Slipped Disks and Helices

WHILE crystallographers still wrestle with the details of tobacco mosaic structure, Klug and his colleagues have forsaken this ungrateful pursuit and have addressed themselves to the question of the mechanism involved in the assembly of proteins and RNA into the functioning virus. The results of their efforts are described in three articles in this issue of *Nature New Biology* (p. 37), whose length is a reflexion of the complexity of the problem. Many macromolecular assembly processes have of course been defined in greater or lesser detail—flagella, actin filaments, collagen and so forth—and the general rule is that there are no general rules, except perhaps for the proposition that the initiation of an aggregated structure from its monomers is entropically grossly unfavourable, and is commonly rate determining. The propagation reaction that follows generally proceeds much faster, so that in sum the formation of ordered aggregates is an autocatalytic process. In tobacco mosaic virus the situation is more complicated, not only because two different molecular species are involved, but also because it has long been recognized that several oligomeric states of the protein exist, and it is not necessarily the single chain that is to be regarded as the effective monomer in the polymerization reaction.

The oligomeric forms of the protein are defined in the first article. The isolated protein at high *p*H and low salt concentration forms a mixture of small aggregates, among which the trimer seems to be an important component; in this species the binding sites for radial and axial association are both presumed to be engaged. When the ionic strength is raised, or the *p*H marginally lowered, a new form makes its appearance: this is the double disk consisting of two rings, seventeen protein molecules in each. At still higher salt concentration and neutral *p*H or above, these disks stack on themselves, but below neutrality the familiar helical aggregates are slowly formed.

The authors evidently surmised at an early stage that the disk is an intermediate occupying a key position in the scheme. Analysis of electron microscope images shows that the reason for the self-sufficiency of the two-ring disk is not that the rings face each other so as to satisfy two identical sets of association sites, but rather that they form a polar structure in which, however, the outer edges of the subunits are axially distorted. The double disk therefore presents a marginally different aspect to the underside of a further disk adding to it. When disks are taken to lower *p*H, aggregation progressively occurs with the appearance of short rods made up of imperfectly meshed sections of two helical turns. After many hours these anneal to give the regular virus-like helices.

The manner in which the transition between the different aggregated forms is brought about is the subject of the second article by Klug *et al.* In the first place, ring closure to the disk forms, though energetically favourable, is, as might be expected, a slow process. Sedimentation equilibrium was used in an effort to show that the protein in the appropriate conditions of *p*H and salt is in equilibrium between states of low molecular weight in which the trimer may predominate, and the fully formed double disk, with no important concentration of intermediates. To draw such conclusions with confidence from sedimentation equilibria may activate the nervous tics of the hydrodynamites, for it is a brave man nowadays who would claim to distinguish an indefinite association from an equilibrium between monomer and one oligomer species, for example, in any but the most favourable cases. The mechanism does not rest solely on the evidence of the ultracentrifuge, however, and in any case the authors disarm the sceptic with a lofty phrase about having exploited available methods "to the full".

One is left then with a scheme in which the preformed disks function as the effective monomer in a condensation polymerization mechanism. A feature of the reaction, however, must be the dislocation of the disks to the open, helical or, as the authors term it, "lock-washer" form. The *p*H profile separating the zones of stability of the stacked disk and helical forms is consistent with the involvement of groups titrating at about *p*H 7, which it is natural to identify with the anomalously titrating groups in the virus, most probably (in the absence of histidines and an α-amino group) carboxyls, discovered long ago by Caspar. In the stacked disk form the ionizing groups in question are normalized, which indicates that the coulombic disturbance caused by the anomalous protonation is linked to the helical dislocation. The same transition can also evidently be induced by the RNA. This mechanism ensures that the protein remains in its disk form in readiness for rapid polymerization when RNA is introduced.

Only when the protein is in the disk form does the RNA indeed trigger rapid polymerization to virus particles. As the third article (p. 47) shows, each RNA molecule causes the condensation of one virus particle. The idea that the protein disk is the nucleus for formation of the helical structure during reconstitution with RNA has proved more profitable to exploit than Klug *et al.* first imagined. Their original idea was that the disk is needed only for initiation, but further experiments suggested that it is also the form used for growth of the nucleoprotein helix. Initiation of nucleoprotein helix growth is fast compared with the addition of further disks to the growing helix, so that the latter may well be the rate determining step.

Partial digestion of the RNA with exonucleases shows that condensation begins at the 5′ end of the RNA. Klug *et al.* suggest that perhaps fifty bases can attach to the first disk, the conformational equilibrium being thereby displaced in favour of the dislocated form. Recognition of this terminal sequence, which must have an exceptionally strong affinity for the protein subunits, provides a way to select the correct polynucleotide. There are now a number of interesting directions in which the system can be further pursued, one of which is to determine the 5′ terminal sequence of the RNA that the protein recognizes, and this project, the authors assure us, is already in hand.