Image Library of Biological Macromolecules
JENA

celebrates

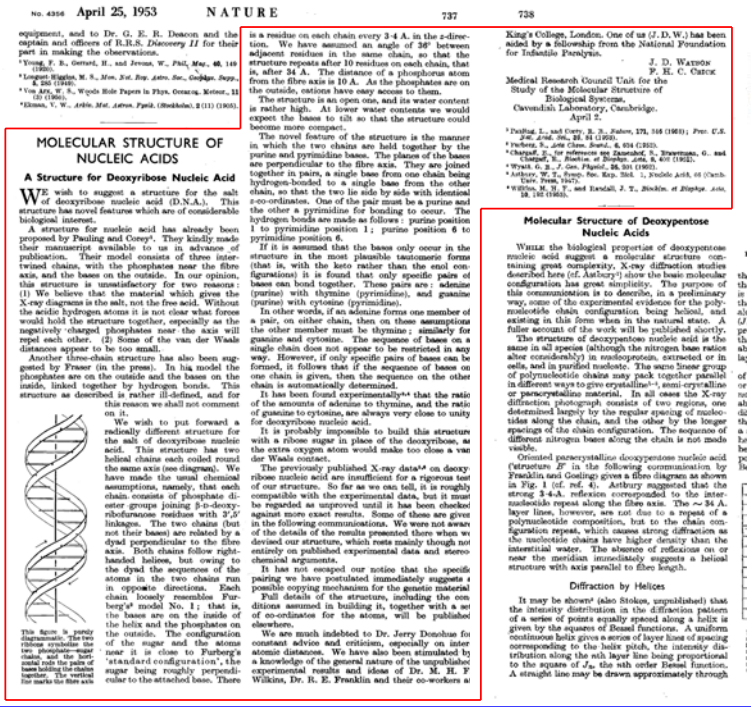
50 Years of DNA Double Helix Structure

On Wednesday, April 2, 1953, **Francis Crick** and **James D. Watson**, then working at the **Medical Research Council Unit** in the **Cavendish Laboratory** of the **University of Cambridge/UK**, submitted a manuscript of about 900 words on the three-dimensional structure of desoxyribonucleic acid (DNA) to the journal **Nature**. The paper did appear on April 25, 1953. This work marks the birthday of **modern molecular biology** as well as of **structural biology**. Together with **Maurice H.F. Wilkins,** whose fiber diffraction data, besides related data by **Rosalind Franklin,** were of utmost importance, they were awarded the **Nobel Prize for Physiology and Medicine** in **1962**. Perhaps, the discovery of the DNA double helix structure, including ist significance for information transfer in biological systems, is one of the most important scientific achievements of the 20th century.

The original diffraction data were obtained from oriented fibres. They can thus only result in an average model structure. In 1981 the first single-crystal B-DNA structure was reported by the group of **R.E. Dickerson** (PDB code: 1bna). This was the first time that sequence-specific geometry parameters of DNA could be obtained.

Currently, the Protein Data Bank (PDB) and the Nucleic Acid Database hold about 1400 entries of three-dimensional DNA structures including DNA-protein complexes. They show a marked geometrical diversity and are by no means that regular as displayed in the original Watson-Crick paper from 1953.

A few examples of the structural diversity of DNA are shown here.


No. 4356 April 25, 1953 NATURE 737

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey. They kindly made their manuscript available to us in advance of publication. Their model consists of three inter-twined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate di-ester groups joining β-D-deoxy-ribofuranose residues with 3',5' linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4 A. in the z-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at King's College, London. One of us (J. D. W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J. D. WATSON
F. H. C. CRICK

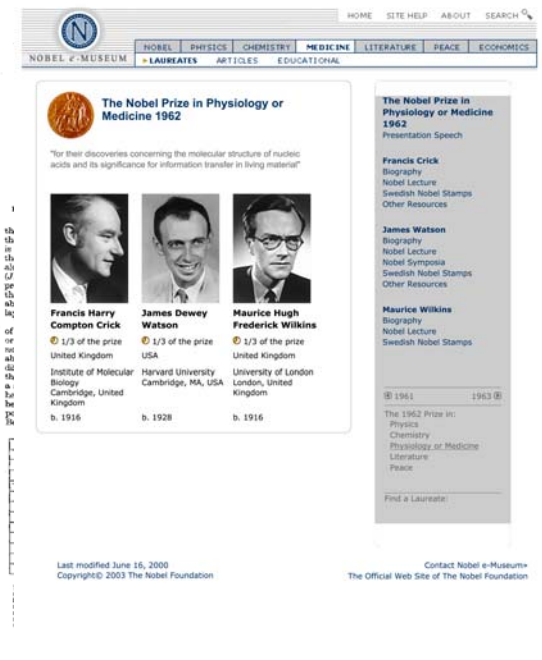
Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems,
Cavendish Laboratory, Cambridge.
April 2.

Molecular Structure of Deoxypentose Nucleic Acids

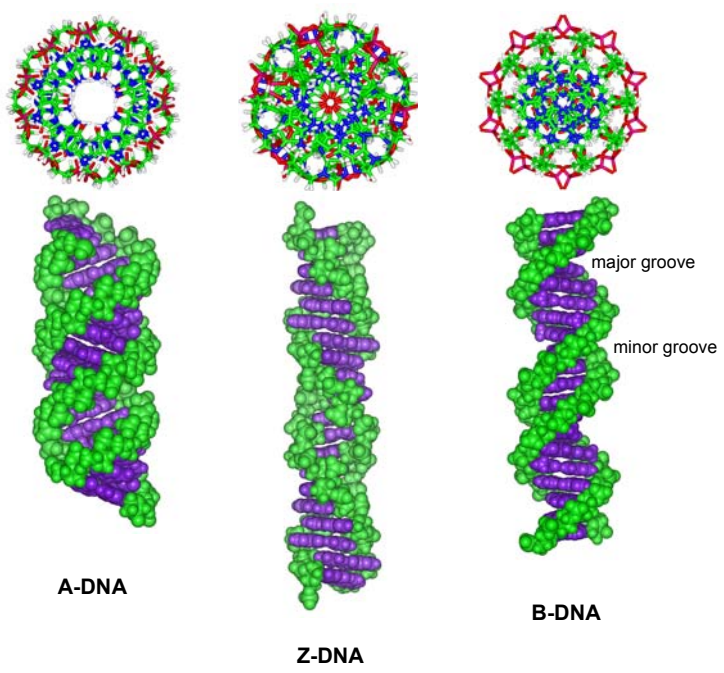
Oriented paracrystalline deoxypentose nucleic acid ("structure B" in the following communication by Franklin and Gosling) gives a fibre diagram as shown in Fig. 1 (cf. ref. 4). Astbury suggested that the strong 3·4-A. reflexion corresponded to the inter-nucleotide repeat along the fibre axis. The ~ 34 A. layer lines, however, are not due to a repeat of a polynucleotide composition, but to the chain configuration repeat, which causes strong diffraction as the nucleotide chains have higher density than the intersticial water. The absence of reflexions on or near the meridian immediately suggests a helical structure with axis parallel to fibre length.

Diffraction by Helices

It may be shown (also Stokes, unpublished) that the intensity distribution in the diffraction pattern of a series of points equally spaced along a helix is given by the squares of Bessel functions. A uniform continuous helix gives a series of layer lines of spacing corresponding to the helix pitch, the intensity distribution along the nth layer line being proportional to the square of J_n, the nth order Bessel function. A straight line may be drawn approximately through

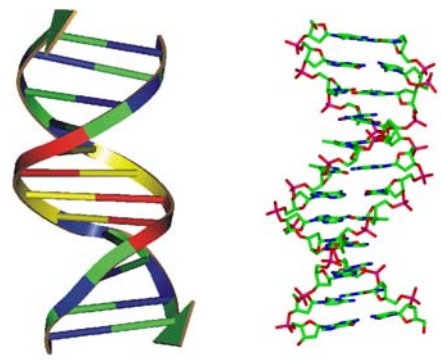


Ideal DNA conformations as top and side views



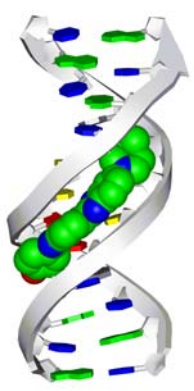
Color code: bases - **magenta**, backbone - **green**.

Two representations of the first single-crystal B-DNA structure
(PDB code: 1bna).

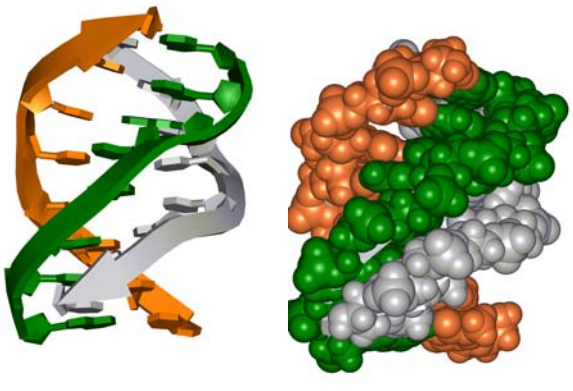


Color code for the left structure (nucleotide specific): **G**, **C**, **A**, **T**.
Color code for the right structure (atom specific): **C**, **N**, **O**, **P**.

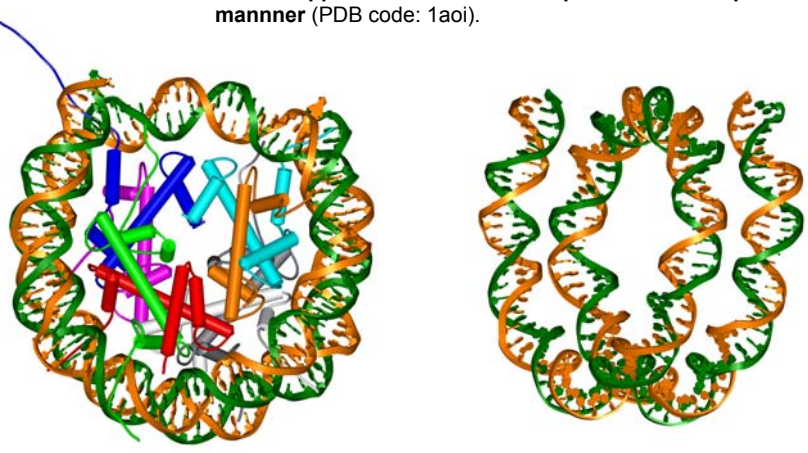
Complex between the minor-groove binding drug Hoechst 33258 and DNA (Code: 128d).



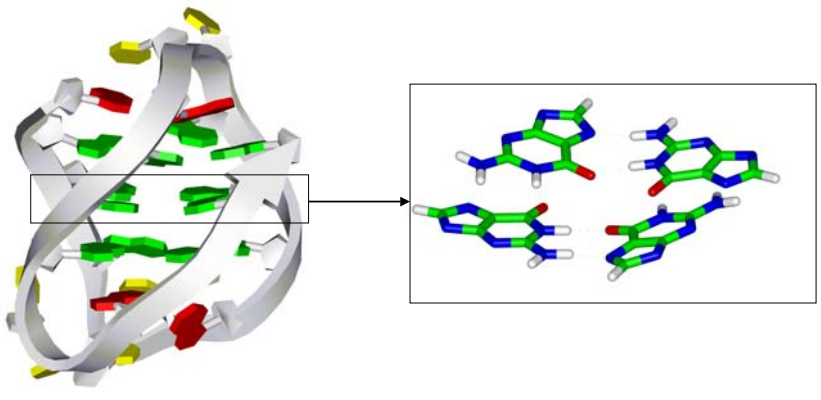
Two representations of DNA-triplex structure. The third strand is used, for example, for gene targeting (PDB code: 1d3x).



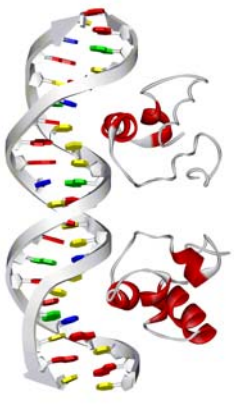
The structure of the nucleosome core particle gives an impression of the spatial arrangement of genomic DNA . The DNA double helix is wrapped around the histone proteines in a superhelical mannner (PDB code: 1aoi).



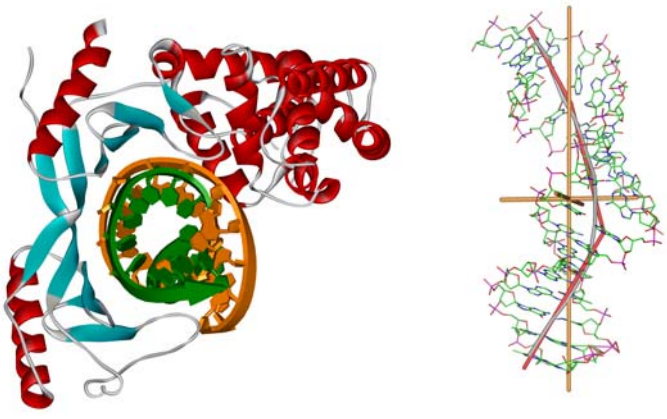
Guanine-rich sequences at the ends of chromosomes, the telomeres, contribute to genome stability and therefore play a role in ageing processes and the development of cancer. They form tetraplex structures. The basic building blocks of the tetraplexes are guanine-tetrads (PDB code: 143d).



Complex between the repressor protein *cro* of bacteriophage 434 and of a DNA with an operator sequence OR1. The protein with a so-called helix-turn-helix motif binds to the major groove of the DNA (PDB code: 3cro).



Top view of a complex between the TATA-box-bindung protein (TBP), transcription factor IIB, and a DNA target.
In addition, results of a quantitative bending analysis for the DNA part are shown. (PDB code: 1ais).



Nucleotides (DNA building blocks): **C**, **G**, **T**, **A**. Usualy two nucleotides form a pair. In normal DNA structures the standard Watson-Crick pairs **C**.**G** und **T**.**A** do occur. A nucleotide consists of a base part, a sugar ring and a phosphate group.