Viral shell game

Roger W. Hendrix
Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania USA 15260

The basic structure of the capsid shells of "spherical" viruses was first described in detail by Caspar and Klug (1), who proposed a geometrical model in which the triangular faces of an icosahedron are imagined to be "triangulated" into smaller triangles, each of which is then fitted with three identical protein subunits to form the lattice of the capsid. Only certain integral multiples of 60 subunits (the "triangulation numbers", T) are allowed. In the resulting structures, the subunits are arranged "quasi-equivalently": that is, each subunit makes approximately the same bonding contacts with its neighbors as do all the other subunits, but there are small distortions in the subunit interfaces or the subunits' shapes to allow them to accommodate to the T geometrically slightly different environments that they find themselves in. An alternative way to visualize the predictions of the Caspar-Klug theory is to think of the shell as having a cyclic pentamer of the subunit at each of the 12 vertices of an icosahedron, with the space between the pentamers filled in by cyclic hexamers, the number and positioning of which is characteristic of the triangulation number.

Over the three decades since this theory was proposed, most of the variations on its basic themes that might have been imagined (as well as one that was probably never imagined) have been documented: some viruses use a separately encoded protein to occupy the fivefold corners; in some viruses a single polypeptide occupies the position expected to be occupied by two quasi-equivalent subunits; and in the beautifully grotesque structure of the papova viruses all the capsomeres (including the 6-coordinated ones) are pentamers equipped with flexible bonding arms to compensate for the broken symmetry (2). Moreover, in most of the available atomic resolution structures of mature viruses, it would be a stretch to call the inter-subunit bonding quasi-equivalent (though since the capsid subunits of many viruses undergo substantial conformational changes following shell assembly, it is not clear that nonequivalences in the mature virion reflect nonequivalences in the initially assembled shell). Despite its weathering over the years, the Caspar-Klug theory remains a useful way to think about virus structure, and there are many examples, including the P22 prohead shells discussed below, which follow the simple Caspar-Klug model in an apparently uncomplicated way.

Inherent in the idea that identical subunits can occupy similar but non-identical positions in the lattice is that those same subunits, with their exact geometries differently assorted, should be able to form shells of a wide variety of sizes and shapes. That protein subunits of viral capsids do indeed have such potential is known from numerous examples. Thus, the capsid subunit of bacteriophage P2, which normally makes $T = 7$ shells (12 pentamers and 60 hexamers), can be efficiently subverted into making $T = 4$ shells (12 pentamers and 30 hexamers) under the genetic direction of a co-infecting bacteriophage P4. Bacteriophage T4, which normally has a prolate capsid based on an elongated $T = 13$ icosahedral structure, can find itself with an isometric "petite" capsid or a super-elongated "giant" in the presence of any of a variety of mutations affecting either the subunit itself or one of several assembly accessory proteins. Despite the potential for incorrect assembly, the frequency of error in assembly during the normal course of an infection by a wild-type virus is remarkably low; evidently the protein subunits are either directed toward the correct structure or away from the many possible incorrect ones.

None of the elegant structural information available for mature virions gives more than an occasional hint about the mechanisms by which protein subunits are assembled into these shells inside the cell. What are the intermediates in assembly: do subunits first assemble into oligomeric units or do they add individually to the growing shell? How does an individual subunit "know" whether it should assemble as part of a five- or sixfold unit, or stated differently, how is it determined where the pentamers are inserted into the lattice of hexamers to determine the size and shape of the shell? Is all the information to specify accurate assembly contained in the subunits themselves, or are other factors (e.g., the viral nucleic acid or other proteins) needed?

The paper by Prevelige, Thomas, and King in this issue takes a small but nicely shaped bite out of this daunting set of questions, using bacteriophage P22 of *Salmonella typhimurium* as the experimental model. P22, as a typical example of the dsDNA phages, has the interesting property that assembly of the shell requires an abundant second protein, the "scaffolding" protein, in addition to the "coat" subunit that will make up the mature shell. Coat and scaffolding subunits co-assemble into a structure called a prohead with a $T = 7$ icosahedral shell of 420 coat subunits surrounding an internal core (the scaffold) of ~300 scaffolding proteins. (The scaffolding subunits subsequently leave the structure as DNA is pumped in, probably by slithering out through holes in the prohead shell.) These authors have shown that P22

0006-3495/93/03/836/02 $2.00

proheads can be assembled in vitro using only purified coat and scaffolding protein (3, 4). Either protein alone is stable in solution as a soluble monomer, but when the two are mixed, they assemble rather efficiently into normal-appearing scaffold-containing proheads. The system is now well enough under control that it is possible to make quantitative kinetic measurements of the assembly process and to search for assembly intermediates, and these studies are described in the current paper.

Measuring shell assembly by turbidity and by electron microscopy, the authors show that the reaction appears to be nucleation limited, with fast completion of assembly following successful nucleation. This means that assembly intermediates are scarce, but it is still possible to identify some probable intermediates. These all fit a picture in which assembly initiates at one point in the shell and the shell grows progressively from that point by addition of subunits to the edge. Significantly, there is no indication of any protein subassemblies that form before adding to the growing shell; most likely the coat subunits add individually to the edge of the shell. The variation of initial rate of assembly with protein concentration shows that nucleation is fifth order in coat subunit and second or third order in scaffolding protein. The simplest interpretation of these numbers is that the nucleation complex is a pentamer of the coat protein (presumably one that will form one of the corners of the assembled shell) stabilized by a few copies of the scaffolding protein.

The fascinating question of how the scaffolding protein helps the coat protein to assemble correctly is still largely unanswered, though the clearer view of the overall process afforded by these studies makes it easier to formulate sensible hypotheses. The coat protein needs to undergo two sorts of conformational adjustments or switches as it goes from the soluble monomeric state to being assembled in a shell: it must change from the unassociable form of the solution to the associable form of the shell, and each subunit must adjust to one of the seven quasi-equivalent positions available in the shell lattice. Previlege et al. suggest that both these changes in the protein may take place at once, at the edge of the shell, under the influence of the scaffolding protein. This view appears to make the best use of the currently available data, though exactly what the scaffolding protein might be doing in this process is still unclear, particularly since the non-integral ratio of coat to scaffolding proteins in the assembled prohead argues against a simple one-to-one interaction between the two proteins. More experiments remain to be done.

Is the P22 assembly mechanism that is gradually emerging a universal mechanism of capsid assembly? Evidently not, since work on various other viruses argue for mechanisms that are clearly incompatible with the P22 data, such as assembly from pentamers or nucleation on the viral nucleic acid (5–7). Certainly the P22 mechanism will be invaluable for understanding how other viruses that use scaffolding proteins achieve accurate assembly. Somewhat surprisingly from the point of view of the P22 work, another lambdoid phage, HK97, appears to assemble its prohead without the assistance of a separate scaffolding protein. Instead, the HK97 subunit has a stretch of 100 amino acids at its amino terminus that is needed for correct assembly and then is cleaved off, and it may provide an analogous assembly-aiding function to that provided by the scaffolding protein in P22. HK97 also differs from P22 in that pentamers and hexamers of the subunit are intermediates in assembly (Xie, Duda, and Hendrix, unpublished results). Apparently these two phages, which are very closely related by other criteria, use quite different mechanisms to assemble very similar structures. Each of the different assembly mechanisms used by different viruses will inform our understanding of the sorts of dances proteins are capable of in achieving these complex and highly ordered structures.

Received for publication and in final form 17 December 1992.

REFERENCES

1. Caspar, D. L. D., and A. Klug. 1962. *Cold Spring Harbor Symp. Quant. Biol.* 27:1–24.
2. Liddington, R. C., Y. Yan, J. Moulai, R. Sahli, T. L. Benjamin, and S. C. Harrison. 1991. *Nature (Lond.).* 354:278–284.
3. Previlege, P., D. Thomas, and J. King. 1988. *J. Mol. Biol.* 202:743–757.
4. Fuller, M. T., and J. King. 1982. *J. Mol. Biol.* 156:633–655.
5. Sorger, P. K., P. G. Stockley, and S. C. Harrison. 1986. *J. Mol. Biol.* 191:639–658.
6. Salunke, D. M., D. L. D. Caspar, and R. L. Garcea. 1986. *Cell.* 46:895–904.
7. Rueckert, R. R. 1991. Fundamentals of Virology. B. N. Fields and D. M. Knipe, editors. Raven Press, New York. 409–450.