Harrington for advice in the procedures and interpretation of viscometry. [Jan. 12

[1] Putnam, F. W., in "The Proteins", edited by Neurath, H., and Bailey, K., **1**, B, 808 (New York, 1953).
[2] Anfinsen, C. B., Harrington, W. F., Hvidt, A., Linderstrøm-Lang, K., Ottesen, M., and Schellman, J. A., *Biochim. Biophys. Acta*, **17**, 141 (1955).
[3] Kunitz, M., *J. Gen. Physiol.*, **22**, 207 (1939).
[4] Lineweaver, H., and Schwimmer, S., *Enzymologia*, **10**, 81 (1941).
[5] Steinhardt, J., *J. Biol. Chem.*, **123**, 543 (1938).
[6] Deutsch, H. F., *Acta Chem. Scand.*, **5**, 1074 (1951).
[7] Viswanatha, T., and Liener, I. E., *J. Biol. Chem.*, **215**, 777 (1955).
[8] Anson, M. L., and Mirsky, A. E., *J. Gen. Physiol.*, **17**, 159 (1933).
[9] Winnick, T., Davis, A. R., and Greenberg, D. M., *J. Gen. Physiol.*, **23**, 275 (1940).
[10] Schwert, G. W., Neurath, H., Kaufman, S., and Snoke, J. E., *J. Biol. Chem.*, **172**, 221 (1948).
[11] Jacobsen, C. F., and Léonis, J., *C.R. Trav. Lab. Carlsberg, Sér. chim.*, **27**, 333 (1951).
[12] Anson, M. L., and Mirsky, A. E., *J. Gen. Physiol.*, **17**, 393 (1933).
[13] Kunitz, M., and Northrop, J. H., *J. Gen. Physiol.*, **17**, 591 (1933).
[14] Linderstrøm-Lang, K., Lane Medical Lectures, 115 (Stanford Univ. Press, 1952).
[15] Northrop, J. H., Kunitz, M., and Herriott, R. M., "Crystalline Enzymes" (2nd edit., New York, 1948).
[16] Anson, M. L., and Mirsky, A. E., *J. Gen. Physiol.*, **17**, 151 (1933).
[17] Dowmont Halsey, Y., and Neurath, H., *J. Biol. Chem.*, **217**, 247 (1955).

STRUCTURE OF SMALL VIRUSES

IT is a striking fact that almost all small viruses are either rods or spheres. The purpose of this article is to explain this observation by means of the following simple hypothesis: a small virus contains identical sub-units, packed together in a regular manner. It has been suggested before[1] that viruses are constructed from sub-units; but the idea has not previously been described in precise terms or put forward as a general feature of all small viruses.

We believe that there is conclusive evidence for this hypothesis in two cases and suggestive evidence in a number of others. As most of the present evidence comes from the plant viruses, we shall restrict our discussion to these, except for a few remarks on animal viruses at the end of the article.

Plant Viruses

Notice first that all plant viruses which have been studied carefully are extremely regular in their shape and size[2]. In electron micrographs their dimensions are constant. One particle of turnip yellow mosaic virus, for example, is the same size as another, to within the errors of measurement. Moreover, the 'spherical' viruses have shapes very close to that of a sphere—there seem to be no ellipsoidal plant viruses. All cases where they have appeared as flattened spheres have been shown to be due to the surface tension caused by drying prior to electron microscope examination. In good photographs there are sometimes suggestions that the 'spheres' are more nearly regular polyhedra, which, as we shall see, is what one might expect.

The great regularity of plant viruses is shown even more strikingly by their ability to form crystals (or paracrystals) which give good X-ray photographs[3], often with reflexions extending to small spacings. From this we can infer that a very high degree of order exists within such viruses, and that, to a resolution almost at the atomic level, one virus particle appears identical, or at least very similar, to all its sister virus particles. A plant virus can thus be considered a 'molecule' in the sense used by protein crystallographers—an entity, the major part of which has its atoms arranged in definite (relative) positions in space.

All known plant viruses consist of two chemical components only: protein and ribonucleic acid. It seems likely that there is a general plan for their relative positions and that the majority of the protein lies on the outside of the virus, surrounding a central core composed largely, if not entirely, of ribonucleic acid. This arrangement is well established for only two viruses—the spherically shaped turnip yellow mosaic virus (by Markham[4]) and the rod-shaped tobacco mosaic virus (by both the Tubingen[5] and Berkeley groups[6])—but we believe that it is likely to apply to all simple viruses. That is, the protein component of a round virus is a spherical shell, and of a rod-shaped virus, a cylindrical shell. Our hypothesis is that in both cases these shells are constructed from a large number of identical protein molecules, of small or moderate size, packed together in a regular manner. Our hypothesis may apply, though in a slightly different form, to the ribonucleic acid component. This is discussed in more detail later.

Tobacco Mosaic Virus

This rod-shaped virus is the best studied and we shall therefore consider its structure in detail.

Tobacco mosaic virus contains 94 per cent protein and 6 per cent ribonucleic acid[7]. The characteristic particle, which is closely connected with the infectivity, has a 'molecular weight' of about 45 million, a length close to 3000 A. and a diameter of about 170 A. The early X-ray work[3] showed clearly that this particle is made up of sub-units of some sort. More recently it was realized that the basic feature of the structure is its helical nature[8]. The protein part of the virus is constructed from a large number of structurally equivalent sub-units (small globular proteins) set in helical array about the central axis. The pitch of the helix is 23 A. The number of sub-units per turn is more difficult to establish—the most probable value (Franklin, R. E., and Holmes, K. C., personal communication) gives a molecular weight for the sub-unit of about 20,000.

A very similar value is suggested by the chemical evidence. Harris and Knight[9] first examined the carboxyl end-groups of the polypeptide chains, and found that the virus particle had about 2,500 terminal groups, all threonine. This suggested that the virus contains 2,500 identical polypeptide chains, an idea which has been further strengthened by the recent work of both the Tubingen[10] and Berkeley[11] groups, who have identified the terminal three residues at the carboxyl end of this polypeptide chain.

Some additional feature is obviously needed to determine the length of the protein shell, and we would guess that in the intact virus this is controlled by the length of the ribonucleic acid core. This would explain why rods of indefinite length are produced when undenatured protein sub-units are re-aggregated in the absence of ribonucleic acid[12]. Moreover, when the re-aggregation occurs in the presence of ribonucleic acid, it is reported by Fraenkel-Conrat and Williams[13] that rods of 3000 A. in length occur very frequently.

The structure of tobacco mosaic virus, then, is based on a helix, or, in other words, it has a screw axis—in this case[8] a non-integer screw axis. This symmetry axis implies that all the protein sub-units in the body of the virus have the same environment. The same contact points between neighbouring

sub-units are used over and over again as we move along the helix. This feature is the clue to the general principle which we can apply whenever, on the molecular level, a structure of a definite size and shape has to be built up from smaller units; namely, that the packing arrangements are likely to be repeated again and again—and hence that the sub-units are likely to be related by symmetry elements.

So far we have been mainly concerned with the protein, and have neglected the ribonucleic acid component of the virus. Is that, too, made up of sub-units? The ribonucleic acid content of tobacco mosaic virus is rather low, and not more than four nucleotides can be associated with a given protein sub-unit. Now if all these groups were identical, the analytical composition of the ribonucleic acid would be based on the number 4, which it certainly is not[14]. Moreover, the ribonucleic acid is probably connected with the genetic properties of the virus, and so its fundamental unit must contain a much larger number of nucleotides.

This does not mean, however, that ribonucleic acid sub-units do not exist, since it is possible that the ribonucleic acid core contains a number of identical strands systematically interacting with the protein shell. The important consideration is that the packing arrangement should be repeated over and over again; and this can be done if the symmetry of the ribonucleic acid is the same as the symmetry of the protein and if the symmetry applies only to the sugar-phosphate backbone and not to the sequence of bases. It remains to be seen whether this type of arrangement can be established experimentally.

Spherical Plant Viruses

We have seen that the rod-shaped helical form of tobacco mosaic virus represents a natural way of constructing a large container from identical much smaller building blocks. The question we must now ask is whether the protein shell of the spherical viruses is likewise constructed by a regular aggregation of one type of small protein molecule, and, if so, how this is done. Unfortunately, there has been, to our knowledge, no systematic chemical search for the presence of sub-units in spherical viruses and so we must rely almost completely on crystallographic evidence.

It has been shown in two cases—bushy stunt virus[15] and turnip yellow mosaic virus[16]—that spherical viruses crystallize in a unit cell which has the shape of a cube; but unfortunately the X-ray photographs did not establish whether the symmetry also was cubic. This is important because, as has been pointed out by Dr. Dorothy Hodgkin[1] and Dr. Barbara Low[1], if the lattice possesses true cubic symmetry so must the virus particle, since there is only one particle in the primitive unit cell.

It has now been clearly established by Caspar (see following communication) that the unit cell of bushy stunt virus has cubic symmetry, and that, in this particular case, the virus has an even higher symmetry than the unit cell. Though this evidence applies to only one virus, we expect that further investigation will show that many small spherical viruses have cubic symmetry, for the reasons given below.

Now a virus possessing cubic symmetry must necessarily be built up by the regular aggregation of smaller asymmetrical building bricks, and this can be done only in a very limited number of ways. Since viruses are made of protein and ribonucleic acid, both of which contain asymmetric carbon atoms of one particular hand only, those symmetry elements (mirror planes and centres of symmetry) which turn a right hand into a left hand are impossible. Thus we can only have rotation axes, and for cubic symmetry this limits us to only three different combinations of symmetry elements.

Table 1. The Three Possible Cubic Point Groups for a Spherical Virus

Crystallographic description	No. and type of rotation axes present	No. of asymmetric units	Platonic solid with these symmetry elements
23	3 2-fold 4 3-fold	12	Tetrahedron
432	6 2-fold 4 3-fold 3 4-fold	24	Cube Octahedron
532	15 2-fold 10 3-fold 6 5-fold	60	Dodecahedron Icosahedron

The number of sub-units will be the same as, or a multiple of, the number of asymmetric units

Each of these three classes must contain at least four three-fold axes and three two-fold axes, arranged as for a tetrahedron. The first class contains no additional type of axis, while the second and third have four- and five-fold axes, respectively. Such an arrangement of symmetry elements is known as a 'point group', in contrast to a space group which applies to a regular arrangement extending to infinity. In Table 1 are listed the three cubic point groups possible for virus particles and also the regular polyhedra which have these symmetry elements (among others). Notice that in all these point groups the minimum number of asymmetric units must be a multiple of 12.

Three further points must be made to prevent misunderstanding. First, it is possible to arrange sub-units in other ways to produce a spherical shell, but the symmetry will not be cubic, and as they are less likely we shall not discuss them further here. Second, the asymmetric unit, upon which the symmetry elements act to build up the spherical shell, may consist of several identical sub-units joined together in some unsymmetrical fashion. This occurs quite often in protein crystals and would not be unexpected. Nor need the sub-unit be a single protein molecule in the chemist's sense of a unit joined together by chemical bonds. Several different protein molecules may aggregate to form the asymmetric unit. Third, our predictions concern the symmetry elements present in a virus particle, not its exact shape. However, this is likely to be approximately spherical, and may, under high resolution, appear polyhedral or perhaps with bumps on, like a rather symmetrical mulberry. Both these forms have been seen in electron micrographs.

It is not easy to explain in a short space why there are so few ways of building a spherical shell, but the reader can soon convince himself that it is difficult by trying to draw identical shapes which completely cover the surface of a tennis ball. It is impossible, for example, to do this entirely with hexagons, even if their shape is irregular. The point is very well stated in D'Arcy Thompson's "On Growth and Form"[17], in which we find "the broad, general principle that we cannot group as we please any number and sort of polygons into a polyhedron, but that the number and kind of facets in the latter is strictly limited to a narrow range of possibilities". The reason is essentially a topological one.

From the present X-ray evidence we are unable to distinguish the respective contributions of the protein and the ribonucleic acid, so we cannot be sure whether the cubic symmetry is perfect and applies strictly to both of them. We cannot tell whether the protein sub-units contain identical sequences of amino-acids, or whether the ribonucleic acid sub-units (if they exist) have identical sequences of nucleotides. It should not be very difficult, by end-group analysis, to decide whether the protein components are all approximately equal. By analogy with tobacco mosaic virus we would guess that this will be found to be the case. With the ribonucleic acid component, however, the problem is more difficult than it was in the case of tobacco mosaic virus, as the number of nucleotides per sub-unit is certainly much larger. (This follows from the higher percentage of ribonucleic acid[7] and the much smaller number of protein sub-units.) Only with a more detailed understanding of the ribonucleic acid core is the problem likely to be settled.

Animal and Other Viruses

For animal viruses we are handicapped because there is no X-ray evidence available so far. However, it is now becoming clear[2] that many of the smaller animal viruses, such as poliomyelitis and the various encephalitic viruses, are morphologically very similar to the spherical plant viruses. Not only are they of similar size (approximately 300 A. diameter); but it has recently been shown[18] that poliomyelitis virus also contains ribonucleic acid and can form crystals which appear as regular as those produced by the plant viruses. We thus think it very probable that cubic symmetry also extends to these animal viruses, and that the soluble antigens[19] (of about 120 A. diameter) frequently observed in infected cells are related to the sub-units normally used in the assembly of the final infective virus.

We also see no reason why our hypothesis should not be valid for viruses containing deoxyribonucleic acid rather than ribonucleic acid. Although the structure of bacteriophages is usually more complex than the smaller viruses discussed here, the fact that their heads appear polyhedral suggests that ideas of this general type may apply to them, too. On the other hand, it is less likely that they will be relevant to the structure of the larger viruses like vaccinia.

Conclusion

We can now describe our hypothesis in a more general manner. We assume that the basic structural requirement for a small virus is the provision of a shell of protein to protect its highly specific packet of ribonucleic acid. This shell is necessarily rather large, and the virus, when in the cell, finds it easier to control the production of a large number of identical small protein molecules rather than that of one or two very large molecules to act as its shell. These small protein molecules then aggregate around the ribonucleic acid in a regular manner, which they can only do in a limited number of ways if they are to use the same packing arrangement repeatedly. Hence small viruses are either rods or spheres. The number of sub-units in a rod-shaped virus is probably unrestricted, but for a spherical virus the number is likely to be a multiple of 12. Every small virus will contain symmetry elements and in favourable cases these can be discovered experimentally

We believe that this hypothesis is likely to apply (in this form or a simple variant of it) to all small viruses which have a fixed size and shape.

F. H. C. Crick
J. D. Watson*
Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems, Cavendish Laboratory, Cambridge
Jan. 23.

* On leave from the Biology Department, Harvard University, and supported by a grant from the National Science Foundation, U.S.A.

[1] Among the more important references are Hodgkin, D. C., *Cold Spring Harbor Symp.*, **14**, 65 (1950). Low, B., in "The Proteins", **1**, 235 (Academic Press, New York, 1953). Schramm, G., *Z. Naturforsch.*, **2b**, 112, 249 (1947).

[2] Williams, R. C., *Cold Spring Harbor Symp.*, **18**, 185 (1953); "Advances in Virus Research", **2**, 184 (Academic Press, New York, 1954).

[3] Bernal, J. D., and Fankuchen, I., *J. Gen. Physiol.*, **25**, 111, 147 (1941).

[4] Markham, R., *Disc. Farad. Soc.*, **11**, 221 (1951). See also Bernal, J. D., and Carlisle, C. H., *Disc. Farad. Soc.*, **11**, 227 (1951), and Schmidt, P., Kaesberg, P., and Beeman, W. W., *Biochim. et Biophys. Acta*, **14**, 1 (1954).

[5] Schramm, G., Schumacher, G., and Zillig, W., *Nature*, **175**, 549 (1955).

[6] Hart, R., *Proc. U.S. Nat. Acad. Sci.*, **41**, 261 (1955).

[7] Knight, C. A., "Advances in Virus Research", **2**, 153 (Academic Press, New York, 1954).

[8] Watson, J. D., *Biochim. et Biophys. Acta*, **13** 10 (1954). Franklin, R. E., *Nature*, **175**, 379 (1955).

[9] Harris, J. I., and Knight, C. A., *Nature*, **170**, 613 (1952); *J. Biol. Chem.*, **214**, 215 (1955).

[10] Schramm, G., Braunitzer, G., and Schneider, J. W., *Nature*, **176**, 456 (1955).

[11] Niu, C. I., and Fraenkel-Conrat, H., *Biochim. et Biophys. Acta*, **16**, 597 (1955); *J. Amer. Chem. Soc.*, **77**, 5882 (1955).

[12] Schramm, G., and Zillig, W., *Z. Naturforsch.*, **10b**, 493 (1955).

[13] Fraenkel-Conrat, H., and Williams, R. C., *Proc. U.S. Nat. Acad. Sci.*, **41**, 690 (1955).

[14] Markham, R., and Smith, J. D., *Biochem. J.*, **46**, 513 (1950).

[15] Bernal, J. D., Fankuchen, I., and Riley, D. P., *Nature*, **142**, 1075 (1938). Carlisle, C. H., and Dornberger, K., *Acta Cryst.*, **1**, 194 (1948).

[16] Bernal, J. D., and Carlisle, C. H., *Nature*, **162**, 139 (1948).

[17] Thompson, D'Arcy, "On Growth and Form", 737 (2nd edit., Camb. Univ. Press, 1952).

[18] Schaffer, F. L., and Schwerdt, C. E., *Proc. U.S. Nat. Acad. Sci.*, **41**, 1020 (1955).

[19] Polson, A., *Nature*, **172**, 1154 (1953). Polson, A., and Selzer, G., *Biochim. et Biophys. Acta*, **15**, 251 (1954). Hampton, J. U. F., *Biochim. et Biophys. Acta*, **18**, 446 (1955).

Structure of Bushy Stunt Virus

Crystals of bushy stunt virus yield detailed X-ray diffraction patterns indicating a high degree of regularity in the virus structure. Bernal, Fankuchen and Riley[1] found that the unit cell is body-centred cubic with one virus particle per primitive lattice point. Previous X-ray photographs[1,2] did not show whether the symmetry as well as the shape of the lattice is cubic. This is of considerable interest, since cubic symmetry for the lattice would imply that the individual virus particles possess cubic symmetry[3]. This is discussed in the preceding article (by Crick, F. H. C., and Watson, J. D.).

Precession photographs have been obtained from single crystals of bushy stunt virus which were grown by Messrs. F. C. Bawden and N. W. Pirie and kindly supplied by Dr. C. H. Carlisle. Two of these photographs, showing the basal reciprocal lattice planes normal to the edge and body diagonal of the cubic unit cell, are illustrated in Figs. 1 and 2. These show two- and three-fold symmetry respectively about the cube edge and cube diagonal, and establish the space group as $I\ 23$ ($a = 386$ A.).

Early in this study, photographs were obtained from a disordered crystal which did not give lattice