Virus crystallography comes of age

from Stephen C. Harrison

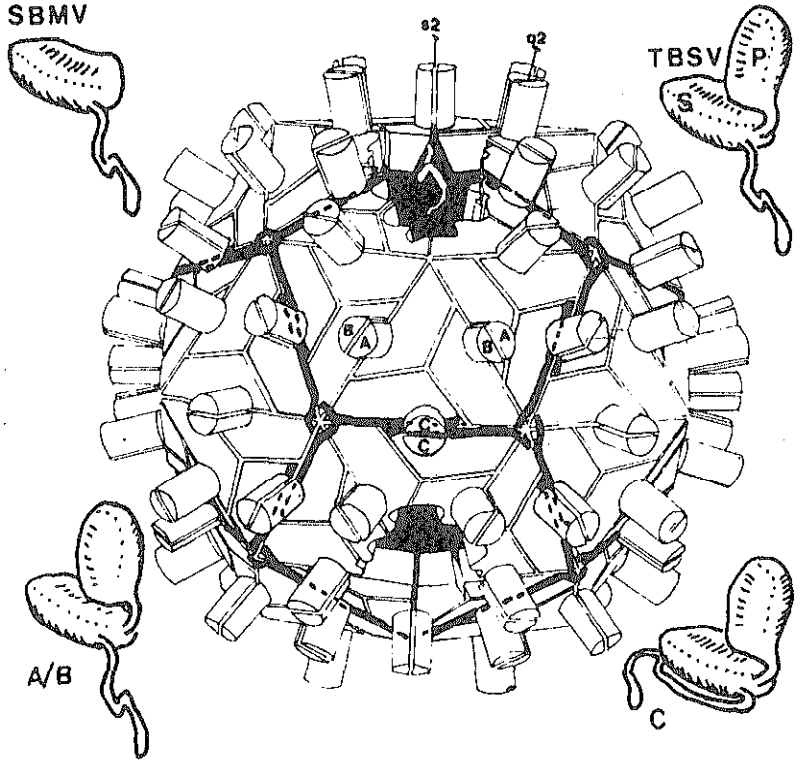
VIRUS particles are simple paradigms for all sorts of sub-cellular assemblies, and detailed views of their organization have long been sought. X-ray diffraction studies of virus crystals began at almost the same time as crystallographic investigation of small proteins[1], and a certain mystique of the 'ultimate' in large unit cells grew up around them. The reports in *Nature* by Abad-Zapatero *et al.* [2] of Southern Bean Mosaic Virus (SBMV) at 2.8 Å resolution and by Unge *et al.*[3] of Satellite Tobacco Necrosis Virus (STNV) at 4 Å resolution follow by less than two years the publication of 2.8 Å resolution structures of Tobacco Mosaic Virus (TMV) disk[4] and of Tomato Bushy Stunt Virus (TBSV)[5]: 'virus crystallography' is no longer a vague promise, and several striking results emerge from comparison of the structures at hand.

All these efforts have relied on some recent advances in technique — particularly in computing. Rossman and Blow[6] observed almost 20 years ago that non-crystallographic symmetry in structures such as viruses could enormously facilitate solution of the X-ray phase problem, but only in the last five years have algorithms, machines and programs been available to carry out the lengthy iterative calculations[7]. In principle, these iterations should converge from any initial approximation to the phases; in practice, however, experimental errors in observed structure amplitudes and approximations in the computational scheme require a non-random starting point. The work on TMV disk, TBSV and SBMV have all relied on the standard method of isomorphous replacement to provide this start. The STNV solution is rather more daring, using isomorphous phases to 10 Å and phase extension stepwise to 4 Å by non-crystallographic symmetry averaging. A more conventional computation was also perfomed, using initial SIR phases to 4 Å, and the reported agreement between this and the phase-extension calculation is excellent. It is difficult to assess the accuracy of a 4 Å map from density alone, since the shape of features depends on local conformation, and a more complete measure of the success of phase extension will be recognition of polypeptide chain and residues when the STNV structure is carried to 3 Å or beyond.

The coat subunit of a spherical virus assembles into a regular, icosahedrally symmetric shell and the shell is a package for nucleic acid[8]. Moreover, in most structures the same subunit is found in several conformationally distinct positions, genetic economy requiring a larger coat than can be constructed from just sixty (i.e. one icosahedral quota). These architectural conditions demand a way of incorporating nucleic acid that adapts an inherently nonrepeating structure (the loops and stems of folded RNA) to a regular surface lattice. In structures with more than 60 subunits, they further demand a subunit that can adopt alternative bonding arrangements and hence also a mechanism for avoiding ambiguity.

In TMV, which has been resolved at 4 Å by diffraction from orientated gels[9], three nucleotides are bound into a groove-like site on each subunit. The coat is an 'unwinding protein' and imposes its own helical order on the RNA backbone. One might imagine by analogy that each protein subunit of a spherical virus could bind to an exact length of nucleic acid. If the binding sites were located on a part of the protein fixed in the surface lattice, the RNA would be piecewise icosahedrally ordered, and portions of it would be visible in a suitable electron density map. This model can be ruled out in TBSV and apparently now in SBMV, where no clear RNA-related features appear. In both cases, however, a significant portion of the protein subunit (about 60 residues) is also invisible in the map, implying absence of a unique spatial relationship to the rest of the chain. This portion lies at the N-terminus, and it is a plausible candidate for a nucleic-acid binding sequence. If so, we can imagine two RNA packing modes that might interact with a flexibly-linked binding domain: (1) Viral RNA *in situ* could have a more or less defined tertiary structure, requiring binding regions on the protein loosely tethered to the rest of the subunit, in order to seek out appropriate sites on a non-icosahedrally-symmetric nucleic-acid structure; (2) RNA stems and loops could be compactly but essentially randomly close-packed in the particle (some 60 to 70% of the RNA in such viruses is probably in double-helical stems[10]), again requiring flexibly-linked protein regions for interaction with appropriate RNA structures or sequences. Crystallographic structure determination necessarily yields an average over the operations of the space group, and in SBMV and TBSV, a unique RNA structure would be blurred by the high symmetry of the lattice postion on which the particle lies. In STNV, the asymmetric unit comprises an entire particle, and Unge *et al.* point out that

Subunits and their arrangement in SBMV and TBSV. The SBMV subunit (upper left) has an N-terminal arm and a single, folded domain, homologous to the S-domain of TBSV (upper right); 180 subunits pack to form a T = 3 viral shell. The arms of 60 of the subunits (positions denoted C) are ordered as shown by heavy lines in the central diagram. The actual folding of an ordered arm is along an inside edge of the S-domain with a loop around the particle threefold axis (see diagram in lower right-hand corner). Three such loops embrace, and the 60 ordered arms form a connected framework. The arms of subunits in positions denoted A and B are not regularly ordered in SBMV or TBSV (lower left). The cylindrical projections in the central diagram represent P-domain dimers in TBSV; they are, of course, absent in SBMV.



Stephen C. Harrison is Professor of Biochemistry at Harvard University.

© 1980 Macmillan Journals Ltd

crystallographic symmetry alone does not exclude determination of RNA configuration. Such a determination would require not only a unique RNA conformation, but also sufficient asymmetry imposed thereby on the shell that consistent orientations appeared throughout the crystal. The present map is ambiguous on that point, but a high degree of order does not appear likely.

The N-terminal 15 residues cannot be traced, and the general appearance of internal density is not immediately recognizable as a single structure. It is useful in this connection to recall that the uniqueness of the secondary structure has not been established for viral RNA, although the correlation of nuclease sites and predicted single-stranded regions in 'flower diagrams' of the MS2 sequence suggest a general regularity[11]. In any one stock of virus, a given RNA molecule is likely to differ from the mean sequence at a number of positions[12]. Moreover, the compactness of the configuration in the virus appears to be imposed by the shell, rather than by its own sequence, since RNA of the size of the genomes in these viruses has a much larger radius of gyration under conditions of physiological ionic strength than that of the cavity into which it is packed[13]. Indeed, the degree of compaction is striking: from figures given by Unge *et al.*, the volume available to STNV RNA implies an inner cavity with space for about 1.2 g H_2O per g RNA; the packing of TBSV RNA is comparably tight. A model of more or less random close packing of roughly parallel RNA stems is appealing, since very efficient packaging can be achieved without any significant demands on RNA structure or sequence. But it will not be simple to prove such a picture.

The most surprising of recent results is probably the similarity of the SBMV subunit to the S-domain of TBSV. Various evolutionary speculations can be made, but from the point of view of assembly and its control, there are two more immediate conclusions that should be emphasized. The first is the identical way in which these two T = 3 structures solve the ambiguity problem alluded to above. A portion of the polypeptide chain, just N-terminal to the compact domain(s), is ordered on sixty subunits and disordered on the remaining 120. The sixty ordered arms effectively serve as an internal, T = 1 framework: they interdigitate around icosahedral twofolds. This scaffold, formed by a flexibly linked part of the subunit chain determines the size and curvature of the particle. The choice between alternative bonding patterns at interfaces that must so adapt is dictated by the symmetry and geometry of the network of arms. One might imagine cases where the network would be formed of a distinct protein — could this be the case with cores in bacteriophage head assembly[14] or VP2 and VP3 in polyoma and SV40[15]? The second direct conclusion that emerges from comparison of TBSV and SBMV is the significance of modularity in the design of assembling subunits. In SBMV as seen in the paper of Abad-Zapatero *et al.*, the subunit has three functional parts: an unseen (RNA-binding?) N-terminal piece, the framework-like arm (including sequences in the intertwined β-annulus), and the major folded domain. In TBSV, there is an additional folded domain (the P-domain). Clearly the P-domain is not essential for particle stability, since it is missing in SBMV. It does appear to be a strong twofold 'clamp', and the hinge between it and the S-domain is required for subunits to conserve these two fold interactions while still adapting to the T = 3 surface lattice.

With these advances now at hand, what about the crystallography of other viruses and large assemblies? SBMV and TBSV show quite dramatically that the spatial arrangement of modules in a subunit protein is characteristic of the assembly, but not necessarily of the isolated polypeptide; the remarkable interdigitation of N-terminal arms could not have been anticipated (or even seen) in the structure of the subunit alone. The possible consequences for those who would like to visualize a ribosome, for example, are sobering: high-resolution structures of large assemblies may not readily be constructed from structures of component parts. There is, however, a converse advantage of flexibly-linked modules: in cases where function is a property of a particular module of a large protein, proteolytic dissection and piecewise crystallization is a fruitful and helpful strategy. This is the case, for example, with influenza virus haemagglutinin[16] and may well be appropriate for the presumptive RNA-binding, N-terminal portions of SBMV or TBSV.

1. Bernal, J. & Fankuchen, I. *Nature* 139, 923 (1939).
2. Abad-Zapatero, C., Abdel-Meguid, S., Johnson, J, Leslie, A., Rayment, I., Rossmann, M., Suck, D. & Tsukihara, T. *Nature* 286, 33 (1980).
3. Unge, T., Liljas, L., Strandberg, B., Vaara, T., Kannan, K., Fridborg, K., Nordman, C. & Lenj, P. *Nature* 285, 373 (1980).
4. Bloomer, A. C., Champness, J. N., Bricogne, G., Staden, R. & Klug, A. *Nature* 276, 362 (1978).
5. Harrison, S. C., Olsen, A. J., Schutt, C. E., Winkler, F. K. & Bricogne, G. *Nature* 276, 368 (1978).
6. Rossman, M. G. & Blow, D. *Acta Cryst.* 16, 39 (1978).
7. *Bricogne, G. Acta Cryst.* A30, 395 (1974). *Acta Cryst.* A32, 832 (1976)
8. Caspar, D. L. D. & Klug, A. *Cold Spring Harb. Symp. quant. Biol.* 27, 1 (1962).
9. Stubbs, G., Warren, S. & Holmes, K. C. *Nature* 267, 216 (1977).
10. Witz, J. & Strazielle, C. in *Subunits in Biological Systems*, B (ed. Fasman, G. & Timashaff, S., Dekkar, New York.)
11. Fiers, W. *et al. Nature* 260, 500 (1976).
12. Domingo, E., Sabo, D., Taniguchi, T. & Weissmann, C. *Cell* 13, 735 (1978).
13. Gesteland, R. & Boedtker, H. *J. mol. Biol.* 8, 496 (1975).
14. Casjens, S. & King, J. *Rev. Biochem.* 44, 555 (1975).
15. Murakami, W. T. & Schaffhausen, B. in *Molecular Studies in Viral Neoplasia 25th M. P. Anderson Symposium*, (Williams & Williams, Baltimore, 1974).
16. Wilson I. A., Skehel, J. J. & Wiley, D. C. in *Structure and Variation in Influenza Virus* (ed. Laver G & Air, G.) (Elsevier/North-Holland, New York, 1980)

0028-0836/80/320559-02$01.00

Open-ended fusion systems

from R. F. Post

JAPAN, among all of the highly industrialized nations, is the most dependent on imported energy. Japan is also a nation that owes its advanced technology both its present economic strength and its hopes for future economic growth. Given these circumstances it is not surprising that nuclear fusion research has been singled out for special attention by the Japanese government. Fusion, for which the primary fuel is heavy hydrogen, extractable at low cost from water, would be an extremely attractive long-term energy source. It is against this backdrop that a specialized scientific meeting dealing with a particular approach to fusion, so-called open-ended magnetic confinement systems was convened in Japan in April.*

Open-ended fusion systems are a class of fusion 'magnetic bottles' in which the field lines of the externally-generated magnetic field (needed to confine the hot fusion plasma) enter at one end of the bottle and exit from the other end. Open-ended magnetic bottles are in contrast to ones such as the tokamak, where the field lines remain entirely within the doughnut-shaped confinement chamber.

R. F. Post is Deputy Associate Director for Physics in the Magnetic Fusion Division of the Lawrence Livermore National Laboratory, University of California.

An example of an open-ended system is the simple mirror machine, in which the field lines form a tubular bundle constricted at both ends (the mirrors). The charged particles of a fusion plasma, once trapped between the mirrors, continue to gyrate back and forth along the field lines, being repeatedly reflected by magnetic mirrors in the same way that charged particles are trapped, between the north and south magnetic poles, in the Van Allen radiation belts that surround the earth.

Present interest in the mirror approach to fusion power stems in part from an inherent advantage that some kinds of mirror fields possess in their ability to suppress the so-called MHD or fluid-like turbulences in magnetically confined plasmas. From this property derives the remarkable ability of specially-shaped mirror fields (called 'magnetic wells') to stably confine very high plasma pressures for a given magnetic field intensity. To this inherent advantage has recently been added a new feature — the tandem mirror idea — an idea that shows promise of greatly improving the economic prospects of mirror fusion systems relative to older mirror approaches. The tandem mirror

The 'International Symposium on Physics in Open Ended Fusion Systems' was held at the new University of Tsukuba, Japan in April 1980.