Preparation and Properties of Recombinant DNA Derived Tobacco Mosaic Virus Coat Protein†

Steven J. Shire,*,‡ Patrick McKay,§ David W. Leung,§ George J. Cachianes,§ Eugene Jackson,‡ and William I. Wood§

Department of Pharmaceutical Research and Development and Department of Animal Health Research, Genentech, Inc., South San Francisco, California 94080

Krishnamurthy Raghavendra, Lamia Khairallah, and Todd M. Schuster*

Department of Molecular and Cell Biology, University of Connecticut, Storrs, Connecticut 06269-3125

Received October 27, 1989; Revised Manuscript Received February 13, 1990

ABSTRACT: Recombinant DNA derived tobacco mosaic virus (vulgare strain) coat protein (r-TMVP) was obtained by cloning and expression in *Escherichia coli* and was purified by column chromatography, self-assembly polymerization, and precipitation. SDS–PAGE, amino terminal sequencing, and immunoblotting with polyclonal antibodies raised against TMVP confirmed the identity and purity of the recombinant protein. Isoelectric focusing in 8 M urea and fast atom bombardment mass spectrometry demonstrated that the r-TMVP is not acetylated at the amino terminus, unlike the wild-type protein isolated from the tobacco plant derived virus. The characterization of r-TMVP with regard to its self-assembly properties revealed reversible endothermic polymerization as studied by analytical ultracentrifugation, circular dichroism, and electron microscopy. However, the details of the assembly process differed from those of the wild-type protein. At neutral pH, low ionic strength, and 20 °C, TMVP forms a 20S two-turn helical rod that acts as a nucleus for further assembly with RNA and additional TMVP to form TMV. Under more acidic conditions, this 20S structure also acts as a nucleus for protein self-assembly to form viruslike RNA-free rods. The r-TMVP that is not acetylated carries an extra positive charge at the amino terminus and does not appear to form the 20S nucleus. Instead, it forms a 28S four-layer structure, which resembles in size and structure the dimer of the bilayer disk formed by the wild-type protein at pH 8.0, high ionic strength, and 20 °C. However, like TMVP at pH 5–6, the recombinant coat protein forms long helical rods which, depending on the pathway of formation, may exhibit packing imperfections. In contrast to TMVP, solutions of r-TMVP helical rods also contain some unpolymerized disklike structures, which suggests either two structures for the 28S aggregates, two parallel polymerization pathways, or both. The influence on the quaternary structure due to the additional charge at the amino terminus of r-TMVP is significant even though this end of the protein chain is on the outside of the protein in the disk, the assembled rod, and the virus.

Self-assembly of tobacco mosaic virus (TMV)[1] from its constituent coat protein (TMVP) and RNA is a model reaction for the study of helical macromolecular assembly and protein–RNA interactions. Although the RNA stabilizes the completely formed virion, the coat protein contains the essential information for assembly into helical structures and can be made to self-assemble into viruslike rods without the RNA (Schramm, 1947; Franklin, 1956; Mandelkow et al., 1981). Physical–chemical studies of the TMVP assembly process have elucidated the effects of solution conditions such as pH, ionic strength, and temperature [for reviews, see Caspar (1963) and Butler and Durham (1977)]. Kinetic studies of TMVP self-assembly have furnished additional information about the mechanism of the process, which is now known to be a nucleation-controlled reaction (Scheele & Schuster, 1974; Schuster et al., 1979; Shire et al., 1979b; Potschka et al., 1988). This information together with the results of X-ray studies of TMVP crystals (Bloomer et al., 1978) and oriented gels of TMV (Stubbs et al., 1977; Namba & Stubbs, 1986; Namba et al., 1989) and in vitro virus reconstitution studies [Shire et al. (1979a, 1981) and Raghavendra et al. (1988); for reviews, see Hirth and Richards (1981); Butler (1984), Stubbs (1984), Bloomer and Butler (1986) and Okada (1986)] have furnished details of the TMV assembly process. However, despite this wealth of information, very little detailed information is available regarding the influence, if any, of specific amino acid residues on the assembly process, especially with respect to the protein–protein or protein–RNA recognition events.

Recombinant DNA technology has made it possible to investigate a variety of structure–function relationships by site-specific mutagenesis (Oxender & Fox, 1987). For example, such studies include the results of using protein engineering to add disulfide bonds (Perry & Wetzel, 1984; Villifranca et al., 1983; Matsumura & Matthews, 1989), modifications to the active sites of enzymes (Wells et al., 1986),

† This research was partially supported by a grant from the National Institutes of Health to T.M.S. (AI 11573) and by the University of Connecticut Analytical Ultracentrifugation Facility, which is supported by a grant from the Division of Instrumentation Resources of the National Science Foundation.

* Author to whom correspondence should be addressed.

‡ Dept. of Pharmaceutical Research and Development.

§ Dept. of Animal Health Research.

[1] Abbreviations: TMV, tobacco mosaic virus (vulgare strain) grown in tobacco plants; TMVP, tobacco mosaic virus coat protein derived from TMV (vulgare strain); TMV RNA, ribonucleic acid of TMV (vulgare strain); r-TMVP, desacetyl amino terminal TMVP (vulgare) derived by recombinant DNA methods, using RNA from the vulgare strain of TMV, and expressed in *Escherichia coli*; Tris-HCl, tris(hydroxymethyl)aminomethane hydrochloride; *I*, ionic strength expressed in molar concentration units; EM, electron microscopy; CD, circular dichroism; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; IEF, isoelectric focusing.

0006-2960/90/0429-5119$02.50/0

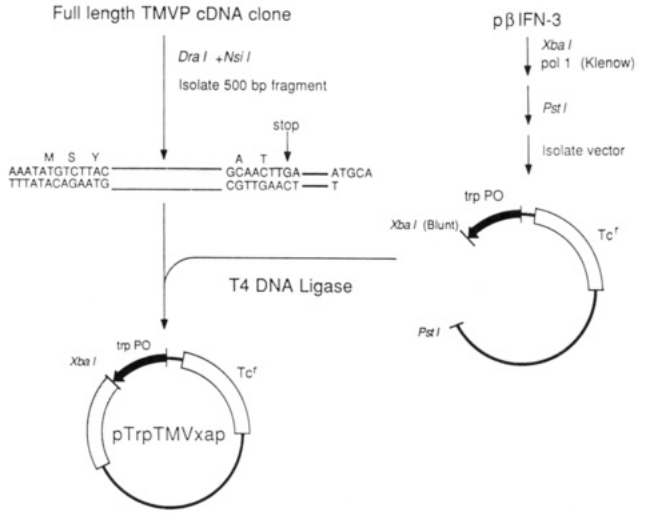

FIGURE 1: Construction of a plasmid coding for the direct expression of tobacco mosaic virus protein (TMVP) in *E. coli*. The full length TMVP cDNA clone was digested with *Dra*I and *Nsi*I restriction enzymes to yield a 500-bp fragment containing the entire TMVP coding region. This fragment was then inserted between the *Xba*I and *Pst*I sites of a Trp expression plasmid, pβIFN3, and the *Xba*I site in the vector was converted to flush ends by DNA polymerase I large (Klenow) fragment. The resultant plasmid, pTrpTMVxap, isolated from *E. coli* strain MM294, was then transformed into *E. coli* strain W3110 for the expression of TMVP.

and effects of amino acid mutations on the reversible dissociation of a dimeric enzyme (Jones et al., 1985). We report here the successful cloning, purification, and characterization of r-TMVP (vulgare strain). Reports of the cloning and expression of TMVP in *Escherichia coli* have been published previously but no detailed characterization of the protein or its assembly has been reported (Goelet & Karn, 1984; Dawson et al., 1986; Meshi et al., 1986; Haynes et al., 1986). We have found, as expected, that the recombinant protein (r-TMVP) expressed in *E. coli* differs from the wild-type protein (TMVP) isolated from the virus grown in tobacco plants in that it lacks an acetyl group on the amino terminus of the polypeptide chain. This subtle modification has allowed us to investigate the effects of charge alteration at the amino terminus on the assembly properties of the coat protein. The characterization of the assembly properties of this protein presented here and in preliminary reports elsewhere (Shire et al., 1987; Raghavendra et al., 1989) suggests an important role for amino terminal acetylation in TMVP assembly.

Materials and Methods

Preparation of TMV, TMVP, and TMV RNA. The original inoculum of TMV (vulgare strain) was obtained from Dr. Knight, University of California, and was used to obtain field-grown TMV on *Nicotiana tabacum* cv. Turkish Samsun at the Connecticut Tobacco Research Station, Windsor Locks, CT. Infected leaves were stored at −20 °C until TMV preparation. Methods of inoculation, growth, and purification of TMV were followed according to previously published procedures (Shire et al., 1979a; Matthews, 1981; Steckert, 1982). TMVP and TMV RNA preparations, analytical methods, and spectroscopic properties have been detailed elsewhere (Shire et al., 1979a; Schuster et al., 1980; Steckert, 1982; Raghavendra et al., 1985a).

Cloning and Expression of TMVP. A cDNA library in vector λgt10 (Huynh et al., 1985) was prepared from purified TMV RNA by priming synthesis with a synthetic oligonucleotide (5′-TGGGCCCCTACCGGG-3′), which hybridizes to the 3′-end of TMV RNA (Goelet et al., 1982). Clones that

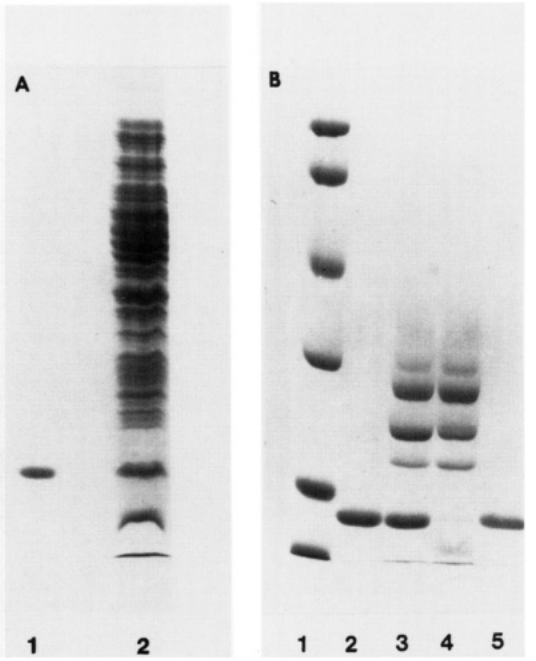

FIGURE 2: A 12.5% PAGE of r-TMVP with 2% SDS in sample preparation buffer and 0.5% in resolving buffer (all samples contained 0.1% 2-mercaptoethanol). (A) Expression of r-TMVP in *E. coli*: lane 1, TMVP standard (17.5 kDa) isolated from the vulgare strain of tobacco mosaic virus; lane 2, supernatant of *E. coli* cell lysate. (B) Selective precipitation of r-TMVP after DE-52 chromatography: lane 1, molecular weight standards; lane 2, TMVP standard (17.5 kDa) isolated from the vulgare strain of tobacco mosaic virus; lane 3, Tris eluant from DE-52 chromatography adjusted to pH 3.5; lane 4, supernatant after centrifugation of sample in lane 3; lane 5, pellet after centrifugation of sample in lane 3. Molecular mass standards (kilodaltons) are lysozyme (14.4), soybean trypsin inhibitor (21.5), carbonic anhydrase (31), ovalbumin (42.7), bovine serum albumin (66.2), and phosphorylase B (97.4).

span the entire TMVP gene were identified by synthetic DNA probes that hybridize to regions flanking the 5′-side (5′-TAGCCTCCGAATCATCATCG-3′) and the 3′-side (5′GCCTGCGGATGTATAT-3′) of the TMVP coding region based on the published sequence of TMV RNA (Goelet et al., 1982). DNA sequence analysis of a full-length TMVP cDNA clone is in complete agreement with the published cDNA and hence TMV RNA sequence (Goelet et al., 1982). Figure 1 summarizes the construction of a plasmid under control of Trp promoter, pTrpTMVxap, that was used to direct the expression of TMVP in *E. coli* (Leung et al., 1984).

Cell Growth and Harvest. E. coli K12 strain W3110 was transformed with plasmid pTrpTMVxap and the TMV coat protein was produced in a 10-L fermenter. The fermentation basal medium was supplemented with casein hydrosylate and yeast extract. Tetracycline was used to maintain the plasmid stability and glucose was fed continuously to the culture during the fermentation. TMV coat protein synthesis was initiated with the addition of 3-β-indoleacrylic acid when the cell density reached ~20 g of dry cell weight per liter of culture fluid. The culture was harvested by continuous centrifugation, resulting in 300–400 g dry weight of cells for each 10-L fermentation. The TMV coat protein was estimated to be 2–3% of the total cell protein by laser densitometry of SDS–PAGE gels (Laemmli, 1970) of cell extracts (Figure 2A).

Analytical Ultracentrifugation. Sedimentation velocity experiments were performed in a Beckman Model E analytical ultracentrifuge equipped with schlieren optics. Sedimentation coefficients were corrected in the standard manner (Schachman, 1959). Weight fractions were determined from the areas under the schlieren peaks by using planimetry of traced

schlieren curves. Other details of solvent density, viscosity corrections, and sample handling are given elsewhere (Raghavendra et al., 1985a).

Mass Spectrometry and Amino Acid Sequence Analysis. The mass spectrum of a tryptic digest of r-TMVP was obtained with a JEOL HX110 mass spectrometer. Fast atom bombardment was performed with 6-keV of xenon atoms. Other details are described in Canova-Davis et al. (1988). The amino acid sequence analysis of the 25 residues of the amino terminus was performed according to Aggarwal et al. (1985).

Isoelectric Focusing. Isoelectric focusing of r-TMVP and TMVP was performed with a Pharmacia PhastGel system. A precast IEF pH 4–6 gel was incubated at room temperature for 20 min in a solution containing 5 mL of distilled water, 4.8 g of ultrapure urea, 1.0 mL of pH 4–6 Pharmalyte ampholyte solution, and 250 μL of 20% NP-40 nonionic detergent (Shell Chemical Co.). The gel was prefocused for 6 min at 25 mA and 2 W at 20 °C and the samples, including pI markers, were then focused at 25 mA and 4 W for ~25 min. The pH gradient was determined from the positions of the markers (Pharmacia).

Electron Microscopy. Electron micrographs were taken on a Philips 300 transmission electron microscope operated at 80 kV. Carbon-coated copper grids were used for sample preparation. Pictures were taken at primary magnifications in the range of 100 000–200 000. Other details are described elsewhere (Schuster et al., 1979; Raghavendra et al., 1986).

Circular Dichroism. Circular dichroism spectra were obtained on r-TMVP and TMVP samples with a Model 60DS AVIV circular dichroism spectrometer outfitted with an on-line computer for data acquisition and reduction. The near-UV CD spectra were recorded from 330 to 240 nm in thermostated (±0.1 °C) 2.00 and 5.00 mm path length cylindrical cells (Helma) for protein concentrations of 3.0 and 1.0 mg/mL, respectively. Far-UV CD spectra were measured from 240 to 185 nm in thermostated 0.10 and 0.20 mm path length cylindrical cells (Helma) for protein concentrations of 3.0 and 1.0 mg/mL, respectively. Data were recorded at 0.2-nm intervals with a 3.0-s time constant and a 1.5-nm constant spectral bandwidth. Each spectrum was an average of three repeated scans, smoothed, normalized for protein concentration and path length, and repeated for two to three protein preparations. The instrument was calibrated with (+)-10-camphorsulfonic acid (Sigma Chemicals Co.) (Adler et al., 1973) and the mean residue ellipticity $[\theta]$ was calculated using a mean residue weight of 111.0 daltons for TMVP and r-TMVP and expressed in deg·cm^2 dmol^{-1}.

Results and Discussion

Isolation and Purification of r-TMVP. Recombinant TMVP was purified at 4 °C from the lysate of transformed *E. coli* cells. After frozen cell paste at −80 °C was thawed at 4 °C, ~200 g was suspended in 10 volumes of lysis buffer (i.e., 0.1 M Tris, 50 mM Na_2EDTA, 0.2 M NaCl, and 0.1% 2-mercaptoethanol, pH 7.2) and mixed well until the suspension was uniformly dispersed in solution. Cells were lysed in a cooled Manton-Gaulin press operated at a pressure of 6000 psi. Centrifugation of the lysate resulted in r-TMVP in the supernatant (Figure 2A, lane 2) and cell debris in the pellet (data not shown). The supernatant was applied to a silica (silica adsorbent, special low fine from Davis Chemical Co.) column and eluted with 20 mM Tris, 0.1% 2-mercaptoethanol, pH 7.2, resulting in approximately a 4-fold purification. The pooled Tris eluant was chromatographed on a DE-52 (DEAE anion-exchange cellulose) column at pH 7.2 and the nonadsorbed protein fractions (Figure 2B, lane 3) contained all the r-TMVP. Dialysis of these flow-through fractions against 50 mM citrate, pH 3.5, resulted in selective precipitation (>95%) of r-TMVP (Figure 2B, lanes 4 and 5). The r-TMVP precipitate was solubilized at pH 8.0 and dialyzed at a concentration of 3–4 mg/mL into potassium phosphate buffer (I = 0.1) at either pH 7.0 or 8.0. This phosphate buffer system was chosen since TMVP self-assembles into intermediate-sized aggregates, which have been studied extensively in this buffer system at a variety of temperatures and pH values (Durham et al., 1971; Butler & Klug, 1971; Richards & Williams, 1976; Schuster et al., 1980). Immunoblotting with polyclonal antibodies to TMVP was used throughout the purification to identify r-TMVP (data not shown). These antibodies were raised in rabbits that were injected with TMV isolated from tobacco plants.

Characterization of the Amino Terminus in r-TMVP. The first 25 amino acid residues of r-TMVP were determined by amino acid sequencing. The only sequence detected was the same as that for the published sequence of TMVP (data not shown). In order to investigate the chemistry of the amino terminus, tryptic digests of TMVP and r-TMVP were analyzed by fast bombardment mass spectrometry. The amino terminal tryptic peptide (residues 1–41) had an observed molecular mass of 4605.4 amu for TMVP and 4563.2 amu for r-TMVP. The difference in molecular mass of 42.2 (±1) amu is consistent with the replacement of an acetyl group by a hydrogen atom. Furthermore, no mass peaks at or near 4605 amu were detected for r-TMVP, demonstrating the absence of acetylated amino termini. This result is consistent with an earlier observation for superoxide dismutase, which is also acetylated on the amino terminus when produced in mammalian cells and not acetylated when expressed in *E. coli* (Hallewell et al., 1985).

The question as to whether the amino terminus is acetylated in r-TMVP was further investigated by isoelectric focusing (IEF) in 8 M urea (Figure 3). TMVP and r-TMVP focused as single bands with pI values of 4.75 and of 5.1, respectively. A free amino terminus without an acetyl group is capable of ionizing and should contribute to the pI of the protein. Theoretical titration curves (not shown here) computed from the amino acid composition of TMVP and r-TMVP (assuming a free amino terminus) and from intrinsic pK values (Shire, 1983) predict a pI difference of ~0.2 pH unit. This difference is in good agreement with the difference obtained by IEF, considering the uncertainties of the intrinsic pK values used for the computations and the uncertainties of the experimental pK values in urea. We assume that the urea effect is the same in both the proteins and the IEF standards used.

Assembly and Structure of r-TMVP. The assembly states of r-TMVP were investigated by analytical ultracentrifugation, electron microscopy, and circular dichroism. Figure 4 presents sedimentation velocity data for r-TMVP. For comparative purposes, results on TMVP are also included. At pH 7.0 and 5 °C, r-TMVP and TMVP behave similarly, sedimenting with $s_{20,w}$ values of 4 S although, approximately 4% of the r-TMVP sediments at 23.3 S. At pH 7 and 8 and at these low temperatures, TMVP exists as limited aggregates in rapid equilibrium with an average size of about a trimer corresponding to a sedimentation coefficient of 4 S.[2] However, depending upon protein concentration and exact solution conditions, the

[2] 4S protein refers to oligomers of TMVP that are in rapid equilibrium between species ranging in size from monomers to about hexamers. They have an average size of trimers and therefore can vary in sedimentation coefficient from 3 to 8 S, depending upon protein concentration, pH, and exact conditions (see text for references).

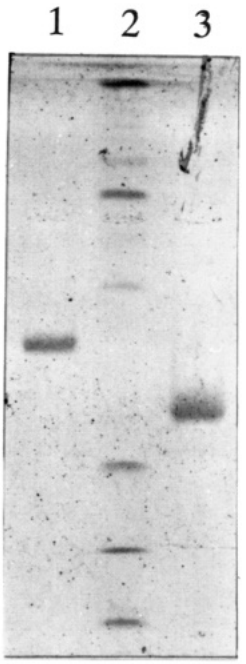

FIGURE 3: Isoelectric focusing of TMVP and r-TMVP in 8 M urea. Lanes 1–3 correspond to r-TMVP, p*I* markers (Pharmacia), and TMVP, respectively. The marker proteins (with the respective p*I* values) from top to bottom are human carbonic anhydrase B (6.55), bovine carbonic anhydrase B (5.85), β-lactoglobulin A (5.20), soybean trypsin inhibitor (4.55), glucose oxidase (4.15), and amyloglucosidase (3.50). p*I* values are not corrected for urea.

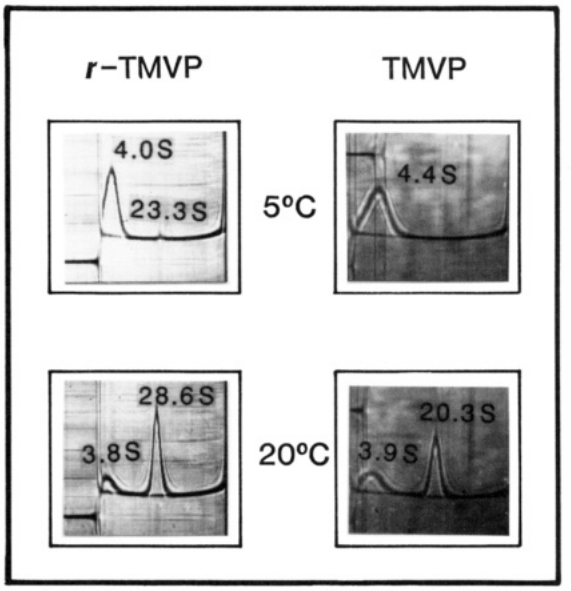

FIGURE 4: Sedimentation patterns of r-TMVP and TMVP aggregates at 5 and 20 °C. Buffer conditions: pH 7.0, $I = 0.1$ potassium phosphate. The concentration of both TMVP and r-TMVP was 3.7 mg/mL. TMVP was previously stored at 0 °C and equilibrated at 5 and 20 °C for 2 and 45 h, respectively, before measurement. r-TMVP was previously stored at 5 °C and was equilibrated at 20 °C for 24 h before measurement. Sedimentation velocity experiments were performed at 40 000 rpm using 12 mm path length cells. All boundaries sediment from left to right with photographs taken at 45 (TMVP) and 50° (r-TMVP) bar angles. The state of aggregation of TMVP at 20 °C is predominantly 20 S (62% of the total protein by weight) whereas it is 28 S (83% of the total protein by weight) for r-TMVP.

s value can vary from 3 to 8 S (Durham et al., 1971; Katzel & Ackers, 1979). Since the major boundary for r-TMVP also sediments with $s_{20,w}$ value of 4 S at pH 7 and 8 (pH 8.0 data not shown), it appears that at low temperatures, r-TMVP also associates into similar limited-size oligomers. Warming of TMVP to 20 °C at pH 7.0 results in the formation of a 20S boundary with a formation half time of 4–6 h depending on the protein concentration. After a 24-h equilibration at 20 °C, 62% by weight of TMVP sediments at 20.3 S and the remainder at 3.9 S. However, after a 24-h equilibration at 20 °C, 83% of the r-TMVP sediments at 28.6 S and 17% at 3.8 S. The r-TMVP sedimenting boundary of 28.6 S is as symmetrical as that seen for the 20S boundary of the wild-type protein, (Figure 4). This faster sedimenting species of r-TMVP is reminiscent of the 28S dimers of bilayer cylindrical disk aggregates formed by wild-type protein in high ionic strength buffers (Durham, 1972; Raghavendra et al., 1988).

Figure 5 shows electron micrographs of aggregates of r-TMVP and TMVP from solutions at 20 °C and pH 7.0, 6.0, and 5.5. Under these conditions, TMVP forms short (pH 7.0) and long (pH <6.5) helical aggregates having a central hole and overall diameter like those of the intact virus. r-TMVP at pH 7.0 (Figure 5A) shows characteristic nonhelical, disklike packing as evidenced by clear striations of uranyl acetate stain between the layers of subunits. Cylindrical disk structures of r-TMVP having 4, 6, 14, and 30 layers of subunits are seen in fields prepared at pH 7.0. However, sedimentation analysis of these same solutions reveal only 4S and 28S aggregates, the latter corresponding to four-layer structures. As has been observed previously with TMVP, the recombinant form of the protein also undergoes further association of preformed aggregates on the EM grid. This aggregation is believed to result from weakly acidic staining and drying artifacts (Raghavendra et al., 1986, 1988). Also, some images of r-TMVP in Figure 5A do not show clear striations between layers of subunits. In such images we cannot distinguish between regions of poor staining and true structural heterogeneity. Examination of a large number of structures, on various grids, reveals that the majority but not all of the aggregates display the striations or banding patterns between layers. In contrast, the naturally occurring TMVP does not form disklike structures under these conditions (Schuster et al., 1980; Correia et al., 1985; Raghavendra et al., 1985a, 1988, 1989). The 20S structures associate on the grid (Figure 5A) to form large, loose aggregates, which do not stain with the characteristic banding pattern exhibited by 28S r-TMVP at low ionic strength and by 28S aggregates and other metastable stacked disks of TMVP at high ionic strength (Durham & Finch, 1972; Raghavendra et al., 1986, 1988, 1989).

At pH 5.5, r-TMVP forms long viruslike rods that are presumably helical aggregates like those formed by TMVP as shown in Figure 5B (Mandelkow et al., 1981). Both rodlike structures have an axial hole that can be penetrated by the stain. While r-TMVP rods show periodic breaks and nicks along their length, TMVP helical rods reveal fewer nicks and are straighter. The diameters of the two, however, are the same at EM resolution, 180 Å, as that of the native virus. The nicks and breaks in the r-TMVP do not "anneal out", even after standing for a few days at 20 °C before grid preparation. Another striking difference seen in Figure 5B is the background. With TMVP the grid surface is quite clean and devoid of nonrodlike polymers and there is also an absence of dissociated 4S protein. This is consistent with sedimentation results, which show only a single fast boundary of >200 S. In contrast, the r-TMVP sample shows a large number of toruslike structures, also evident at pH 7.0 (Figure 5A), with a well-stained central hole and well-defined subunit detail on the outer circumference. The structures that fill the background field between the r-TMVP rods are similar to the 28S r-TMVP four-layer disklike aggregates seen at pH 7.0 (Figure 5A).

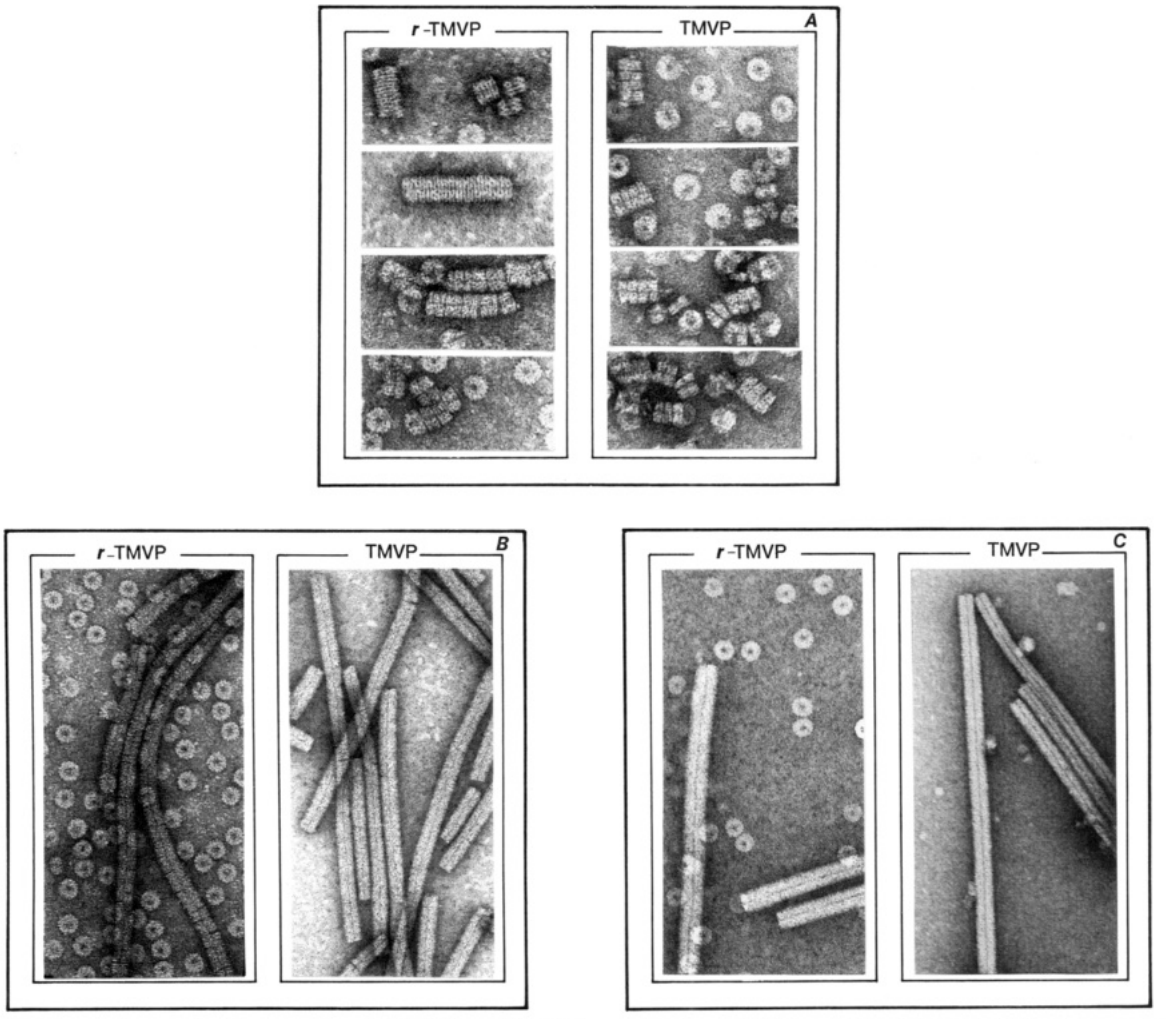

FIGURE 5: Electron micrographs of aggregates of r-TMVP and TMVP from solutions at 20 °C and (A) pH 7.0 (I = 0.1 potassium phosphate) as in Figure 4, (B) pH 5.5 (0.05 M sodium acetate), and (C) pH 6.0 (I = 0.1 potassium phosphate). In (A), a majority of the images of r-TMVP show characteristic striations of cylindrical disk packing of layers of subunits. Representative fields show the basic four-layer, 28S structure seen in solution (Figure 4) and larger aggregates not seen in solution by sedimentation analysis. Some images that do not show disklike clear striations are also observed. TMVP does not display disk packing with strong banding patterns under these conditions. Instead, out-of-register extended aggregates of >20 S are observed. However, only 20S aggregates are seen in solution (Figure 4). Circular, flat toruslike images in r-TMVP and TMVP samples are from aggregates positioned flat on the grid surface. Note resolution of subunits on outer circumference in such r-TMVP but not TMVP aggregates (see text). Long helical aggregates of r-TMVP and TMVP at pH 5.5 in (B) were obtained from (A) by dialyzing at 20 °C for 48 h and equilibrating for 48 h at 20 °C after dialysis. Note in (B) the periodic breaks or nicks along the length of the helical rods and the significant amount of 28S aggregates in the background in the r-TMVP sample but not in the TMVP sample. Long helical aggregates at pH 6.0 in (C) were obtained by rapidly heating to 20 °C (at 1.0–5.0 °C s^{-1}) a cold solution of 1.0 mg/mL protein (0 °C for TMVP or −7 °C for r-TMVP). The resulting viruslike helical rods are completely annealed without breaks or nicks in both proteins unlike those in (B). However, the helical rods of r-TMVP in (C) are shorter than those formed by TMVP, some of which reach lengths of 0.9–1.0 μm (3 times the TMV length). The long straight rod on the left-hand side of the TMVP image is one such specimen but is only partially shown here because of enlargement to show fine structure. The r-TMVP background consists of some undissociated 28S aggregates whereas the TMVP background is devoid of such aggregates. Separate sedimentation experiments were performed under these final conditions to verify the presence of undissociated 28S aggregates.

When TMVP at pH <7.2 is heated, a well-known endothermic polymerization is induced. With rapid heating at pH <6.8, the protein undergoes a nonequilibrium nucleation-controlled polymerization of virtually all of the 4S protein, which results in an overshoot in polymerization (Scheele & Schuster, 1974; Schuster et al., 1979; Shire et al., 1979b; Raghavendra et al., 1985b, 1985c, 1989; Potschka et al., 1988). The formation of very long helical rods at pH 6.0 obtained by such a process is shown in Figure 5C (right-hand panel). These TMVP rods show the tight subunit packing characteristic of the helical array found in the whole virus. The length of the overshoot rods can be as much as 3 times greater than the whole virus, as shown by the vertical single rod in the right-hand panel in Figure 5C, only a portion of which is seen in this enlargement. Figure 5C also depicts overshoot helical rods formed by r-TMVP under the same conditions. The r-TMVP overshoot rods are shorter than those formed from TMVP and the background on the r-TMVP grid contains a considerable amount of unpolymerized 28S r-TMVP, which appears to be the same as the background material in Figure 5B (sedimentation data not shown). Both pH 5.5 and 6.0 are strong polymerization conditions for TMVP. Yet some fraction of the r-TMVP appears to be incapable of polymerizing beyond the 28S state by either method of polymerization, i.e., rapid heating or slow dialysis (see Figure 5 legend).

Far-UV CD spectra of TMVP and r-TMVP, which represent the gross secondary structure of the proteins, are shown in Figure 6A. The spectra recorded at 5 and 20 °C for both TMVP and r-TMVP indicate that the overall folding of r-TMVP is not significantly different from that of TMVP (Figure 6, parts A and B) and both spectra are dominated by α-helix contributions (Vogel & Jaenicke, 1976; Schubert &

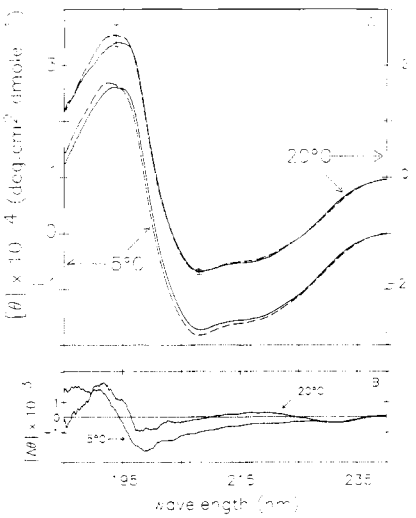

FIGURE 6: (A) Far-UV CD spectra at pH 7.0 of r-TMVP (—) and TMVP (---) at 5 and 20 °C and (B) difference CD spectra, TMVP minus r-TMVP, at both 5 and 20 °C. Conditions and aggregation states are the same as in Figure 4. Spectra were measured at protein concentrations of 3.0 and 1.0 mg/mL by using optical path lengths of 0.10 and 0.20 mm, respectively. Error bar represents uncertainty in $[\theta]$ (in units of deg·cm^2 dmol^{-1}) of ±2000 at 193 nm and ±500 at 208 nm as a result of different path lengths, protein concentrations, and protein preparations. Note in (A) that 5 and 20 °C spectra are offset along the ordinate and correspond to left and right ordinates, respectively, as indicated by arrows. The temperature difference spectra, 5 °C–20 °C, for TMVP and r-TMVP show no significant differences (data not shown). Also (B) shows that there is no significant difference between the spectra of the two proteins at either 5 or 20 °C, respectively.

Krafczyk, 1969). In addition, the different states of aggregation of the proteins at 5 and 20 °C do not lead to any detectable changes in the far-UV CD spectra of TMVP and r-TMVP (see Figure 6A). These results indicate that the amino terminal differences between TMVP and r-TMVP do not appear to influence the overall backbone conformation of the proteins. Since each protein undergoes extensive self-assembly upon changing the temperature from 5 to 20 °C, these far-UV CD spectra suggest that the gross secondary structure of the protein in 4S, 20S, and 28S aggregates is similar.

However, the near-UV CD spectra at 5 and 20 °C of r-TMVP (Figure 7A) reveal subtle differences when compared with the spectra of TMVP. $[\theta]$ values of the recombinant protein are slightly lower in the spectral region 330–240 nm of the aromatic amino acids than is found for the plant-derived coat protein. Also, both the 4S oligomer (5 °C) and 28S four-layer aggregates (20 °C) of r-TMVP exhibit a small red shift in their 287-nm negative dichoric bands. The 287-nm negative dichroic bands contain contributions from both tyrosine and tryptophan, whereas the 295-nm bands are dominated by trytophan chromophores. The positive CD bands between 250 and 270 nm are due to phenylalanine chromophores (Budzynski, 1972; Vogel & Jaenicke, 1976; Strickland, 1976; Schuster et al., 1980; Reed & Kinzel, 1984; Raghavendra et al., 1985a). The difference spectra in Figure 7B, showing the 5 °C difference spectrum of the two 4S proteins and the 20 °C difference spectrum of the more highly assembled proteins, i.e., 20S and 28S, appear to be largely due to tyrosine changes since the changes at 295 nm are not significant. Since the CD differences observed in Figure 7B are independent of the state of association, we assume that they are the result of an electrostatic effect caused by the addition of a charge at the amino terminal group on the adjacent amino acid, tyrosine 2. The temperature difference spectra (i.e., 5 °C minus 20 °C spectrum for each protein) in Figure 7C on the other hand suggest contributions from tryptophan in addition to tyrosine for both the proteins. We note that there is only one ionizable side chain (tyrosine 2) in the first 18 amino acid residues, counting from the amino terminus. For TMVP, this entire segment would be uncharged at neutral pH. Also, the 13 amino acid segment at the carboxyl end of the protein chain contains no ionizable side chains (Altschuh et al., 1987). Further, on the basis of a hydrophobicity analysis of the amino and carboxyl terminal segments (Kyte & Doolittle, 1982), it appears that these segments of the protein encompassing the outer radius of the protein may have a low dielectric resulting in an enhancement of all electrostatic interactions.

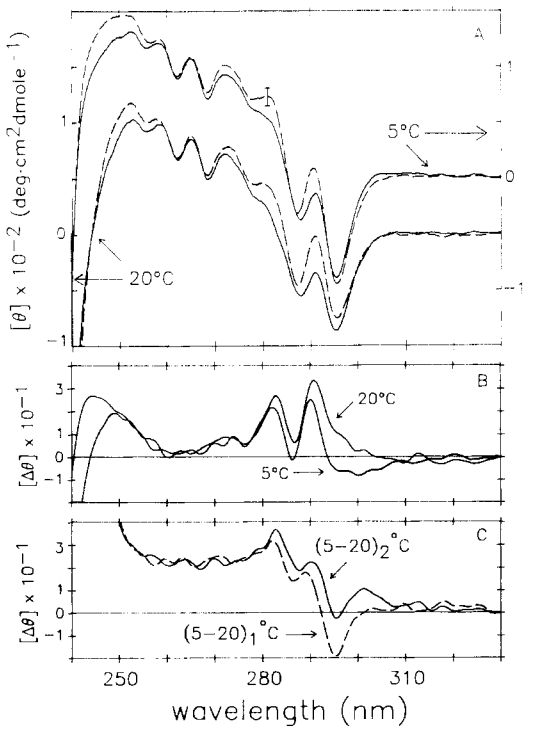

FIGURE 7: (A) Near-UV CD spectra of r-TMVP (—) and TMVP (---) at both 5 and 20 °C; (B) difference CD spectra, TMVP minus r-TMVP, at 5 and 20 °C, and (C) temperature difference CD spectra (5–20)$_1$ °C for TMVP and (5–20)$_2$ °C for r-TMVP. Solution conditions and aggregation states are the same as in Figure 6 except that the spectra were measured by using 2.0- and 5.0-mm optical path lengths at protein concentrations of 3.0 and 1.0 mg/mL, respectively. The error bar represents uncertainty in $[\theta]$ (in units of deg·cm^2 dmol^{-1}) of ±8 at 282 nm, as a result of different path lengths, protein concentrations, and protein preparations. Note in (A) that 5 and 20 °C spectra are offset along the ordinate and correspond to right and left ordinates, respectively, as indicated by arrows. The spectra of r-TMVP are slightly more negative than those of TMVP and show a red shift in the 287-nm dichroic band at both temperatures. Difference CD spectra, TMVP minus r-TMVP, at 5 and 20 °C in (B) indicate aggregation independent changes since these difference spectra are essentially the same at both temperatures even though the states of aggregation are very different at these temperatures (see Figure 4). Also, the temperature-induced, aggregation-dependent, difference CD spectra in (C) i.e., (5–20)$_1$ °C for TMVP and (5–20)$_2$ °C for r-TMVP, are similar. The temperature difference spectra have maxima at 282 and 290 nm as do the difference spectra in (B), but the negative temperature difference at 295 nm is not seen in the difference spectra in (B); see text for discussion.

Summary and Conclusions

Although cloning and expression in *E. coli* of the gene for the coat protein of a positive-strand RNA virus, TMV, has been reported previously, very little characterization beyond electron micrographs of helical rods has been presented (Haynes et al., 1986). We have been able to express and purify

adequate amounts of r-TMVP in *E. coli* for physical studies. A purification scheme is described that takes advantage of the pH dependency of coat protein assembly. The ability of the recombinant protein to assemble into long helical rods at low pH and subsequently depolymerize at higher pH values demonstrates that most of the coat protein expressed and purified out of *E. coli* behaves in a fashion similar to that of protein isolated from virus with regard to pH dependency of self-assembly. Characterization by mass spectrometry and isoelectric focusing of the recombinant DNA produced protein reveals that this protein lacks an amino terminal acetyl group when expressed in *E. coli* and therefore carries an additional positive charge, near neutral pH, compared to the tobacco plant produced viral protein. This fact has allowed us to investigate the effect of the amino terminal cotranslational modification on the self-assembly properties of the coat protein. The charge modification in r-TMVP results in a significant change in its self-assembly properties. In particular, the intermediate-size aggregates at pH 7.0 and 20 °C no longer sediment as a 20S boundary but rather as a boundary with a sedimentation coefficient of 28 S. EM results show clearly that the major component of the 28S TMVP boundary looks like the high ionic strength disk aggregate formed by TMVP (Raghavendra et al., 1990a). Polymerization of 28S r-TMVP at pH 5.5 results in helical rods that exhibit imperfections, nicks, and dislocations that do not anneal, unlike the helical rods produced from 20S TMVP. In contrast, helical rods produced from 4S subunits of r-TMVP resemble the TMVP helical rods more closely. However, in both cases a portion of the r-TMVP does not form helical rods but either remains as, or forms new, 28S disklike aggregates. These results suggest that the acetylation at the amino terminus of TMVP influences the kinetic pathway of assembly and in particular allows the TMVP to assemble mainly into helical rather than disklike structures at neutral pH (Raghavendra et al., 1990b). Although r-TMVP can reconstitute with TMV RNA, it does so at a reduced rate compared to the reconstitution with TMVP (Raghavendra et al., 1989). This suggests that one function of the acetylation is to allow the protein to form the correct nucleating structure, resulting in a more efficient reconstitution process. However, clarification of specific interactions involving free amino terminus in r-TMVP has to await structural analysis of the disk structure of r-TMVP. Initial crystallography studies on high-salt single crystals of r-TMVP reveal the same cell constants as found for TMVP (Raghavendra et al., 1989).

The occurrence of amino terminal acetylation in proteins is widespread (Persson et al., 1985). Examples are several members of the toba movirus group, cytochrome *c*, egg albumin, histone IV, tropomyosin, apoferritin, and superoxide dismutase (Driessen, et al., 1985; Augen & Wold, 1986). The functional significance of amino terminal acetylation is not fully understood. Recent results implicate acetylation as a factor in the regulation of ubiquitin-dependent protein degradation (Mayer et al., 1989). In the present study we have found that a nonacetylated amino terminus does not significantly alter the overall secondary structure but results in extensive changes in the assembly properties of a plant virus coat protein.

Acknowledgments

The facilities provided by the University of Connecticut Electron Microscopy Facility, Tobacco Research Station at Windsor Locks (CT) for growing TMV-infected plants and the technical assistance of Carol Masheter are gratefully acknowledged. We thank Drs. Alexander Miller III and Gerald Stubbs for careful reading of the manuscript. We also gratefully acknowledge the technical assistance of Bill Kohr for amino acid analysis and sequencing, Dr. John A. Chakel for mass spectrometric analysis, and Dr. Norman Lin for fermentation and harvesting of *E. coli* with r-TMVP plasmid.

References

Ackers, G. K., & Katzel, L. (1979) *J. Supramol. Struct. 10*, 216 (abstract).

Adler, A. J., Greenfield, N. J., & Fasman, G. D. (1973) *Methods Enzymol. 27*, 675–735.

Aggarwal, B. B., Henzel, W. J., Moffat, B., Kohr, W. J., & Harkins, R. N. (1985) *J. Biol. Chem. 260*, 2334–2344.

Altschuh, D., Lesk, A. M., Bloomer, A. C., & Klug, A. (1987) *J. Mol. Biol. 193*, 693–707.

Augen, J., & Wold, F. (1986) *Trends Biochem. Sci. (Pers. Ed.) 11*, 494–497.

Bloomer, A. C., & Butler, P. J. G. (1986) in *The Plant Viruses* (Van Regenmortel, M. H. V., & Fraenkel-Conrat, H., Eds.), Vol. 2, pp 19–57, Plenum, New York.

Bloomer, A. C., Champness, J. N., Bricogne, G., Staden, R., & Klug, A. (1978) *Nature (London) 276*, 362–368.

Budzynski, A. Z. (1972) *Biochim. Biophys. Acta 251*, 292–302.

Butler, P. J. G. (1984) *J. Gen. Virol. 65*, 253–279.

Butler, P. J. G., & Klug, A. (1971) *Nature (London), New Biol. 229*, 47–50.

Butler, P. J. G., & Durham, A. C. H. (1977) *Adv. Protein Chem. 31*, 187–251.

Canova-Davis, E., Chloupek, R. C., Baldonado, I. P., Battersby, J. E., Spellman, M. W., Basa, L. J., O'Connor, B., Pearlman, R., Quan, C., Chakel, J. A., Stults, J. T., & Hancock, W. S. (1988) *Am. Biotechnol. Lab.* May, 8–17.

Caspar, D. L. D. (1963) *Adv. Protein Chem. 18*, 37–121.

Correia, J. J., Shire, S., Yphantis, D. A., & Schuster, T. M. (1985) *Biochemistry 24*, 3292–3297.

Dawson, W. O., Beck, D. L., Knorr, D. A., & Grantham, G. L. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1832–1836.

Driessen, H. P. C., deJong, Tesser, G. I., & Bloemendal, H. (1985) *CRC Crit. Rev. Biochem. 18*, 281–325.

Durham, A. C. H. (1972) *J. Mol. Biol. 67*, 289–305.

Durham, A. C. H., & Finch, J. T. (1972) *J. Mol. Biol. 67*, 307–314.

Durham, A. C. H., Finch, J. T., & Klug, A. (1971) *Nature (London), New Biol. 229*, 37–42.

Englander, S. W., & Epstein, H. T. (1957) *Arch. Biochem. Biophys. 68*, 144–149.

Franklin, R. E. (1956) *Nature (London) 177*, 928–930.

Goelet, P., & Karn, J. (1984) *Gene 29*, 331–342.

Goelet, P., Lomonosoff, G. P., Butler, P. J. G., Abam, M. E., Gait, M. J., & Karn, J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 76*, 5818–5822.

Hallewell, R. A., Masiarz, F. R., Najarian, R. C., Puma, J. P., Quiroga, M. R., Randolph, A., Sanchez-Pescadow, R., Scandella, C. J., Smith, B., Setimer, K. S., & Mullenbach, G. T. (1985) *Nucleic Acids Res. 13*, 2017–2034.

Haynes, J. R., Cuningham, J., Von Seefried, A., Lennick, M., Garvin, R. T., & Shen, Shi-Hsiang (1986) *Biotechnology 4*, 637–641.

Hirth, L., & Richards, K. E. (1981) *Adv. Virus Res. 26*, 145–199.

Huynh, T., Young, R., & Davis, R. (1985) in *DNA Cloning, Practical Approaches in Biochemistry* (Grover, D., Ed.) Vol. I, pp 49–78, IRL Press, Oxford.

Jones, D. H., McMillan, A. J., Fersht, A. R., & Winter, G. (1985) *Biochemistry 24*, 5852–5857.

Kyte, J., & Doolittle, R. (1982) *J. Mol. Biol. 157*, 105–132.

Laemmli, B. B. (1970) *Nature (London) 277*, 680.
Leung, D. W., Capon, D. J., & Goeddel, D. V. (1984) *Biotechnology 2*, 458–464.
Mandelkow, E., Stubbs, G., & Warren, S. (1981) *J. Mol. Biol. 152*, 375–386.
Matthews, R. E. F. (1981) *Plant Virology*, pp 287–300, Academic Press, New York.
Matsumura, M., & Matthews, B. W. (1989) *Science (Washington, D.C.) 243*, 792–794.
Mayer, A., Siegel, N. R., Schwartz, A. L., & Ciechanover, A. (1989) *Science (Washington, D.C.) 244*, 1480–1483.
Meshi, T., Ishikawa, M., Motoyoshi, F., Semba, K., & Okada, Y. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5043–5047.
Namba, K., & Stubbs, G. (1986) *Science (Washington, D.C.) 231*, 1401–1406.
Namba, K., Pattanayek, R., & Stubbs, G. (1989) *J. Mol. Biol. 208*, 307–325.
Okada, Y. (1986) *Adv. Biophys. 22*, 95–145.
Oxender, D. L., & Fox, C. F., Eds. (1987) *Protein Engineering*, Alan R. Liss, Inc., New York.
Perry, L. J., & Wetzel, R. (1984) *Science 226*, 555–557.
Persson, B., Flinta, C., von Heijne, G., & Jornvall, H. (1985) *Eur. J. Biochem. 152*, 523–527.
Potschka, M., Koch, M. H. L., Adams, M. L., & Schuster, T. M. (1988) *Biochemistry 27*, 8481–8491.
Raghavendra, K., Adams, M. L., & Schuster, T. M. (1985a) *Biochemistry 24*, 3298–3304.
Raghavendra, K., Adams, M. L., & Schuster, T. M. (1985b) *Biophys. J. 47*, 51a (abstract).
Raghavendra, K., Adams, M. L., & Schuster, T. M. (1985c) *Book of Abstracts: Fourth Conversation in Biomolecular Stereodynamics* (Sarma, R. H., Ed.) p 312, Institute for Biomolecular Stereodynamics, SUNY Albany, Albany, NY.
Raghavendra, K., Salunke, D. M., Caspar, D. L. D., & Schuster, T. M. (1986) *Biochemistry 25*, 6276–6279.
Raghavendra, K., Kelly, J. A., Khairallah, L., & Schuster, T. M. (1988) *Biochemistry 27*, 7583–7588.
Raghavendra, K., MacDonald, G., Schuster, T. M., Shire, S. J., Leung, D. W., Cachianes, G. J., Jackson, E., Wood, W. I., McKay, P. (1989) *Biophys. J. 55*, 156a (abstract).
Raghavendra, K., Prasad, B. V. V., Marietta, E., Chiu, W., & Schuster, T. M. (1990a) *Biophys. J. 57*, 256a (abstract).
Raghavendra, K., Pattanayek, R., Stubbs, G., & Schuster, T. M. (1990b) *Biophys. J. 57*, 257a (abstract).
Reed, J., & Kinzel, V. (1984) *Biochemistry 23*, 1357–1362.
Richards, K. E., & Williams, R. C. (1976) *Compr. Virol. 6*, 1–37.
Schachman, H. (1959) *Ultracentrifugation in Biochemistry*, Academic Press, New York.
Scheele, R. B., & Schuster, T. M. (1974) *Biopolymers 13*, 275–288.
Schramm, G. (1947) *Z. Naturforsch. 2B*, *112*, 249.
Schubert, D., & Krafczyk, B. (1969) *Biochim. Biophys. Acta 188*, 155–157.
Schuster, T. M., Scheele, R. B., & Khairallah, L. H. (1979) *J. Mol. Biol. 127*, 461–486.
Schuster, T. M., Scheele, R. B., Adams, M. L., Shire, S. J., Steckert, J. J., & Potschka, M. (1980) *Biophys. J. 32*, 313–329.
Shire, S. J. (1983) *Biochemistry 22*, 2664–2671.
Shire, S. J., Leung, D. W., Cachianes, G., Jackson, J., Wood, W. I., Raghavendra, K., Schuster, T. M., & McKay, P. (1987) *Biophys. J. 51*, 91a (abstract).
Shire, S. J., Steckert, J. J., Adams, M. L., & Schuster, T. M. (1979a) *Proc. Natl. Acad. Sci. U.S.A. 76*, 2745–2749.
Shire, S. J., Steckert, J. J., & Schuster, T. M. (1979b) *J. Mol. Biol. 127*, 487–506.
Shire, S. J., Steckert, J. J., & Schuster, T. M. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 256–260.
Steckert, J. J. (1982) Ph.D. Thesis, University of Connecticut.
Strickland, E. H. (1976) in *CRC Handbook of Biochemistry and Molecular Biology, Proteins* (Fasman, G. D., Ed.) Vol. III, pp 141–166, CRC Press, Cleveland, OH.
Stubbs, G. (1984) in *Biological Macromolecules and Assemblies* (Jurnak, F. A., & McPherson, A., Eds.) Vol. I, pp 149–202, Wiley, New York.
Stubbs, G., Warren, S., & Holmes, K. (1977) *Nature (London) 267*, 216–221.
Villifranca, J. E., Howell, E. E., Voet, D. H., Strobel, M. S., Ogden, R. C., Abelsen, J. N., & Kraut, J. (1983) *Science 222*, 782–788.
Vogel, D., & Jaenicke, R. (1976) *Eur. J. Biochem. 61*, 423–431.
Wells, J. A., Cunningham, B. C., Graycar, T. P., & Estell, D. A. (1986) *Philos. Trans. R. Soc. London A 317*, 415–423.