RNA-PROTEIN INTERACTIONS IN THE ASSEMBLY OF TOBACCO MOSAIC VIRUS

P. J. G. Butler and G. P. Lomonossoff, *Medical Research Council, Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 2QH, United Kingdom*

ABSTRACT Assembly of tobacco mosaic virus is initiated by the binding of a specific loop of the RNA into the central hole of the disk aggregate of protein subunits. Since the nucleation loop is located about five-sixths along the RNA molecule, subsequent elongation must be bidirectional. We have now measured the rates of elongation in the two directions by determining the lengths of RNA protected from nuclease digestion at different times and using either intact TMV RNA, or RNA with most of the longer tail removed. Comparison of the rates with the protein supplied as either a mixture of disks with A-protein (a mixture of less aggregated states) or just A-protein, shows that different mechanisms and protein aggregates are used for the most rapid growth. When disks are present, they add more rapidly along the longer RNA tail but do not appear to add directly on the shorter tail. In contrast, smaller aggregates (A-protein) can add at both ends of the rod, but do so more slowly. Mechanisms for these processes are discussed. Preliminary results on the binding of the specific hexanucleotide AAGAAG to the disk are given and compared with the known changes on binding nonspecific hexanucleotides or the trinucleotide AAG.

INTRODUCTION

Tobacco mosaic virus (TMV) is a simple virus having a single type of coat protein subunit and a single-stranded RNA genome ~6,400 nucleotides long. The coat protein forms a helical aggregate, with 16⅓ subunits per turn, and the RNA is packaged by being intercalated between the turns (for review, see reference 1). The overall length of the rod-shaped virion is determined by the length of the RNA, which packs with 3 nucleotides per protein subunit, with the growing rod elongating until the RNA molecule is completely coated to give the 300-nm rod containing ~2,100 protein subunits.

With such a simple structure, assembly might also be quite simple, with single subunits adding onto the "step" at the growing end of the particle where the RNA would protrude, thereby regenerating the step each time until all of the RNA was incorporated (2). This hypothesis received substantial support from the classic experiments of Fraenkel-Conrat and Williams (3) showing that the virus could reassemble from its isolated protein and RNA, thus for the first time demonstrating the process of "self-assembly" in a biological system. Moreover, self-assembly to regenerate faithfully the original structure was found to be both fairly specific for the conditions under which it occurred and highly specific for the RNA which could be incorporated (4,5). This specificity very strongly favors assembly upon the homologous viral RNA, even against RNA from different strains of TMV.

This straightforward picture of the assembly of TMV overlooks a major difficulty — how does the nucleation occur to start the process? To form the smallest stable nucleoprotein helix, about 20 protein subunits would have to align themselves along the RNA, bonding only to a single neighbor on each side until the growing aggregate was large enough to form more than a complete turn of helix, so that interaction in an axial direction could occur to stabilize the

structure. Single subunits, each interacting with only 3 nucleotides, could not give the high specificity for the RNA which is observed.

The isolated coat protein shows polymorphic aggregation driven by entropic effects (reviewed in reference 6), with the mode of aggregation controlled largely by pH (7). Below pH 7, the protein forms a helix which is very similar to that in the virus, while at higher pH's it occurs as a mixture of small aggregates known as "A-protein." In the region around pH 7, however, a specific aggregate is found, known as the "disk"; of which the structure has now been solved to atomic resolution (8, 9). The disk has two rings of 17 subunits, giving a diameter similar to that of the helix, to which it is readily and rapidly converted when the pH is lowered (7). The subunit packing within each of the rings is very similar to that in the helix, although the contacts between rings are quite different (8).

The occurrence of disks as the main (~80%) component of the protein equilibrium around neutral pH and at 20°C — conditions which have been found to favor reassembly (4) — led us to postulate that they might be essential for overcoming the problems of nucleation by interacting with the RNA to form the first turns of the nucleoprotein helix (10). This proved to be the case. Moreover, it was shown that the nucleation reaction is highly specific for the viral RNA (10) and occurs at a unique site upon the RNA which has been isolated (11). This region of the RNA has been sequenced (12, 13) and the smallest "core" binding to the protein disk during nucleation has been shown to be able to form a hairpin loop of ~50 nucleotides, with the special sequence AGAAGAAGUUGUUGAUGA at the open end (12). Since 3 nucleotides bind per subunit, this repeat pattern with G in every third position strongly suggested that this open loop was the actual origin of assembly. Surprisingly, however, the nucleation region was found to be internal in the RNA, ~1,000 nucleotides from the 3′-end (14), a result confirmed recently by sequence analysis (15).

The structure of the disk, as determined by x-ray crystallography (8, 9), shows the two rings of subunits to be in contact at their outer ends, but opened apart onto the central hole, so that there is ready access from this hole to the RNA binding site between the subunits. Taken together with the probable structure for the nucleation region on the RNA, this led us to propose (16) that nucleation might be occurring by the insertion of the RNA loop into the central hole of the disk to reach the binding site between the protein subunits. As the single-stranded loop of the RNA interacted with the protein binding site, the base paired stem of the loop could melt, allowing further RNA binding around the complete ring. The interaction could then trigger the dislocation of the disk, causing it to form a short protohelix containing the first turn of the RNA. Because of the internal location of the nucleation region on the RNA, such a growing helix would necessarily have both RNA tails protruding from the same end of the rod, one of them being looped back down the central hole. Such a structure has been found (17, 18). Furthermore, it has been shown that the longer (5′) tail was the one which was looped back, and that this unusual structure appeared necessary for the rapid elongation of the growing particles (17).

The Outstanding Problems

There is widespread agreement on the requirement of disks for the nucleation of assembly of TMV (10, 19, 20) and even on the kinetic order of the reaction, with a single disk interacting with each RNA molecule in the rate limiting step (21, 22). However, the nature of the protein species involved in the subsequent elongation is still a matter of controversy (reviewed with full discussion of the literature in references 23 and 24, and more recently in references 25–28), which has proved difficult to settle experimentally. The major obstacle is the very

nature of the disk preparations. Under optimal conditons for the reassembly (5) the equilibrium mixture of protein aggregates consists of ~80% disks and 20% A-protein (29) with a rapid microequilibrium of subunits between the states, though a somewhat slower bulk equilibrium (see reference 24 for discussion). It is thus possible to obtain preparations of A-protein alone, but the normal equilibrating disk preparations always contain some A-protein as well as the two-layer disks. In an attempt to avoid this problem some experiments have been conducted under conditons where the "disks" are metastable (28, 30). However, these metastable aggregates were not adequately characterized to show that they were really two-turn disks rather than, say, short helical segments, while their metastability itself shows that they are kinetically "locked" in some way, unlike the normal freely-equilibrating disks (see discussion below). Such results cannot therefore be validly compared with those obtained with equilibrium disks, and their usefulness in understanding the normal elongation of TMV is questionable.

Subsidiary complications arise from the inherent problem in measuring the elongation of particles up to sizes of 300 nm. The interpretation of average properties (e.g., turbidity) requires assumptions about the nucleation of the particles and their competence for further elongation, while apparently direct methods (in particular elecron microscopy) may require specimen preparation techniques that perturb the distribution being measured. Moreover, the internal site of nucleation results in bidirectional elongation, and it has been shown that growth occurs in both directions simultaneously (27). Since the same protein source may not be responsible for growth in each direction, the optimum technique for studying the elongation should be capable of resolving which direction is dominant under the conditions employed.

We have recently found that measurements of the RNA protected from nuclease attack provide a reliable method for determining the rate of elongation and, from the protection of oligonucleotides in known locations along the RNA, its main direction. We have already reported some results with this technique, studying reassembly upon intact TMV RNA with the protein as a disk preparation (26, 27). Under these conditions the elongation is mainly in the major (3′ to 5′) direction and much slower in the minor (5′ to 3′) direction (11, 12, 18). However, stripping of TMV under mildly alkaline conditions removes protein specifically from the 5′-end of the particles, so the partially stripped virus (PSV) isolated contains partial RNA molecules with intact 3′-termini but shortened 5′-tails (31). RNA molecules with very short tails to the 5′-side of the nucleation region show much less rapid elongation than intact molecules (32) but, as expected, reaction does still occur (33). Taking advantage of this, we have been able to extend our observations to the rates of elongation specifically along the minor (3′) tail of the RNA and we report these in this paper.

Another problem of continuing interest is the detailed interactions between the RNA and protein. The location of the RNA within the virus has been determined from the 0.4nm resolution electron density map obtained from x-ray diffraction studies on oriented gels of TMV (34), but this can only show an "average nuclotide" in each position and, moreover, is not yet at a resolution to show any detailed contacts. Difference maps of a specific trinucleoside diphosphate (ApApG) from the nucleation region of the RNA, bound into the disks in the crystal, have been obtained to a resolution of 0.5 nm (35) and show clear movements of protein and, possibly, the site of the nucleotide binding, although again not at an adequate resolution to study details. We are currently extending these studies to a resolution higher than 0.3 nm using the oligonuclotide AAGAAG (again corresponding to part of the origin of initiation). The preliminary results are described here.

MATERIALS AND METHODS

TMV, TMV protein (as both A-protein and a disk preparation), and TMV RNA were prepared and their concentrations determined as previously reported (11). Partially stripped virus (PSV) (36) was prepared as previously reported (27), but with dialysis for 24 h, RNA tails being removed with micrococcal nuclease (31), and the RNA prepared as usual. The RNA was fractionated in 15-40% (wt/vol) linear sucrose gradients containing 0.1% (wt/vol) sodium dodecyl sulphate for 24 h at 25,000 r/min and 20°C; the gradients were collected with an Isco gradient fractionator (Isco Manufacturing Co., Inco, Kansas City, Mo.). Fractions containing appropriate length RNA were taken and the RNA was precipitated with ethanol.

All reassembly experiments were carried out in sodium phosphate buffer, pH 7.0, ionic strength 0.1 M at 20°C. Partially assembled rodlets were prepared by reaction with appropriate amounts of a disk preparation (see below) and, when necessary, fractionated in sucrose step gradients as before (17), except for being in assembly buffer to minimize the time required for subsequent dialysis (4 h).

Assembly was assayed by turbidity increase at 310 nm (10), measured with a Unicam SP 8-100 (Pye-Unicam Ltd., Cambridge, England) or a Gilford 2400-2 spectrophotometer (Gilford Instrument Laboratories, Inc., Oberlin, Ohio), with an RNA concentration of 0.05 mg/ml. Alternatively, the length of RNA protected from digestion by micrococcal nuclease was determined as before (26, 27), using agarose/acrylamide gel electrophoresis. Densitometry of the negatives from photography of gels stained with ethidium bromide was with a Joyce Loebl densitometer (Joyce Loebl Ltd., Newcastle-upon-Tyre, England). Analysis of coating of specific oligonucleotides was by ribonuclease T_1 digestion and fingerprinting as previously described (27).

RESULTS AND DISCUSSION

Elongation towards the 3'-Terminus

To make a substrate for studies of reassembly in the 3' direction we have partially reassembled PSV RNA preparations with a protein disk preparation to coat completely the 5'-tail but still to leave a substantial 3'-tail. Such rodlets can be used to investigate the reassembly using different protein sources.

A PSV RNA preparation containing molecules 2,200 to 3,200 nucleotides long was reacted with 12.5 times its weight of protein and shown, by ribonuclease T_1 digestion and fingerprinting, to be fully coated to its 5'-end but still to have most of the 3'-tail uncoated. After sedimentation through a sucrose step gradient to remove any unreacted RNA or protein (17), the rodlets were rapidly dialyzed back into reassembly buffer and used to measure the kinetics of elongation, assayed by the rate of turbidity increase. Although the absolute rates of change will depend upon the RNA lengths, the relative rates between different protein species and concentrations will be strictly comparable despite the heterogeneity in original RNA length.

The relative rates of elongation (Fig. 1) show a saturation with increasing concentrations of protein, irrespective of the form in which this was added. The maximum rates are the same when adding either A-protein or a disk preparation, but at lower concentrations, the rate is always faster when adding A-protein rather than a disk preparation (containing ~ 20% A-protein). This suggests that the elongation in this 5' to 3' direction is not occurring directly from disks, but rather from some component of the A-protein in the disk preparation, so that a higher total protein concentration is required to achieve the same rate.

A more sensitive assay for elongation is to measure the length distribution of the RNA at increasing times. To maximize resolution, we employed the shortest PSV RNA which reproducibly contains the nucleation region (1,700–2,200 nucleotides long) and reacted this with a tenfold weight excess of protein to prepare rodlets. This again gave essentially complete coating to the 5'-ends, while leaving the 3'-tails uncoated. Since the measurements were to be of the longest protected fragments, possible unreacted RNA was not removed, as it

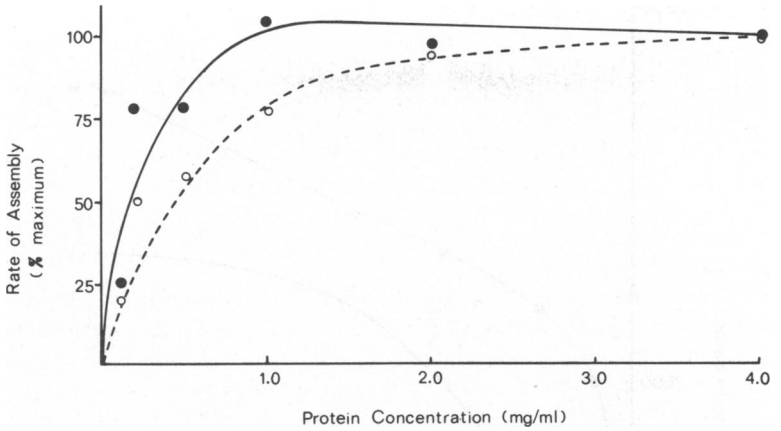

Figure 1 Effects of protein concentration and aggregation state on rate of elongation in 5′ to 3′ direction. Rates are expressed as percent maximum to allow data from different experiments to be shown together. Partially reassembled rodlets with PSV RNA were reacted further with either A-protein (-●-) or a disk preparation (-O-).

could not affect the measurements and any extra preparative steps might cause some damage to the RNA tails of the rodlets. Such measurements will be dominated by the longest RNA molecules present in the preparation and so will overcome any complicating effect of length heterogeneity at the 5′-ends of the PSV RNA. Elongation was carried out with either a disk preparation or A-protein added to 1 mg/ml, since the turbidity measurements had shown that this gave significant rates of elongation and also a clear difference between the protein species (Fig. 1).

Unlike the RNA protected during reassembly of intact TMV RNA with a disk preparation, which shows discrete bands (26 and below), the RNA protected in the rodlet preparation migrated as a single broad band in gels, due in large measure to the variable lengths at the 5′-ends of the PSV RNA. With the protein concentration below saturation, the spread will increase with time. To eliminate these problems, the length of the longest class of molecules was determined by measuring the position of the trailing edge of the RNA peak in the gel (taking the half-height). After densitometry of photographs of the stained RNA, this trailing edge of the peak was found to be consistent, unlike the edge representing the smaller material, which showed increasing dispersity with time due to damaged RNA molecules which would become fully coated but lack the 3′-terminus.

The time-course of elongation (Fig. 2) again shows the more rapid extension when the protein was added as A-protein. Extrapolation to zero time of elongation gives an estimate that the longest rodlets contained ~ 1,600 protected nucleotides of RNA. The actual elongations with A-protein are ~ 280 nucleotides in the first minute and 130 nucleotides in the second minute; the fall in rate probably reflects the completion of coating on some of the RNA tails, since the longest rods are within the total size range of the PSV RNA by the end of the first minute. With a disk preparation these rates are 40 and 73 bases per minute during the first and second minutes, respectively. The increase is probably due to error in extrapolation to zero time.

The extent of protection of the RNA can also be followed by the recovery of known oligonucleotides (27); since these are specifically located in the sequence, there can be no possible effect of the 5′-terminal heterogeneity. Since the sequence of ~ 1,000 nucleotides at

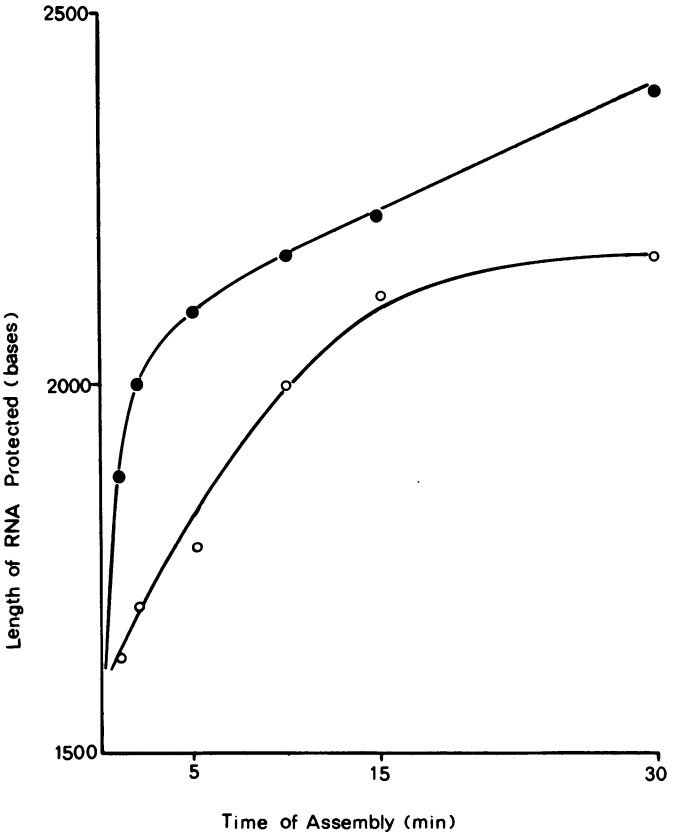

Figure 2 Time course of protection of RNA 3'-tails with different protein aggregates. Lengths of the longest protected RNA molecules were determined (see text) during elongation of partially assembled rodlets containing PSV RNA, with A-protein (-●-) or a disk preparation (O), thus allowing estimation of maximal rates of reassembly.

the 3'-end of TMV RNA is known (15), we know the location of the characteristic ribonuclease T_1 nucleotides chosen: "spot 4" between 563 and 579 nucleotides in and "spot 5" between 510 and 525 nucleotides in (numbering from ref 27). The time courses for protection of these nucleotides (Fig. 3) again show the more rapid elongation in the 5' to 3' direction from A-protein than from disks, although since the fractional protection will measure the state of the average particle, the actual rates cannot be compared directly with those measured above.

Elongation towards the 5-Terminus

The nucleation region is ~ 1,000 nucleotides from the 3'-end of TMV RNA; hence over 5,000 nuclotides have to be coated in the 3' to 5' direction. Elongation in this direction is much more rapid on intact RNA (11, 18), thus dominating the overall kinetics of assembly, so that most studies apply mainly to this direction. We have found that during assembly with protein supplied as a disk preparation, the protected RNA lengths are "quantized" over the entire range from 470 to 2,750 nucleotides (the limits of resolution of the gels used), with a step size of either 50 or 100 nucleotides, corresponding rather precisely to one or two turns of RNA incorporated, as expected if subunits were adding directly from disks (26). We have now extended this study to include the effects of adding A-protein.

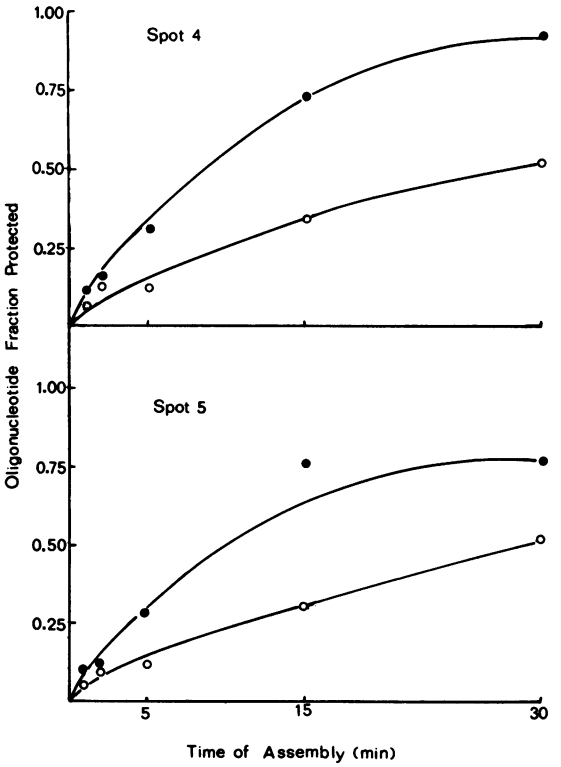

Figure 3 Extent of protection of specific oligonucleotides in the 3'-tail of PSV RNA during elongation with different protein aggregates. The recovery of specific oligonucleotides in the protected RNA was measured during the elongation of partially assembled rodlets containing PSV RNA, with A-protein (-●-) or a disk preparation (-O-), giving a measure of the average extent of coating at each time.

Since disks are needed for nucleation, it is not possible to observe assembly simply from A-protein. We therefore compared the lengths protected during reassembly of RNA (0.2 mg/ml) with (*a*) a disk preparation (4 mg/ml); (*b*) with a disk preparation plus A-protein (each at 4 mg/ml); (*c*) A-protein (4 mg/ml) after 10 min reaction with a disk preparation (0.2 mg/ml) to allow nucleation (average length of rodlet 1/20th that of TMV). The results (Fig. 4) show the more rapid increase in length of protected RNA whenever disks are present, although, not surprisingly, A-protein can add in this direction. By following the longest molecules, any possible complication of partial nucleation in the initial stages is eliminated. (The RNA markers show the expected log/linear relationship from 870 to 3,375 nucleotides, but, as is usual in acrylamide gel electrophoresis (37), this relationship ceases to be linear at the highest molecular weights, in this case between 3,375 and 6,400 nucleotides long.)

The addition of extra A-protein together with the disk preparation has little effect; if anything, it slows the elongation compared with addition of disk preparation alone. However, elongation from A-protein alone, after nucleation with a disk preparation, is markedly slower. As shown in Fig. 4 and other time courses, the first full length RNA molecules are protected within 6 min with a disk preparation and within 15 min with A-protein, corresponding to elongation rates of 1,067 and 427 nucleotides per minute, respectively. Measurements made during the second minute (to offset any effect of nucleation; see Fig. 4) give rates of 1,200 and 400 nucleotides per minute for the disk preparation and A-protein, respectively. We therefore conclude that the overall rate of elongation is two-and-a-half to three times faster from a disk

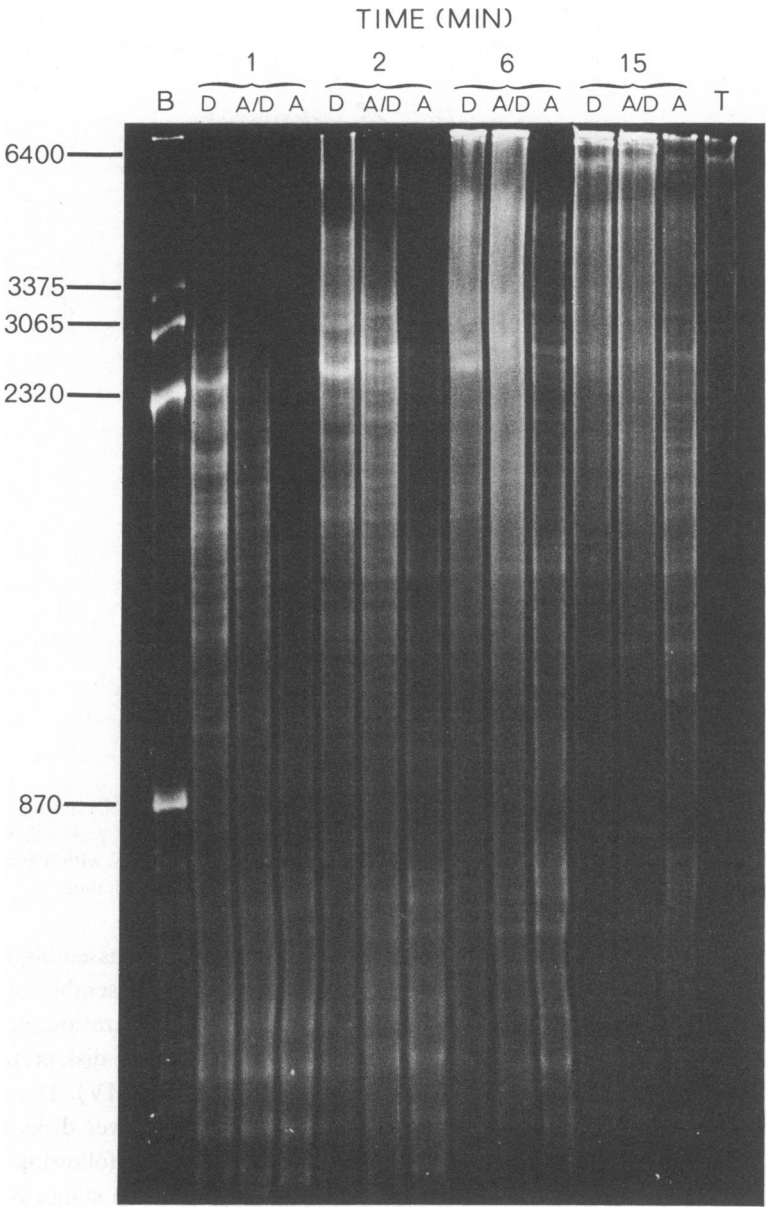

Figure 4 Agarose/acrylamide gel to determine sizes of protected RNA during assembly with different protein aggregates. Assembly was carried out with TMV RNA and either a disk preparation (D), A-protein together with a disk preparation (A/D), or A-protein after nucleation with limited amounts of a disk preparation (A) (see text for details). Uncoated RNA tails were removed by digestion with micrococcal nuclease and the protected RNA extracted after inactivation of the nuclease. The lengths of markers of brome mosaic virus RNA (track B) and the original TMV RNA (track T) are shown in nucleotides.

preparation than from A-protein. Furthermore, since we have shown by the same method that elongation in the minor (5′ to 3′) direction is faster from A-protein, growth in the major (3′ to 5′) direction must be favored from disks by more than this factor.

The one feature which might be in any way discordant with this picture is the occurrence of a banding pattern during elongation from A-protein. For the shorter bands this is the direct

result of the nucleation and initial elongation from the disk preparation giving such a banded pattern, and the pattern after 1 min elongation from A-protein is almost indistinguishable from that of just the nucleated RNA (data not shown). The persistence of some specific bands when only A-protein is present suggests that their structure presents some barrier to addition of subunits from the A-protein, a result consistent with the earlier observations of the lack of effect of A-protein upon the rate of turbidity increase from a disk preparation (22, 38). The longer bands which appear only during the extension do not correlate in detail with those from a disk preparation and, in particular, do not occur at the spacing of 50 or 100 nucleotides. Thus it is likely that they are a consequence of regions of the RNA with sequences unfavorable for coating either because of unfavored RNA-protein interactions or strong base pairing of the uncoated RNA. Such sequences would have much less effect upon elongation from disks because the cooperative addition of many subunits would more readily overcome such inhibition (10), although a number of bands do still persist in an anomalous fashion.

Comparison of Rates of Elongation

The most efficient elongation of the TMV particle is clearly that which occurs most rapidly. Changes in conditions or in the form in which protein is supplied, which result in slowing down the rate, may well be achieving this effect by altering the mechanism of assembly; in particular, they might possibly be preventing more complex assembly directly from disks while still allowing elongation by the simpler mechanisms from A-protein. Since these experiments are frequently intended to distinguish between these mechanisms, any experimental technique that requires or involves a lowering of the rates must be of doubtful value.

We consider only experiments under optimal conditions and have compared the rates of elongation we have measured by various techniques (Table I). For the sake of comparability, these rates have been taken either from experiments where a high protein concentration was used (at or above 4 mg/ml), or extrapolated to the "maximum rate" from the data available at various protein concentrations and assuming that saturation kinetics will apply (as observed, 39, 41), thus obtaining rates with nonlimiting protein concentrations. All rates are expressed in common units (subunits per second), converting turbidity changes as previously described (39), length changes by taking 7.1 subunits/nm (i.e., 2,130 subunits in a 300nm rod), and RNA lengths by using 3 nucleotides bound per subunit. It has not been possible to make a detailed comparison with the results of other workers because the necessary details for the conversions are often not available to us, and also because of differing conditions for the experiments, which in some cases cause large effects, or changes in the overall effect being measured (e.g. infectivity rather than elongation) (see reference 24 for discussion).

Measurements 1 to 5 were made either by following the elongation of the most rapidly growing rods (2, 4, and 5) or during the initial stages of growth when the effects of any damaged RNA will be minimal (1 and 3), and will therefore give estimates of the maximum rate under these conditions. The average rates are 6.5 subunits/s (SD 0.9) and 2.1 subunits/s (SD 0.3) from disk preparations and A-protein, respectively. In each case the standard deviation of the measurements is less than 15% of the rate, despite the different measurement techniques employed, suggesting that all of them are giving reasonable estimates of the true rate. Since experiments 1 and 3 involved the nucleation of assembly, no measurement was possible with A-protein alone (i.e., in the absence of a disk preparation). In all of the measurements upon the longest particle (2, 4, and 5), there can be no possible effect due to continuing nucleation, as has been suggested for the average rates (19, 20), for the reason already discussed.

TABLE I

COMPARISON OF THE RATES OF ASSEMBLY OBSERVED WITH DIFFERENT TECHNIQUES AT NONLIMITING PROTEIN CONCENTRATIONS

Technique employed	Rate		Reference
	Disk preparation	A-protein	
	(*subunits per second*)		
1 Turbidity with free RNA	5.4	—	39
2 Full length rods in e/m	7.1	1.8	40
3 Pulse-chase of labeled protein	7.6	—	22
4 Protection of full-length RNA	5.9	2.4	—
5 Increase in protected RNA length	6.7	2.2	—
6 Average growth rate in e/m	2.9	0.62	40
7 Turbidity with prenucleated rodlets*			
a	3.6	0.67	41
b	0.85	0.37	—
8 Rate of protection of 3′-tails of RNA with prenucleated rodlets‡			
a	0.22	1.6	—
b	0.41	(0.70)§	—

When available, separate rates are shown for elongation with a disk preparation or A-protein as the protein source. In the latter case, nucleation had to be carried out beforehand with a disk preparation. Assembly is at 20°C in sodium phosphate buffer, pH 7.0, ionic strength 0.1 M.

*Absolute rates vary between preparations, depending upon the free RNA tails still available (see text), but comparisons between protein sources (across line) are valid.

‡Assembly was prenucleated with limited amounts of a disk preparation upon RNA lacking most of the tail to the 5′-side of the nucleation region (see text). Elongation was then measured on the sole (3′) tail. The rates estimated during the first and second minutes are given.

§Value in second minute is unreliable, as rods are running out of RNA (see text).

Experiments 6 and 7 measure quantities which depend upon all the particles present in the solution during the experiment and are therefore rather slower and more variable than the measurements upon the fastest growing particles, probably due to differing degrees of damage to the RNA tails, particularly when partially assembled rodlets are used (cf. 7*a* and *b*). However, the comparative picture is again similar, with elongation from a disk preparation 3–4 times faster than that from A-protein.

An opposite effect is seen when the rates of elongation along the minor RNA tail are compared (line 8). In this case A-protein gives a rate significantly higher than that from the disk preparation, suggesting that it alone can add in this direction (see above). The rapid fall-off in the rate with A-protein between the first and second minutes (a and b) is probably due to the RNA tails on the longest rods, which are those being followed, becoming fully coated. Comparison of the fastest rates obtained in this direction with those for overall growth shows that the two rates are similar from A-protein. Since the overall rate includes elongation in both directions, this suggests that the 5′ to 3′ elongation is somewhat inhibited during the overall reaction, probably because the "doubled back" tail down the central hole gets in the way.

RNA/Protein Interaction in the Disk

To date, very little information is available about the detailed atomic interactions between any protein and nucleic acid chain. We have found that specific oligonucleotides can be bound into

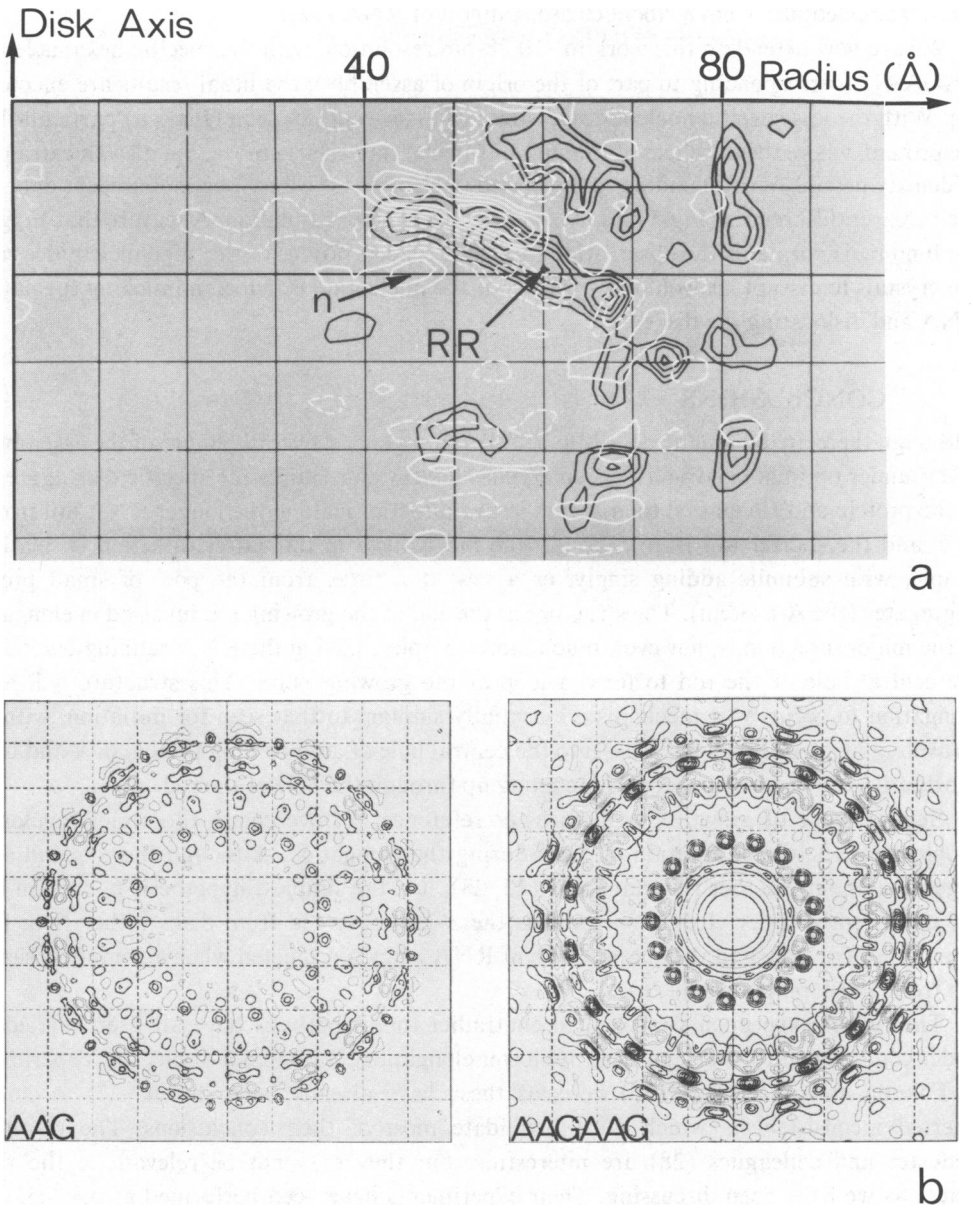

Figure 5 Electron density difference maps with nucleotides bound to TMV protein disks. (*a*) Vertical section containing the radius passing through the right radial helix (marked RR; nomenclature of reference 8). The native electron density is shown in black and the difference density due to binding of mixed hexanucleotide in white, with negative contours dotted. The upward movement of the helix is visible, together with a positive peak (marked n) without any corresponding negative peak, which is tentatively identified as the bound nucleotide (from reference 31). (*b*) Projection down the disk axis, comparing the effects of the trinucleotide AAG (similar to mixed hexanucleotides as in *a*) and the specific hexanucleotide AAGAAG. Positive differences are shown by heavy contour lines and negative differences by faint contour lines.

the protein disk of TMV in the crystal, producing differences in the x-ray diffraction pattern, and have obtained electron density difference maps to ~ 0.5-nm resolution with both AAG and hexanucleotides from a ribonuclease A digest of RNA (32).

We are now extending this work to ~ 0.28-nm resolution, with the specific hexanucleotide AAGAAG (corresponding to part of the origin of assembly); the intial results are encouraging. With the less specific nucleotides, distinct changes could be seen (Fig 5 *a*) particularly in the protein, where the α-helices of the subunits in one ring clearly move, but also an extra peak of density possibly corresponding to the bound nucleotide is visible. The more recent data give even clearer differences (Fig.5 *b*) and also extend to higher resolution. A feature that suggests the binding is normal is the observation that high concentrations of the oligonucleotides cause the crystals to disrupt, as would be expected if the nucleotide is indeed mimicking the natural RNA and dislocating the disks.

CONCLUSIONS

Although there are still many details to be filled in, a general overall picture of the assembly of TMV under optimal conditions has emerged. Nucleation requires the specific disk aggregate of the protein and elongation then occurs in two directions: along the longer RNA tail from 3′ to 5′ and the shorter tail from 5′ to 3′. The mechanism in this latter direction is relatively simple, with subunits adding singly, or a few at a time, from the pool of small protein aggregates (the A-protein). The structure at the end of the growing rod involved in elongation in the major direction is, however, much more complex, having the RNA running back down the central hole of the rod to form a loop at the growing point. This structure will allow elongation to occur by a mechanism essentially similar to that seen for initiation, with the "traveling loop" of RNA inserting into the central hole of an incoming protein disk and being constantly renewed by further RNA coming up through the hole of the rod.

Such a mode of growth (for picture see reference 42) overcomes the major topological problem for the direct addition of disks during the elongation. Although there is still some argument about whether this can occur (25, 28), it is the simplest hypothesis to explain both the consistently faster elongation towards the 5′-end observed from disks rather than from A-protein (Table I), and the quantization of RNA lengths protected when assembly is carried out from a disk preparation (26).

The recent experiments invoked to contradict this hypothesis have all been carried out under conditions involving significantly slower elongation. We will not discuss the experiments of Fukuda and co-workers (25) here, as these have already been shown (27) to contain internal inconsistencies which must invalidate most of their conclusions. The results of Schuster and colleagues (28) are interesting, but they may not be relevant to the rapid reactions we have been discussing. Their experiments have been performed at pH 6.5, ionic strength 0.1 M, and 6.5°C, conditions which were deliberately chosen to give metastable 20-S protein aggregates. Unfortunately, these are also conditions in which the protein has been shown to start forming short helical nuclei which have sedimentation coefficients about 20-S (43) and so are impossible to distinguish from disks on the basis of sedimentation analysis. The assumption that the metastable 20-S aggregates are two-layer disks is therefore unsafe. Moreover, the metastability of the aggregates invalidates their use in drawing conclusions about the possible involvement of the freely equilibrating aggregates, since their half-life of 12 d (cf. normal disks in reference 44) shows that their structure is "locked." Interestingly, the actual rate found is only ~ 6% of that under more usual conditions (pH 7, 20°C), and therefore well within the rate observed for A-protein alone. It is thus highly probable that the

metastable aggregates take no part in the reaction, not because disks cannot participate (as concluded), but rather because of their specific metastability, while elongation from A-protein alone simply occurs at a rather slow rate.

The major outstanding problems concern the structure of the virus and the interactions between the RNA and protein. While the structure of the disk is known at atomic resolution (9), work is still in hand to solve the virus structure in such detail and to elucidate the interactions involved both in causing the disk to dislocate into a short helix and in stabilizing the resulting helix. Despite the magnitude of these problems, considerable progress is being made and one can hope to see soon the atomic details of the interaction between a protein and an RNA chain to complement our knowledge of their kinetic behavior.

We thank Doctors A. C. Bloomer, S. Hovmöller, and J. van Boom for allowing us to quote their unpublished results on AAGAAG binding to the disk.

Dr. Lomonossoff thanks the Medical Research Council for a Scholarship for Training in Research Methods.

Received for publication 29 November 1979.

REFERENCES

1. Caspar, D. L. D. 1963. Assembly and stability of the tobacco mosaic virus particle. *Adv. Protein Chem.* **18**:37–121.
2. Watson, J. D. 1954. The structure of tobacco mosaic virus. I. X-ray evidence of a helical arrangment of subunits around the longitudinal axis. *Biochim Biophys. Acta.* **13**:10–19.
3. Fraenkel-Conrat, H., and R. C. Williams. 1955. Reconstitution of active tobacco mosaic virus from its inactive protein and nucleic acid components. *Proc.Natl. Acad. Sci. U.S.A.* **41**:690–698.
4. Frankel-Conrat, H., and B. Singer. 1959. Reconstitution of tobacco mosaic virus. III. Improved methods and the use of mixed nucleic acids. *Biochim. Biophys. Acta.* **33**:359–370.
5. Frankel-Conrat, H., and B. Singer. 1964. Reconstitution of tobacco mosiac virus. IV. Inhibition by enzymes and other proteins, and use of polynucleotides. *Virology* **23**:354–362.
6. Lauffer, M. A. 1975. Entropy Driven Processes in Biology. Springer-Verlag, Berlin and New York.
7. Durham, A. C. H., and A. Klug. 1971. Polymerization of tobacco mosaic virus protein and its control. *Nat. New Biol.* **229**:42–46.
8. Champness, J. N., A. C. Bloomer, G. Bricogne, P. J. G. Butler, and A. Klug. 1976. The structure of the protein disk of tobacco mosaic virus to 5Å resolution. *Nature (Lond.)* **259**:20–24.
9. Bloomer, A. C., J. N. Champness, G. Bricogne, R. Staden, and A. Klug. 1978. Protein disk of tobacco mosaic virus at 2.8Å resolution showing the interactions within and between subunits. *Nature (Lond.).* **276**:362–368.
10. Butler, P. J. G., and A. Klug. 1971. Assembly of the particle of tobacco mosaic virus from RNA and disks of protein. *Nat. New Biol.* **229**:47–50.
11. Zimmern, D., and P. J. G. Butler. 1977. The isolation of tobacco mosaic virus RNA fragments containing the origin for viral assembly. *Cell.* **11**:455–462.
12. Zimmern, D. 1977. The nucleotide sequence at the origin of assembly on tobacco mosaic virus RNA. *Cell.* **11**:463–482.
13. Jonard, G., K. E. Richards, H. Guilley, and L. Hirth. 1977. Sequence from the assembly nucleation region of TMV RNA. *Cell.* **11**: 483–493.
14. Zimmern, D., and T. M. A. Wilson. 1976. Location of the origin for viral reassembly on tobacco mosaic virus RNA and its relation to stable fragment. *Fed. Eur. Biochem. Soc. Lett.* **71**:294–298.
15. Guilley, H., G. Jonard, B. Kukla, and K. E. Richards. 1979. Sequence of 1,000 nucleotides at the 3'-end of the tobacco mosaic virus RNA. *Nuc. Acids. Res.* **6**:1287–1308.
16. Butler, P. J. G., A. C. Bloomer, G. Bricogne, J. N. Champness, J. Graham, H. Guilley, A. Klug, and D. Zimmern. 1976. Tobacco mosaic virus assembly. Specificity and the transition in protein structure during RNA packaging. *In* Structure-Function Relationship of Proteins. R. Markham and R. W. Horne, editors. 3rd John Innes Symposium, North Holland Publishing Company, Amsterdam. 101–110.
17. Butler, P. J. G., J. T. Finch, and D. Zimmern. 1977. Configuration of tobacco mosaic virus RNA during virus assembly. *Nature (Lond.).* **265**:217–219.
18. Lebeurier, G., A. Nicolaieff, and K. E. Richards. 1977. Inside-out model for self-assembly of tobacco mosaic virus. *Proc. Natl. Acad. Sci. U.S.A.* **74**:149–153.
19. Richards, K. E., and R. C. Williams. 1972. Assembly of tobacco mosaic virus *in vitro*: effect of state of polymerization of the protein component. *Proc. Natl. Acad. Sci. U.S.A.* **69**:1121–1124.

20. Okada, Y., and T. Ohno. 1972. Assembly mechanism of tobacco mosaic virus particle from its ribonucleic acid and protein. *Mol. Gen. Genet.* **114**:205–213.
21. Richards, K. E., and R. C. Williams. 1973. Assembly of tobacco mosaic virus rods *in vitro*. Elongation of partially assembled rods. *Biochemistry*. **12**:4574–4581.
22. Butler, P. J. G. 1974. Structures and roles of the polymorphic forms of tobacco mosaic virus protein. IX. Initial stages of assembly of nucleoprotein rods from virus RNA and the protein disks. *J. Mol. Biol.* **82**:343–353.
23. Richards, K. E., and R. C. Williams. 1976. Assembly of tobacco mosaic virus *in vitro*. *In* Comprehensive Virology. H. Frankel-Conrat and R. R. Wagner, editors. Plenum Press, New York **6**:1–37.
24. Butler, P. J. G., and A. C. H. Durham. 1977. Tobacco mosaic virus protein aggregation and the virus assembly. *Adv. Protein Chem.* **31**:187–251.
25. Fukuda, M., T. Ohno, Y. Okada, Y. Otsuki, and I. Takebe. 1978. Kinetics of the biphasic reconstitution of tobacco mosaic virus *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.* **75**:1727–1730.
26. Butler, P. J. G., and G. P. Lomonossoff. 1978. Quantized incorporation of RNA during assembly of tobacco mosaic virus from protein disks. *J Mol. Biol.* **126**:877–882.
27. Lomonossoff, G. P., and P. J. G. Butler. 1979. Location and encapsidation of the coat protein cistron of tobacco mosaic virus. A bidirectional elongation of the nucleoprotein rod. *Eur. J. Biochem.* **93**:157–164.
28. Shire, S. J., J. J. Steckert, M. L. Adams, and T. M. Schuster. 1979. Kinetics and mechanism of tobacco mosaic virus assembly. Direct measurement of relative rates of incorporation of 4S and 20S protein. *Proc. Natl. Acad. Sci. U.S.A.* **76**:2745–2749.
29. Durham, A. C. H. 1972. Structures and roles of the polymorphic forms of tobacco mosaic virus protein. I. Sedimentation studies. *J. Mol. Biol.* **67**:289–305.
30. Ohno, T., H. Inoue and Y. Okada.1972. Assembly of rod-shaped virus *in vitro*: reconstitution with cucumber green mottle virus protein and tobacco mosaic virus RNA. *Proc. Natl. Acad. Sci. U.S.A.* **69**:3680–3683.
31. Perham, R. N. , and T. M. A. Wilson. 1976. The polarity of stripping of coat protein subunits from the RNA in tobacco mosaic virus under alkaline conditions. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* **62**: 11–15.
32. Wilson, T.M.A., R. N. Perham, and P. J. G. Butler. 1978. Intermediates in the disassembly of tobacco mosaic virus at alkaline pH. Infectivity, self-assembly, and translational activities. *Virology*. **89**:475–483.
33. Pelcher, L. E., and M. C. Halasa. 1979. Reassembly of particles using the RNA from partially disassembled tobacco mosaic virus. *Virology*. **97**:488–491.
34. Stubbs, G., S. Warren, and K. C. Holmes. 1977. Structure of RNA and RNA binding sites in tobacco mosaic virus from 4Å map calculated from X-ray fibre diagrams. *Nature (Lond.)*. **267**:216–221.
35. Graham, J., and P. J. G. Butler. 1979. Binding of oligonucleotides to the disk of tobacco mosaic virus protein. *Eur. J. Biochem.* **93**: 333–337.
36. Perham, R. N., and T. M. A. Wilson. 1978. The characterization of intermediates formed during the disassembly of tobacco mosaic virus at alkaline pH. *Virology*. **84**:293–302.
37. Guilley, H., G. Jonard, B. Kukla, and K. E. Richards. 1979. Sequence of 1,000 nucleotides at the 3′ end of tobacco mosaic virus RNA. *Nucleic Acids Res.* **6**:1287–1308.
38. Butler, P. J. G., and A. Klug. 1972. Assembly of tobacco mosaic virus *in vitro*: effect of state of polymerization of the protein component. *Proc. Natl. Acad. Sci. U.S.A.* **69**:2950–2953.
39. Butler, P. J. G. 1972. Structures and roles of the polymorphic forms of tobacco mosaic virus protein. VI. Assembly of the nucleoprotein rods of tobacco mosaic virus from the protein disks and RNA. *J. Mol. Biol.* **72**:25–35.
40. Butler, P. J. G., and J. T. Finch. 1973. Structures and roles of the polymorphic forms of tobacco mosaic virus protein. VII. Lengths of the growing rods during assembly into nucleoprotein with the viral RNA. *J. Mol. Biol.* **78**:637–649.
41. Butler, P. J. G. 1974. Structures and roles of the polymorphic forms of tobacco mosaic virus protein. VIII. Elongation of nucleoprotein rods of the virus RNA and protein. *J. Mol. Biol.* **82**: 333–341.
42. Butler, P. J. G., and A. Klug. 1978. The assembly of a virus. *Sci. Am.* **239**: No. 5:62–69.
43. Butler, P. J. G., A. C. H. Durham, and A. Klug. 1972. Structures and roles of the polymorphic forms of tobacco mosaic virus protein. IV. Control of mode of aggregation of tobacco mosaic virus protein by proton binding. *J.. Mol. Biol.* **72**:1–18.
44. Butler, P. J. G. 1976. Assembly of tobacco mosaic virus. *Philos. Trans.R. Soc. Lond. B. Biol. Sci.* **276**:151–163.

DISCUSSION

Session Chairman: Victor Bloomfield *Scribe*: Jill Taylor

J. KING: The curves of assembly vs protein concentration (Fig. 1 in the text) look like Michaelis-Menten curves. Asakura interpreted similar data with flagellin assembly to mean that the rate-limiting step in assembly was the