Nucleic Acid

26 November 1952.

Astbury says eq. spacing 16.2Å. $a_0 = 18.1$Å

$16.2^2 \cdot 2/\sqrt{3} = 303$ Å^2. eq. diam. 19.6Å

R.W. = 330. Density 1.62 ∴ M.V. = 338 Å^3.

∴ 1.12 Å per residue.

Observed 3.34 Å.

Perhaps we have a triple-chain structure!

Size of molecule. Yesterday (25 Nov. 1952) in a biology seminar Robley Williams showed a slide of sodium thymonucleate and said that the small fibrils have diameter ~ 15 Å.
I asked ~~ag~~ the size, and he repeated 15Å, and discussed the difficulty of measuring such small objects. Polystyrene spheres with
? diameter ~250Å were ~20 times larger.

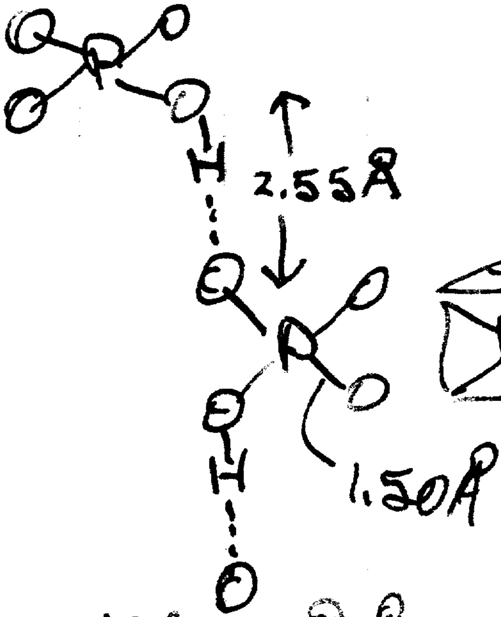
Only one diameter; ∴ cylindrical.

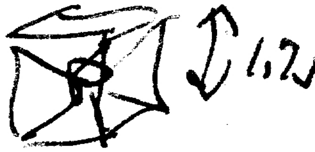
∴ 1.1 Å per residue.

See also Haurowitz book: gives references for 18Å to 20Å diameter.

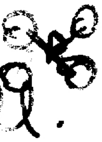
②

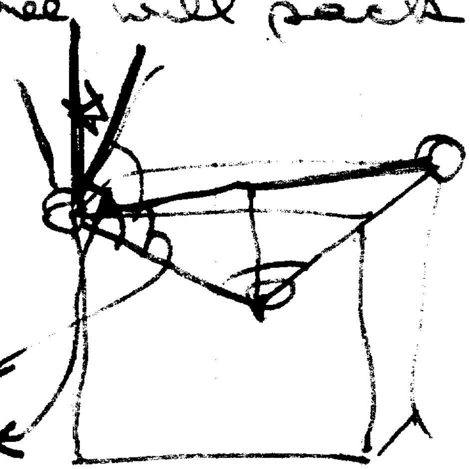
Let us form helixes with O-H···O nearly along axis





$$\begin{array}{r} 2.55\text{Å} \\ 1.73 \\ \hline 4.28\text{Å} \end{array}$$ max.

I assume PO₄ ~~oriented~~ , so that three will pack together well.



$$\begin{array}{r} 180 \\ -180 \\ \hline 270 \\ 35^\circ 16' \\ 90 \\ \hline 125^\circ 16' \end{array}$$

To get ∠P-O-H = 110°, H-O would bend back somewhat.

If it bent back in plane , angle would be 15°, and component 2.55 × cos 15° = 2.47 Å

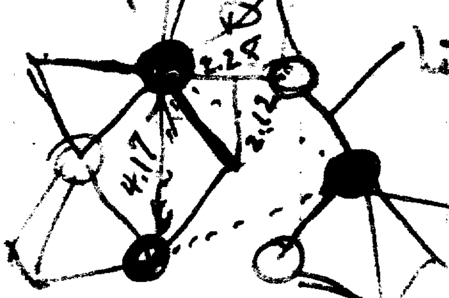
$$\begin{array}{r} 2.47\text{Å} \\ 1.73 \\ \hline 4.20\text{Å} \end{array}$$

I assume component $\mathrm{O{-}H\cdots O}$ = 2.45 Å;

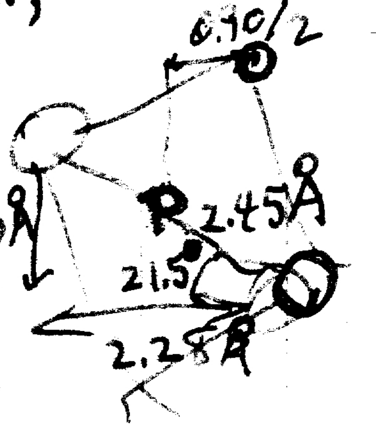
then OPO ↕ = 3.35 − 2.45 = <u>0.90 Å</u>

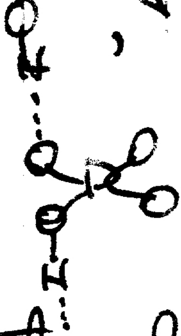
Phosphate is rotated 23.5°.

$$\begin{array}{r} 2.25 \\ 508 \\ \hline 7.36 \end{array}$$



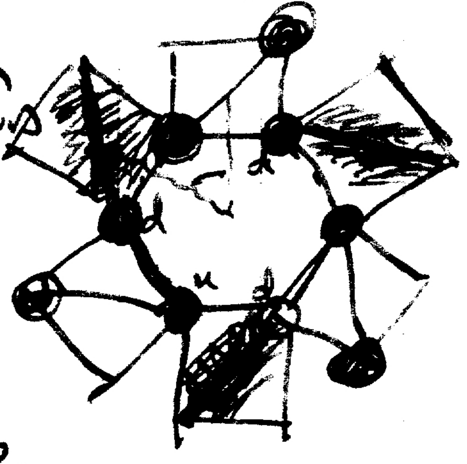
Let this be 2.30 Å
Horizontal 2.54 Å
Then ●—● = 4.17 Å — too large.



I conclude that the hydrogen bonds are not located there  , or that my assumption about bond angle is in error.

I abandon assumption about P-O-H angle.

Let us pack three PO4's in one horizontal plane.

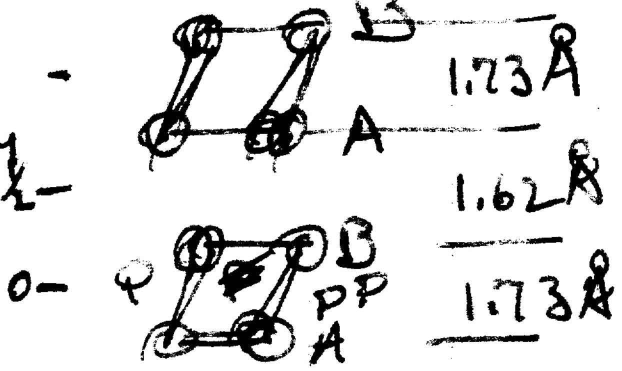


3.35
−1.73
1.62

The next would pack above if rotated ~60°

We put O-H···O bonds between A+B, either at z=0 (connecting PO4's in same layer) or at z=½ (connecting layers).

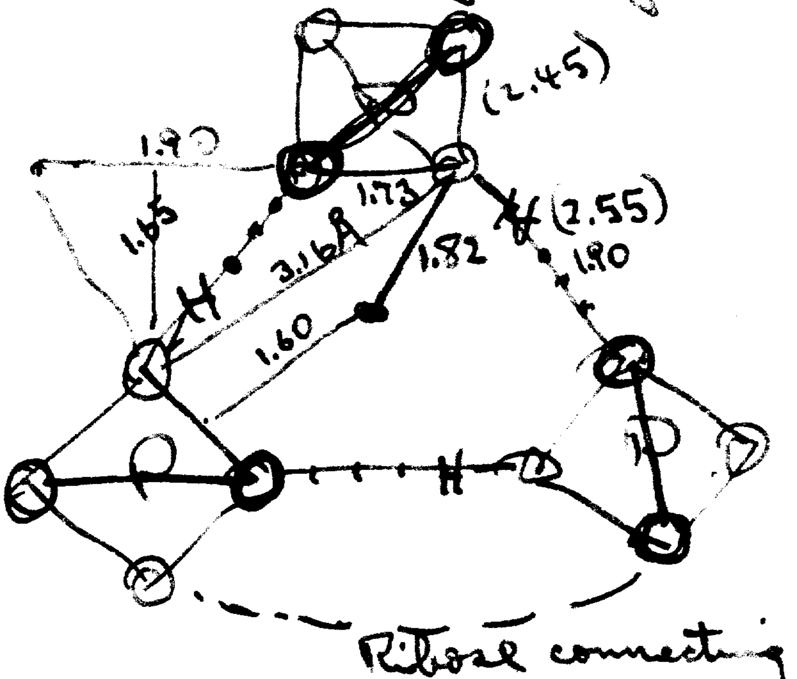
This is wrong. They pack by having PO4's nearly over one another.



The $(PO_4)_3$ rings are nearly over one another.

The P–O–H bond angles require that H be at z=0, i.e., on the level of P. Thus there are $(PO_4{}'H)_3$ rings:

$$3.80 + 1.73 = 5.53 \qquad 3.19 = \frac{5.53}{\sqrt{3}}$$



I assume O layers 1.70 Å apart, both with + without P.

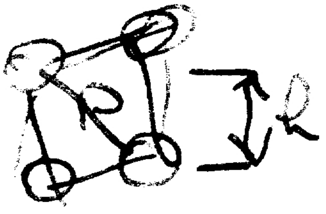
This makes O–O = 2.45 Å in PO_4
2.55 Å in O–H···O
3.16 Å in basal plane
2.49 Å between $(PO_4)_3$ layers.

The last is too small. Probably PO_4's are tipped,

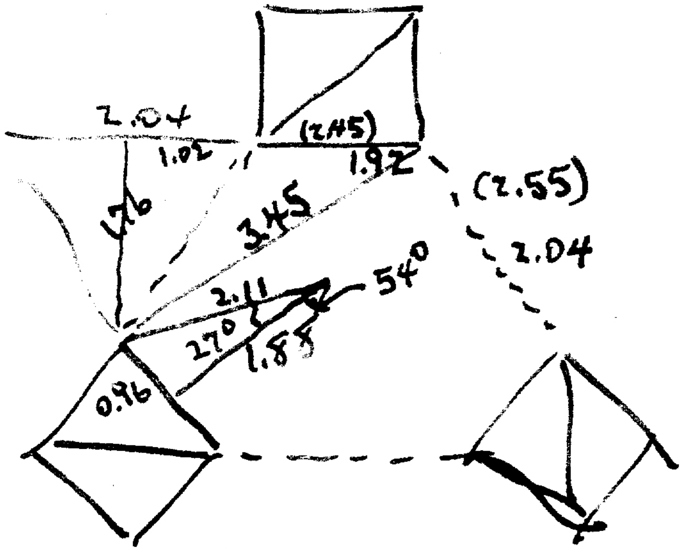
Tipping of PO_4's.

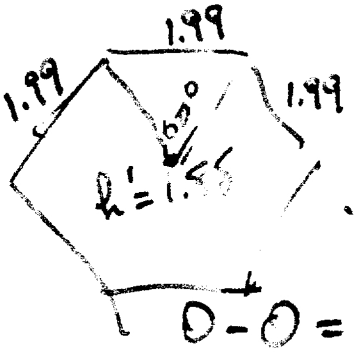
Measurements of a model gives

$h = 1.52$ Å. h' (between $(PO_4)_3$'s) = 1.88 Å



4.08
2.45
6.53

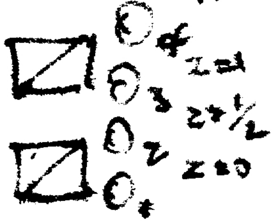




between $(PO_4)_3$'s, if rotated.

O-O = 2.74 Å

Hence with $h = 1.52$ Å, $h' = 1.88$ Å we have



O-O = 2.45 Å in PO_4
= 2.55 Å for O-H···O (diagonal, z=0)
= 3.45 Å in basal plane
= 2.74 Å diagonal, z = ½.

If O_3 & O_2 make 60° angles in projection, rotation from z=0 to z=1 is 6°; i.e., repeat after 60 layers (really 20).

If ~~rotation~~ rotation were 12° (repeat in 30 layers - (really 10) - O-O would be 2.69 Å and 2.77 Å. These are quite all right.

Hence we cannot predict number of residues per turn very closely. We do predict

60 (lead 60 × 3.4 Å, pitch = c_0 = 20 × 3.4 = 68 Å)

but considerably smaller or larger would be acceptable

The PO_4 groups are rotated 6.65°.
Inner O's at $z = \pm$ 0.76 Å, outer O's at $z =$ 0.96 Å

Note that each chain has $\Delta\varphi = \sim 113°$ or $\sim 127°$; that is, roughly 3 residues per turn. There are three chains closely intertwined, and held together by hydrogen bonds between PO_4's.

The ribose residues connect upper O at $z = 0.76$ Å, $\varphi \cong 0$ with lower O, $z = \frac{1}{2} - 0.76$ Å, $\varphi \cong 120°$.

The $(PO_4)_3$ may be either ~~right~~ right-handed or left-handed (screw determined by O-H⋯O).

The ribose groups will hook across either. With my choice for ribose, the right-handed looks better. For it the ribose plane is at about 45° with basal plane, and 3.4 Å along z might permit packing; whereas for the other the ribose plane is at 90° to basal plane.

$\rho \cong 5$ Å for center of ribose ring. $\Delta\varphi = 12°$ would give displacement of $\rho\cdot\Delta\varphi =$ 1 Å – not enough for rings to clear one another.

I haven't checked the strain in the ribose ring strung across between $(PO_4)_3$ very carefully. There may be some trouble, although the van der Waals contacts seem

Assuming α-ribose

With right-handed $(HPO_4)_3$ the purine-pyrimidines stick out at ~ 80° to axis, which agrees with large negative birefringence that is observed. They have first atom (N) at $\rho \cong 7$ Å.

With left-handed $(HPO_4)_3$ they are at ~ 55°, which is probably incompatible with birefringence.

29 November 1952

Helical arrangements of PO_4 tetrahedra.

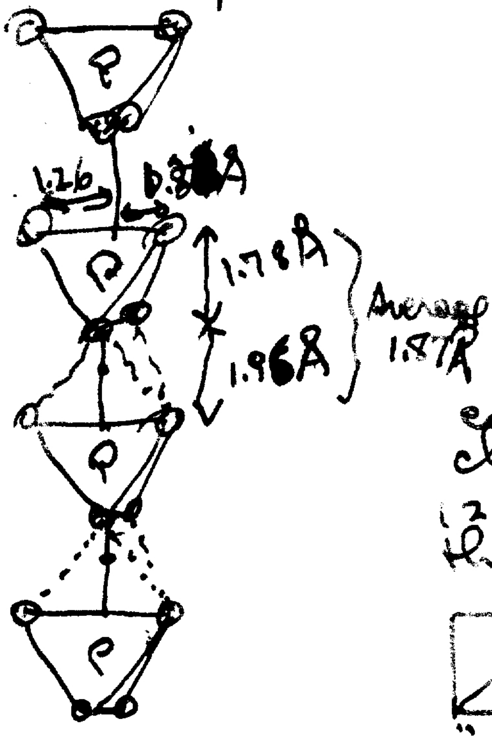
A straight column. Assume P-O = 1.54Å

Then O-O = 2.52Å in tetrahedron

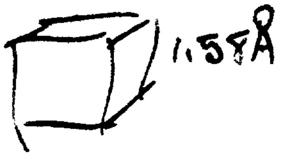
1.75 = height

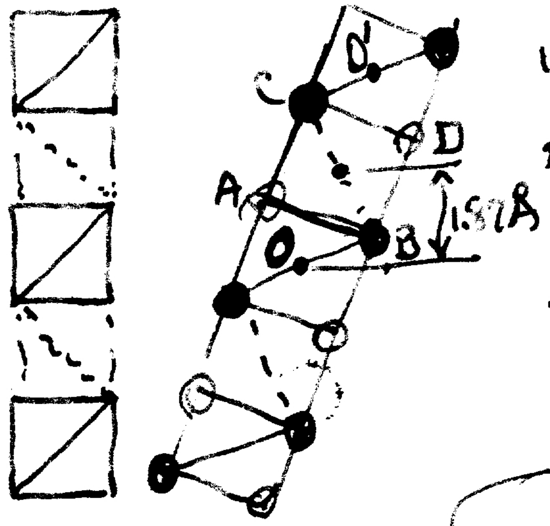
Assume O-O = 2.65 between tetrahedra

Then vertical distance = 1.96Å:

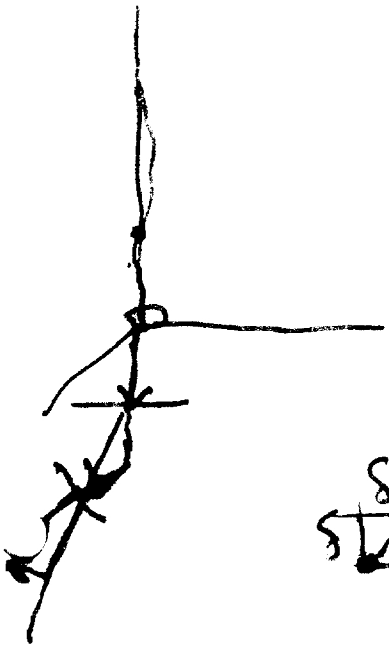


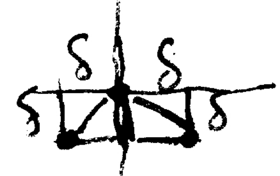
The column of linked tetrahedra (PO_4 and empty, the latter not regular) can be deformed into a helix without change in the contact O...O distances (2.52Å & 2.65Å) by rotating around the horizontal O⊥O axes!

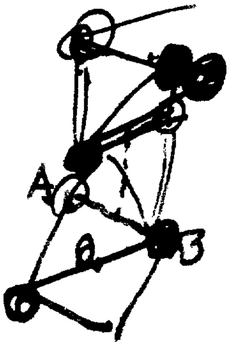


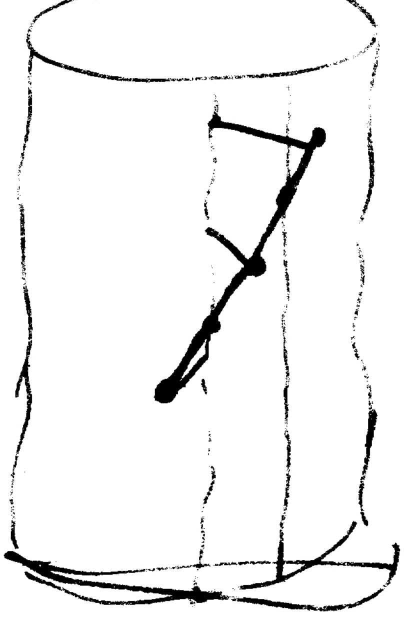


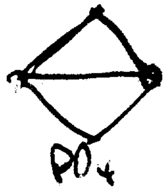
1. Tip column to right around O.
2. Bend top forward around AB (and to left).
3. Bend top forward & to right around CD.

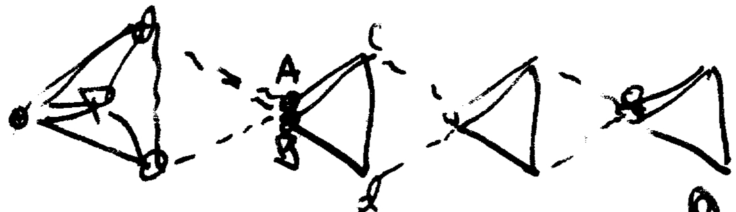




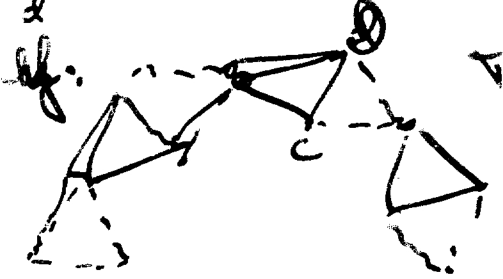








Bend around AB only:

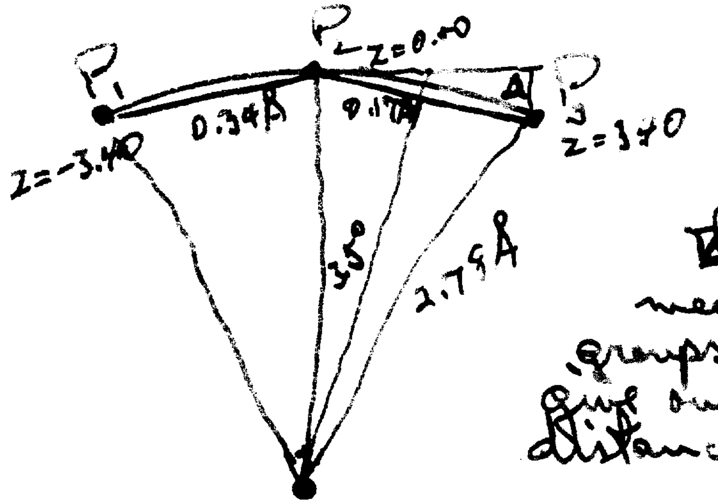


This gives a ring.

Now bend around CD: This gives a helix.

I assume 7° as angle of operation (from Astbury's 17).

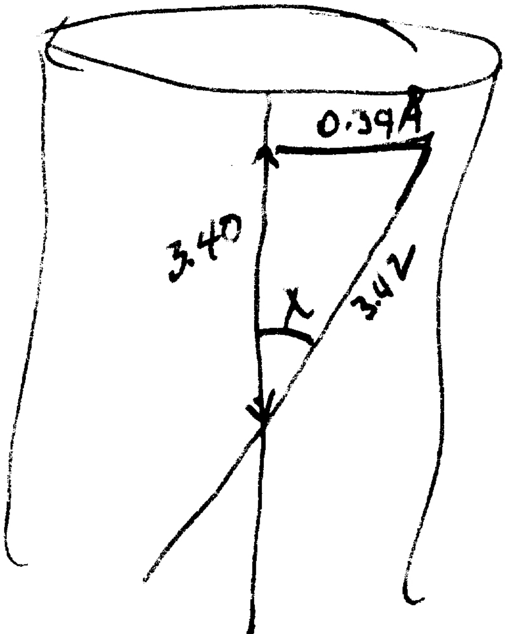
Radius to P = 2.78 Å. Translation along z = 3.40 Å



$P_1 - P_2 = 3.42$ Å

$\Delta = 0.34 \cdot \sin 3.5^\circ = 0.02$ Å

Δ seems awfully small. This means that very little tip of PO4 groups around O-O axis is enough to give our helix, and contact O--O distances change very little.



$\chi = 5.8^\circ$ (at radius of P atom)

This is the angle through which the PO_4 tetrahedra are to be rotated.

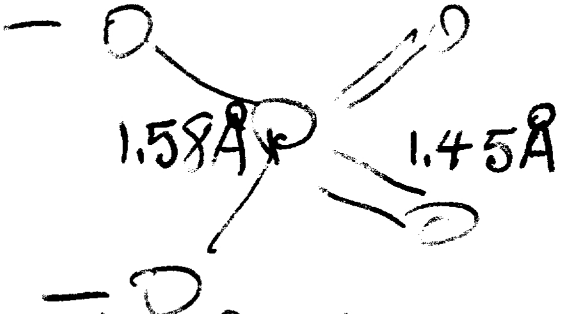
On p. 6 we give $\chi = 6.65^\circ$.

Why are the PO_4 in a column so close together.

They are at 3.42 Å (if 58° = χ is right), whereas we predict 3.74Å, assuming $O \cdots O = 2.65$ Å.

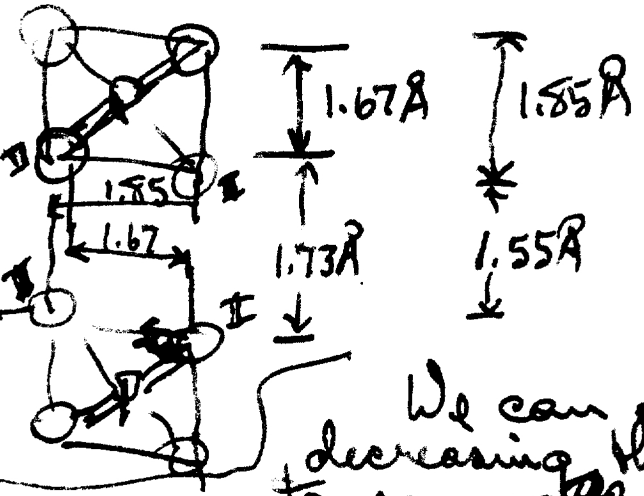
Answer: The PO_4 tetrahedra are not regular, with $P-O = 1.54$ Å.

We have here a diester, for which interatomic distances can be predicted (from S_3O_9, P_4O_{10}, etc.).



As a guess now (to be corrected later) I select $P=O = 1.45$ Å and $P-O = 1.60$ Å.

Tetrahedron in normal orientation:



With no rotation:
$O_I \cdots O_I = 2.41$ Å
$O_{II} \cdots O_{II} = 2.41$ Å
assuming tetrahedral angles.

We can gain a little by decreasing the O_I-P-O_I angle, to say 103°, increasing $O_{II}=P=O_{II}$ to 116°. Then all 12 $O \cdots O$ distances are nearly equal:

$O_I \cdots O_I = 2.50$ Å
$O_I \cdots O_{II} = 2.49$ Å
$O_{II} \cdots O_{II} = 2.46$ Å

We can gain a little by decreasing ∠O_I-P-O_I to 102° and decreasing O_{II}-P-O_{II} to 102 (some evidence exists for this – O_{II} has only 2 unshared pairs).

109.5
-102
7.5
3.75
109.5
113.25

$O_I \cdots O_I = 2.49$ Å
$O_{II} \cdots O_{II} = 2.26$ Å
∠O_I-P-O_{II} = 113.3°
$O_I \cdots O_{II} = 2.55$ Å

Because of uncertainty about ∠O=P=O, I leave it at 109°28′, and reduce ∠O_I-P-O_I ~~until $O_I \cdots O_I = O_{II} \cdots O_{II}$~~ to 100°.

∠O_{II}-P-O_{II} = 110° $O_{II} \cdots O_{II}$ = 2.38 Å
∠O_I-P-O_I = 100° $O_I \cdots O_I$ = 2.45 Å
∠O_I-P=O_{II} = 113.75° $O_I \cdots O_{II}$ = 2.53 Å

219°
-105
114°
109.5
4.25
117.25

$\frac{2.41}{\sqrt{2}} = 1.70$ Å. Hence contact distance between tetrahedra, assuming 3.40 Å, is 2.47 Å. This is still too small.

Perhaps Astbury's 17 is really 12 in 3 turns (that is, 3×17 = 51 residues in 2 turns of a chain). Then rotation is by 14°. This doesn't help much.

109.5
4.75
114.25

Let us put $\angle O_I - P - O_I = 100°$
$\angle O_{II} = P = O_{II} = 100°$
$\angle O_I - P = O_{II} = 114.2°$

$O_I \cdots O_I = 2.45$ Å
$O_{II} \cdots O_{II} = 2.22$ Å
$O_I \cdots O_{II} = 2.56$ Å

$\sqrt{2} = 1.73$
$\sqrt{2} = 1.57$
} 0.5

3.40
-1.65
1.75

With P-P = 3.40 Å, contact O…O ≅ **2.52 Å**

Hence by trying hard we get O…O nearly large enough.

If Astbury's value 3.34 Å ~~was~~ a bit low all would be well.

Perhaps the angles are still larger. (Check S_3O_9).

Let $O_I - P - O_I$ ~~=~~ 90° $O_{II} = P = O_{II} = 100°$
$O_I - P = O_{II} = 118.7°$ }

$O_I \cdots O_I = 2.26$ Å
$O_{II} \cdots O_{II} = 2.22$ Å
$O_I \cdots O_{II} = 2.60$ Å

$\sqrt{2} = 1.60$
$\sqrt{2} = 1.57$
} 0.7

3.40
-1.58
1.82

$O_I \cdots O_{II} =$ **2.59 Å** This is all right

Moreover, the pentose links from column to column are tight ones — it is ~~barely~~ possible ~~to~~ hold to the accepted ~~structural~~ parameters, when the phosphate core is very compact. The whole structure must be unusually rigid. The small z translation 3.40 Å is required by the short length of the pentose.

Williams EM of Na·RNA shows stiffness of molecules. Minimum radius of curvature ~ 1000 Å?

MHF

Na nucleate - Wilkins, R.G. Gosling, + W.E. Seeds,
Nature 167 759 (1951)
Nucleic Acid: an Extensible Molecule?

Fibers neck down on stretching & become opt. pos. [mics now at 45°.] Break at double length.

X-ray spacing does not change on drying.

MJ Fraser + RDB Fraser ibid p. 761.

Deoxyribonucleic acid & polarized infra red.

Astbury observed merid! arcs at 3.34 Å & 27 Å.

Chemistry

A Proposed Structure for the Nucleic Acids

By Linus Pauling and Robert B. Corey

Gates and Crellin Laboratories of Chemistry,* California
Institute of Technology, Pasadena 4, Calif.

(Communicated December 1952)

The nucleic acids seem to be comparable in importance to the proteins, as constituents of living organisms. There is evidence that they are involved in the processes of cell division and growth, ~~and~~ that they participate in the transmission of hereditary characters, and that they ~~seem~~ are ~~to be~~ important constituents of viruses, ~~as well as of bacteria.~~ An understanding of the molecular structure of the nucleic acids should ~~pro-vide~~ be of value in the effort to understand the fundamental ~~biological~~ ~~processes~~ phenomena of ~~biology~~ life.

Only recently has reasonably complete information been gathered about the chemical nature of nucleic acids. ~~They consist of~~ The nucleic acids are giant molecules, composed of complex units. Each unit consists of a phosphate ion, HPO_3^{--}, a sugar (ribose in the ribonucleic acids, deoxyribose in the deoxyribonucleic acids), and a purine or pyrimidine side chain (adenine, guanine, thymine, cytosine, uracil, 5-methylcytosine). The purine or pyrimidine is attached to carbon atom 1′ of the sugar. ~~Only~~ recently, through the investigations of Todd and his collaborators;[1] ~~has~~ Good evidence been obtained as to the nature of the linkage between the sugar and the phosphate; it seems likely that the phosphate ester links involve carbon atoms 3′ and 5′ of the ribose or deoxyribose.

X-ray photographs have been made of sodium thymonucleate and other preparations of nucleic acids by Astbury and Bell[2] and, more recently, by [3]. Some information about the nature of the structures has been obtained from these photographs, but it has not been found possible to derive detailed structures from the x-ray data.

We have now formulated a promising structure, by making use of the general principles of molecular structure and the available information about the nucleic acids. This structure is not a vague one, but is precisely predicted; atomic coordinates for the principal atoms are given in ~~the following pages~~ Table 1. This is the first precisely described structure for the nucleic acids that has been suggested. The structure accounts for some of the features of the x-ray photographs; ~~no~~ but detailed intensity calculations have not as yet been made,

~~however,~~ and the structure cannot be considered to have been proved to be correct.

The formulation of the structure. - The most important configuration of polypeptide chains in proteins is the α helix.[4] In this structure the amino-acid residues are equivalent (except for differences in the side chains); there is only one type of relation between a residue and neighboring residues, one operation which converts a residue into a following residue. Through the continued application of this operation, a ~~rotatory~~ rotation-translation, the α helix is built up. It seems not unlikely that a single general operation is also involved in the construction of nucleic acids from their asymmetric fundamental units. The general operation involved would be a rotation-reflection, and its application would lead to a helical structure. We assume, accordingly, that the structure to be formulated is a helix. The giant molecule would thus be cylindrical, with approximately circular cross section.

Some evidence in support of this assumption is provided by the electron micrographs of preparations of sodium thymonucleate described by Williams.[5] The preparation seen in the shadowed electron micrographs is clearly fibrous/ in nature. The small fibrils or molecules seem to be circular in cross section, ~~that~~ and their diameter is apparently constant; there is no evidence that the molecules are ribbon-like. The diameter as estimated from the length of the shadow ~~is~~ has been estimated at 15 or 20 A.

The x-ray photographs of sodium thymonucleate show a strong equatorial reflection at 16.2 A. If it is assumed that this is due to a hexagonal packing of cylindrical molecules, the diameter of the molecules is 18.7 A. From the average residue weight of sodium thymonucleate, about 330, and the density, about 1.62 g cm^{-3}, we calculate that the volume per residue is 338 A^3. The cross-sectional area ~~is~~ per residue is 303 A^2; accordingly the length per residue along the fiber axis is about 1.12 A.

The x-ray photographs show a very strong meridional reflection, with spacing about ~~[illegible]~~ 3.40 A. This reflection corresponds to a distance along the fiber axis equal to three times the distance per residue. Accordingly, the reflection is to be attributed to three residues.

If the molecule of a nucleic acid is a single helix, the reflection at 3.4 A would have to be attributed to a regularity in the purine-pyrimidine sequence - that is, to a regular sequence of nucleotides, involving repetition of a structural unit of three nucleotides. It seems unlikely that the nucleotides repeat in this regular way; it is likely instead that the nucleic acids, like the proteins (insulin), involve a less regular sequence of the fundamental units. The alternative explanation of the strong 3.4-Å meridional reflection is that the cylindrical molecule is formed of three chains, which are coiled about one another. The structure described below is a three-chain structure, each chain being a helix with fundamental translation equal to 3.4 A. Each of the three helical chains is tightly coiled, with a little more ~~or a little less~~ than three residues per turn of the helix. The pitch of the helix representing a single chain is ~~approximately~~ ~~10~~ 11.65 A. The three chains interpenetrate, in such a way that the ~~lead~~ pitch of the triple helix is ~~about~~ ~~3.4~~ 3.88 A.

The first question to be answered is that as to the nature of the core of the three-chain helical molecule - the part of the molecule closest to the axis. It is important for stability of the molecule that atoms be well packed together, ~~It seems likely~~ and the problem of packing atoms together is a more difficult one to solve in the neighborhood of the axis than ~~of~~ at a distance away from the axis, where there is a larger distance between an atom and the equivalent atom in the next unit. (An example of a helical structure which seems to ~~be~~ satisfy all of the structural requirements except that of close packing of atoms in the region near the helical axis is the 5.2-residue helix (the γ helix) of polypeptide chains. This structure~~s~~ seems not to be represented in proteins, whereas the ~~closely~~ similar α helix, in which the atoms are packed in a satisfactorily close manner about the axis, is an important structure.) There are three possibilities as to the composition of the core; it may consist of the purine-pyrimidine groups, the sugar residues, or the phosphate groups ~~phosphoric acid~~. It is found by trial that ~~Because of their varied nature, it is improbable that~~ the purine-pyrimidine groups cannot ~~could~~ be packed along the axis of the helix in such a way that suitable bonds can ~~could~~ be formed between the sugar and the phosphate groups ~~phosphoric acid~~; this ~~possibility~~ choice is accordingly eliminated. It is also unlikely that the sugar groups constitute the core of the molecule; The shape of the ribose molecule and the deoxyribose molecule is such that close packing of these molecules along a helical axis is difficult, and no satisfactory way of packing them has been found. An example ~~of~~ that shows the difficulty of achieving close packing is provided by the polysaccharide starch, which forms helixes with a hole along the axis, into which iodine molecules can fit. We conclude that the core of the molecule is probably formed of the phospho~~ric acid~~ate groups.

A close-packed core of phosphoric acid residues, HPO_4^{--}, can easily be formed. At each level along the fiber axis there are three phosphate groups. These can be packed together in the way shown in Figure 1. Two oxygen atoms of each tetrahedral phosphate group form an octahedron, the trigonal axis of ~~each tetrahedral phosphate~~ which is the ~~fiber~~ axis of the three-chain helical molecule. A similar complex of three phosphate tetrahedra can be superimposed on this one, with ~~little or no~~ only a small change in azimuthal orientation. The neighborhood of the axis of the molecule is then filled with oxygen atoms, arranged in groups of three, which change their azimuthal orientation by about 60° from layer to layer, in such a way as to produce approximate closest packing of these atoms.

The altitude of a phosphate ion ~~(P-O = 1.56 A)~~ is about 1.7 A. If the same distance were preserved between the next oxygen layers, the basal-plane distance along the fiber axis would be 3.4~~0~~ A. This value is ~~close to~~ the spacing observed for the principal meridional reflection. It is to be expected that the outer oxygen atoms of the complex of three phosphate groups would be attached to the ribose or deoxyribose, and that the hydrogen atom of the HPO_4^{--} residues would be attached to an inner oxygen atom, and presumably would be involved in hydrogen-bond formation with another of the inner oxygen atoms. The length of the O-H...O bond should be close to that observed in KH_2PO_4, 2.55 A. The angle P-O-H should be approximately the tetrahedral angle. It is found that the spacing 3.4 A is not preserved, with this bond angle, if the hydrogen bonds are formed between one $(HPO_4)_3$ group and the group above or below. Accordingly we assume that hydrogen bonds are formed between the oxygen atoms of the phosphate groups in the same basal plane, as indicated in Figure 1.

If the bond angle P-O-H is assumed to be the tetrahedral angle, and the hydrogen bonds O-H···O are assumed to be linear, the phosphate groups must be rotated by 6.7°, in such a direction as to bring the plane of the inner oxygen atoms closer to the plane of the phosphorus atoms. The *z* parameter of the inner oxygen atoms then becomes ± 0.76 A, with that of P equal to 0.00 A. The *z* parameter of the outer oxygen atoms is ± 0.96 A.

The radius of the inner oxygen atoms (the distance from the axis of the molecule) is found to be 2.11 A, assuming the values given above for the P-O and O-H•••O distances. The parameters of the phosphorus atom and the outer oxygen atoms are easily calculated, and are given in Table 1.

If the oxygen atoms in the next layer are placed at equal distances from those in the first layer, it is found that the group of three tetrahedra is to be rotated through 6°, ~~x~~ while being translated by 3.40° along the z axis. The oxygen-oxygen contact distances are 2.45 A (in the phosphate tetrahedron), 2.55 A (O-H•••O distance), 3.45 A (in the basal plane), and 2.74 A (diagonal distance, between $(HPO_4)_3$ groups). It is found that a ribose residue may be bridged across between the upper oxygen atom of a tetrahedron and the lower oxygen atom of the tetrahedron above it, and rotated by approximately 120° (114° or 126°) in azimuth. The bridging may be achieved for either the right-handed screw arrangement of phosphate tetrahedra, shown in Figure 1, or the left-handed screw, the mirror image of this. However, the right-handed screw seems to be the better, in several respects. In order to form ester linkages with carbon atom 2 and carbon atom 5 of the ribose residue, with the furanose-ring configuration, the plane of the 5-membered ribose ring must be placed nearly at right angles to the basal plane (perpendicular to the axis of the nucleic acid molecule), if the left-handed configuration is used for the phosphate complex. There then occurs steric hindrance between the ribose residue and the similar residue almost directly above it - the rotation by 6° corresponds to a lateral translation, at the radius (about 6 A) of the center of the ribose ring, of only about 1 A, which is not enough to permit the atoms of the two residues to clear one another. For the right-handed configuration of the phosphate complex the plane of the ribose ring is at about 45° with the basal plane, and satisfactory packing of the sugar residues is achieved.

Also, the angle between the ~~C-N axis~~ C1-N axis, where N is the nitrogen atom of the purine or pyrimidine group, and the basal plane is about 25° for the left ~~right~~-handed phosphate complex, and about 10° for the right-handed complex. The nucleic acids are observed to have strong negative birefringence. This anisotropy in ~~optical~~ index of refraction is to be attributed ~~largely to~~ almost entirely to the purine and pyrimidine planes, and it provides strong evidence that the planes of these conjugated systems are nearly parallel to the basal plane of the molecule. The structure involving right-handed phosphate complexes accordingly provides a more satisfactory explanation of the birefringence than does the other structure.

Coordinates of the atoms of the ribose ~~[illegible]~~ residue and of the nitrogen atom of the purine or pyrimidine group are given in Table 1. These coordinates are subject to greater uncertainty than those for the phosphate groups. The way in which the ribose residue bridges the region between one ~~Description of the structure~~ phosphate group and the next in the nucleic acid chain is shown in Figure 2.

<u>Description of the structure</u>. - In the proposed structure each nucleic acid chain forms a tightly coiled helix, with approximately three ribose-phosphate residues per turn of the helix. The lead of the helix (the distance along the fiber axis from one ~~turn of the helix~~ position on the chain to the corresponding position on the same chain after one complete turn) is approximately 10 A. Three identical chains are intertwined. A single chain is represented in Figure 3. In the complete molecule three identical chains are intertwined, as shown in Figure 4, to form a closely packed three-chain helical molecule. The three chains are attached to one another by lateral

hydrogen bonds between the inner oxygens of the phosphate groups.

The diameter of the three-chain molecule, taking into consideration the size of the purine-pyrimidine groups, is about 20 A.

If the oxygen atoms in one phosphate complex are equidistant from neighboring oxygen atoms in the next complex, the ~~rotational~~ basic angle of rotation of the helix differs from 120° by 6°. ~~The~~ An individual chain accordingly has an identity distance of approximately 60 times the axial length per residue, 3.4 A. The identity distance of the three-chain molecule is predicted to be one third as great, about 20 times 3.4 A = 68 A. ~~This~~ The identity distance cannot be predicted very accurately, however, because a considerable change can be made without causing the oxygen-oxygen contact distances to be unsatisfactory. If the rotation differed from 120° by 12°, rather than by 6°, the oxygen-oxygen contact distances would be 2.69 A and 2.77 A, respectively. These values are acceptable, and accordingly the identity distance might be that corresponding to a 12° rotation, which is 10 times 3.4 A = 34 A. The x-ray photographs indicate an identity distance along the fiber axis of approximately 50 A.

We plan to make a detailed comparison of intensities and other features of x-ray photogrpphs of nucleic acid preparations, and the ~~predictions~~ calculated values for the proposed structure. It should be possible to eliminate the structure, or to obtain further support for it.

This investigation was aided by grants from The National Foundation for Infantile Paralysis and The Rockefeller Foundation.