26 November 1952.

Nucleic Acid

Lipman says e.g. up to 16.2 Å,

\[ 16.2^2 \cdot \frac{2}{\sqrt{5}} = 303 \text{ Å}^2 \]

\[ ho \text{ W.} = 330, \text{ Density 1.62} \quad \therefore \text{ M.V.} = 355 \text{ Å}^3. \]

\[ 1.12 \text{ Å per residue.} \]

Observed: 3.34 Å.

Perhaps we have a triple-chain structure!

Size of molecule. Yesterday (25 Nov. 1952)
in a biochemistry seminar Rosalyn Williams showed a slide of a nucleic acid molecule and said that the small filaments have a diameter of 15 Å.

I asked about the size, and she repeated 15 Å, and discussed the difficulty of measuring such small objects. Ribozyme spheres with diameter 250 Å were 20 times larger. Only one diameter; i.e., cylindrical.

\[ 1.12 \text{ Å per residue.} \]

See also transcript book; quite references for 15 Å, 20 Å diameter.
Let us form helices with \( O-H-O \) nearly along axis
\[
\frac{2.55}{1.71} = 1.50 \text{ Å max.}
\]

I assume \( PO_4 \) oriented \( \beta \), so that three will pack together well.

To get \( \angle P-O-H = 110^\circ \),
\( H-O \) would bend back somewhat.

If it bent back in plane \( \alpha \), angle would
be \( 15^\circ \), and component
\[
2.55 \times \cos 15^\circ = 2.97 \text{ Å}
\]
\[
\frac{1.71}{+2.00} \text{ Å}
\]

I assume component \( \beta \) = 2.45 Å;
then \( \Phi P \alpha = 3.35 - 2.45 = 0.90 \text{ Å} \)

Phosphete is rotated 23.5°.

Let this be 2.30 Å—
horizontal 2.54 Å—
then \( O-O = 4.17 \text{ Å} \) —
too large.
I conclude that the hydrogen bonds are not located there, so that my assumption about bond angle is in error. I abandon assumption about $\theta - \theta - \theta$ angle.

Let us pack three PO$_4$'s in one horizontal plane. The next would pack above if rotated $90^\circ$.

We first O-H...O-bonds between A+B, either at $z = 0$ (connecting PO$_4$'s in same layer) or at $z = \frac{1}{2}$ (connecting layers).

This is wrong. They pack by having PO$_4$'s nearly over one another.
The $(\text{PO}_4)_3$ rings are nearly over one another.

The $\text{P-O-H}$ bond angles require that $H$ be at $z=0$, i.e., on the level of P. Thus there are $(\text{PO}_4 + H)_3$ rings:

\[
\begin{align*}
\frac{3.67}{3.83} &= 0.96
\end{align*}
\]

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

This makes $O-O = 2.45\text{Å}$ in $\text{PO}_4$

2.55Å in O-H...O

3.16Å in basal plane

2.49Å between $(\text{PO}_4)_3$ layers.

The last is too small. Probably PO$_4$'s are tipped,
Flipping of PO₄'s.

Measurement of a model gives
L = 1.52 Å. L' between (PO₄)'s = 1.88 Å

Hence with L = 1.52 Å, L' = 1.88 Å we have

\[ \text{O-O} = 2.45 \text{ Å in PO₄} \]
\[ = 2.55 \text{ Å for O-H...O (diagonal, } z=0) \]
\[ = 3.45 \text{ Å in basal plane} \]
\[ = 2.74 \text{ Å diagonal, } z=\frac{1}{2} \]

If O₃ + O₂ make 60° angles in projection, rotation from \( z=0 \) to \( z=1 \) is 60°; i.e., repeat after 60 layers (really 20).

If notation were 12° repeat in 30 layers (really 12) - 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 x 3.4 Å, pitch = \( \frac{2\pi}{5} = 20 x 3.4 = 68 \text{ Å} \))

but considerably smaller or larger would be most unlikely.
The PO₄ groups are rotated 6.65°.
Inner O's at \( z = \pm 0.76 \text{ Å} \), outer O's at \( z = 0.96 \text{ Å} \)

Note that each chain has \( \varphi = 113° \) or \( 127° \)
that is, roughly 3 residues per turn. There are three chains closely intertwined, and held together by hydrogen bonds between PO₄'s.

The ribose residues connect upper O at \( z = 0.76 \text{ Å} \), \( \varphi = 0° \) with lower O, \( z = 2 - 0.76 \text{ Å} \), \( \varphi = 120° \).

The (PO₄) may be either right-handed or left-handed (screw determined by O-H...O).

The ribose groups will loop across either.

With my chiral for ribose, the right-handed looks better. For if 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.

\( \varphi = 5 \text{ Å} \) for center of ribose ring,
\( \alpha = 12° \) would give displacement \( 1 \text{ Å} \), not enough for rings to clash with another.

I haven't checked the chain for the ribose ring string across between PO₄'s very carefully. There may be some trouble, although the Van der Waals constants seem to be...
With right-handed (Na, D) the maximum stick-out at \( \phi = 350^\circ \) to axis, which agrees with large negative bindings, that is observed. They have first at \( \phi = 78^\circ \).

With left-handed (HPO_4) the peak at \( \phi = 55^\circ \), which is probably in agreement.
29 November 1952

Helical arrangements of PO₄ tetrahedra.

A straight column. Assume P-O = 1.54 Å
Then O-O = 2.54 Å in the column
1.75 Å = height
Assume O-O = 2.65 Å between tetrahedra
Then vertical distance = 1.96 Å.

The columns of linked tetrahedra can be deformed into a helix without change in the distances D...O distances (2.54 Å, 2.65 Å) by rotating around the horizontal O1 O axis.

1. Column to right around D
2. Hand to forward around AB (and to left)
3. Hand to forward + tonight around D.
Now bend around E0: This gives a helix.
Assume 7° as angle of rotation (from Astbury's 17).
Radius to \( P = 2.78 \) Å. Translation along \( z = 3.40 \) Å

\[ \Delta = 0.34 \times 2 \times 3.5^\circ = 0.42 \] Å

Seems a little small. This means that very little (7°) in the 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 0.35 Å)

Which is the angle through which the tetrahedra are to be rotated.

On p.6 we give \( \chi = 6.65^\circ \).
Why are the P-O in a column 20 close together? They are at 3.42 Å (if 58° = X, is a right), whereas we predict 3.74 Å, assuming O...O = 2.65 Å.

Answer: The P-O tetrahedra are not regular, with P-O = 1.54 Å.
We have here a diester, for which intraatomic distances can be predicted (from O3O9, P+O10, 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:
O1 - O1 = 2.41 Å
O1 - O1 = 2.41 Å
assuming tetrahedral angles.

We can gain a little by decreasing the O1 - P - O1 angle to say 105°, increasing O1 = P = O1 to 116°. Then all 12 O...O distances are nearly equal:

O1...O1 = 2.56 Å
O1...O1 = 2.49 Å
O1...O1 = 2.46 Å
We can gain a little by decreasing $\angle D_2 - D_3 - D_4$ to $102^\circ$ and decreasing $\angle D_3 - D_4 - D_3$ to $10^\circ$ (some evidence exists for this -- $D_4$ has only 2 unshared pairs).

\[
\begin{align*}
D_3 - D_1 &= 2.40 \text{ Å} \\
D_3 - D_4 &= 2.26 \text{ Å} \\
D_4 - D_5 &= 1.13 \text{ Å} \\
D_1 - D_4 &= 2.55 \text{ Å}
\end{align*}
\]

Because of uncertainty about $\angle D_1 - D_3 - D_2$, I leave it at $109.25^\circ$ and reduce $\angle D_3 - D_4 - D_5$ to $100^\circ$.

\[
\begin{align*}
\angle D_3 - D_4 &= 110^\circ \\
\angle D_4 - D_5 &= 100^\circ \\
\angle D_1 - D_4 &= 113.75^\circ \\
\angle D_1 - D_3 &= 2.53 \text{ Å}
\end{align*}
\]

\[
\frac{2.41}{\sqrt{2}} = 1.70 \text{ Å}.
\]

Using contact distance between tetrahedra, assuming $3.40 \text{ Å}$, is $2.47 \text{ Å}$. This is still too small.

Perhaps placing 17 is really 12 in 3 turns (that is, 3 x 17 = 51 residues in 2 turns of 3 residues). This doesn't help much.
Let us put $\angle O_1 - P - O_2 = 120^\circ$
$\angle O_2 - P - O_3 = 120^\circ$
$\angle O_1 - P - O_3 = 114.2^\circ$

$O_1 \ldots O_5 = 3.45 \text{ Å}$
$O_2 \ldots O_5 = 2.22 \text{ Å}$
$O_2 \ldots O_4 = 2.56 \text{ Å}$

With $O_1 - P = 3.40 \text{ Å}$, contact $O_1 \ldots O_2 = 2.52 \text{ Å}$

Hence from trigonometry we get $O_1 \ldots O_2$ nearly large enough.

If Sutherland's value $3.34 \text{ Å}$ were a bit less, all would be well.

Perhaps the angles are still larger.

(2.3 $O_9$)

Le $O_1 - P - O_5 \quad \angle O_1 - P - O_5 = 108^\circ$
$O_1 - P - O_3 = 118.7^\circ$

$O_1 \ldots O_5 = 2.24 \text{ Å}$
$O_2 \ldots O_5 = 2.22 \text{ Å}$
$O_3 \ldots O_5 = 2.60 \text{ Å}$
$O_3 \ldots O_4 = 2.59 \text{ Å}$

This is all right.

Moreover, the pentose units from column to column are tight, since it is claimed possible to hold to the accepted structural parameters, where the phosphate core is very compact. The whole structure must be uniformly rigid. The small $z$ translation $3.40 \text{ Å}$ is required by the short length of the pentose.

Williams EM 13 Na-PNA shows stiffness of molecules.
Minimum radius of curvature = 1000 Å?
MHF


Nucleic Acids on Entrapped Fibers?

- Fiber neck down on stretching and become opt. axis.
- Break at double length.

X-ray spacing does not change on drying.

MH Fraser & RDB Fraser, 'bid, p. 761.

Dempierhamelid acid & polarized infra red.

Ashby observed merid. axes at 3.34Å & 2.7Å.
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. They are to be important constituents of viruses, as well as of bacteria. An understanding of the molecular structure of the nucleic acids should be of value in the effort to understand the fundamental biological phenomena of biological life.

Only recently has complete information been gathered about the chemical nature of nucleic acids. The nucleic acids are giant molecules, composed of complex units. Each unit consists of a phosphate ion, $\text{HPO}_4^{2-}$, a sugar (ribose in the ribonucleic acids, deoxyribose in the
deoxyribonucleic acids), and a purine or pyrimidine side chain (adenine, guanine, thymine, cytosine, uracil). The purine or pyrimidine is attached to carbon atom 1' of the sugar. Only recently, through the investigations of Todd and his collaborators, 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, 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. The structure is not a vague one, but is precisely predicted; atomic coordinates for the principal atoms are given in the following page. The structure accounts for some of the features of the x-ray photographs; no intensity calculations have as yet been made.
The formulation of the structure. - The most important configuration of polypeptide chains in proteins is the a helix. 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 rotation the continued application of this operation, a translation, the a helix is built up. It seems not unlikely that a single general operation asymmetric is also involved in the construction of nucleic acids from their/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. 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 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 has been estimated at 15 or 20 Å.

The x-ray photographs of sodium thymonucleate show a strong equatorial reflection at 16.2 Å. If it is assumed that this is due to a hexagonal packing of cylindrical molecules, the diameter of the molecules is 18.7 Å. 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 Å³. The cross-sectional area per residue is 303 Å²; accordingly the length per residue along the fiber axis is about 1.12 Å.

The x-ray photographs show a strong meridional reflection, with spacing about 3.40 Å. This reflection corresponds to a distance 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 Å 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 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-A 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 than three residues per turn of the helix. The pitch of the helix representing a single chain is approximately 10 A. The three chains interpenetrate, in such a way that the core of the triple helix is about 3.8 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, and the problem of packing atoms together is a more difficult one to solve in the neighborhood of the axis than 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 \textit{im} satisfy all of the structural require-
ments except that of close packing of atoms in the region near the helical
axis is the 5.2-residue helix (the \( \gamma \) helix) of polypeptide chains. This
structure seems not to be represented in proteins, whereas the closely
packed in a
similar \( \alpha \) helix, in which the atoms are \textit{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 \textit{purine-pyrimidine}
groups, the sugar or the phosphorus. Because of their varied nature,
it is improbable that the purine-pyrimidine groups could be packed along the
axis of the helix in such a way that suitable bonds could be formed between
the sugar and the phosphorus; this possibility choice is accordingly
eliminated. It is unlikely that the sugar groups \textit{are} constitute the core
of the molecule. The shape of the ribose molecule and the deoxyribose mole-
cule is such that close packing of these molecules along a helical axis is
difficult. An example \textit{of} that shows the difficulty of achieving close
packing is provided by the polysaccharide starch, which forms helixes with
a hole \textit{along} the axis, into which iodine molecules can fit. We conclude
that the core of the molecule is probably formed of the phosphoric acid groups.

A close-packed core of phosphoric acid residues, $\text{HPO}_4^{2-}$, 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 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 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 $60^\circ$ from layer to layer, in such a way as to produce closest packing of these atoms.

The altitude of a phosphate ion ($\text{PO}_4^{3-}$) is $1.75$ Å. If the same distance were preserved between the next oxygen layers, the basal-plane distance along the fiber axis would be $3.48$ Å. This value is close to the spacing observed for the principal meridional reflection, suggesting...

It is to be expected that the outer oxygen atoms of the complex of three phos-
phosphate groups would be attached to the ribose or deoxyribose, and that the hydrogen atom of the $\text{HPO}_4^{3-}$ 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 $\text{O}-\text{H}...\text{O}$ bond should be close to that observed in $\text{KH}_2\text{PO}_4$, 2.55 Å. The angle $\text{P}-\text{O}-\text{H}$ should be approximately the tetrahedral angle. It is found that the spacing 3.4 Å is not preserved, with this bond angle, if the hydrogen bonds are formed between one $(\text{HPO}_4)_3$ 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 $\text{P}-\text{O}-\text{H}$ is assumed to be the tetrahedral angle, and the hydrogen bonds $\text{O}-\text{H}...\text{O}$ 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 $\pm 0.76$ Å, with that of P equal to 0.00 Å. The $z$ parameter of the outer oxygen atoms is $\pm 0.96$ Å.
The radius of the inner oxygen atoms (the distance from the axis of the molecule) is found to be 2.11 Å, assuming the values given above for the P-O and O-H\cdots 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 60°, while being translated by 3.40° along the z axis. The oxygen-oxygen contact distances are 2.45 Å (in the phosphate tetrahedron), 2.55 Å (O-H\cdots O distance), 3.45 Å (in the basal plane), and 2.74 Å (diagonal distance, between (HPO₄)₂ 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 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 simi-
lar residue almost directly above it - the rotation by 60° corresponds to a
lateral translation, at the radius (about 6 Å) of the center of the ribose
ring, of only about 1 Å, 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 Cl-N axis, where N is the
nitrogen atom of the purine or pyrimidine group, and the basal plane is about
left 25° for the 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 the index of refraction is to be at-
tributed 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 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 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 phosphate group and the next in the nucleic acid chain is shown in Figure 2.

Description of the structure. - 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 position on the chain to the corresponding position on the same chain after one complete turn) is approximately 10 Å. Three 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 Å.

If the oxygen atoms in one phosphate complex are equidistant from basic neighboring oxygen atoms in the next complex, the rotation of the helix differs from 120° by 6°. An individual molecule accordingly has an identity distance of approximately 60 times the axial length per residue, 3.4 Å. The identity distance of the three-chain molecule is predicted to be one third as great, about 20 times 3.4 Å = 68 Å. 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 Å and 2.77 Å, respectively. These values are acceptable, and accordingly the identity distance might be that corresponding to a 12° rotation, which is 10 times 3.4 Å = 34 Å. The x-ray photographs indicate an identity distance along the fiber axis of approximately 50 Å.
We plan to make a detailed comparison of intensities and other features of x-ray photographs of nucleic acid preparations, and the predicted 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.