quite common to explain complex biological processes, like transcription, replication and recombination, through simple drawings. However, as pointed out recently by Alberts⁴⁰, many more

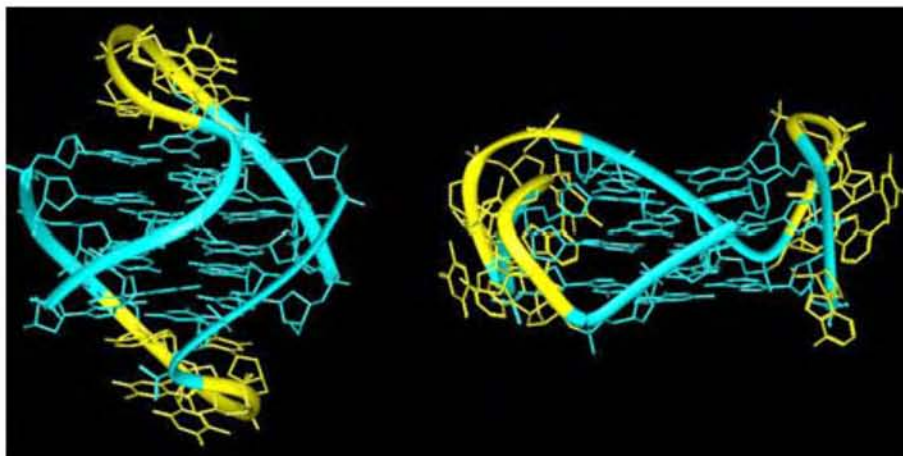


Figure 6. A G-quadruplex structure, formed by an association of two hairpins³¹ (on left) and a parallel G-quadruplex with TTA loops³² (on right). In both structures, the guanine tetrads are shown in cyan, with the loops shown in yellow.

3-D structures, new biophysical information (in the form of rate constants for each reaction and interaction) and genetic experimentation is required to fully understand the thermodynamics of these processes.

One of the major unsolved problems in structural molecular biology has been to explain the forces behind the unwinding/untwisting of long DNA duplexes during replication. A large number of enzymes (such as topoisomerases, helicases, etc.) have been identified, which seem to facilitate everything from removal of DNA from histone core, to its unwinding, unzipping and the actual replication or transcription. It is clear that such enzymes must generate considerable force to unzip the two parental strands in the duplex. Recently some novel experiments, performed by manipulating single molecules of DNA, have provided new insights about how this may be achieved through external forces and torques. When a DNA molecule is subjected to a tension > 60 pN, it undergoes a reversible transition to an extended form that is about 70% longer than B-DNA and has a much lower twist value. Whether this overstretched DNA is simply an extended Watson–Crick paired duplex⁴¹ or two independent strands⁴² in which the bases on the two strands may be partially intercalated remains unclear. These structures also seem capable of taking up an inside-out helical structure DNA^{6,43}, with sugar-phosphate chains coiling around each other to form a close-packed core and the bases exposed on the outside. This type of structure is similar to the original model for DNA, proposed in 1953 by Linus Pauling⁴⁴ (hence referred to as P-DNA) and subsequently abandoned in favour of the Watson–Crick model. These single molecule studies seem to indicate that DNA transcription and replication may involve a tugging or stretching force which could facilitate unwinding and make the bases accessible for further interaction to take place.

Double-stranded structures have also been proposed, wherein the two strands are not wound/twisted around each

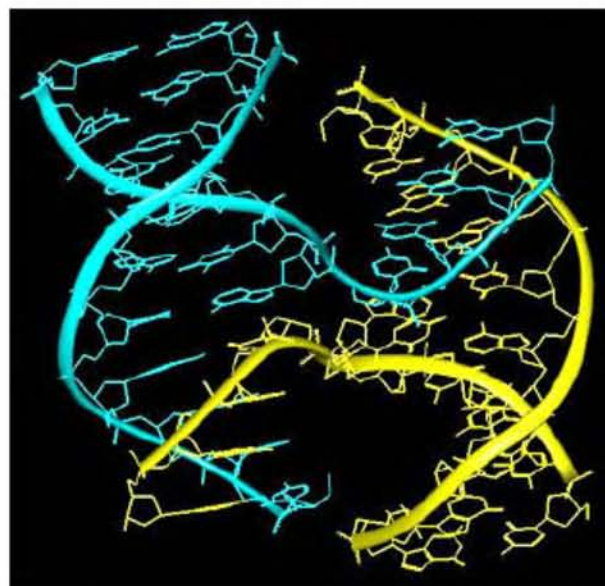


Figure 7. A Holliday junction structure for a decamer³⁸ with the inverted repeat sequence d(CCGGTACCGG) with one duplex in cyan and the other in yellow. The strand exchange between the two duplexes occurs at the AC step, with very little disruption of the base pairs.

other, but contain helical fragments with alternating right and left handedness (known as the side-by-side or S–B–S models) potentially facilitating uncoiling of the double helix during DNA replication^{45–47}. Such models have not yet been confirmed by independent structural studies but still engage the attention of many theoretical biologists^{48,49} and can readily explain some of the observed phenomenon involving circular duplex, as well as long linear DNA without invoking the idea of topological linkage between the DNA strands, as is the case with the canonical Watson–Crick type structure.

Thus after half-century of research on DNA, it may appear that everything about its structure is known. This

is only partly true, since only the essential features of various forms of the structure and the mechanical properties of the molecule have been well characterized. However, the dynamic changes required for the DNA molecule, which in the cell is generally present in the form of a tightly packed entity, closely associated with several proteins, to unravel, unwind and perform its diverse functions, remain unclear. The interaction of DNA with other cellular components and the elucidation of various transition state structures may yet reveal new features about the DNA structure itself. The ability of DNA molecules to assume various forms is already being used to construct nanomachines and to direct the assembly of highly complex structured materials⁵⁰. DNA is also being considered as a medium for processing complex information, viz. serve as a tool in computation⁵⁰. In future, new revelations about the properties of this versatile molecule, as well as novel uses for it, may very well come from non-biological experiments.

1. Watson, J. D. and Crick, F. H. C., Molecular structure for nucleic acids: A structure for deoxyribose nucleic acid. *Nature*, 1953, **171**, 737–738.
2. Franklin, R. E. and Gosling, R., Molecular configuration in sodium thymonucleate. *Nature*, 1953, **171**, 740–741.
3. Franklin, R. E. and Gosling, R., The structure of sodium thymonucleate fibres. I. The influence of water content. *Acta Crystallogr.*, 1953, **6**, 673–677.
4. Crick, F. H. C., Wang, J. C. and Bauer, W. R., Is DNA really a double helix? *J. Mol. Biol.*, 1979, **129**, 449–461.
5. Olby, R., Quiet debut for the double helix. *Nature*, 2003, **421**, 402–405.
6. Bansal, M., DNA structure: Yet another avatar? *Curr. Sci.*, 1999, **76**, 1178–1181.
7. Saenger, W., In *Principles of Nucleic Acid Structure* (ed. Cantor, C. R.), Springer-Verlag, 1983.
8. Bansal, M. and Sasisekharan, V., Molecular model-building of DNA: Constraints and restraints. In *Theoretical Chemistry of Biological Systems* (ed. Naray Szabo), Elsevier, Amsterdam, 1986, pp. 127–218.
9. Hoogsteen, K., The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Crystallogr.*, 1963, **16**, 907–916.
10. Fuller, W., Wilkins, M. H. F., Wilson, H. R. and Hamilton, L. D., The molecular configuration of deoxyribonucleic acid. iv. X-ray diffraction study of the a form. *J. Mol. Biol.*, 1965, **12**, 60–80.
11. Bingman, C. A., Zon, G. and Sundaralingam, M., Crystal and molecular structure of the A-DNA dodecamer d(CCGTACGTACGG). Choice of fragment helical axis. *J. Mol. Biol.*, 1992, **227**, 738.
12. Langridge, R., Marvin, D. A., Seeds, W. E., Wilson, H. R., Hooper, C. W., Wilkins, M. H. F. and Hamilton, L. D., The molecular configuration of deoxyribonucleic acid: Molecular models and their Fourier transforms. *J. Mol. Biol.*, 1960, **2**, 38–64.
13. Drew, H. R., Wing, R. M., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R. E., Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. USA*, 1981, **78**, 2179.
14. Marvin, D. A., Spencer, M., Wilkins, M. H. F. and Hamilton, L. D., The molecular configuration of deoxyribonucleic acid. III. X-ray diffraction study of the C form of the lithium salt. *J. Mol. Biol.*, 1961, **3**, 547–565.
15. Arnott, S., Chandrasekaran, R., Hukins, D. W. L., Smith, P. J. C. and Watts, L., Structural details of double-helix observed for DNAs containing alternating purine and pyrimidine sequences. *J. Mol. Biol.*, 1974, **88**, 523–533.
16. Wang, J. C., Helical repeat of DNA in solution. *Proc. Natl. Acad. Sci. USA*, 1979, **76**, 200–203.
17. Ghosh, A. and Bansal, M., A glossary of DNA structures from A to Z. *Acta Crystallogr.*, 2003, **D59**, 620–626.
18. Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Kim, J. J. P. and Rich, A., RNA double-helical fragments at atomic resolution. I. The crystal and molecular structure of sodium adenylyl-3',5'-uridine hexahydrate. *J. Mol. Biol.*, 1976, **104**, 109–144.
19. Rosenberg, J. M., Seeman, N. C., Day, R. O. and Rich, A., RNA double-helical fragments at atomic resolution. II. The crystal structure of sodium guanylyl-3',5'-cytidine nonahydrate. *J. Mol. Biol.*, 1976, **104**, 145–167.
20. Vishwamitra, M., Kennard, O., Jones, P. G., Sheldrick, G. M., Salisbury, S., Falvello, L. and Shakked, Z., DNA double helical fragment at atomic resolution. *Nature*, 1978, **273**, 687–688.
21. Wang, A. H. J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., VanBoom, J. H., Marel, G. V. D. and Rich, A., Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature*, 1979, **282**, 680–686.
22. Ban, C., Ramakrishnan, B. and Sundaralingam, M., Crystal structure of the self-complementary 5'-purine start decamer d(GCGCGCGCGC) in the Z-DNA conformation. I. *Biophys. J.*, 1996, **71**, 1215.
23. Abrescia, N. G., Thompson, A., Huynh-Dinh, T. and Subirana, J. A., Crystal structure of an antiparallel DNA fragment with Hoogsteen base pairing. *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 2806–2811.
24. Arnott, S. and Selsing, E., Structures for the polynucleotide complexes poly(dA) with poly(dT) and poly(dT) with poly(dA) with poly(dT). *J. Mol. Biol.*, 1974, **88**, 509–521.
25. Vlieghe, D., Van Meervelt, L., Dautant, A., Gallois, B., Precigoux, G. and Kennard, O., Parallel and antiparallel (G.GC)₂ triple helix fragments in a crystal structure. *Science*, 1996, **273**, 1702.
26. Radhakrishnan, I. and Patel D. J., Solution structure of a purine. pyrimidine DNA triplex containing G.GC and T.A.T triples. *Structure*, 1993, **1**, 135–152.
27. Htun, H. and Dahlberg, J. E., Topology and formation of triple-stranded H-DNA. *Science*, 1989, **243**, 1571–1576.
28. Blackburn, E. H. and Szostak, J. W., The molecular structure of centromeres and telomeres. *Annu. Rev. Biochem.*, 1984, **53**, 163–194.
29. Mohanty, D. and Bansal, M., Conformational polymorphism in G-tetraplex structures: Strand reversal by base flipover or sugar flipover. *Nucleic Acids Res.*, 1993, **21**, 1767–1774.
30. Horvath, M. P. and Schultz, S. C., DNA G-quartets in a 1.86 Å resolution structure of an oxytricha nova telomeric protein-DNA complex. *J. Mol. Biol.*, 2001, **310**, 367–377.
31. Haider, S. M., Parkinson, G. and Neidle, S., Crystal structure of the potassium form of an Oxytricha nova G-quadruplex. *J. Mol. Biol.*, 2002, **320**, 189.
32. Parkinson, G. N., Lee, M. P. H. and Neidle, S., Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, 2002, **417**, 876.
33. Gehring, K., Leroy, J. L. and Gueron, M., A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature*, 1993, **363**, 561–565.
34. Chen, X., Ramakrishnan, B., Rao, S. T. and Sundaralingam, M., Binding of two distamycin A molecules in the minor groove of an alternating B-DNA duplex. *Nature Struct. Biol.*, 1994, **1**, 169–175.
35. Stettler, U. H., Weber, H., Koller, T. and Weissmann, C., Preparation and characterization of form V DNA, the duplex DNA result-

- ing from association of complementary, circular single-stranded DNA. *J. Mol. Biol.*, 1979, **131**, 21–40.
36. Holliday, R., A mechanism for gene conversion in fungi. *Genet. Res. Camb.*, 1964, **5**, 282–304.
37. Ariyoshi, M., Nishino, T., Iwasaki, H., Shinagawa, H. and Morikawa, K., Crystal structure of the holliday junction DNA in complex with a single RuvA tetramer. *Proc. Natl. Acad. Sci. USA*, 2002, **97**, 8257–8262.
38. Eichman, B. F., Vargason, J. M., Mooers, B. H. and Ho, P. S., The Holliday junction in an inverted repeat DNA sequence: sequence effects on the structure of four-way junctions. *Proc. Natl. Acad. Sci. USA*, 2000, **97**, 3971–3976.
39. Heinemann, U., Alings, C. and Bansal, M., Double helix conformation, groove dimensions and ligand binding potential of a G/C stretch in B-DNA. *EMBO J.*, 1992, **11**, 1931–1939.
40. Alberts, B., DNA replication and recombination. *Nature*, 2003, **421**, 431–435.
41. Cluzel, P., Lebrun, A., Heller, C., Lavery, R., Viovy, J. L., Chatenay, D. and Caron, F., DNA: an extensible molecule. *Science*, 1996, **271**, 792–794.
42. Williams, M. C., Rouzina, I. and Bloomfield, V., Thermodynamics of DNA interactions from single molecule stretching experiments. *Acc. Chem. Res.*, 2002, **35**, 159–166.
43. Allemand, J. F., Bensimon, D., Lavery, R. and Croquette, V., Stretched and overwound DNA forms a Pauling-like structure with exposed bases. *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 14152–14157.
44. Pauling, L. and Corey, R. B., A proposed structure for the nucleic acids. *Proc. Natl. Acad. Sci. USA*, 1953, **39**, 84–96.
45. Sasisekharan, V. and Pattabiraman, N., Double-stranded polynucleotides: Two typical alternative conformations for nucleic acids. *Curr. Sci.*, 1976, **45**, 779–783.
46. Sasisekharan, V., Pattabiraman, N. and Gupta, G., Some implications of an alternative structure for DNA. *Proc. Natl. Acad. Sci. USA*, 1978, **75**, 4092–4096.
47. Rodley, G. A., Scobie, R. S., Bates, R. H. and Lewitt, R. M., A possible conformation for double-stranded polynucleotides. *Proc. Natl. Acad. Sci. USA*, 1976, **73**, 2959–2963.
48. Biegeleisen, K., Topologically non-linked circular duplex DNA. *Bull. Math. Biol.*, 2002, **64**, 589–609.
49. Delmonte, C. S. and Mann, L. R. B., *Curr. Sci.*, 2003, **85**, (this issue).
50. Seeman, N. C., DNA in a material world. *Nature*, 2003, **421**, 427–431.

ACKNOWLEDGEMENTS. This work was supported by financial assistance from CSIR, New Delhi. I thank Mr Paramsvam for preparation of the figures.

Received 13 October 2003; accepted 14 October 2003