Tobacco Mosaic Virus Protein: Sedimentation Equilibrium Studies of the Initial Stages of Polymerization†

Charles J. Westover, Jr.,‡ and Charles L. Stevens*

ABSTRACT: The lowest stages of polymerization of tobacco mosaic virus protein were studied by means of high-speed sedimentation equilibrium experiments. Several distinct modes of polymerization were found. At pH 7.1 the expected monomer-trimer-higher polymer equilibrium was observed—very little dimer was detected at this pH. At pH 7.5, however, a strong dimerization was observed—neither monomer nor trimer was detected at this pH. An octamer appeared to be the only species present other than the dimer. When 0.01 M β-mercaptoethanol was added to the solvent at pH 7.5, the dimer was dissociated, resulting in a monomer-trimer association. The dimerization may be the basis for the larger "doubled" polymers formed by the protein at alkaline pH, while the octamer may correspond to the 8S peak frequently observed in sedimentation velocity experiments at alkaline pH. On the other hand, the monomer-trimer-higher polymer equilibrium may correspond to the single helix formed by the protein at slightly acid pH and to the combination of 4S and 20S peaks seen in sedimentation velocity experiments at slightly acid pH.

Tobacco mosaic virus (TMV)[1] has long been studied as a model for the process by which macromolecules interact to form larger functional units. As a result many details are known concerning both the finished structure of the virus and the intermediate stages of its assembly in vitro.

Previous studies, although both thorough and extensive, have been unable to examine directly the equilibria existing at the lowest stages of polymerization. Caspar (1963) suggested that the electrophoresis experiments of Sarkar (1960) indicated that a cyclic trimer (three subunits arranged at the corners of a triangle) would be the first stable intermediate. Further support for the trimer was provided by Banerjee and Lauffer (1966), who obtained number average molecular weights, by osmometry, which extrapolated at infinite dilution to the weight expected for a trimer. Caspar (1963) also suggested that an 8S peak which had been seen in sedimentation velocity experiments (Schramm and Zillig, 1955) might correspond to a cyclic heptamer (six subunits hexagonally close-packed around a central one).

The method of sedimentation equilibrium is ideally suited for studying the initial stages of polymerization of the protein. The experimental parameters are readily varied, enabling a broad range of molecular weights and concentrations to be studied—molecular weights can be obtained at concentrations as low as 0.1 mg/mL. An important advantage over the earlier methods is that, in principle, an unlimited number of average molecular weights can be determined, although in practice the accuracy of the data generally allows only the first few average weights to be calculated.

With these advantages in mind it is particularly interesting to investigate the state of the protein in the region around pH 7.3, for at this pH the protein specifically recognizes its own RNA and copolymerizes with it to form the virus rod (Fraenkel-Conrat and Singer, 1964).

Materials and Methods

A. *Preparation of Virus and Protein.* TMV was grown in Turkish tobacco plants. Purification was accomplished by repeated cycles of high and low speed centrifugation (Boedtker and Simmons, 1958), alternated with incubations at 37 °C in 0.1 M ethylenediaminetetraacetate (EDTA) (Ginoza et al., 1954). Virus concentrations were determined from absorbance at 260 nm (Englander and Epstein, 1957).

Protein was prepared from the virus by the acetic acid method of Fraenkel-Conrat (1957). It was then dialyzed against the solvent used in the equilibrium runs (0.001 M potassium phosphate buffer with either 0.1 M or 0.15 M potassium chloride added to provide a moderate ionic strength and a slight stabilizing density gradient). The protein was generally used within a few days of its preparation, although protein kept 4 months in the cold gave results identical with those of freshly prepared protein.

Protein concentrations were determined from absorbance at 280 nm (Fraenkel-Conrat and Williams, 1955). The purity of the protein (i.e., the absence of RNA) was determined from the ratio of absorbance at 280 nm to that at 250 nm, with a ratio of 2.5 being considered essentially free of RNA (Ansevin, 1958).

B. *Sedimentation Equilibrium Experiments.* Sedimentation equilibrium experiments were carried out according to the high-speed, meniscus-depletion method of Yphantis (1964). The conditions chosen were 33 450 rpm, 5-mm height of solution column (meniscus to base), and 0.1–0.4 mg/mL loading concentration. Equilibrium times calculated for these conditions were approximately 30 h (Yphantis, 1964), but, in practice, sedimentation was generally carried out for at least 40 h and equilibrium was considered to have been attained when the concentration distribution remained unchanged within experimental error over a period of 6 to 12 h.

A Beckman analytical ultracentrifuge (Model E), equipped with Rayleigh interference optics and a mechanical speed control, was used for the experiments. The optics were aligned by the procedures given in the Beckman manual and the camera lens was focused on the $\frac{2}{3}$ level of the cell (Yphantis, 1964). A symmetrical-slit aperture was used on the condensing lens. From a synthetic boundary cell experiment, it was found

† From the Department of Life Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260. *Received August 25, 1977.* This work was supported by Public Health Service Grants GM 10403 and 22558.

‡ Present address: Department of Human Genetics, University of Michigan, Ann Arbor, Michigan 48108.

[1] Abbreviations used are: TMV, tobacco mosiac virus; EDTA, ethylenediaminetetraacetate; M^{-2}, liters^2/moles^2.

that a protein concentration of 1 mg/mL corresponded to a Rayleigh fringe shift of 4.13 fringes (or 1160 μm).

For a typical run a cell was assembled and loaded in the cold, using sapphire windows and a 12-mm, filled-epon centerpiece. The solution was then sedimented to equilibrium in a precooled AN-D rotor. After the equilibrium photograph was taken, the rotor temperature was slowly increased several degrees while centrifugation continued. Photographs were taken until equilibrium was established at the new temperature. By repeating this procedure, temperature polymerization of the protein could be studied without requiring unduly long runs at each new temperature. For each experiment solvent blanks were run without disassembling the cell; the solution sector was rinsed thoroughly before running the solvent blank. These blanks were generally reproducible from run to run and any small differences between them were attributed to the process of disassembling and reassembling the cell.

The Rayleigh fringe patterns were recorded on Kodak 2G spectroscopic plates and measured on a Nikon 6C profile projector. The vertical displacements of five fringes were read and averaged at selected points along the horizontal (i.e., radial) axis. Fringe displacements could be determined to within 2 μm (i.e., 0.002 mg/mL) near the meniscus (where the fringes were flat) and to within 2% of the total displacement near the base of the cell (where the slope of the fringes became very steep). Optical distortions were corrected in all runs by subtracting the readings of the solvent blank from those of the solution.

C. *Analysis of the Data.* In order to obtain reliable molecular weights in experiments such as these, it is extremely helpful if the absolute level of concentration in the cell can be determined *experimentally* to within a few microns. This was no problem in the present experiments, because the long solution column (5 mm from meniscus to base) ensured that the fringe patterns from all the runs had an extensive flat region near the meniscus. This flat region was taken as the zero level of concentration and all displacements were measured relative to it. The zero level could generally be determined to within ±0.001 mg/mL. This degree of accuracy is important—it was found that (arbitrarily) introducing an error of only 0.01 mg/mL into the zero level resulted in errors of up to 20% in the number average weights.

Once the zero level was determined, values of the apparent number average (M_n) and weight average (M_w) molecular weights were calculated at each radial distance (r) in the cell according to the equations of Yphantis (1964):

$$M_n(r) = \left\{ c(r) \Big/ \int_a^r c(r)\mathrm{d}[r^2/2] \right\} \{RT/(1 - \bar{V}\rho)\omega^2\} \quad (1)$$

$$M_w(r) = \{\mathrm{d}[\ln c(r)]/\mathrm{d}[r^2/2]\} \{RT/(1 - \bar{V}\rho)\omega^2\} \quad (2)$$

Here $c(r)$ is concentration (in arbitrary units) at the point r; a is the value of r at the meniscus; and R, T, $\bar{V}$, ρ, and ω have their customary meanings of the gas constant, absolute temperature, partial specific volume, solution density, and angular velocity, respectively. The value of $\bar{V}$ for TMV protein was taken as 0.728 cm^3/g (Jaenicke and Lauffer, 1969), while solution densities were approximated by values for potassium chloride solutions taken from Weast and Selby (1966) and Timmermans (1960). Equation 1 holds if the concentration at the meniscus can be neglected. The integral in eq 1 was approximated by trapezoidal summation and the differential in eq 2 was approximated by the slope of a five-point least-squares line calculated about each point of a plot of ln $c(r)$ vs. $r^2/2$.

Two methods were used to analyze these average weights in terms of the individual molecular species present. The first is termed a "two-species" plot (Sophianopoulos and Van Holde, 1964; Roark and Yphantis, 1969; Teller et al., 1969). For a system containing only two species with molecular weights M_1 and M_2, the following equation is derived:

$$M_k(r) = M_1 + M_2 - M_1M_2(1/M_{k-1}(r)) \quad (3)$$

where $M_k = M_n, M_w, M_z, \ldots$ as $K = 0, 1, 2, \ldots$ respectively. Thus, if only two species are present, a plot of $M_k(r)$ vs. $1/M_{k-1}(r)$ yields a straight line with a slope of $-M_1M_2$ and an intercept of $M_1 + M_2$. This line passes through the points corresponding to M_1 and M_2 on the plot and the experimental data are contained on that segment of the line (the "tie-line") which lies between M_1 and M_2 (e.g., see Figure 3 or Figure 4b).

This analysis was supplemented by the following analysis, which takes advantage of the fact that at equilibrium each molecular species must be distributed exponentially throughout the cell (see also Haschemayer and Bowers, 1970). At each point in the cell the following relations hold:

$$\sum_i y_i = 1 \quad (4)$$

$$\sum_i (y_i/M_i) = 1/M_n \quad (5)$$

$$\sum_i (M_iy_i) = M_w \quad (6)$$

where y_i is the weight fraction of the ith species present at the point. If particular values are assigned to the M_i's (i.e., if particular species are assumed to be present), this system of equations can be solved at each point in the cell for as many y_i's as there are equations. The concentration of the ith species at any point r is then obtained by multiplying y_i by the total concentration at the point, $c(r)$. Since the average molecular weights calculated at any point in the cell are (mathematically) independent of those calculated at any other point, there is no reason a priori to expect any particular correlation to hold between $c_i(r)$ and r. But, if the "correct" choice of species has been made, a plot of ln $c_i(r)$ vs. $r^2/2$ for each species will be linear and the slope of the plot will match the molecular weight of the chosen species. These slopes were generally within ±10% of the expected values. Such plots provide an additional test for the "correct" analysis of the data (e.g., see Figure 2). Both methods of analysis were greatly facilitated by the fact that the molecular weight of the monomer—17 530—is accurately known from the amino-acid composition (Anderer, 1963).

Equilibrium constants and free energies were obtained from these plots in the following manner. At any given point in the cell the concentrations of the jth and kth species can be taken from their respective lines and inserted in the equation

$$\Delta F^\circ = -RT \ln K = -RT[\ln(m_j) - (j/k) \ln m_k] \quad (7)$$

where m_j and m_k are the molar concentrations of the jth and kth species (here the indices j and k have a particular meaning in that they represent the number of monomer units contained in the given polymer), and K, R, T, and ΔF° are the equilibrium constant, the gas constant, the absolute temperature, and the standard free energy, respectively. In practice the concentrations were generally taken from the central portions of the lines, which correspond to values of $c(r)$ of approximately 0.3–0.6 mg/mL. This was done because the greatest weight had been given to these points in drawing the lines—points at lower concentrations often showed large deviations due to random errors of measurement, while points at higher concentrations sometimes showed systematic deviations dependent

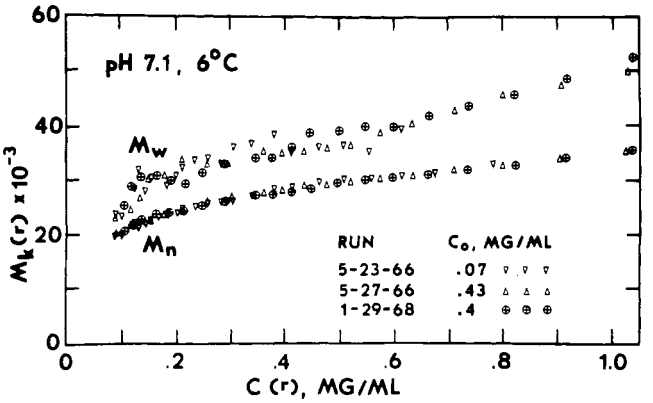

FIGURE 1: $M_w(r)$ and $M_n(r)$ for three runs at pH 7.1, 6 °C. c_0 is initial loading concentration. Other conditions of centrifugation are as indicated in Materials and Methods. Data shown in this and succeeding figures have not been smoothed in any way.

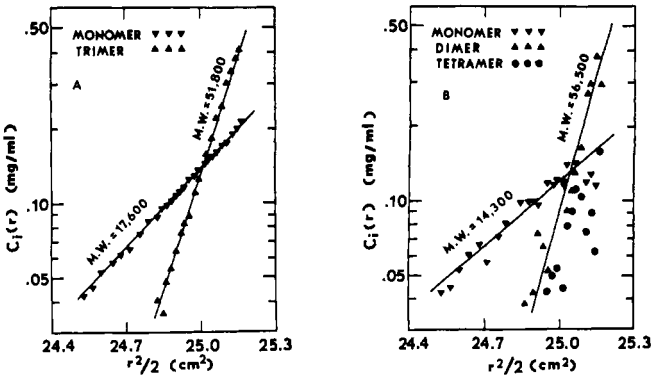

FIGURE 2: Supplemental analysis of a monomer-trimer run. (A) Presents the data of run 5-23-66 (data also shown in Figure 1) when analyzed according to eq 4-6 with monomer, trimer, and hexadecamer as selected species. Points for hexadecamer are not shown since the concentrations calculated for hexadecamer were less than 0.01 mg/mL. (B) Presents the same data as A, also analyzed according to eq 4-6, but here using monomer, dimer, and tetramer as the selected species. Molecular weights shown in A and B were calculated from the slopes of the fitted lines.

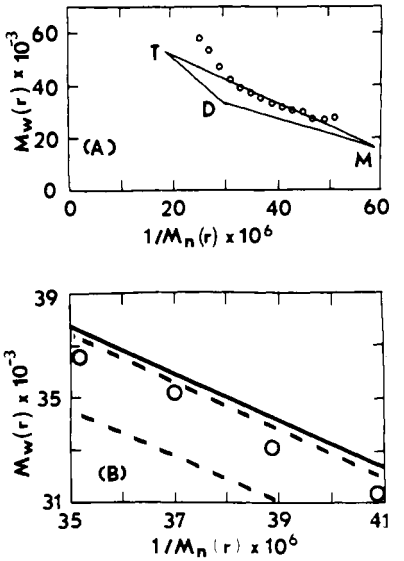

FIGURE 3: Two-species plot (eq 3) of monomer-trimer runs. (A) Presents the averaged data from a dozen monomer-trimer runs when analyzed according to eq 3. For this plot the data from all 12 runs at values of $c(r)$ greater than 0.1 mg/mL were grouped in intervals according to their value of $1/M_n(r)$. Each circle in A represents the mean value of $M_w(r)$ for a given group plotted against the mean value of $1/M_n(r)$ for that group. The points M, D, and T are respectively the points at which the monomer, dimer, and trimer of TMV protein would occur in this plot. (B) Presents an enlarged view of the central portion of A. The solid line is the monomer-trimer line of A, while the upper and lower dashed lines present respectively a monomer-dimer-cyclic trimer equilibrium (Caspar, 1963) and a monomer-dimer-linear trimer equilibrium (Flory, 1936).

on the choice of the third species. Thus, whenever the slope of the fitted line did not exactly match the molecular weight of the chosen species, these central points were the ones which were least in error.

Results

A. *Results at pH 7.1.* Figure 1 presents the results of three runs at pH 7.1, 6 °C, on two samples of protein prepared 2 years apart. This figure illustrates three points. The first is that the data are highly reproducible from one run to another and from one protein preparation to another. The second is that the age of the protein sample had no effect on these results—the protein used for runs 5-23-66 and 5-27-66 had been prepared 4 months prior to the runs, whereas the protein used for run 1-29-68 was less than 2 days old. The third is that a true chemical equilibrium exists, since two of the runs were made with the same protein preparation, but with a sixfold difference in loading concentration (runs 5-23-66 and 5-27-66 at 0.07 and 0.43 mg/mL, respectively). If the solution had contained species which were not in equilibrium with each other, then the data from these two runs should not have coincided when plotted as in Figure 1 (Yphantis, 1964).

Both methods of analysis indicate that these data can be interpreted as a monomer-trimer-higher polymer association. Figure 2 presents the data of run 5-23-66 (the data shown in Figure 1) when treated according to eq 4-6. For the plot of Figure 2A it was assumed that only monomer, trimer, and hexadecamer were present. It can be seen that the concentrations calculated for monomer and trimer fall on straight lines whose slopes yield the correct molecular weights for the chosen species. The concentration of hexadecamer was essentially zero for total protein concentrations less than 0.4 mg/mL and increased only slightly at higher concentrations; i.e., only monomer and trimer were needed to fit the data at low concentrations. Under the conditions of these experiments, it was not possible to define the third species any more exactly than as that species which resulted in the straightest lines for the monomer and trimer species. The need for a higher polymer in fitting the data may actually reflect the presence of a mixture of higher polymers.

Figure 2B presents for comparison a "poor" choice of species for the data of Figure 2A. Here it was assumed that monomer, dimer, and tetramer were present. The greater scatter in the calculated points and the incorrect slopes for the chosen species were typical of such "poor" choices.

Figure 3A presents a two-species plot (eq 3) of the averaged data from a dozen runs classed as monomer-trimer runs. For values of $1/M_n(r)$ greater than 32×10^{-6} (corresponding to protein concentrations less than approximately 0.4 mg/mL), the points fall very close to the line expected for a monomer-trimer association, while for values of $1/M_n(r)$ less than this, they fall above the line. This elevation of the points above the monomer-trimer line is consistent with the presence of species larger than trimer at concentrations greater than 0.4 mg/mL. This is the same concentration above which the analysis of Figure 2A requires that species larger than trimer be taken into account. Thus the two analyses both indicate that monomer and trimer are the predominant species at low concentration, while larger species need to be considered at higher concentrations.

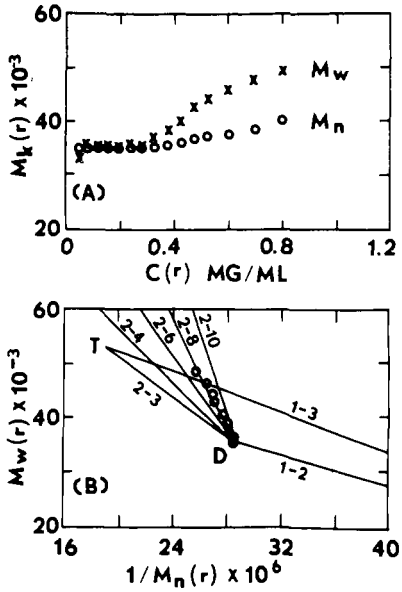

FIGURE 4: Averaged data of runs at pH 7.5, 6 °C. Values of $M_n(r)$ and $M_w(r)$ from five runs at pH 7.5, 6 °C, were grouped in intervals according to the associated value of $c(r)$. In A, the mean values of $M_n(r)$ and $M_w(r)$ for each group are plotted vs. the mean value of $c(r)$ for the group. (B) Presents a two-species plot (eq 3) of the data of A. Lines marked 2-4, 2-6, 2-8, and 2-10 are the lines expected for two-species systems of dimer-tetramer, dimer-hexamer, dimer-octamer, and dimer-decamer, respectively. The points D and T correspond to dimer and trimer, respectively. The points for monomer, tetramer, hexamer, octamer, and decamer lie beyond the limits of the figure.

Figure 3B presents an enlarged view of the central portion of Figure 3A. It is apparent that the data lie slightly below the monomer-trimer line rather than lying exactly on it. It was found that this effect could be simulated by assuming that the solutions contained a small amount of dimer (approximately 0.1 weight fraction of the protein) in addition to monomer and trimer. Although this amount of dimer is too small to be detected with certainty in the data of a single run, it is detectable in the averaged data of these dozen runs.

B. *Results at pH 7.5.* The small change in pH from 7.1 to 7.5 produced a striking change in the polymerization of the protein. Whereas a monomer-trimer equilibrium was observed at pH 7.1 with only a very small amount of dimer being detected there, at pH 7.5 the protein was almost entirely in the form of dimer with neither monomer nor trimer being detected here.

Figure 4A presents the averaged data from five equilibrium runs at pH 7.5, 6 °C. All runs were made within a few days of preparation and the protein was never allowed to warm up above 6 °C. Four different protein preparations were used for these runs.

In the concentration range below 0.35 mg/mL the values of $M_n(r)$ and $M_w(r)$ are identical and constant, indicating that in this concentration range essentially only one molecular species is present. In this range there are 70 values of $M_n(r)$ and $M_w(r)$ available in the ungrouped data of the five runs. These 70 values were averaged, yielding a mean M_n of 35 020 ± 100 and a mean M_w of 35 060 ± 200. These molecular weights are in excellent agreement with the value of 35 060 expected for a dimer of TMV protein. There does not appear to be any tendency for the molecular weights to drop below that of dimer even at concentrations as low as 0.05 mg/mL. Thus monomer is undetectable at pH 7.5.

When 0.01 M β-mercaptoethanol was added to the solvent at pH 7.5, however, the dimer was dissociated, resulting in a monomer-trimer association. A similar result was obtained when 0.01 M EDTA was added along with the 0.01 M β-mercaptoethanol. In each case a protein preparation which had originally resembled Figure 4 now resembled Figure 3. The only difference between the two cases lies in the equilibrium constant of the resulting monomer-trimer equilibrium; this was 12.1×10^9 M^{-2} in the first case and 4.93×10^9 M^{-2} in the second case. This latter value agrees well with the equilibrium constants found for the monomer-trimer equilibrium at pH 7.1, 6 °C; e.g., for the three runs presented in Figure 1, the equilibrium constants were 5.11, 5.11, and 4.55×10^9 M^{-2}.

The increasing molecular weights seen in Figure 4A at concentrations greater than 0.35 mg/mL indicate that a polymer larger than dimer must be present at these higher concentrations. A two-species plot (eq 3) was made of the data in this concentration range and the results are shown in Figure 4B. The data fall along the dimer-octamer line. As confirmation of this, a least-squares line was calculated for the ungrouped data from the five runs of Figure 4. 112 points were available for this calculation; 70 of these points superimpose on the dimer point (D) in Figure 4B. Physically this corresponds to the fact that the dimer is by far the major species present in the cell. Mathematically it has the effect of forcing the calculated line to pass through the dimer point. The remaining 42 points are then the ones primarily responsible for the direction of the line. The equation calculated for this line is $M_w(r) = -4.85 \times 10^9 \times 1/M_n(r) + 1.74 \times 10^5$ with a standard error of estimate of ±600 (the size of the circles in Figure 4B approximates this error). Although the data cover only a very short segment of the line that extends from dimer to octamer, the lengthy extrapolation of this regression line to the value of $1/M_n(r)$ which corresponds to the octamer (i.e., to $1/M_n(r) = 7.13 \times 10^{-6}$) yields a value of $M_w(r) = 139\,400$. This is within 1% of the molecular weight of an octamer of TMV protein (i.e., 140 240). Thus at pH 7.5 it appears that there is no significant amount of any species present other than dimer and octamer.

This dimerization occurred repeatedly at pH 7.5, even though the protein was always kept cold and was always used within a few days of its preparation. In contrast, one protein preparation, which had been stored for several months in the cold in 0.033 M phosphate buffer, pH 7.2, displayed a monomer-trimer equilibrium with almost no dimer present when run at pH 7.1 (runs 5-23-66 and 5-27-66, Figures 1-3). Dimerization was seen in every run at pH 7.5, involving a total of four separate protein preparations, but was never seen at pH 7.1 (other than the small amount of dimer indicated by Figure 3B) in runs involving three separate protein preparations. Furthermore, a single protein preparation could display either the monomer-trimer or the dimer-octamer mode of association, depending on whether it was dialyzed into pH 7.1 buffer (run 1-29-68, Figure 1) or pH 7.5 buffer (run 2-6-68, one of the dimer runs incorporated in Figure 4).

Discussion

The present experiments, in which molecular weights were obtained at concentrations as low as 0.1 mg/mL, complement and extend studies carried out by other techniques at higher concentrations and higher degrees of polymerization. The results range from the expected monomer-trimer equilibrium to the unexpected dimer-octamer association. Numerous parallels may be drawn, however, between the observations reported here and those seen in other studies.

A. *The Monomer-Trimer Equilibrium at pH 7.1.* The monomer-trimer equilibrium observed at pH 7.1 agrees with the prediction of Caspar (1963) that a cyclic trimer would be

the first stable intermediate in the polymerization of the protein. Although sedimentation equilibrium experiments cannot distinguish a cyclic from a linear trimer, the conclusion that the trimer observed here is cyclic rather than linear may be inferred from the observation that very little dimer is present at pH 7.1 (Figure 3B), much less than would be expected if the trimer were formed by a linear condensation polymerization (Flory, 1936; Caspar, 1963).

B. *The Dimerization at pH 7.5.* The sharp transition from a monomer–trimer association at pH 7.1 to a dimer–octamer association at pH 7.5 was unexpected. This transition has several parallels, however. One is found in the work of Finch et al. (1966), Carpenter (1970), and Durham et al. (1972), which showed that the protein forms "doubled" structures—i.e., crystals of double disks, double helix, and "figure-8" polymers—at slightly alkaline pH, in contrast to the single helix which it forms at slightly acid pH. The other is found in the frequent observation of an 8S peak in sedimentation velocity experiments at alkaline pH, in contrast to the combination of 4S and 20S peaks generally seen at slightly acid pH (Stauffer et al., 1970; Carpenter, 1970; Durham, 1972; Lonchampt et al., 1972; Rodionova et al., 1973; Vogel and Jaenicke, 1974). Of particular interest for the present studies, Durham (1972) notes that the transition between the two sedimentation velocity patterns occurs at pH 7.2.

Although a structural relation is easily imagined between the dimer and the larger "doubled" polymers (i.e., that the dimer forms the basic unit for these polymers), a structural relation with the 8S peak is perhaps not so readily perceived. It should be noted, therefore, that the 8S sedimentation rate approximates the rate that might be expected for an octamer of TMV protein. Caspar (1963) originally proposed that the 8S peak might be due to a heptamer—six subunits hexagonally close-packed around a central one. His calculated sedimentation rates depend, however, on the shape and the degree of hydration of the polymer. Thus, the 8S peak could as easily be due to an octamer as a heptamer. Furthermore it seems unlikely that an odd-numbered polymer such as a heptamer would be present in our solutions at pH 7.5, because at this pH practically all the protein was in the form of dimer and there was no detectable monomer.

What forces are involved in the dimerization? The pH dependence of the transition would suggest that the titration of some group on the protein is responsible for dimerization. Yet the titration experiments of Scheele and Lauffer (1967) show that under the conditions of these experiments (6 °C, 0.1 ionic strength) there are no groups on the protein which titrate in the pH range 7.1 to 7.5. The ability of 0.01 M β-mercaptoethanol to dissociate the dimer would suggest that a disulfide bond may be involved in dimerization. Yet the interpretation of x-ray diffraction experiments on crystals of double disks places the single sulfhydryl of the protein in the middle of the molecule, about 13 Å distant from both its upper and lower surfaces (Gilbert and Klug, 1974; Champness et al., 1976). Because these are the surfaces where the "doubling" interactions occur in the doubled polymers, it seems that a disulfide bond could not be involved in their formation.

Although the forces involved in dimerization remain uncertain, its discovery nevertheless sheds new light on the results of other experiments on TMV protein. For instance, Lonchampt et al. (1972) observed that the 8S sedimentation velocity pattern (obtained in sodium pyrophosphate buffer at pH 7.25, 0.1 ionic strength) could be converted to either a 4S or a 25S pattern merely by dialyzing against lower or higher ionic strength, respectively. They interpreted this result as indicating that the 8S pattern may represent an intermediate state in the formation of the 25S pattern from the 4S pattern. We suggest, however, that this result may represent, instead, a change in the free energy of dimerization and that the 8S and 25S patterns may thus reflect *alternate* pathways of polymerization for the monomer (rather than a single, *sequential* pathway), depending on whether dimers or trimers, respectively, are the predominant product of the first step of polymerization. We suggest the control point for these two pathways lies in the dimerization rather than the trimerization, because (as noted in Results, section B) trimer forms just as readily at pH 7.5 (in the presence of β-mercaptoethanol) as it does at pH 7.1. The inability to detect trimer at pH 7.5 is probably due to the low concentration of monomer at that pH, vitually all the monomers having been incorporated into dimers as a result of the increased free energy of dimerization.

This view (that there are alternate paths of polymerization for the monomer) is in accord with the results of Sperling and Klug (1975). In studying the protein of the Dahlemense strain of TMV they found that, as ionic strength was raised at pH 8, an 8S sedimentation pattern was transformed first into a 4S pattern which then gave way to a 20S pattern at higher ionic strength. They also observed a 30S pattern which was correlated with the occurrence of "figure-8" polymers (so-called because of their appearance in electron micrographs). These polymers seem to be a dimer of double disks, i.e., two double disks interlocked approximately side-by-side, an arrangement similar to the interaction which occurs in crystals of double disks (Durham et al., 1972). "Figure-8" polymers are also formed by the protein of the common strain of TMV (Durham et al., 1972). Sperling and Klug (1975) proposed that the 30S and 20S patterns represented alternate paths of polymerization, because the 20S pattern seemed to lead to the formation of helical polymers, while the 30S pattern did not.

Thus it may be that the dimer–octamer association, the 8S sedimentation pattern, the "figure-8" (30S) polymer, and the crystals of double disks are steps along a path which does not lead to the formation of the single helix found in the virus, while the monomer–trimer–higher polymer association and the 4S and 20S patterns are steps along a path which does lead to the formation of single helix.

Finally, it is worth reemphasizing that the transition from the monomer–trimer association at pH 7.1 to the dimer–octamer association at pH 7.5 neatly brackets the pH optimum of 7.3 for reconstitution of the protein with the viral RNA, thereby suggesting that the transition may have biological relevance.

References

Anderer, F. A. (1963), *Adv. Protein Chem. 18,* 1.

Ansevin, A. T. (1958), M.S. Thesis, University of Pittsburgh, Pittsburgh, Pennsylvania.

Banerjee, K., and Lauffer, M. A. (1966), *Biochemistry 5,* 1957.

Boedtker, H., and Simmons, N. S. (1958), *J. Am. Chem. Soc. 80,* 2250.

Carpenter, J. M. (1970), *Virology 41,* 603.

Caspar, D. L. D. (1963), *Adv. Protein Chem. 18,* 37.

Champness, J. N., Bloomer, A. C., Bricogne, C., Butler, P. J. G., and Klug, A. (1976), *Nature (London) 259,* 20.

Durham, A. C. H. (1972), *J. Mol. Biol. 67,* 289.

Durham, A. C. H., Finch, J. T., and Klug, A. (1972), *J. Mol. Biol. 72,* 1.

Englander, S. W., and Epstein, H. T. (1957), *Arch. Biochem. Biophys. 68,* 144.

Finch, J. T., Leberman, R., Yu-Shang, C., and Klug, A. (1966), *Nature (London) 212,* 349.

Flory, P. D. (1936), *J. Am. Chem. Soc. 58*, 1877.
Fraenkel-Conrat, H. (1957), *Virology 4*, 1.
Fraenkel-Conrat, H., and Singer, B. (1964), *Virology 23*, 354.
Fraenkel-Conrat, H., and Williams, R. C. (1955), *Proc. Natl. Acad. Sci. U.S.A. 41*, 690.
Gilbert, P. F. C., and Klug, A. (1974), *J. Mol. Biol. 86*, 193.
Ginoza, W., Atkinson, D. E., and Wildman, S. G. (1954), *Science 119*, 269.
Haschemayer, R., and Bowers, W. (1970), *Biochemistry 9*, 435.
Jaenicke R., and Lauffer, M. A. (1969), *Biochemistry 8*, 3083.
Lonchampt, M., Lebeurier, G., and Hirth, L. (1972), *FEBS Lett. 22*, 297.
Roark, D. D., and Yphantis, D. A. (1969), *Ann. N.Y. Acad. Sci. 164*, 245.
Rodionova, N. P., Vesenina, N. E., Atabekova, T. I., Dzavakhia, V. G., and Atabekov, J. G. (1973), *Virology 51*, 24.
Sarkar, S. (1960), *Z. Naturforsch. B 15*, 778.
Scheele, R. B., and Lauffer, M. A. (1967), *Biochemistry 6*, 3076.
Schramm, G., and Zillig, W. (1955), *Z. Naturforsch. B 10*, 493.
Sophianopoulos, A. J., and Van Holde, K. E. (1964), *J. Biol. Chem. 239*, 2516.
Sperling, R., and Klug, A. (1975), *J. Mol. Biol. 96*, 425.
Stauffer, H., Srinivasan, S., and Lauffer, M. A. (1970), *Biochemistry 9*, 193.
Teller, D. C., Horbett, T. A., Richards, E. G., and Schachman, H. K. (1969), *Ann. N.Y. Acad. Sci. 164*, 66.
Timmermans, J. (1960), Physico-Chemical Constants of Binary Systems, Vol. III, New York, N.Y., Interscience.
Vogel, D., and Jaenicke, R. (1974), *Eur. J. Biochem. 41*, 607.
Weast, R. C., and Selby, S. M. (1966), in Handbook of Chemistry and Physics, 47th ed, Cleveland, Ohio, Chemical Rubber Co., p D 139.
Yphantis, D. A. (1964), *Biochemistry 3*, 297.

Anticoagulant Properties of Bovine Plasma Protein C following Activation by Thrombin†

Walter Kisiel,*,‡ William M. Canfield,§ Lowell H. Ericsson, and Earl W. Davie

ABSTRACT: Protein C is a vitamin K dependent glycoprotein which exists in bovine plasma as a precursor of a serine enzyme. Incubation of this plasma protein with α-thrombin at an enzyme-to-substrate weight ratio of 1:50 resulted in the cleavage of an Arg-Ile bond between residues 14 and 15 of the heavy chain of the molecule and the formation of activated protein C. The heavy chain of this serine enzyme contains the active-site sequence of -Leu-Cys-Ala-Gly-Ile-Leu-Gly-Asp-Pro-Arg-Asp-Ala-Cys-Gln-Gly-Asp-*SER*-Gly which is homologous with the corresponding regions of a number of plasma serine proteases. Activated protein C markedly prolongs the kaolin-cephalin clotting time of bovine plasma but not that of human plasma. The anticoagulant effect was totally obviated by prior incubation of the enzyme with diisopropyl phosphorofluoridate or phenylmethanesulfonyl fluoride. Incubation of activated protein C with bovine factor V resulted in a time- and temperature-dependent inactivation of this clotting factor, and this reaction was dependent on the presence of phospholipid and calcium ions. Activated protein C had no effect on the coagulant activity of factor XII, factor XI, factor X, factor IX, factor VII, or prothrombin. These data provide evidence for a mechanism for the activation of protein C in plasma and a potential role for this enzyme in blood coagulation and hemostasis.

Protein C is a vitamin K dependent protein present in bovine plasma (Stenflo, 1976). It is a glycoprotein composed of a heavy chain (mol wt 41 000) and a light chain (mol wt 21 000) held together by a disulfide bond(s). The light chain of protein C is homologous in its amino-terminal sequence with that of the four vitamin K dependent coagulation factors. The latter four proteins exist in plasma as zymogens and are converted to serine proteases by limited proteolysis (Davie and Fujikawa, 1975).

Present evidence indicates that protein C, like the four vitamin K dependent coagulation factors, exists in plasma as a precursor to a serine enzyme (Kisiel et al., 1976a; Esmon et al., 1976). Furthermore, protein C is converted to a serine amidase by a protease from Russell's viper venom or by trypsin. In this reaction, a tetradecapeptide is cleaved from the amino terminus of the heavy chain of the precursor protein with the concomitant formation of a diisopropyl phosphorofluoridate (iPr_2PF)[1] sensitive serine enzyme. This cleavage occurred at a specific Arg-Ile bond in the heavy chain of the molecule and resulted in the formation of a new amino-terminal sequence of Ile-Val-Asp-Gly-. Activated protein C did not exhibit any procoagulant activity. On the contrary, a significant inhibition of the partial thromboplastin time of bovine plasma was observed in the presence of activated protein C.

The goal of the present investigation was to determine

† From the Department of Biochemistry, University of Washington, Seattle, Washington 98195. *Received June 27, 1977*. This work was supported in part by Research Grants HL 16919 and GM 15731 from the National Institutes of Health.

‡ An Established Investigator of the American Heart Association.

§ Predoctoral Fellow, Medical Scientist Training Program, National Institutes of General Medical Science (GM 07266).

[1] Abbreviations used are: iPr_2PF, diisopropyl phosphorofluoridate; $PhCH_2SO_2F$, phenylmethanesulfonyl fluoride; EDTA, ethylenediaminetetraacetic acid.