

FINAL REPORT TO TURNER AND NEWALL FELLOWSHIP COMMITTEE

Rosalind E. Franklin,

Birkbeck College,
University of London,
21 Torrington Square,
London, W.C.1.

The work done during the period January 1st, 1951 to January 1st, 1954, has been described in a series of publications which were presented with last year's report. A list of references to these is appended. One further publication on the same subject has now been sent to press. The title is "The 3-Dimensional Patterson F_0 function of Desoxyribose Nucleic Acid". The content of this paper has been described in previous reports.

THE STRUCTURE OF TOBACCO MOSAIC VIRUS

Research carried out in 1954 has been concerned with the structure of tobacco mosaic virus. Since the previous report was written greatly improved X-ray diffraction photographs have been obtained, and studied in detail. The best photographs were from a gel of about 25% by weight TMV, spontaneously orientated in a fine glass capillary (Fig. 1). Since, in such a gel, the virus particles are fully dispersed in the plane perpendicular to their long axes and are in random rotation about these axes (Bernal and Fankuchen, 1942) such a photograph represents the continuous Fourier transform of a single virus particle. This has a repeat period of 69 Å along the c-axis, but no other lattice periodicity.

Improved measurements of the position of the innermost

intensity maximum on each layer-line confirm the conclusion of Watson (1954) that the virus particle is a giant helical molecule, the repeat period of 69 Å containing 3 turns of the helix, on which lie $3n+1$ building units. The value of n is most probably 12 (Watson suggests that n is 10; but the strong maximum on the 31st layer line which he believed to be meridional has now been clearly resolved into two maxima well removed from the meridian. The maximum on the 37th layer line has not, so far, been so resolved).

CALCULATION OF THE CYLINDRICALLY AVERAGED PATTERSON FUNCTION

The intensity distribution on some 25 layer lines was measured with a recording microphotometer, and used to calculate the cylindrically averaged Patterson function of the virus particle.

Since the intensity distribution within a layer-line is continuous, with rapid fluctuations but no regular lattice repeat, it could not be satisfactorily represented by means of sampling at regular intervals. Instead, therefore, for the purpose of computing the Patterson function, the intensity on each layer line was decomposed into a series of Gaussians, $A_n e^{-a_1^2 r^2}$, the value of a_1 being constant for a given layer line, and chosen to correspond with the width of the narrowest peak on the layer line l . The Bessel function transform $\sum_n A_n J_0(2\pi r z_n)$ was then carried out for a series of peak functions A_n for values of r from 0 to 160 Å, and the resulting function multiplied by the transform Gaussian $e^{-(\sqrt{\pi}/a_1)^2 r^2}$ to obtain the transform of the continuous intensity distribution of the layer line. After carrying out this operation on each of the layer lines, a series of cosine transforms gives the cylindrically

averaged Patterson function (Fig. 2a). For large values of γ , the calculation was repeated using a very strong "artificial temperature factor" (Fig. 2b).

The Equator. The first few equatorial maxima correspond closely, in both relative intensity and in position, to those which would be given by a uniform-density rod of diameter 140-150 A. The first maximum, at about 90 A, is by far the strongest of the diagram. Its inclusion in the Patterson computation imposes a strong oscillation of period 90 A on the whole Patterson function. This is obviously spurious, and is due to the omission of the non-observable central maximum. To eliminate the effect with a minimum of arbitrariness, the following device was used.

The intensity and position of the first two equatorial maxima were used to determine the intensity of the non-observable central peak which would be given by uniform-density rods of diameter 150 A scattering independently. The shape of this theoretical peak is known. Its transform was calculated, and added to the Patterson function. The resulting complete Patterson function of the virus particle no longer shows the 90 A oscillation, but has the inconvenience of rising very steeply at small values of γ . To eliminate this, the Patterson function of structureless rods of the same mean density (and diameter 150 A) was subtracted. The final Patterson function, shown in Fig. 2, thus relates to the virus particles embedded in a structureless medium of density equal to the mean density of the virus.

INTERPRETATION OF THE PATTERSON FUNCTION

The experience of others who have studied the Patterson

functions of large molecules and in particular of proteins, shows that great caution must be exercised in attempting an interpretation. In the present case, in the cylindrically averaged Patterson function of such a complex molecule it is most probable that important features of the structure are entirely obscured, and others distorted. Nevertheless, two features of Fig. 2 appear to be worthy of attention.

The series of peaks indicated by crosses in Fig. 2a suggests an end-on view of a double layer of rods of diameter 11 A. This is consistent with Watson's tentative suggestion that each turn of the helix contains a double layer of α -helices lying perpendicular to the axis of the virus particle. The linearity of the series of Patterson peaks indicates that the 11 A diameter rods lie in a tangential rather than a radial direction with respect to the particle axis (Fig. 3).

The peak at 140-150 A in Fig. 2b coincides with the known diameter of the particle. The shape of this peak with peaks on the base line and at the pitch height and bulging symmetrically on either side, can best be explained by supposing that the particle has an external groove, the groove following the line of the helix. Further evidence of such a groove is given below.

FOURIER PROJECTION OF THE VIRUS PARTICLE

Since the number of units on the helix in each repeat period in the axial direction is large (probably 37), the projection of the structure on a plane perpendicular to the axis of the virus particle is virtually centro-symmetric. The strong periodicity of

intensity over a considerable region of the equator, and the close resemblance between the central maxima and those which would be given by structureless rods of the same diameter, together make it possible to determine the signs of the first 12 equatorial maxima. Hence, with the aid of a strong "artificial temperature factor" to eliminate the effect of diffraction at larger angles, the Fourier projection of the virus was calculated.

The most striking feature of the Fourier projection is a shell of high density at a radius of about 55 A. One possible explanation of this is that the virus RNA lies in this shell. Alternatively, the high density may simply correspond to a shell of low hydration.

FURTHER INFORMATION DERIVED FROM X-RAY DIAGRAM

It is not, in general, possible to attribute certain parts of an X-ray diagram to particular features of the diffracting structure. However, in the case of helical structures, once the helical parameters have been established there are certain restrictions on the contribution of parts of the structure to the near-meridional reflections. This is because knowledge of the helical parameters determines which is the lowest order, n , of Bessel Function J_n contributing to each layer line. The value of $J_n(x)$ is close to zero for $x < x_n$. Since the contribution of an atom at a radius r to the amplitude scattered at ξ is proportional to $J_n(2\pi r \xi)$, then if the first maxima on any layer line occurs at the position ξ_1 , the atoms which make the principal contribution to this maxima lie on a radius r such that

$$2\pi r \xi_1 = x_n$$
$$r = \frac{x_n}{2\pi \xi_1}$$

There can be no appreciable contribution from atoms lying on a smaller radius.

Analysis of the positions of the first 4 maxima on the third layer line shows that those can only come from a J_1 contribution from a cylindrical shell of diameter 150 A. This means that at a diameter corresponding to the diameter of the virus particle there are stray density fluctuations having a period equal to the pitch of the helix. The simplest explanation of this is again that the virus particle has a screw-like contour, the groove of the screw corresponding to the helical arrangement of the protein. A semi-quantitative calculation based on the intensity of the third layer line maxima relative to that of the theoretical origin peak suggests that the depth of the groove is of the order of 20 A.

A similar analysis of the 6th layer line shows that the innermost maxima are due to a J_2 contribution from a shell of radius 55 A. This coincides with the high-density shell found in the Fourier projection and thus shows that this high-density shell has a strong periodicity of 11 A in the axial direction.

COMPARISON WITH CHEMICAL EVIDENCE

If the virus particle is considered as a rod of diameter 150 A length 3000 A and molecular weight 5×10^7 , and there are 37 protein building units on 3 turns of the helix in the repeat period of 69 A, then the molecular weight of the building unit is 29,000. This is about twice the molecular weight found by chemical methods (end-group analysis and amino-acid content). It has been shown that the X-ray results suggest that a turn of the helix contains a double layer of protein molecules, so that it seems that each

building unit of the helix consists of two chemically equivalent or near-equivalent parts.

The external grooving which gives the particle a rather large surface, may perhaps account for the surprising variety and extent of the chemical modifications which is possible to make in tobacco mosaic virus without breaking up the particle, and, in some cases, without destroying its infectivity.

There is thus reason to hope that further chemical work on the accessibility of the different amino-acids together with X-ray studies of the chemically modified forms may help to elucidate not only the configurations of the protein chains, but also the positions of certain amino acids in the virus.

AMERICAN JOURNEY, 20th August - 20th October, 1954

As a result of an invitation to read a paper at the Gordon Research Conference on Coal and Related Substances (at New Hampton, N.H., 22nd-27th August) two months this summer were spent in visiting American laboratories. The first part of the journey was mainly concerned with research on coal and carbon, both in universities and in industrial organisations.

The second half of the journey most of the time was spent talking with people interested in the structure of biological materials. Visits at St. Louis, Pasadena and Berkeley proved particularly valuable. In these places biologists and chemists working on tobacco mosaic virus produced a wealth of interesting information, much of it unpublished, and also showed themselves willing and anxious to send their products for X-ray investigation. Some of these have already been received.

FUTURE WORK

After the termination of the Turner and Newall Fellowship (31st December, 1954) the work on tobacco mosaic virus will be continued in this laboratory under a grant from the Agricultural Research Council. This grant also provides for the appointment of a junior research worker, an assistant, and a technician, to assist in the work and to enable the programme to be expanded.

DESCRIPTION OF DIAGRAMS

- Fig. 1. X-ray diagram of TMV gel. Orientated specimen obtained from N.W. Pirie's preparation of TMV, Rothamsted Culture, incubated with proteolytic enzymes.
- Fig. 2. Cylindrical Patterson function of tobacco mosaic virus. For. Fig. 2b (large values of ρ) a strong "artificial temperature factor" has been used.
- Fig. 3. Schematic representation of proposed arrangement of protein in TMV. N.B. This drawing is inserted for the sake of clarity; it inevitably involves great over-simplification, as well as the insertion of detail which is of no quantitative significance.
- a) View of a short length of virus particle, showing units and sub-units on 6 turns of the helix.
 - b) Axial section of a short length of virus particle. Shaded region represents RNA.
 - c) Transverse section of virus rod, showing 12 units in one turn of the helix. Arrows indicate principal direction of protein chains in each unit. RNA not shown.

REFERENCES

J.D. Bernal and I. Fankuchen, *J.Gen.Physiol.*, 25, 111 (1941)

J.D. Watson, *Biochem.Biophys. Acta*, 13, 10 (1954).

PUBLICATIONS FOR PERIOD 1st January 1951 - 1st January, 1954

1. Molecular Configuration in Sodium Thymonucleate, Nature, 171, 742, (1953) (With R.G. Gosling)
2. Evidence for 2-Chain Helix in Crystalline Structure of Sodium Desoxyribonucleate, Nature, 172, 156, (1953) (with R.G. Gosling)
3. The Structure of Sodium Thymonucleate Fibres: I. The Influence of Water Content. Acta. Cryst., 6, 678 (1953) (with R.G. Gosling)
4. The Structure of Sodium Thymonucleate Fibres: II. The Cylindrically Symmetrical Patterson Function. Acta Cryst., 6, 678 (1953). (with R.G. Gosling).