A new analysis of the scattering data for the protein particle associated with turnip yellow mosaic virus (Schmidt *et al.*, 1954) has been made using the methods outlined above. At the time the data were published, hollow-sphere functions corrected for slits of infinite height and negligible width had not been calculated, and so the data were interpreted by extrapolation from available calculations, giving a sphere radius of 140 Å, and an h of 0·75. Applying the same correction for the effects of the finite slit widths as was used by Schmidt *et al.* for turnip yellow mosaic nucleoprotein, use of Tables 1 and 2 and the methods described above essentially confirms the previous results.

The author wishes to express his gratitude to the University of Missouri for financial support, to Dr W. W. Beeman for suggesting problems leading to this paper, to Dr Bernard Goodman, Dr N. S. Gingrich and other members of the University of Missouri Physics Department for many helpful discussions and suggestions during the writing of the manuscript, and to Mr J. B. Combs for aid with the numerical calculations.

References

ANDEREGG, J. W. (1952). Thesis, University of Wisconsin. (The material in this thesis dealing with slit corrections is being prepared for publication.)

COMPTON, A. H. & ALLISON, S. K. (1935). *X-rays in Theory and Experiment*, p. 134. New York: Van Nostrand.

FOURNET, G. & GUINIER, A. (1950). *J. Phys. Radium* **11**, 516.

GUINIER, A. (1939). *Ann. Phys., Paris*, **12**, 161.

GUINIER, A. & FOURNET, G. (1947). *J. Phys. Radium* **8**, 345.

KRATKY, O., POROD, G. & KAHOVEC, L. (1951). *Z. Elektrochem.* **55**, 53.

LEONARD, B. R., ANDEREGG, J. W., SHULMAN, S., KAESBERG, P. & BEEMAN, W. W. (1953). *Biochim. Biophys. Acta*, **12**, 499.

LOWAN, A. N. & ABRAMOWITZ, M. (1943). *J. Math. Phys.* **22**, 1.

POROD, G. (1948–9). *Acta Phys. Austriaca*, **2**, 255.

RITLAND, H. N., KAESBERG, P. & BEEMAN, W. W. (1950). *J. Chem. Phys.* **18**, 1237.

SCHMIDT, P. W. (1955). *J. Math. Phys.* To be published.

SCHMIDT, P. W. KAESBERG, P. & BEEMAN, W. W. (1954). *Biochim. Biophys. Acta*, **14**, 1.

Acta Cryst. (1955). 8, 777

The Splitting of Layer Lines in X-ray Fibre Diagrams of Helical Structures: Application to Tobacco Mosaic Virus

BY ROSALIND E. FRANKLIN AND A. KLUG

Birkbeck College Crystallography Laboratory (University of London),
21 Torrington Square, London W.C. 1, England

(*Received* 13 *June* 1955)

The layer lines in X-ray fibre diagrams of tobacco mosaic virus gel have been observed to be split, the extent of the splitting varying with the strain of the virus. This effect is interpreted in terms of the helical arrangement of the protein sub-units about the long axis of the particle, and has stimulated some general remarks on diffraction by structures of this type.

Introduction

X-ray fibre diagrams of highly orientated preparations of tobacco mosaic virus (TMV) were first obtained by Bernal & Fankuchen (1941). Watson (1954) observed that the diagrams contained prominent features characteristic of helical structures, and suggested that the virus particle, of diameter 150 Å and length 3000 Å (Williams & Steere, 1951) was in fact one giant helical molecule, identical protein units being set in helical array around the long axis. He showed that in the axial repeat period of 69 Å there were $3n+1$ such protein units distributed over 3 turns of the helix. It is not possible to determine unequivocally, from high-angle meridional reflexions, whether n is 10 (Watson, 1954) or 12 (Franklin, 1955), but a recent interpretation (to be published) of certain other features of the X-ray diagram favours the value 10, giving 31 protein units in 3 turns of the helix.

We have now observed that in fibre diagrams of TMV the intensity maxima do not lie *exactly* on a set of equally spaced layer lines. If one chooses the set of equally spaced layer lines which gives the best fit with the diagram as a whole, then one finds that the layer lines whose order is a multiple of 3 ($l = 3n$) have maxima lying exactly on them, whereas for the layer lines $l = 3n+1$ and $l = 3n+2$ the maxima are displaced to a small distance on either side of the mean layer-line position. The extent of the effect varies with the strain of the virus. This phenomenon is readily explained in terms of the suggested helical arrangement of the protein units of which the virus particle

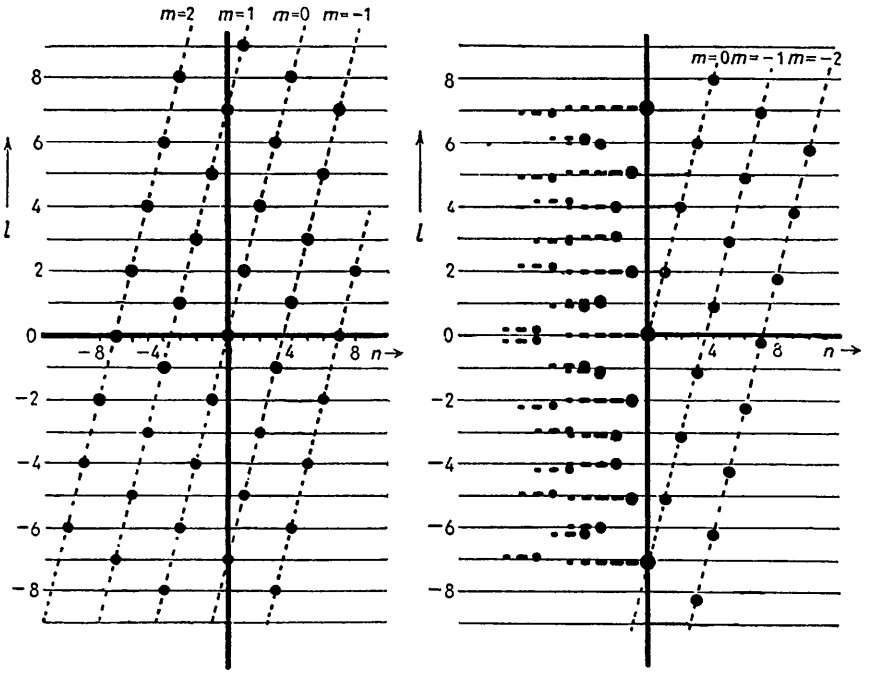

Fig. 1. (*a*) Plot of the orders n of the Bessel functions contributing to the layer lines l for the case of a periodic repeat containing $u = 7$ units in $t = 2$ turns of a helix. (These numbers are illustrative only and are not those appropriate to TMV for which t is 3 and u probably 31.)

The family of broken lines satisfy the equations $l = 2n + 7m$ and show how the diagram may be built up by repeating the line through the origin (transform of a smooth helix) at a series of new origins, $m = \ldots, -1, 0, 1, 2$, etc.

(*b*) The right-hand side of the diagram is an (n, l) plot as in (*a*), for the case when there are not exactly, but slightly more than, 7 units in 2 turns of a helix.

The left-hand side is meant to illustrate the general appearance of the actual fibre diagram that would result. It is obtained by putting a plane of symmetry in the (n, l) plot and indicating successive maxima of the Bessel-function contributions.

is built. The variation observed among the different strains of TMV examined gives some indication of the way in which the protein units are bound to one another.

Further, a theoretical consideration of the phenomenon has led to some new results and general observations on diffraction by complex helical structures.

Theoretical

1. *The splitting of the layer lines*

It will first be necessary to recapitulate briefly the theory of diffraction by helical structures (Cochran, Crick & Vand, 1952; referred to subsequently as C. C. V.). We shall adhere as far as possible to the notation in C. C. V.; (r, φ, z) and (R, ψ, ζ) are cylindrical co-ordinates in real and reciprocal space respectively.

In a helical molecule the atoms may be considered to be arranged on a number of coaxial helices of varying radii, each set of structurally equivalent atoms in the molecule being uniformly distributed along one such helix. If the molecule is built up of sub-units which are chemically and structurally identical and lie in helical array around the molecular axis, there will be one such constituent helix for each atom of the sub-unit. It is important to note that, in this case, all the constituent helices must be of the same pitch and bear the same number of atoms. The azimuth and z co-ordinate of the first atom on each helix will, of course, vary.

The Fourier transform of a set of point atoms lying along a helix of radius r and pitch P and having a separation of p in the z direction is obtained by putting down the transform of a smooth helix of the same radius and pitch with its origin at a series of points separated by a distance $1/p$ along the ζ axis in reciprocal space. Each of these origins thus has associated with it a set of levels $0, 1, \ldots, n, \ldots$ of spacing $1/P$ on each of which the amplitude is given by the corresponding Bessel function $J_n(2\pi Rr)$. If P/p is expressible as a rational fraction u/t, where u and t are integers, then the structure repeats exactly after a distance $c = up = tP$ in which there are u units (or atoms) and t turns of the helix. There will then be layer lines of spacing $1/c$ in reciprocal space and each layer line will contain a contribution of one Bessel function from each origin. However, the further away an origin is from the layer line considered, the higher will be the order of the Bessel function and so the less

important the contribution. This follows since high-order Bessel functions contribute only to the scattering at high angles and are in any case weak.

The orders n of the Bessel functions occurring on the lth layer-line are given by those values of n satisfying the relation (see C. C. V.)

$$l = tn+um\,, \tag{1}$$

where m is an integer and specifies the origin from which the Bessel function considered emanates. This is illustrated schematically in Fig. 1(a) for a hypothetical structure in which there is a repeat after two turns containing seven units. (These numbers have been chosen rather than those obtaining in TMV for the sake of greater clarity in the illustration.)

If P/p is not rational then there is no true axial repeat and theoretically there should be no layer lines, the whole of reciprocal space being filled. But, since one can always obtain a close approximation to P/p as a rational fraction if one goes to high enough integers, the transform will be effectively confined to layer lines, or very nearly so (see C. C. V.). The departure from an exact integral ratio will manifest itself as a *splitting* of the layer lines; that is to say, the Bessel-function contributions from different origins will no longer lie at exactly the same level, but will lie on either side of the mean layer lines that correspond to the rational approximation to P/p. This effect is illustrated in Fig. 1(b) for the case of 7·1 units in 2 turns, in contrast to Fig. 1(a) where P/p is strictly rational and equal to 7/2.

This is the effect which is observed in TMV. Its detailed interpretation is discussed later.

2. *Non-interference of Bessel functions of different order*

When layer-lines show a splitting of this kind it is clear that Bessel functions of different order cannot interfere in the theoretical expression for the diffracted intensity, since they fall at different levels in reciprocal space. We shall now show that this non-interference still holds even when the structure has a true axial repeat and any one layer line contains contributions from Bessel functions of more than one order.

The structure factor of the repeating unit is for the lth layer-line (see C. C. V.)

$$F(R, \psi, l/c) = \sum_j \sum_n J_n(2\pi R r_j) \times \exp\left[i\{n(\psi+\tfrac{1}{2}\pi)-n\varphi_j+l.2\pi z_j/c\}\right]. \tag{2}$$

For any one atom there is a summation over the orders n of Bessel functions determined by the selection rule (1), and there is a further summation over all the atoms in the unit, their co-ordinates being r_j, φ_j, z_j.

The three-dimensional intensity is given by the square of this expression and will contain products of Bessel functions of different order. In a fibre diagram, however, only the cylindrically averaged intensity is relevant; that is, the theoretical intensity is averaged over ψ. The cross-terms containing ψ in the exponential disappear under this operation, and we are left with the result

$$\langle F^2(R, l/c)\rangle_\psi = \sum_n \sum_j J_n^2(2\pi R r_j) + 2\sum_n \sum_{i>j} J_n(2\pi R r_i) J_n(2\pi R r_j) \times \cos\left[n(\varphi_i-\varphi_j)+2\pi(l/c)(z_j-z_i)\right] \tag{3}$$

subject to the selection rule (1) for n.

In this expression products of Bessel functions of different order do not occur; that is atoms on different radii interfere only through the same J_n's. Now, as stated above, the Bessel functions of different order on a layer line come from different origins in reciprocal space. Any deformation or perturbation of the helix shifts these origins and their associated Bessel functions to new positions, the original layer lines being split. Since there is always non-interference between Bessel functions of different order, such a perturbation does not introduce any discontinuity in the intensity pattern.

Digressing for the moment, we also note that there is an interesting connection between the expression (3) and the usual result for diffraction by a two-dimensional periodic structure. The products of Bessel functions, which produce the characteristic oscillatory variation in the intensity along a layer line, have weights given by a cosine term very like a two-dimensional structure factor. The order n of the Bessel function plays the role of an index for rotational periodicity just as does the ordinary index l for translational periodicity. This relates to a result, first pointed out by Crick (1953), for the case when all the atoms of the helix lie on one radius. If the positions of the atoms are marked by points on the cylindrical surface in which they lie and the latter is then unrolled flat, a pattern of points characteristic of the helix will be formed. This pattern will be reciprocal to the array of points formed by plotting the points (n, l) satisfying equation (1) in a Cartesian frame (Fig. 1(a) is an example of such an array).

Experimental

Measurements were made on the following three strains of TMV:

(a) *Rothamsted strain*, prepared by Mr N. W. Pirie, purified by incubation with commercial trypsin.

(b) *U*1, a 'normal' strain.

(c) *U*2, a 'mild' strain.

*U*1 and *U*2 were prepared by Dr A. Siegel, University of California, Los Angeles; their preparation and properties have been described in detail (Siegel & Wildman, 1954).

The splitting of the layer lines is most marked in the strain *U*2. The central region of the fibre diagram of this strain is shown in Fig. 2. It is clear that the separation between the inner regions of layer lines

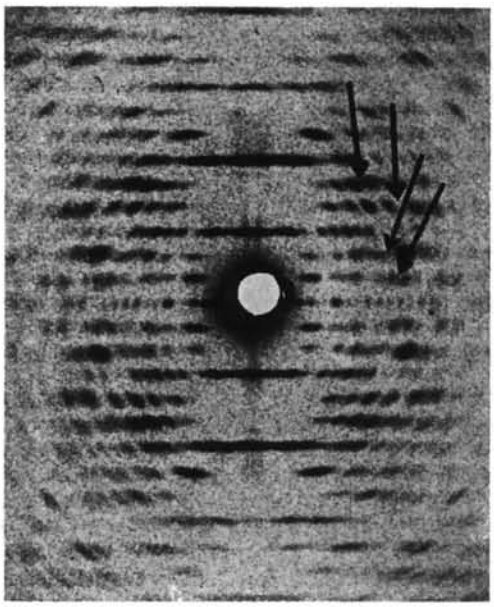

Fig. 2. X-ray fibre diagram of orientated gel of strain *U*2 of tobacco mosaic virus. Splitting of the layer lines is clearly visible. The pairs of layer lines 1 and 2, and 3 and 4, are closer than the average in their inner regions and more widely separated in their outer regions Arrows indicate the points at which the reversal occurs.

1 and 2 is less than that between layer lines 0 and 1 and that between layer lines 2 and 3. Similarly, the separation between the inner regions of layer lines 4 and 5 is small, while that between 3 and 4 and between 5 and 6 is large. A little further from the centre of the diagram this effect is reversed; the points at which this reversal takes place are indicated in Fig. 2 by arrows.

The most accurate measurements of the degree of displacement of the diffraction maxima from the mean layer-line positions are those made on the inner regions of the first and second layer lines of the strain *U*2. Using the fact that there are very nearly $3n+1$ units in 3 turns, we find that in *U*2 there are $31{\cdot}05\pm0{\cdot}01$ protein units in 3 turns of the helix, if $n = 10$. In the Rothamsted strain and also in the strain *U*1, the extent of the splitting effect is only about one-third of that in TMV and in the same sense. In these strains therefore, there are approximately 31·02 units in 3 turns of the helix. These strains thus differ from the strain *U*2 in having 0·03 fewer units in 3 turns, or 0·01 fewer units in one turn of the helix.

Discussion

The above results lead to the conclusion that the structural difference between one strain of TMV and another consists, in part, in a small shift of any one protein building unit with respect to its neighbours immediately above and below, the extent of the shift being 0·01 of the extent of the protein unit, measured along any arc about the axis of the particle. Schematically, the units may be considered to occupy a wedge-shaped volume (Franklin, 1955, see diagram) and the shift will, of course, be greatest at the widest part of the wedge—that is, at the outermost shell of the virus particle. If there are very nearly $10\frac{1}{3}$ units in one turn of the helix, each unit will occupy a length of 46 Å on the circumference of radius 75 Å. The variation in the relative positions of neighbouring units, on this radius, is therefore about 0·5 Å. On smaller radii the shift is correspondingly less.

Although the absolute value of the difference in relative positions of neighbouring units turns out to be small, the fact that such a difference can occur is probably significant. It suggests again that there may be little specific chemical bonding between protein units and their neighbours on the turns of the helix immediately above and below (see Franklin, 1955; Franklin & Commoner, 1955). The bonding is therefore presumably strongest between neighbours along the same turn of the helix, indicating that the helical arrangement of the protein sub-units is of structural, and not merely geometrical, significance.

Finally we may briefly consider the shape of the protein sub-unit. X-ray data cannot provide direct evidence as to the shape of the true chemical entity: it may be anything from the schematical wedge mentioned above to a thin disc. In the latter case successive discs would have to be rotated by an angle $(t/u)\times2\pi = (3/31)\times2\pi = 34{\cdot}8°$ with respect to one another, in order to account for the helical arrangement. The thickness of a disc would, however, be only $69/31 = 2{\cdot}23$ Å, so that it is clearly an impossible shape for the real chemical unit (although it may be useful formally to describe the unit and its repetition as a pile of rotated discs). The actual protein units are probably much more similar to the schematical wedge, which has a thickness of 23 Å.

We are greatly indebted to Mr N. W. Pirie and to Dr A. Siegel for supplying us with purified virus solutions. We wish to express our thanks to Prof. J. D. Bernal for his constant interest, and to acknowledge the financial support of the Agricultural Research Council (R. E. F.) and the Nuffield Foundation (A. K.).

References

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

Cochran, W., Crick, F. H. C. & Vand, V. (1952). *Acta Cryst.* **5**, 581.

Crick, F. H. C. (1953). Thesis, Cambridge.

Franklin, R. E. (1955). *Nature, Lond.* **175**, 379.

Franklin, R. E. & Commoner, B. (1955). *Nature, Lond.* **175**, 1077.

Siegel, A. & Wildman, S. G. (1954). *Phytopathology*, **44**, 277.

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

Williams, R. C. & Steere, R. L. (1951). *J. Amer. Chem. Soc.* **73**, 2057.