Acta Cryst. (1955). 8, 151

The Structure of Sodium Thymonucleate Fibres. III. The Three-Dimensional Patterson Function*

BY ROSALIND E. FRANKLIN† AND R. G. GOSLING

Wheatstone Physics Laboratory, King's College, London, W.C. 2, England

(*Received* 29 *October* 1954)

The three-dimensional Patterson function of the crystalline form of sodium thymonucleate has been calculated, using the same intensity data as for the cylindrical Patterson function already described. Further evidence was thus obtained in support of the modified form of the Watson & Crick model previously deduced from the cylindrical function. In addition, the use of three-dimensional data made it possible to determine the orientation of the helical molecule in the unit cell.

It is shown that the size and shape of the unit cell is such as to reduce to a minimum the distance between certain phosphate groups in neighbouring molecules.

Introduction

The molecular structure of sodium thymonucleate (SDN)‡ fibres is very sensitive to the humidity of the surrounding atmosphere (Franklin & Gosling, 1953*a*). By varying the water content of highly orientated fibres we have shown, from the changes in the X-ray diagram, that SDN may exist in two different structural states, which we have called *A* and *B*. The change $A \rightleftharpoons B$ is, in general, readily reversible. *B* is the equilibrium form at high relative humidity and *A* is a crystalline form containing about 40% by weight of water, best obtained in fibres at a relative humidity of about 75%. At lower relative humidity the structure becomes (reversibly) disordered.

Watson & Crick (1953*a*) have proposed a structure for DNA in which the phosphate-desoxyribose backbone chains form two coaxial helical strands related by a diad axis. This structure must be considered in relation to our structure *B*, in which each helical molecule may be assumed to be shielded from the deforming influence of neighbouring molecules by a sheath of water. The molecule is then free to take up its least-energy configuration and the X-ray diagram represents the continuous transform of a single molecule. We have shown (Franklin & Gosling, 1953*b*; Gosling, 1954) that the X-ray diagram of structure *B*, taken in conjunction with measurements of density and water content, lends support to the general features of the Watson & Crick model, but indicates that the molecule is rather more compact than was suggested by these authors, the phosphorus atoms lying on coaxial helices of radius about 8·5 Å rather than 10 Å.

If a two-strand helical molecule exists in structure *B*, then it follows, from the ease of transition from one structure to the other, that the molecule in structure *A* must also be a two-strand helix; the structural modifications induced by the close proximity of neighbouring molecules in the crystalline form must be relatively minor ones. Therefore, although structure *B* is undoubtedly the more closely related to DNA in its natural state, a more detailed study of structure *A* is justified, both by the close relationship which it must bear to structure *B* and by the greater wealth and precision of the diffraction data available.

The measurement of intensities and *R*-space parameters for structure *A*, and the use of these to calculate the cylindrically symmetrical Patterson function, have already been described (Franklin & Gosling, 1953*c*). The existence of the two-strand helix in structure *A* was clearly revealed in the Patterson function, and it was shown that in changing from *B* to *A* the number of residues per turn of each strand increases from ten to eleven and the pitch of the helix decreases from 33·1 Å to 28·1 Å (Franklin & Gosling, 1953*d*). In structure *A* we have suggested that the phosphorus atoms lie equispaced on a two-strand helix of radius close to 9 Å, with a separation of $\frac{1}{2}c$ between the strands, the phosphorus atoms in one strand being directly above those in the other strand in the direction of the axis of the helix, with the sugar and base rings turned inwards towards that axis. This model is examined in this paper with the aid of the three-dimensional Patterson function of structure *A*.

Calculation of the three-dimensional Patterson function

The experimental data used for the calculation of the three-dimensional Patterson function were the same as those used to obtain the cylindrical function previously described (Franklin & Gosling, 1953*c*). We

---

* The work presented in this paper forms part of a thesis presented by R.G.G. for the Ph.D. degree of the University of London.

† Now at Birkbeck College Crystallography Laboratory, 21 Torrington Square, London W. C. 1, England.

‡ In previous papers we have used the notation NaDNA, but SDN (as used by American authors) is clearly preferable.

Table 1. *Data used in calculation of the three-dimensional Patterson*

Layer	$(\xi/\lambda)\times 10^2$ Obs.	$(\xi/\lambda)\times 10^2$ Calc.	*hkl*	$I_{\text{corr.}}$
0	5·25	5·21	110	9
	8·80	8·81	130	104
	9·00	9·14	200	30
	10·00	10·04	040	3
	10·50	10·42	220	4
	13·25	13·35	150	13
	13·8	13·57	240	5·5
		13·94	310	5·5
	15·75	15·64	330	5
	17·55	17·62	260	9
	18·20	18·16	170	12
		18·28	400	9
	18·85	18·72	350	4
		18·96	420	8
	20·00	20·08	080	7
	20·80	20·86	440	5
	22·10	22·07	280	6
		22·29	370	6
	22·85	22·99	510	10·5
		23·04	190	10·5
	24·00	25·34	530	14
1	4·90	4·86	$11\bar{1}$	22
	5·50	5·57	111	11
	8·60	8·61	$13\bar{1}$	39
	9·05	9·02	131	23
	10·10	10·08	$22\bar{1}$	10
	13·10	13·22	$15\bar{1}$	16
	15·30	15·29	$33\bar{1}$	7
	16·10	16·00	331	7
	22·60	22·59	$51\bar{1}$	6
		22·54	371	4
	23·50	23·38	$46\bar{1}$	4
		23·67	$53\bar{1}$	5
2	4·55	4·53	$11\bar{2}$	52
	4·90	5·08	022	32
	5·90	5·93	122	29
	9·80	9·73	$22\bar{2}$	20
	11·15	11·14	222	19
	13·00	13·05	$24\bar{2}$	14
		13·14	$31\bar{2}$	14
		13·18	$15\bar{2}$	14
	17·20	17·21	$26\bar{2}$	31
	18·05	18·00	$35\bar{2}$	9
		17·97	$17\bar{2}$	9
		18·05	262	9
	20·00	20·10	082	2·5
		20·16	$44\bar{2}$	2·5
	21·9	21·80	$37\bar{2}$	6
		21·75	$28\bar{2}$	6
	22·90	22·79	372	5
		22·90	$19\bar{2}$	5
		23·08	$46\bar{2}$	5
	26·60	26·63	$48\bar{2}$	1
		26·86	552	1
		26·45	$2,10,\bar{2}$	1
		26·62	$60\bar{2}$	1
	27·90	27·63	$1,11,\bar{2}$	1·5
		28·20	$57\bar{2}$	1·5
3	12·80	12·80	$24\bar{3}$	3·5
		12·76	$31\bar{3}$	3·5
	14·55	14·41	243	4
		14·60	$33\bar{3}$	3
	14·95	15·11	063	4
		15·12	313	4
	19·80	19·82	$44\bar{3}$	2
	21·50	21·60	$28\bar{3}$	2
4	12·35	12·56	$31\bar{4}$	1·5
		12·37	$24\bar{4}$	1·5
	14·10	14·16	154	10
		14·26	$33\bar{4}$	9
	14·95	14·70	244	13·5
		15·15	064	13·5
	16·90	16·84	$26\bar{4}$	14
		17·06	334	4
	21·45	21·34	$37\bar{4}$	2·5
		21·43	$28\bar{4}$	2·5
	22·75	22·77	284	2
		22·78	$19\bar{4}$	3
5	11·90	11·98	$31\bar{5}$	12
	12·40	12·32	$24\bar{5}$	6
		12·22	225	6
	14·0	12·34	155	3
6	8·40	8·40	$22\bar{6}$	12
	10·4	10·33	046	14
		10·26	136	14
	11·70	11·59	$31\bar{6}$	12
	12·75	12·93	$15\bar{6}$	12
	13·50	13·58	$33\bar{6}$	21
7	7·90	7·79	117	4
		7·74	$13\bar{7}$	5
	11·0	11·19	$31\bar{7}$	11
	12·1	11·94	207	5
		11·87	$24\bar{7}$	5
	13·2	13·26	$33\bar{7}$	4
		12·95	227	6
8	10·80	10·81	$31\bar{8}$	10
		10·82	138	6
	12·50	12·83	$15\bar{8}$	9
		12·34	208	6
	15·70	15·89	$42\bar{8}$	3
9	5·7	5·54	$20\bar{9}$	2

repeat here, for convenience, that the unit cell parameters are

$$a = 22{\cdot}0,\ b = 39{\cdot}8,\ c = 28{\cdot}1\ \text{Å};\ \beta = 96{\cdot}5^\circ,$$

with c coincident with the fibre axis. The space group is $C2$ and an 'artificial temperature factor' of

$$\exp\left[-(4{\cdot}56)^2\times(2\sin\theta/\lambda)^2\right]$$

was applied to all intensities. The methods used to index the observed reflexions have already been described (Franklin & Gosling, 1953*c*).

In Table 1 we list the observed and calculated ξ values with their indices and observed (corrected) intensities for each layer line. It will be seen from the table that for the larger values of θ no reflexion could be indexed unambiguously. In cases where the possible indices included *hkl* and $hk\bar{l}$ the photograph showing double orientation (Franklin & Gosling, 1953*a*) was used as a guide in distributing the ob-

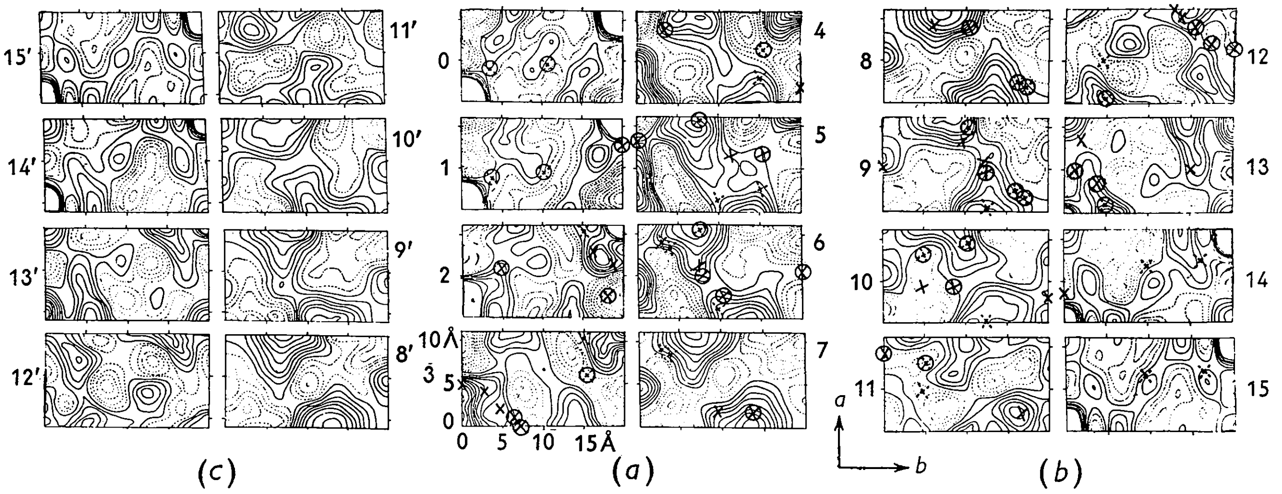


Fig. 1. (*a*), (*b*) The three-dimensional Patterson function of SDN, shown as 16 *a*–*b* quarter-cell sections from $z = 0$ to $z = \frac{1}{2}c$. × : *Intra*-strand P–P vector; ⊗: *Inter*-strand P–P vector.

These are the *intra*-helical P–P vectors to be expected from the proposed two-strand helical model. A broken cross denotes a vector peak close to a half-section level in *Z*.

(*c*) The *a*–*b* sections of Fig. 1(*b*) reflected across a diad axis at the quarter cell height. The similarity between these sections and Fig. 1(*a*) suggests that part of the structure repeats in half the cell period.

served intensity. Where no external guide was available the intensity was equally divided among the possible reflexions.

The three-dimensional Patterson function (Fig. 1 (*a*, *b*)) was calculated at intervals of $a/30$, $b/60$ and $c/30$ in the form of 16 *a*–*b* sections.

Interpretation of the Patterson function

We have shown previously from a study of both the X-ray diagram of structure *B* (Franklin & Gosling, 1953*b*) and the cylindrical Patterson function of structure *A* (Franklin & Gosling, 1953*d*) that the principal diffraction features are due to phosphate–phosphate vectors. This is to be expected since these are the heaviest groups in the structure. If *z* is the atomic number of any atom, the value of Σz^2 for the group PO_4Na is 602 and that for the remainder of the nucleotide is only 685 (cytidine) to 819 (guanine). Thus the PO_4Na group, which is highly compact, may be expected to act as a 'heavy atom'. Moreover, since the P atom lies at the centre of the $-PO_4$ group, the phosphate–phosphate vectors may be grossly identified with the P–P vectors. We shall therefore seek to interpret the three-dimensional Patterson function in terms of P–P vectors.

(i) *Orientation of the helix*

In order to determine the P–P vector peaks to be expected from the model outlined above it was first necessary to determine the orientation, with respect to the cell parameters, of the arrangement of eleven phosphorus atoms forming one turn of each strand of the helical molecule. The axis of the helix is known to coincide with the *c* axis of the unit cell and the pitch of the helix is equal to *c*. The density indicates that only one helical molecule traverses each primitive unit cell. Therefore, if the helix can be right-handed or left-handed, there are four possible arrangements of the eleven phosphorus atoms about a lattice point that could comply with the space-group symmetry. These are shown in Fig. 2 as projections of a turn of one strand of the helix on to a plane at right-angles to the *c* axis through the level $z = 0$. Since the phosphate–sugar backbone chain is non-centric, the symmetry axes along *b* cannot pass through a chain, but must relate one chain, *n*, to the other, *n'*. Thus in arrangements (i) and (ii) P_0 must be placed at $z = \frac{1}{4}c$ and so be related by the diad axis at $\frac{1}{2}c$ to the atom

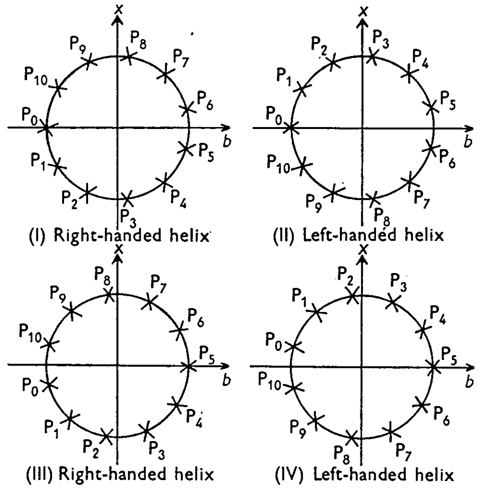


Fig. 2. Projections of the four possible helical arrangements of the eleven phosphorus atoms in one turn of one strand of structure *A*.

P'_0 at $z = \frac{3}{4}c$. In arrangements (iii) and (iv) P_5 must be considered to be at $z = \frac{1}{4}c$. Since from the Patterson function it is not possible to distinguish between a right-handed and a left-handed helix, we have to decide only between two possible configurations. From model-building, Watson & Crick (1953*b*) have suggested that the DNA helix could only be right-handed. Adopting this suggestion, the possible configurations are Fig. 2(i) and (iii). In both Fig. 2(i) and Fig. 2(iii) one vector joining neighbouring phosphorus atoms, P_n–P_{n+1}, lies in the *a–c* plane. Consideration of this vector enables us to choose between the two possible orientations of the helix. In Fig. 2(i), P_5–P_6 has $x = 5{\cdot}07$ Å, $z = 2{\cdot}55$ Å. In Fig. 2(iii) P_{10}–P_0 has $x = -5{\cdot}07$ Å, $z = 2{\cdot}55$ Å. The corresponding Patterson peaks will lie between sections 2 and 3 of Fig. 1(*a*), and it is clear that a choice must be made in favour of the orientation of Fig. 2(i). A diagram of this arrangement of phosphorus atoms projected on to the *a–c* plane is shown in Fig. 3.

(ii) *Intra-helical* P–P *vectors*: P_n–$P_{(n+m)}$ and P_n–$P_{(n'+m)}$

The *intra*-helical phosphorus vectors for one strand P_n–$P_{(n+m)}$, and the vectors from one strand to the next, P_n–$P_{(n'+m)}$, were determined graphically for $n = 0$–10 and $m = 1$–5, using that configuration of P atoms shown in Figs. 2(i) and 3. Since the proposed pattern of P atoms repeats at $z = \frac{1}{2}c$, it was not necessary to calculate the vector series due to values of m from 6 to 10. The theoretical positions of peaks due to P_n–$P_{(n+m)}$ vectors are marked with a cross in Fig. 1(*a, b*); where the position lies nearly half-way between two sections, it is indicated on both sections by a broken cross. The P_n–$P_{(n'+m}$ vectors are similarly marked by a full or broken cross ringed by a full line. Since there is only one helix associated with each lattice point, the *inter-helical* P–P vectors give rise to no further Patterson peaks.

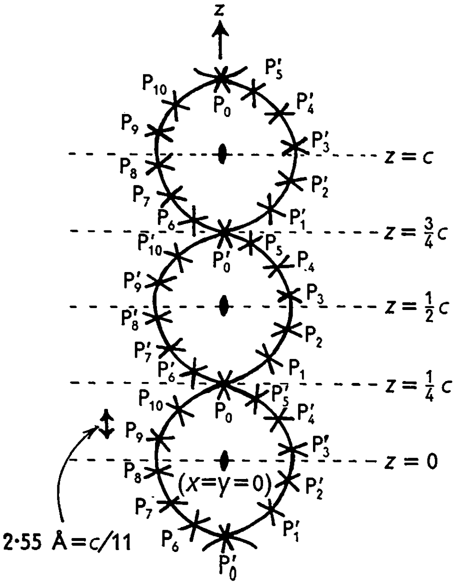


Fig. 3. Projection on the *a–c* plane of the configuration of phosphorus atoms proposed for one helical unit of structure *A*.

The complete array of P–P vectors indicated in Fig. 1(*a, b*) clearly accounts well for a substantial part of the Patterson function.

(iii) *Superposition method*

Before confronting the three-dimensional data with the proposed helical model, extensive attempts were made to use the superposition method (Buerger, 1951) to obtain a partial solution of the problem without introducing any hypothesis. This method is strictly applicable only when a well-isolated non-multiple peak can be identified in the Patterson function. Although this condition was not fulfilled in the present case, we thought it possible that phosphate–phosphate vectors might predominate to such an extent that certain reasonably well defined peaks might be effectively entirely due to them, and that the use of such peaks in a series of three-dimensional superpositions might give some indication of the arrangement of these groups in the structure. In spite of rather numerous attempts, using a variety of possible peaks as displacement vectors, and searching for the common elements in the resulting superposition functions, no progress was made by this method before a model structure became available.

After the development of the two-strand helical model described above, we returned to the superposition method to look for confirmation of the model. For this purpose superposition functions were traced using displacement vectors corresponding to certain of the P–P vectors of the model, with the object of investigating whether or not the other P positions of the model appeared as prominent features of the resulting functions.

In the proposed model, if the phosphorus atoms only are considered, the fibre-axis repeat period is halved, pseudo-diad axes being present at heights $c = \frac{1}{4}$ and $c = \frac{3}{4}$ in the complete unit cell. It follows that that part of the Patterson function due to P–P vectors must also show this extra symmetry. That this pseudo-halving of c is a strong feature of the Patterson function may be seen from Fig. 1(*c*). This figure consists of sections 15–8 rotated about a diad axis parallel to b at height $\frac{1}{4}c$. There is clearly a strong resemblance between this set of reflected sections and the sections 0–7 of Fig. 1(*a*).

In order to remove those peaks not related by this additional symmetry, the two sets of sections shown in Fig. 1(*c*) and Fig. 1(*a*) were superimposed and the common positive regions were traced. Peaks in negative regions were also indicated. The eight Patterson sections resulting from this procedure should contain all the P–P vectors of the proposed model.

Using these sections, three superposition functions were constructed using the three displacement vectors

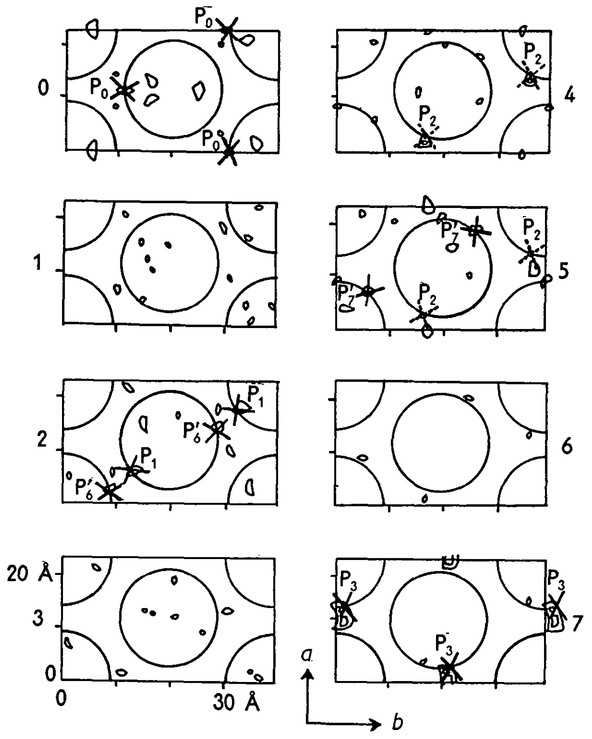


Fig. 4. Superposition function prepared from the Patterson sections, using some of the P–P distances given by the proposed helical structure.

P_0–P_3, P_0–P_4, P_3–P_4 obtained from the model (projections shown in Figs. 2(i) and 3). These three functions were then superimposed on one another using the spatial relationships of P_0, P_3 and P_4 given by the model, and the part common to all three superposition functions was traced.

The sections of this final superposition function are shown in Fig. 4. On each section there are also shown circular traces of 9 Å radius, on which the P atoms of the model structure would lie. (These traces should be elliptical owing to the slope of the *a*–*b* section, but the difference between the major and minor axes is only 0·06 Å and can therefore be neglected.) The section numbers at which the phosphorus atoms should occur are listed in Table 2 (P_3, P_4 and P_5 are related by a pseudo-diad axis to P_8', P_7' and P_6').

Table 2

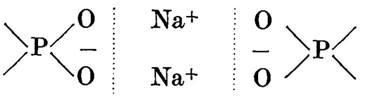
P_n	P_0	P_6'	P_1	P_2	P_7'	P_3	P_8'
Section	0	1·7	2·1	4·5	4·9	7·1	7·9

The expected position of these P atoms is marked on the appropriate sections with a full cross, except for P_2, the coordinates of which are indicated on both sections 4 and 5.

It seems that Fig. 4 provides substantial confirmation of the proposed helical array of P atoms. Only two of the atoms P_0, P_3 and P_4 were introduced into each of the three primary superposition functions, and the final superposition function (Fig. 4) contains not only peaks corresponding to all three of these atoms, but, in addition, the remaining P atoms, P_1, P_2 and P_5 (P_6'), all appear distinctly in their correct positions. Moreover, the cylinders of radius 9 Å, on which all the P atoms lie, are seen to be sufficiently free from extraneous peaks for these results to be significant.

Inter-helical bonding

We have suggested previously that neighbouring molecules of SDN in the crystalline structure would be most probably linked to one another by ionic bonds represented schematically by



This suggestion finds support in the structure which we have described here.

From the sections of the helices shown in Fig. 4, strong inter-helical phosphate–phosphate bonds, of the type suggested above, might be expected between P_1 and P_3 in one helix and P_6' and P_8' respectively, in adjoining helices. The inter-helical P_1–P_6' distance is 4·9 Å and the P_3–P_8' distance is 4·4 Å. For any one helix there are *eight* positions of the P_1–P_6' type and *four* positions of the P_3–P_8' type in each complete period of the structure. Now the size and shape of the monoclinic face-centred unit cell are such that the *c*-axis displacement of the face-centring helix with respect to those at the corners of the cell is $\pm c/22$, i.e. this *c*-axis displacement is exactly half that between neighbouring P atoms on any given helical strand. With the two strands of the helical molecules equally spaced on the helical axis, and orientated as described above, this has the effect of bringing two P atoms of neighbouring helices into the closest possible proximity for the given values of *a* and *b*. The two P

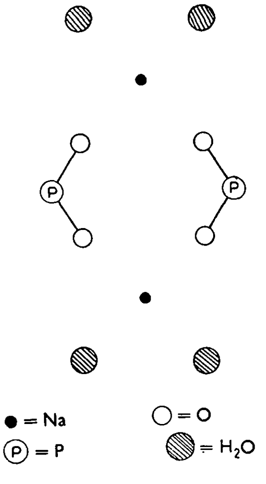


Fig. 5. Schematic representation of phosphate–phosphate inter-helical bonding involving octahedrally co-ordinated sodium ions as intermediaries. (Other water molecules lie above and below the plane of the diagram.)

atoms which are brought close together in this way lie on non-equivalent helical strands; that is, they lie on strands of the type which are related to one another by a true diad in the full structure.

These P–P distances are at first sight surprisingly small. They are too small to permit the Na^+ ions to be situated directly between the two PO_4^- groups. It seems reasonable to suppose, however, that the bonding would be affected by the two Na^+ ions situated just outside the phosphate groups, their coordination (probably 5- or 6-fold) being completed by the water molecules known to be present. This is shown schematically in Fig. 5.

If the O–Na distance is taken as 2·32 Å, the lowest value found by Beevers & Cochran (1947) for an Na–(OH) bond in an octahedral configuration in sucrose sodium bromide dihydrate, and the O–P–O angle is taken as 116° with the P–O bond of length 1·5 Å, then the P–P distance for the planar arrangement shown in Fig. 5 is 4·87 Å. This value agrees well with the estimated P_1–P_6' distance of 4·9 Å. That the P_3–P_8' distance is apparently as low as 4·4 Å (Fig. 4, section 7) suggests that either the helices may be slightly flattened in the *a* direction, or that in this case the two $>PO_2^-$ groups are not coplanar.

If the Na^+ are 6-coordinated this type of linkage would determine, for half the nucleotides, the position of four of the average eight oxygen atoms that the water-content measurements indicate are associated with each nucleotide in structure *A*.

Inspection of sections 4 and 5 of Fig. 4 shows that the phosphorus atom P_2 lies very close to the adjoining helix. The nearest phosphorus atom on the neighbouring helix is P_8' at a distance of 7·8 Å. A phosphate–phosphate linkage between these groups would therefore seem unlikely. It seems possible, therefore, that in this case there may be a base-phosphate bridge between neighbouring helices. Such additional *inter*-helical bonding close to the *a* direction may in part explain why the separation of the helices is least in that direction.

Discussion

The combined study of the cylindrical and three-dimensional Patterson function has enabled us to suggest a detailed picture of the arrangement of the phosphate groups in the crystalline structure (*A*) of SDN. This is a modified form of the two-strand helix proposed by Watson & Crick. The phosphorus atoms of two polynucleotide chains form two coaxial helical strands of 9 Å radius and 28·1 Å pitch, with 11 phosphorus atoms spaced equally along each turn of each strand. The separation of these helical strands in the direction of their common axis is 14 Å. Corresponding successive phosphorus atoms in each strand, P_n and P_n', have the same coordinates in a plane at right-angles to the axis of the helix. The orientation of the helix and the shape of the unit cell are such as to reduce to a minimum the distance between certain pairs of P atoms on non-equivalent neighbouring chains. This model lends support to our previous suggestion that ionic links between phosphate groups are primarily responsible for maintaining three-dimensional order in the crystalline material.

The X-ray data have not given any direct evidence as to the position of the sugar and base rings in the structure, but it is now fully established that this part of the molecule is turned inwards towards the helical axis. The evidence for this has been discussed in previous papers, to which we have referred.

In conclusion, it seems of interest to recapitulate the part played in establishing the proposed structure by each of the interpretive methods which we have used. Briefly, the cylindrical Patterson function gave us, with fairly high accuracy, the nature of the helix and its parameters. The three-dimensional Patterson function enabled us to determine the orientation of the helix in the unit cell. The application of the superposition method gave some further confirmation of the correctness of the proposed model.

The relative roles of the cylindrical and three-dimensional Patterson functions are such as would be expected. The cylindrical Patterson function is especially well-suited to the determination of helical parameters, since any set of vectors related by the helical axis in the structure appears as a single vector on the cylindrical Patterson function. The cylindrical Patterson function is therefore a more powerful tool than the three-dimensional for detecting helical features of the structure. But for determining the orientation of the helix in the unit cell it is obviously necessary to use three-dimensional diffraction data.

The authors are grateful to Prof. J. T. Randall for his constant interest. This work was carried out during the tenure of a Turner and Newall Fellowship (R.E.F.) and latterly with the aid of a grant from the British Empire Cancer Campaign (R.G.G.). The work was completed in 1953, before the more accurate intensity measurements of Wilkins, Stokes & Wilson (to be published) were made.

References

Beevers, C. A. & Cochran, W. (1947). *Proc. Roy. Soc.* A, **190**, 257.

Buerger, M. J. (1951). *Acta Cryst.* **4**, 531.

Franklin, R. E. & Gosling, R. G. (1953*a*). *Acta Cryst.* **6**, 673.

Franklin, R. E. & Gosling, R. G. (1953*b*). *Nature, Lond.* **171**, 742.

Franklin, R. E. & Gosling, R. G. (1953*c*). *Acta Cryst.* **6**, 678.

Franklin, R. E. & Gosling, R. G. (1953*d*). *Nature, Lond.* **172**, 156.

Gosling, R. G. (1954). Ph.D. thesis, London.

Watson, J. D. & Crick, F. H. C. (1953*a*). *Nature, Lond.* **171**, 737.

Watson, J. D. & Crick, F. H. C. (1953*b*). Private communication.