The April 25 1953 issue of the journal *Nature* carried three classic papers collectively entitled 'Molecular Structure of Nucleic Acids' published back to back. The paper by Rosalind Franklin and R G Gosling was preceded by the paper of Watson and Crick that announced their famous model of DNA and one from the Wilkins group. For most biologists, the paper by Watson and Crick is very lucid, as it is not constrained by hard data and interpretations, whereas the one by Franklin and Gosling is elegant but technical. All the same, among the three, only Franklin's had real data relevant to the model, her beautiful X-ray photograph of B-DNA and the parameters of the double helix calculated from it. Yet, it is ironic that the paper gives the impression of being just an afterthought, with the statement "Thus, our general ideas are not inconsistent with the model proposed by Watson and Crick in the preceding communication". With her insight and original data, Rosalind Franklin would have been the most likely person who would have solved the structure of DNA if Watson and Crick had not seen her data and come up with the model first. But her approach would have been more formal, deducing it from first principles using Fourier transforms, as she was not in favour of short-cut methods such as model building. The full text of her classic paper is reproduced below.

S Mahadevan

Molecular Configuration in Sodium Thymonucleate

Rosalind E Franklin* and R G Gosling
Wheatstone Physics Laboratory, King's College, London
April 2
* Now at Birkbeck College Research Laboratories, 21 Torrington Square, London, W.C.1

Sodium thymonucleate fibres give two distinct types of X-ray diagram. The first corresponds to a crystalline form, structure *A*, obtained at about 75 per cent relative humidity; a study of this is described in detail elsewhere [1]. At higher humidities a different structure, structure *B*, showing a lower degree of order , appears and persists over a wide range of ambient humidity. The change from *A* to *B* is reversible. The water content of structure *B* fibres which undergo this reversible change may vary from 40-50

Reproduced from *Nature*, April 25, Vol.171, p.740, 1953.

per cent to several hundred per cent of the dry weight. Moreover, some fibres never show structure A, and in thew structure B can be obtained with an even lower water content.

The X-ray diagram of structure B (see photograph[+]) shows in striking manner the features characteristic of helical structures, first worked out in this laboratory by Stokes (unpublished) and by Crick, Cochran and Vand [2]. Stokes and Wilkins were the first to propose such structures for nucleic acid as a result of direct studies of nucleic acid fibres, although a helical structure had been previously suggested by Furberg (thesis, London, 1949) on the basis of X-ray studies of nucleosides and nucleotides.

While the X-ray evidence cannot, at present, be taken as direct proof that the structure is helical, other considerations discussed below make the existence of a helical structure highly probable.

Structure B is derived from the crystalline structure A when the sodium thymonucleate fibres take up quantities of water in excess of about 40 per cent of their weight. The change is accompanied by an increase of about 30 per cent in the length of the fibre, and by a substantial re-arrangement of the molecule. It therefore seems reasonable to suppose that in structure B the structural units of sodium thymonucleate (molecules on groups of molecules) are relatively free from the influence of neighbouring molecules, each unit being shielded by a sheath of water. Each unit is then free to take up its least-energy configuration independently of its neighbours and, in view of the nature of the long-chain molecules involved, it is highly likely that the general form will be helical [3]. If we adopt the hypothesis of a helical structure, it is immediately possible, from the X-ray diagram of structure B, to make certain deductions as to the nature and dimensions of the helix.

The innermost maxima on the first, second, third and fifth layer lines lie approximately on straight lines radiating from the origin. For a smooth single-strand helix the structure factor on the nth layer line is given by:

$$F_n = \mathcal{J}_n(2\pi rR) \exp i\, n\, (\psi + 1/2\, \pi),$$

where $\mathcal{J}_n(u)$ is the nth-order Bessel function of u, r is the radius of the helix, and R and ψ are the radial and azimuthal co-ordinates in reciprocal space [2]; this expression leads

[+] Reproduced on the cover page.

to an approximately linear array of intensity maxima of the type observed, corresponding to the first maxima in the functions $\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3$, etc.

If, instead of a smooth helix, we consider a series of residues equally spaced along the helix, the transform in the general case treated by Crick, Cochran and Vand is more complicated. But if there is a whole number, m, of residues per turn, the form of the transform is as for a smooth helix with the addition, only, of the same pattern repeated with its origin at heights mc^*, $2mc^*$ . . ., etc. (c is the fibre-axis period).

In the present case the fibre-axis period is 34 A. and the very strong reflexion at 3.4 A. lies on the tenth layer line. Moreover, lines of maxima radiating from the 3.4-A. reflexion as from the origin are visible on the fifth- and lower layer lines, having a $\mathcal{J}_5$ maximum coincident with that of the origin series on the fifth layer line. (The strong outer streaks which apparently radiate from the 3.4-A. maximum are not, however, so easily explained.) This suggests strongly that there are exactly 10 residues per turn of the helix. If this is so, then from a measurement of R_n the position of the first maximum on the nth layer line (for n 5≤), the radius of the helix, can be obtained. In the present instance, measurements of R_1, R_2, R_3, and R_5, all lead to values of r of about 10 A.

Since this linear array of maxima is one of the strongest features of the X-ray diagram, we must conclude that a crystallographically important part of the molecule lies on a helix of this diameter. This can only be the phosphate groups or phosphorus atoms.

If ten phosphorus atoms lie on one turn of a helix of radius 10A., the distance between neighbouring phosphorus atoms in a molecule is 7.1 A. This corresponds to the P ... P distance in a fully extended molecule, and therefore provides a further indication that the phosphates lie on the outside of the structural unit.

Thus, our conclusions differ from those of Pauling and Corey [4], who proposed for the nuoleic acids a helical structure in which the phosphate groups form a dense core.

We must now consider briefly the equatorial reflexions. For a single helix the series of equatorial maxims should correspond to the maxima in $\mathcal{J}_0(2\pi rR)$. The maxima on our photograph do not, however, fit this function for the value of r deduced above. There is a very strong reflexion at about 24 A. and then only a faint sharp reflexion at 9.0 A. and two diffuse bands around 5.5 A. and 4.0 A. This lack of agreement is, however, to be expected, for we know that the helix so far considered can only be the most important member of a series of coaxial helices of different radii ; the non-phosphate parts of the

molecule will lie on inner co-axial helices, and it can be shown that, whereas these will not appreciably influence the innermost maxima on the layer lines, they may have the effect of destroying or shifting both the equatorial maxima and the outer maxima on other layer lines.

Thus, if the structure is helical, we find that the phosphate groups or phosphorus atoms lie on a helix of diameter about 20 A., and the sugar and base groups must accordingly be turned inwards towards the helical axis.

Considerations of density show, however, that a cylindrical repeat unit of height, 34 A. and diameter 20 A. must contain many more than ten nucleotides.

Since structure *B* often exists in fibres with low water content, it seems that the density of the helical unit cannot differ greatly from that of dry sodium thymonucleate, 1.63 gm./cm^3 [1,5], the water in fibres of high water-content being situated outside the structural unit. On this basis we find that a cylinder of radius 10 A. and height 34 A. would contain thirty-two nucleotides. However, there might possibly be some slight inter-penetration of the cylindrical units in the dry state making their effective radius rather less. It is therefore difficult to decide, on the basis of density measurements alone, whether one repeating unit contains ten nucleotides on each of two or on each of three co-axial molecules. (If the effective radius were 8 A. the cylinder would contain twenty nucleotides.) Two other arguments, however, make it highly probable that there are only two co-axial molecules.

First, a study of the Patterson function of structure *A*, using superposition methods. has indicated [6] that there are only two chains passing through a primitive unit cell in this structure. Since the $A \rightleftharpoons B$ transformation is readily reversible, it seems very unlikely that the molecules would be grouped in threes in structure *B*. Secondly, from meaurements on the X-ray diagram of structure *B* it can readily be shown that, whether the number of chains per unit is two or three, the chains are not equally spaced along the fibre axis. For example, three equally spaced chains would mean that the nth layer line depended on $\mathcal{J}_{3n}$, and would lead to a helix of diameter about 60 A. This is many times larger than the primitive unit cell in structure *A*, and absurdly large in relation to the dimensions of nucleotides. Three unequally spaced chains, on the other hand, would be crystallographically non-equivalent, and this, again, seems unlikely. It therefore seems probable that there are only two co-axial molecules and that these are unequally spaced along the fibre axis.

Thus, while we do not attempt to offer a complete interpretation of the fibre-diagram of structure *B*, we may state the following conclusions. The structure is probably helical. The phosphate groups lie on the outside of the structural unit, on a helix of diameter about 20 A. The structural unit probably consists of two co-axial molecules which are not equally spaced along the fibre axis, their mutual displacement being such as to account for the variation of observed intensities of the innermost maxima on the layer lines; if one molecule is displaced from the other by about three-eighths of the fibre-axis period, this would account for the absence of the fourth layer line maxima and the weakness of the sixth. Thus our general ideas are not inconsistent with the model proposed by Watson and Crick in the preceding communication.

The conclusion that the phosphate groups lie on the outside of the structural unit has been reached previously by quite other reasoning [1]. Two principal lines of argument were invoked. The first derives from the work of Gulland and his collaborators [7], who showed that even in aqueous solution the –CO and $-NH_2$ groups of the bases are inaccessible and cannot be titrated, whereas the phosphate groups are fully accessible. The second is based on our own observations [1] on the way in which the structural units in structures *A* and *B* are progressively separated by an excess of water, the process being a continuous one which leads to the formation first of a gel and ultimately to a solution. The hygroscopic part of the molecule may be presumed to lie in the phosphate groups [$(C_2H_5O)_2$ PO_2Na and $(C_3H_7O)_2PO_2Na$ are highly hygroscopic [8]], and the simplest explanation of the above process is that these groups lie on the outside of the structural units. Moreover, the ready availability of the phosphate groups for interaction with proteins can most easily be explained in this way.

We are grateful to Prof. J. T. Randall for his interest and to Drs. F. H. C. Crick, A. R. Stokes and M. H. F. Wilkins for discussion. One of us (R.E.F.) acknowledges the award of a Turner and Newall Fellowship.

[1] Franklin, R. E., and Gosling, R.G. (in the press).
[2] Cochran, W., Crick, F. H. C., and Vand, V., *Acta. Cryst.*, **5**, 501 (1952).
[3] Pauling, L., Corey, R.B., and Bransom, H.R., *Proc. U.S. Nat. Acad. Sci.*, **37**, 205 (1951).
[4] Pauling, L., and Corey, R. B., *Proc. U.S. Nat. Acad. Sci.*, **39**, 84 (1953).
[5] Astbury, W. T., Cold Spring Harbor Symp. on Quant. Biol., **12**, 56 (1947).
[6] Franklin, R. E., and Gosling, R. G. (to be published).
[7] Gulland, J. M., and Jordan D. O., Cold Spring Harbor Symp. on Quant. Biol., **12**, 5 (1947).
[8] Drushel, W. A., and Felty, A. R., *Chem. Zent.*, **89**, 1016 (1918).