Proc. Nat. Acad. Sci. USA
Vol. 69, No. 10, pp. 2950–2953, October 1972

Assembly of Tobacco Mosaic Virus *In Vitro*: Effect of State of Polymerization of the Protein Component

(RNA–protein interactions/A-protein/disk state/subunits)

P. J. G. BUTLER AND A. KLUG

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England

Communicated by M. F. Perutz, August 7, 1972

ABSTRACT **We previously proposed that the assembly of tobacco mosaic virus is initiated by disks of 34 protein subunits attaching to the RNA, after which the particle grows by the addition of further disks. Other workers have reported growth from "A-protein" instead. We now report experiments that confirm our previous results and show that the contrary findings are largely due to a nonequilibrium form of A-protein that has a "memory" for the disk state and rapidly reaggregates, either in solution or on the growing particle.**

Some time ago we reported that both the initiation and the subsequent growth of tobacco mosaic virus (TMV) occur from a protein aggregate consisting of two rings, each of 17 protein subunits and known as the "disk" (1). Our conclusion that the growth, as well as the initiation, of the nucleoprotein helix occurs from the disk form of the protein, rather than from the mixture of small aggregates known as the "A-protein," has recently been questioned by Richards and Williams (2). We had initially been disturbed (1, 3) by the "implicit topological complexity" (2) of a process in which a two-ring disk is incorporated into a one-start helical structure, but the force of the data drove us to accept the fact that the disks, and not the A-protein, are the source of the protein for the rapid growth of the nucleoprotein helix. Moreover, with hindsight we have come to recognize the inherent advantages of use of a preformed aggregate rather than single subunits (1, 4, 5).

In this paper, we re-examine this question, and show that the apparently contradictory results of Richards and Williams, which indicate growth from A-protein, largely stem from the use in their experiments of a nonequilibrium aggregate of the protein, which we designate A*-protein. The A*-protein is freshly generated from disks and has the property of reaggregating more rapidly than normal A-protein, suggesting that it has a memory for the disk state. The experiments of Richards and Williams (2), therefore, do not show that the A-protein is directly incorporated into the TMV helix, but rather can be interpreted as revealing a further remarkable property of TMV protein.

We were compelled to accept that the growth occurred from disks, rather than from A-protein, mainly by the observation (ref. 1, Fig. 1, lower two curves) that the addition of A-protein at a concentration of 1.0 mg/ml to a reconstitution mixture containing 1.3 mg/ml of a disk preparation did not affect the rate of the reconstitution, whereas a change in the concentration of disks in this same concentration range had a marked effect upon the rate (4, 5). Although the disk preparations used in all experiments contain about 20% A-protein as well as the disks, it is implausible to suggest that growth occurs from this A-protein, since this suggestion would imply that a large change in a reactant concentration has no effect upon the rate of the reaction. On the other hand, when the protein concentration varied is that of the disks, a marked effect is found. A quantitative example is shown in Fig. 1, where the curve is the rate of reconstitution based on other experiments with a disk preparation as the sole protein source (4, 5), while the points are the actual rates measured for various disk concentrations with the addition of 1.0 mg/ml of A-protein freshly generated by the disaggregation of protein helix (6, 7). This result clearly shows that this A-protein has no effect upon the overall rate.

A weaker piece of evidence, but one that nevertheless supports the hypothesis that growth occurs from disks and not from A-protein, is the observation that reaction of a disk preparation with an excess of TMV RNA largely abolishes the 20S disk peak, while having no obvious effect upon the 4S A-protein peak (see ref. 5, Plate 1, for data). If the main source of the protein for the growth were indeed the A-protein, it might be expected that it would be used up during the reaction and the equilibrium between disks and A-protein disturbed in the direction of a reduction in the fraction as A-protein, rather than vice versa.

Source of protein for reconstitution

The conflict about the state of the protein necessary for reconstitution arises from the interpretation given by Richards and Williams (2) to their experiments on growing rods that have been initiated with disks and then temporarily cooled to disaggregate the remaining disks into "A-protein," so that the "A-protein" is the only species of protein available for subsequent growth. In their words, "The results show that the amount of protein converted to rods, by addition of protein to previously initiated rods, is as great when A-protein alone is the protein supply as it is when disk protein is the supply throughout," and they go on to interpret this as meaning that A-protein is the sole source of protein for growth. The similarity they found in growth from their disk protein or their "A-protein" seemed surprising to us if they were distinct species. The experiments suggested to us that the same protein species must be adding to the growing nucleoprotein rods in each case. Moreover, this species could not be A-protein since the increase in its concentration, by a factor of about five, consequent to depolymerizing the disks was having no effect upon the rate.

Our explanation of these experiments is as follows. It is possible that TMV protein has a "memory" for its previous state of aggregation when its temperature is changed, and it has also been shown that the aggregation of A-protein into

Abbreviation: TMV, tobacco mosaic virus.

disks is considerably faster in pyrophosphate than in phosphate buffer (8). We have confirmed the combination of these effects, namely that 20S material is regenerated in minutes when a solution of disks in pyrophosphate buffer is cooled for a short period and then rewarmed. In phosphate buffer, the rate of regeneration is slower, but again much faster than the *de novo* formation of disks from A-protein generated either by disaggregation of the protein helix or by equilibration over 48 hr at 4°.

Therefore, we thought that the rapid growth found by Richards and Williams in pyrophosphate buffer, no matter whether their reaction mixture was or was not cooled temporarily after initiation, was probably due to the addition of disks, rather than of A-protein. To test this hypothesis, we performed some reconstitutions from TMV RNA and a very limited supply of disks, but still sufficient for nucleation, and with added A-protein generated in three different ways: (*i*) from helix by disaggregation upon raising the pH from 5.0 to 7.3 for the reconstitution reaction, (*ii*) by equilibration at pH 7.0, ionic strength 0.1 and 4° for 48 hr, and (*iii*) by the technique of Richards and Williams of disaggregating disks by cooling for about 10 min on ice. The results are shown in Fig. 2 and confirm that A-protein, whether freshly generated from helix or at equilibrium, has an insignificant effect on growth compared with that of disks (broken line in Fig. 2*d*). Therefore, the kinetic results (e.g., Fig. 1), which show a rate of reconstitution that increases with the disk concentration, cannot be interpreted simply as an effect of the disk concentration on the initiation rate, with subsequent growth from the A-protein present in the disk preparation.

Fig. 2 also demonstrates the distinct difference between the behavior of the "A-protein" generated by the method of Richards and Williams and that of equilibrium A-protein. This nonequilibrium material gives a much faster growth than the equilibrium A-protein, although still slower than the same

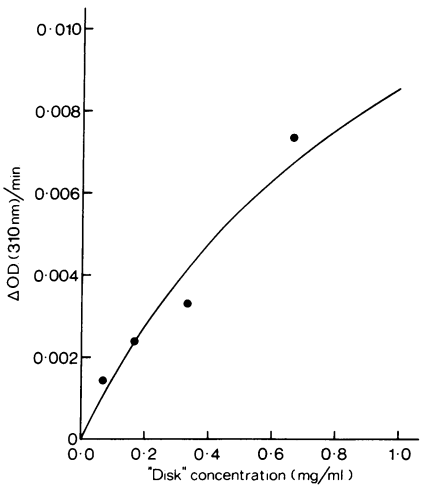

FIG. 1. Effect of addition of A-protein on the rate of reconstitution from disks and RNA. The *curve* is calculated for disks alone (4, 5), whereas each *point* is the average of two determinations made with a mixture of disks, at the concentration shown, and of A-protein (1.0 mg/ml) freshly generated from helix by raising the pH to 7.3 (6, 7). The *abscissa* gives initial rates of reconstitution followed by increase in optical density at 310 nm. Reconstitution is with TMV RNA (50 μg/ml) in 0.1 M sodium pyrophosphate, pH 7.3, at 25°.

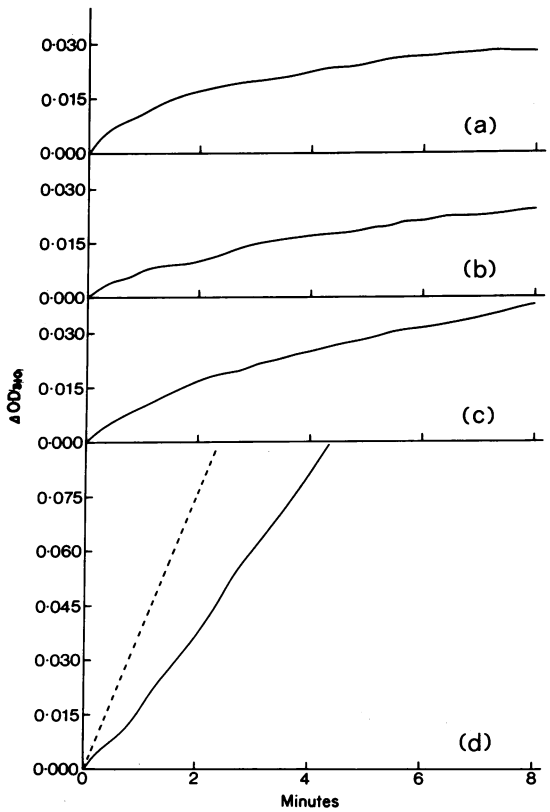

FIG. 2. Effect of state of protein aggregation upon reconstitution of TMV particles. Reconstitution with 0.2 mg/ml TMV RNA in 0.1 M sodium pyrophosphate, pH 7.3, at 25° followed by the increase in optical density at 310 nm. (*a*) Control, 0.2 mg/ml disk preparation, without addition of A-protein. Other curves with A-protein added to 1.0 mg/ml: (*b*) A-protein generated by disaggregation of helix immediately before experiment: (*c*) A-protein at equilibrium at 4°, in sodium phosphate, pH 7.0, ionic strength 0.1, immediately before experiment; (*d*) "A-protein" generated by the method of Richards and Williams (2), i.e., by cooling the equilibrium disk preparation on ice immediately before the experiment. The *broken line* in (*d*) shows the reconstitution found in experiments in which all the protein present (1.2 mg/ml) is added as a disk preparation (4, 5).

concentration of disks. To convey its distinctness from the equilibrium material, we have called the protein freshly depolymerized from disks "A*-protein."

The rapid growth found by Richards and Williams, taken by them to be from A-protein as the protein source, in fact occurred from A*-protein, which reaggregates more rapidly than ordinary A-protein. Moreover, this rate may also be increased by a catalytic effect at the growing end of the nucleoprotein rod, which might act as a surface upon which the disk can reform. Under these conditions complete disks may not be needed for incorporation into the rod. It is not yet possible to distinguish between a direct addition of A*-protein and a limiting rate of aggregation of A*-protein to disks. If A*-protein adds directly, then the growing end of the nucleoprotein rod may be acting as a sensor for the disk-like state of the A*-protein. Even so, the experiments of Richards and Williams show that A*-protein is not sufficiently disk-like to nucleate the assembly.

When ordinary A-protein is available for growth, there is the possibility that protein subunits are slowly added either

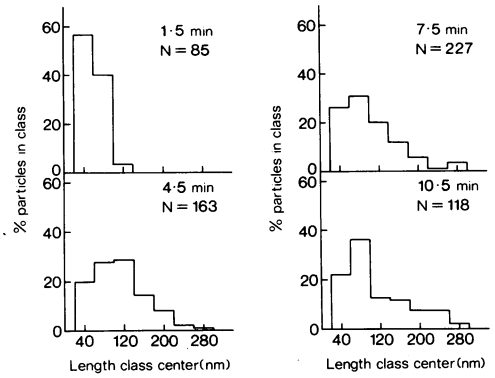

FIG. 3. Distribution of particle lengths during reconstitution. Reconstitution in sodium phosphate, pH 7.0, ionic strength 0.1, at 20° from TMV RNA (0.2 mg/ml) and disks (8.0 mg/ml). Samples were applied to electron-microscope grids, rinsed with water after about 5 sec, stained with uranyl acetate, and then photographed in an electron microscope. Lengths were measured in classes of 40 nm, and the shortest length observed was about 20 nm. N is the total number of particles measured at each time. Electron micrographs by Dr. J. T. Finch.

singly or a few at a time. Fig. 2 shows that the reaction in curve (*a*), with protein added only as a limited quantity of disks (about four disks per RNA on a molar basis), has terminated at a turbidity that correlates closely with that predicted (5) for nucleoprotein rods containing the equivalent of four disks, whereas in the presence of A-protein (curves *b* and *c*) the reaction is continuing at a slow rate after this point. Whether this slow rate is due to the direct addition of A-protein or is a measure of the rate of conversion of A-protein to disks, which then add rapidly, is not clear. Whatever the mechanism, such addition is so slow that, in the presence of an adequate supply of disks, its contribution to the overall growth is insignificant.

Effect of reaction conditions for reconstitution

A comparison of the results obtained by Richards and Williams (2) for reconstitution under different conditions raises a further question. Their data yield a rate of growth from disks in phosphate buffer, pH 7.0, 20°, that is only one-quarter of that in pyrophosphate buffer, pH 7.3, 25°. This difference cannot be due to the different temperatures, as their Table 1 shows the temperature effect to be small. In fact, we have found by light scattering no significant change in rate between the two different sets of conditions, the rate in phosphate being within 10% of that in pyrophosphate. Moreover, the rate of growth in pyrophosphate determined by light scattering (5) corresponds closely to that in phosphate determined by electron microscopy (J. T. Finch and P. J. G. Butler, manuscript in preparation). Since Richards and Williams find a similar rate in pyrophosphate, their rate in phosphate is anomalously slow. This could be due to interference by the bentonite present in their reaction mixtures, possibly through its adsorbing either the RNA or the protein and only liberating these for reaction at a slow rate. One would expect this effect to be less marked in the pyrophosphate buffer because of the higher concentration used, and also because the more highly charged ionic species competes more effectively with the reactants for binding to the bentonite.

The experiments of Finch and Butler also show a distinct difference between growth from disks and from A*-protein generated in phosphate buffer. Using Richards and Williams' technique for initiation with disks and then growth from A*-protein, these experiments show a much slower rate of growth from A*-protein than from disks (about one-fifth). Indeed, a repetition of the experiments shown in Fig. 2, using light scattering to follow the growth, but in phosphate buffer rather than pyrophosphate, suggests that, in reconstitution in phosphate, A*-protein may not be so very different from A-protein in its behavior. Clearly the determination of the lifetime and properties of A*-protein in different buffers needs fuller investigation.

We should emphasize that pyrophosphate buffer was chosen for historical reasons (9), while we subsequently used the phosphate buffer because it is optimal for disks in the protein equilibrium (7). No attempt has been made to find the best conditions for reassembly from disks and TMV RNA.

Rate of initiation of nucleoprotein helices

There is another divergence between our views not emphasized by Richards and Williams. This concerns the time taken for the initiation, although here the role of disks is not in dispute. We had concluded that, with TMV RNA, the initiation of nucleoprotein rods from disks is relatively rapid and, hence, that the rate of the overall assembly is limited by growth (1). On this basis we have analyzed the kinetics of assembly in detail (4, 5). Our conclusion was based upon the length distribution of growing rods shown by electron microscopy and also upon our discovery of a special sequence of nucleotides at the 5′-hydroxyl end of the TMV RNA that makes for a much more rapid assembly of the nucleoprotein from TMV RNA than from other RNAs.

Nucleoprotein rods are formed from disks (at 1 mg/ml) and polyadenylic acid [poly(A)] at about 1% of the rate from TMV RNA; because of the special initiation sequence on the TMV RNA, we had reasoned that this slower rate was due to a reduced rate of initiation of particles rather than to a significant change in the rate of growth. Butler (4, 5) found the dependence of the rate of assembly with poly(A) upon the disk concentration to have a quite different character from that with TMV RNA. Assuming that with poly(A) initiation, and not growth, is the rate-limiting step, he concluded that the initiation rate depended upon the second power of the disk concentration. This conclusion is compatible with an obvious model for the initiation reaction, in which the initiation sequence at the 5′-hydroxyl end of the RNA is sandwiched between two protein disks to start the nucleoprotein helix. By contrast, the kinetics for the overall reassembly with TMV RNA show a saturation of the rate with increasing disk concentrations, similar to a Michaelis–Menten type of curve for an enzyme–substrate reaction with a K_m of 1 mg/ml (1, 4). We conclude that this saturation is due to the rate of growth ceasing to be collision-limited above a certain protein disk concentration, but requiring a finite time for the rearrangement of the protein subunits from a disk into the nucleoprotein helix. This time has been estimated (5) at about 6 sec, giving a minimum time for the completion of a TMV particle under these conditions of just over 6 min.

In all of these kinetic experiments the rates of reconstitution were measured as initial rates; if it is not to influence the overall rate, the initiation with TMV RNA must, therefore,

be rapid relative to growth. The effect of initiation cannot be detected, even at low protein concentrations, since the measured points lie close to the rectangular hyperbola theoretically expected for a single rate-limiting process that saturates. Further evidence for rapid initiation comes from curves such as Fig. 2*a*. This figure shows that the optical density expected for rods with two disks added to each RNA molecule is reached within the first minute, even at such low initial protein concentrations as 0.2 mg/ml and at an RNA-to-protein ratio such that half the protein will have reacted by this stage, thus significantly lowering the reactant concentration during the experiment. Since the initiation appears to be second order in the disk concentration, at higher concentrations this reaction will be so rapid that it is observable only by rapid reaction techniques.

The results of Richards and Williams (2) that indicate a relatively long time for initiation might be explained by their technique for preparing the samples for electron microscopy. They stopped the reaction by diluting the reconstitution mixture either 200-fold or 50-fold into ice-cold 2 mM NH_4HCO_3 (pH 7.8), having observed that this procedure "had no discernible effect on native TMV, but that it depolymerized RNA-free protein disks and larger aggregates to material not detectable in the electron microscope." We suggest that some of the shorter, and therefore less-highly stabilized, nucleoprotein particles might also be depolymerized by this procedure and, hence, not be counted. A depolymerizing effect of this kind has indeed been observed (J. T. Finch and P. J. G. Butler, manuscript in preparation); when a preparation of partially reconstituted rods was cooled the average length was significantly shorter after cooling than it had been beforehand. Since such depolymerization would probably have more effect upon the shorter rods than upon the longer ones, it would render the measurement of average rod lengths unreliable while most of the rods were too short to be stable.

Comparison of techniques for examining particle assembly

Light scattering (measured as turbidity) and particle measurement by electron microscopy have both been used to follow the assembly of the TMV nucleoprotein. While electron microscopy appears to be the more direct, in that the actual lengths of the particles on the microscope grids are determined, the specimen preparation may, as we have seen, cause either partial or complete depolymerization of some rods. We have used the simple technique of allowing the particles to adsorb onto the carbon substrate of the grid from a drop of the sample solution, then rinsing with water and negatively staining with uranyl acetate. This seems to be satisfactory because it gives results consistent with those found from light scattering experiments, although the absolute number of particles is not determined.

Fig. 3 shows histograms of particle lengths, which were obtained from some preliminary experiments; these agree with later and fuller results (J. T. Finch and P. J. G. Butler, manuscript in preparation). Here the class with a center at 280 nm includes full-length TMV particles of about 300 nm; such particles appear by 4.5 min and have reached their final fraction of the total molecules by 7.5 min. This time of completion compares with the 6 min predicted for a saturating protein concentration, by extrapolation of the initial rates of assembly measured by light scattering over a wide range of disk concentrations. The wide distribution of particle lengths and the small fraction of full-length particles even in the final state is due to the presence of degraded RNA molecules, and the final distribution is a function of the particular RNA preparation used. (This effect can be corrected for, and a detailed analysis will be presented in a subsequent fuller paper.)

By the appropriate choice of technique and also of conditions, both methods give a consistent picture of the growth of the TMV particle from initiation to the completion of the full-length rods that agrees with our earlier conclusions. The two techniques are complementary in that light scattering is convenient for studying the early stages of the process, when the rods are too short to be reliably examined by electron microscopy, while the microscope is useful for studying the later stages, by which time the rods are too long for the aggregation to be followed quantitatively by light scattering.

We thank Drs. Williams and Richards for sending us a copy of their manuscript before publication and for further written discussion.

1. Butler, P. J. G. & Klug, A. (1971) "Assembly of the Particle of Tobacco Mosaic Virus from RNA and Disks of Protein," *Nature New Biol.* **229**, 47–50.
2. Richards, K. E. & Williams, R. C. (1972) "Assembly of Tobacco Mosaic Virus *In Vitro:* Effect of State of Polymerization of the Protein Component," *Proc. Nat. Acad. Sci. USA* **69**, 1121–1124.
3. Butler, P. J. G. (1971) "Assembly of Tobacco Mosaic Virus Particle," *Nature* **233**, 25–27.
4. Butler, P. J. G. (1971) "The Mechanism and Control of the Assembly of Tobacco Mosaic Virus from Its RNA and Protein Disks," *Cold Spring Harbor Symp. Quant. Biol.* **36**, 461–468.
5. Butler, P. J. G. (1972) "Structures and Roles of the Polymorphic Forms of Tobacco Mosaic Virus Protein: 6. The assembly of the nucleoprotein rods of tobacco mosaic virus from the protein disks and RNA," *J. Mol. Biol.*, in press.
6. Durham, A. C. H. & Klug, A. (1971) "Polymerization of Tobacco Mosaic Virus Protein and its Control," *Nature New Biol.* **229**, 42–46.
7. Durham, A. C. H. (1972) "Structures and Roles of the Polymorphic Forms of Tobacco Mosaic Virus Protein: 1. Sedimentation Studies," *J. Mol. Biol.* **67**, 289–305.
8. Lonchampt, M., Lebeurier, G. & Hirth, L. (1972) "Polymerization of Tobacco Mosaic Virus Protein in Na Pyrophosphate," *FEBS Lett.* **22**, 297–300.
9. Fraenkel-Conrat, H. & Singer, B. (1964) "Reconstitution of Tobacco Mosaic Virus IV. Inhibition by Enzymes and Other Proteins, and Use of Polynucleotides," *Virology*, **23**, 354–362.