

- Ferguson-Miller, S., Brautigan, P. L., & Margoliash, E. (1978) *J. Biol. Chem.* 253, 149.
- Fox, T. (1979) *Proc. Natl. Acad. Sci. U.S.A.* 76, 6534.
- Fox, T. D., & Leaver, C. I. (1981) *Cell (Cambridge, Mass.)* 26, 315-323.
- Fuller, S. F., Capaldi, R. A., & Henderson, R. (1979) *J. Mol. Biol.* 134, 305.
- Fuller, S. D., Darley-Usmar, V. M., & Capaldi, R. A. (1981) *Biochemistry* 20, 7046-7053.
- Geren, L. M., & Millett, F. S. (1981) *J. Biol. Chem.* 256, 4851.
- Jentoft, N., & Dearborn, D. G. (1979) *J. Biol. Chem.* 254, 4359-4365.
- Kang, C. H., Brautigan, D. L., Osheroff, N., & Margoliash, E. (1978) *J. Biol. Chem.* 253, 6502.
- Laemmli, U. K. (1970) *Nature (London)* 227, 680.
- Moreland, R. N., & Dockter, M. E. (1981) *Biochem. Biophys. Res. Commun.* 99, 339.
- Poulos, T. L., & Kraut, J. (1980) *J. Biol. Chem.* 255, 10, 322.
- Prochaska, L. I., Bisson, R., Capaldi, R. A., Steffens, G. L. M., & Buse, G. (1981) *Biochim. Biophys. Acta* 637, 360-373.
- Rieder, R., & Bosshard, H. R. (1980) *J. Biol. Chem.* 255, 4732.
- Salemme, F. R. (1976) *J. Mol. Biol.* 102, 563.
- Seiter, C. H. A., Margalit, R., & Ferreault, R. A. (1979) *Biochem. Biophys. Res. Commun.* 86, 473.
- Sheenan, J. C., Cruickshank, P. A., & Boshart, G. L. (1961) *J. Org. Chem.* 26, 2525-2528.
- Smith, H. T., Staudenmayer, N., & Millett, F. (1977) *Biochemistry* 16, 4971.
- Smith, H. T., Ahmed, A. J., & Millett, F. (1981) *J. Biol. Chem.* 256, 4984.
- Smith, M. B., & Millett, F. (1980) *Biochim. Biophys. Acta* 626, 64.
- Speck, S. H., Ferguson-Miller, S., Osheroff, N., & Margoliash, E. (1979) *Proc. Natl. Acad. Sci. U.S.A.* 76, 155.
- Steffens, G., & Buse, G. (1979) *Hoppe-Seyler's Z. Physiol. Chem.* 360, 613.
- Stonehuerner, J., Williams, J. B., & Millett, F. (1979) *Biochemistry* 18, 5422.
- Takano, T., & Dickerson, R. E. (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77, 6371.
- Thomas, P. E., Ryand, D., & Levin, W. (1976) *Anal. Biochem.* 75, 168-176.
- Timkovich, R. (1977) *Anal. Biochem.* 79, 135-143.
- Towbin, H., Staehelin, T., & Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A.* 76, 4350-4354.
- Webb, M., Stonehuerner, J., & Millett, F. S. (1980) *Biochim. Biophys. Acta* 543, 290.
- Wilson, M. T., Lalla-Maharajh, W., Darley-Usmar, V. M., Bonaventura, J., Bonaventura, C., & Brunori, M. (1980) *J. Biol. Chem.* 255, 2722.

## Proteolytic Dissection of Turnip Crinkle Virus Subunit in Solution<sup>†</sup>

John S. Golden and Stephen C. Harrison\*

**ABSTRACT:** Turnip crinkle virus (TCV) can be dissociated under mild conditions, yielding dimers of the protein subunit, as shown by chemical cross-linking and by gel filtration. These dimers have defined regions of proteolytic sensitivity, corre-

sponding to the arm and hinge of the folded polypeptide chain. The results are discussed in terms of the known high-resolution structure of TCV and with regard to implications for the pathway of viral assembly.

**D**etermination of the complete three-dimensional structure of tomato bushy stunt virus (TBSV) has revealed a number of remarkable features with implications for mechanisms of assembly (Harrison et al., 1978; Harrison, 1980). The related turnip crinkle virus (TCV) has also been visualized by high-resolution X-ray diffraction, and the structure of its protein shell is, as expected, substantially the same as that of TBSV (J. Hogle and S. C. Harrison, unpublished results). Success in reversible disassembly of TCV makes it a better choice than TBSV for mapping an assembly pathway. As an initial step toward this mapping, we have carried out chemical cross-linking and proteolytic dissection studies of the isolated TCV subunit.

The rationale of these experiments depends on the general architecture of TBSV and TCV (shown diagrammatically in Figure 1). The particle contains 180 coat protein subunits

( $M_r \sim 40\,000$ ), probably one chain of a  $M_r$  80,000 protein, and a molecule of single-stranded RNA (4800 nucleotides) (Ziegler et al., 1974). The coat subunit, containing about 390 amino acids, folds into distinct regions: a projecting domain (P), a domain forming a tightly connected shell (S), an arm (a), and an internal domain (R). The three symmetrically distinct environments for this subunit are labeled A, B, and C. The polypeptide accommodates to these three packing modes by flexion at the hinge (h) between S and P and by an ordering or disordering of the arm, a. Units at positions A and B (60 of each) have one hinge configuration, and the entire N-terminal region (R and a) appears to be spatially disordered. Subunits at positions C (60 in all) have another hinge position, and the arm is folded in an ordered way along the inner side of the S domain. The R region is spatially disordered in C subunits as well as in A and B. "Spatially disordered" is used here to mean merely that the segment is not fixed with respect to the outer shell. The R domain itself may be precisely folded but flexibly tethered to the rest of the subunit such that it can adopt a variety of positions vis-à-vis the S domain and hence present no strong features in a high-resolution electron density map. The entire RNA molecule is spatially disordered in this sense.

<sup>†</sup> From the Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138. Received January 14, 1982. This work was supported in part by the National Institutes of Health (Grant CA-13202), by the National Science Foundation (Grant PCM 79-22159), and by an Alfred R. Sloan fellowship to S.C.H. The experiments form part of a Senior Honors Thesis in Biochemical Sciences, Harvard University.

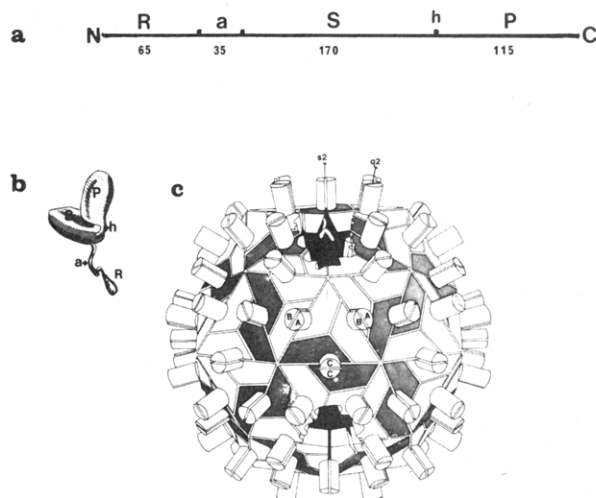


FIGURE 1: Architecture of TCV particle. (a) Order of domains in polypeptide chain N terminus to C terminus. The approximate number of residues in each segment is indicated below the line. The letters indicate R domain (possible RNA binding region), arm a (the connector that forms the  $\beta$  annulus and extended arms structures on C subunits and that remains disordered on A and B), S domain, hinge, and P domain. (b) Schematic view of folded polypeptide chain. (c) Arrangement of subunits of particle. A, B, and C denote distinct packing environments for the subunit; outer surfaces of C-subunit S domains are shaded. S domains of A subunits pack around 5-fold axes; S domains of B and C alternate around 3-fold axes. The local 3-fold axis relating to S domains of an ABC trimer is nearly parallel to the adjacent strict 2-fold axis (s2), across which P domains of C subunits are paired. Trimers present a rather flat surface across the strict dyad and a distinctly sharper dihedral angle ( $\sim 40^\circ$ ) across the quasi-dyad (q2). A and B P domains, paired across q2, therefore have a hinge angle with respect to their S domain that differs by  $\sim 20^\circ$  from the angle on C.

#### Experimental Procedures

**Virus Purification.** TCV, propagated in Chinese cabbage (*Brassica chinensis* var. crispy choy), was isolated by techniques described elsewhere (Leberman, 1966; J. Hogle and S. C. Harrison, unpublished results). Virus concentrations were determined spectrophotometrically with  $E_{260}^{1\%} = 50$ . Virus was stored as a 3–4% solution in distilled water containing 0.01% sodium azide.

**Dissociation and Expansion.** Virus was dissociated by incubation for 1 h at  $4^\circ\text{C}$  in 0.5 M NaCl (or KCl), 0.1 M Tris-HCl,<sup>1</sup> and 0.01 M EDTA, pH 8.3 (Leberman, 1966). Triethanolamine buffers were used for cross-linking experiments. Expanded, undissociated TCV was obtained by omitting 0.5 M salt.

**DMS Cross-Linking (Davies & Stark, 1970).** A 10-mg sample of dimethyl suberimidate (DMS) (Pierce) was dissolved in 1 mL of 0.2 M triethanolamine, 0.5 M NaCl, and 0.01 M EDTA adjusted with HCl to between pH 7.2 and pH 8.0. Equal volumes of this solution and of dissociated virus (2 mg/mL in the same buffer) were combined, and the pH was corrected when necessary. Aliquots were taken at appropriate times, the reaction was quenched with 1  $\mu\text{L}$  of 1 M ammonium acetate, and the samples were boiled with equal volumes of electrophoresis sample buffer.

**Proteolytic Cleavage.** Cleavage was initiated by the addition of protease in ratios of 1–10  $\mu\text{g}/\text{mg}$  of virus. Samples were incubated either at room temperature or at  $37^\circ\text{C}$ , and aliquots were removed at specific times (10 s to 2.5 h), quenched with addition of excess phenylmethanesulfonyl fluoride (PMSF)

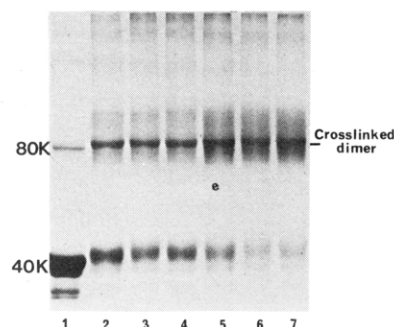


FIGURE 2: Polyacrylamide gel (7.5%; Laemmli, 1970) showing DMS cross-linking of TCV 40K protein. For conditions of cross-linking, see Experimental Procedures. (Lane 1) Untreated TCV. The lane has been overloaded to show the minor 80K protein (one chain per particle; Ziegler et al., 1974). (Lanes 2–4) Cross-linking at pH 7.6 for 1, 3, and 4 h; (lanes 5–7) cross-linking at pH 7.8 for 1, 3, and 4 h.

in ethanol, and boiled immediately with electrophoresis sample buffer containing 5 mg of NaDodSO<sub>4</sub>/mg of viral protein.

#### Results

**Dissociation of TCV.** TCV was dissociated by incubating at  $4^\circ\text{C}$  for 1 h in 0.5 M NaCl, 0.1 M Tris, and 0.01 M EDTA, pH 8.5. Buffers other than Tris are also effective: triethanolamine (TEA) was used for DMS cross-linking experiments. At lower ionic strength (below 0.3 M NaCl) the virus expands but does not dissociate; at room temperature, dissociation is extremely slow. RNA and protein are completely separated under these conditions, and protein may be isolated essentially RNA free by gel filtration. Calibration of an Ultrogel ACA-54 column with protein standards suggests that TCV coat protein is a dimer under the dissociating conditions described.

**DMS Cross-Linking of Dissociated TCV.** TCV, dissociated at  $4^\circ\text{C}$  in 0.5 M KCl, 0.2 M TEA, and 0.01 M EDTA, pH 8, was cross-linked for 1 h with DMS at various pHs from 7.2 to 8.0. Gel electrophoresis (Figure 2) showed a prominent band at 80K but no other discrete species. Similar results were obtained at pH 9. Since the protein elutes as a relatively homogeneous peak from a gel filtration column, we conclude that dissociation of TCV indeed yields dimers of coat protein.

**Proteolytic Cleavage of TCV Protein.** Treatment of dissociated TCV with chymotrypsin leads to very rapid cleavage of the 40K subunit (Figure 3). The principal initial cleavage product is a species migrating at about 29K (actually, a doublet). A 34K product is also regularly observed, but it is rapidly "chased" into the 29K band. The site and rate of cleavage are unaffected by the presence of RNA. RNase treatment of the dissociation mixture or column purification of the coat protein did not alter the course of chymotryptic digestion (Figure 3). Elastase produced a fragment similar in most respects to that obtained with chymotrypsin: major loci of cleavage thus appear to be determined primarily by conformational accessibility (Figure 4). The 29K fragment was isolated by gel filtration on Ultrogel ACA-54. Its elution rate suggested that it forms a dimer, a conclusion confirmed by DMS cross-linking (below).

Figure 4 illustrates another important property of the proteolytic cleavages: TCV protein in expanded but not dissociated virus is cleaved similarly (though incompletely) to dissociated protein. The low ionic strength conditions used for samples in lanes 1, 3, 5, and 7 produce an expanded TCV particle from which no subunits are lost (P. K. Sorger, P. G. Stockley, I. K. Robinson, and S. C. Harrison, unpublished results). There is similarity of cleavages in odd- and even-

<sup>1</sup> Abbreviations: Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; NaDodSO<sub>4</sub>, sodium dodecyl sulfate.

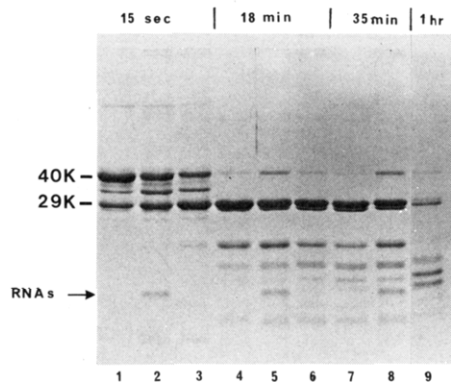


FIGURE 3: Polyacrylamide gel showing chymotryptic cleavage of TCV protein: (lanes 1, 4, and 7) dissociated TCV, prepared by preliminary incubation in dissociation buffer (see Experimental Procedures), in 0.5 M KCl, 0.1 M Tris, and 0.01 M EDTA, pH 8.5; (lanes 2, 5, and 8) samples as in lanes 1, 4, and 7 but with addition of RNase A (0.025  $\mu\text{g}/\mu\text{g}$  of virus) and in reduced ionic strength (0.125 M KCl); (lanes 3, 6, and 9) reduced ionic strength but no RNase. Chymotrypsin (1  $\mu\text{g}/300 \mu\text{g}$  of dissociated virus) was added at room temperature, and aliquots were taken for electrophoresis on a 12% gel (Laemmli, 1970) after 15 s (lanes 1–3), 18 min (lanes 4–6), 35 min (lanes 7 and 8), and 1 h (lane 9).

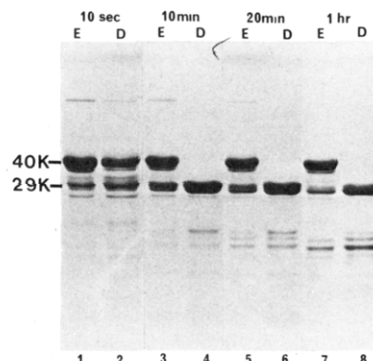


FIGURE 4: Polyacrylamide gel showing elastase cleavage of expanded TCV and of TCV protein: (lanes 1, 3, 5, and 7) TCV in 0.1 M Tris-HCl-0.01 M EDTA, pH 8.5 (expanded virus); (lanes 2, 4, 6, and 8) TCV in 0.5 M KCl, 0.1 M Tris-HCl, and 0.01 M EDTA, pH 8.5 (dissociated virus). TCV (200  $\mu\text{g}$  in 100  $\mu\text{L}$  of buffer) was incubated at 4 °C for 1 h; elastase (0.25  $\mu\text{g}$ ) was then added at room temperature. Aliquots (10  $\mu\text{L}$ ) were removed at appropriate times, quenched with 1  $\mu\text{L}$  of 50 mM PMSF in ethanol, boiled with equal volumes of sample buffer, and subjected to electrophoresis on a 12% polyacrylamide gel (Laemmli, 1970): (lanes 1 and 2) 10 s; (lanes 3 and 4) 10 min; (lanes 5 and 6) 20 min; (lanes 7 and 8) 1 h.

numbered lanes, although only about one-third of the subunits in expanded TCV are actually susceptible.

Effective further cleavage of the 29K fragment required partial denaturation. Figure 5 shows cleavage of dissociated TCV in the presence of urea by chymotrypsin and by elastase. The 29K species is rapidly converted to fragments migrating at 19.5K and 9.5K, respectively, with variable subsequent degradation. Elastase was also used to produce similar fragments from column-purified 29K protein (Figure 6), showing that the 9.5K species indeed derives from this part of the molecule.

**DMS Cross-Linking of Fragments.** Cross-linking of elastase-cleaved TCV in 0.5 M NaCl at alkaline pH produces a band migrating at about 60K, consistent with dimeric association of the 29K species (Figure 7). No evidence was obtained for any cross-linking of the 19.5K fragment.

## Discussion

**Assignment of Cleavage Sites.** A schematic diagram of the TCV subunit derived from high-resolution X-ray crystallog-

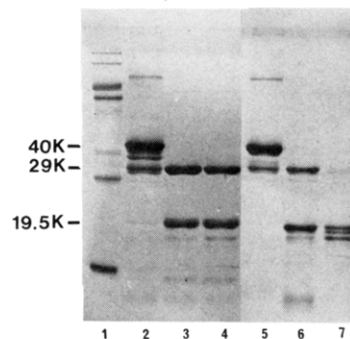


FIGURE 5: Polyacrylamide gel showing proteolytic cleavage in 1 M urea. TCV (200  $\mu\text{g}$ ) was dissociated by incubation for 1 h at 4 °C in 100  $\mu\text{L}$  of the usual buffer with addition of 1 M urea. Chymotrypsin (0.33  $\mu\text{g}$ ) or elastase (0.25  $\mu\text{g}$ ) was added and incubation continued at 4 °C. Aliquots were taken at appropriate times, quenched with PMSF, boiled with sample buffer, and subjected to electrophoresis on a 12% polyacrylamide gel (Laemmli, 1970): (lane 1) standards ( $\beta$ -galactosidase, phosphorylase *a*, bovine serum albumin, catalase, aldolase, chymotrypsinogen, cytochrome *c*); (lanes 2, 3, and 4) chymotryptic digestion for 30 s, 15 min, and 1 h; (lanes 5, 6, and 7) elastase digestion for 30 s, 15 min, and 45 min.

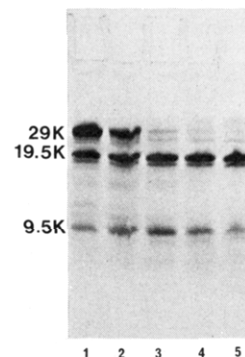


FIGURE 6: Polyacrylamide gel showing TCV 29K fragment digested with elastase (1  $\mu\text{g}/300 \mu\text{g}$  of virus) in 0.5 M NaCl, 0.1 M Tris-HCl, 0.01 M EDTA, and 2.5 M urea, pH 8.5. Aliquots taken at appropriate times were quenched with PMSF and subjected to electrophoresis in a 12% acrylamide gel containing 6 M urea in Tris-phosphate buffer (Swank & Munkres, 1971). Lanes 1, 2, 3, 4, and 5 correspond to 30 s, 1 min, 3 min, 5 min, and 7 min, respectively.

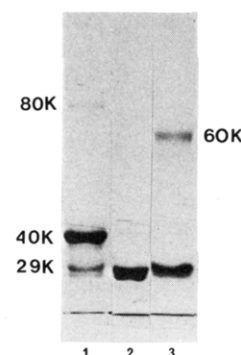


FIGURE 7: Polyacrylamide gel showing DMS cross-linking of TCV 29K fragment. TCV (200  $\mu\text{g}$  in 100  $\mu\text{L}$  of 0.5 M KCl, 0.2 M TEA, and 0.01 M EDTA, pH 8.0) was dissociated by incubation at 4 °C for 1 h and cleaved by addition of 0.25  $\mu\text{g}$  of elastase (20 min at room temperature before addition of 5  $\mu\text{L}$  of 50 mM PMSF). An aliquot was taken for electrophoresis and an equal volume of DMS (10 mg/mL in the same buffer) added (see Experimental Procedures for further details of cross-linking and analysis): (lane 1) TCV; (lane 2) cleaved TCV; (lane 3) cleaved TCV cross-linked with DMS for 1 h at room temperature.

raphy of TBSV and TCV is shown in Figure 1, along with a representation of its packing in the virus particle. An ambiguity arises in assigning the initial cleavage site since scission

Table I: Assignment of Fragments<sup>a</sup>

$M_r$ from mobility in gel	structural assignment	size of corresponding TBSV fragment
40K	total subunit (R + a + S + P)	42K
34K	a + S + P	35K
29K	S + P	31K
19.5K	S	19K
9.5K	P	12K

<sup>a</sup> R, S, P, and a represent R, S, and P domains and the connector between R and S that forms (on C subunits) the  $\beta$  annulus and extended arm. The sequence of TCV protein is not yet known, but from the TCV electron density map we do not expect differences between TCV and TBSV of more than  $\pm 10$  residues in the length of each region. The size of TBSV fragments is estimated from numbers of amino acid residues in each region of the map and from the preliminary amino acid sequence (R. Sauer, personal communication).

at either end of the S domain would generate fragments of about 30K and 12K. The chymotryptic cleavage is sufficiently precise to enable some sequence information to be obtained at its N terminus. The 8–10 residues identified are consistent with a sequence at the N terminus of the S domain visualized in the 3-Å electron density map, but the comparison is too limited to be definitive (data not shown). We can use an indirect argument, however, to establish that the cleavage is indeed immediately N terminal to the folded S domain. This argument depends on comparison of chymotryptic cleavage patterns of *expanded* TCV and TBSV with the cleavage of dissociated TCV protein. Characterization of the expanded structures will be published elsewhere (P. K. Sorger, P. G. Stockley, I. K. Robinson, and S. C. Harrison, unpublished results), but the conclusion relevant to the present argument is that both expanded viruses are cleaved homologously. Moreover, comparison of odd- and even-numbered lanes in Figure 4 shows that the sites of cleavage are the same in expanded and dissociated TCV. Direct sequence information has established unambiguously that the point of cleavage in expanded TBSV is just N terminal to the S domain (R. Sauer, personal communication). We therefore conclude that the 29K fragment of TCV protein, like its TBSV homologue, comprises P and S domains, initial dissection having removed the R domain and arm. The 34K fragment, seen after brief chymotryptic cleavage of TCV protein (Figure 3, lanes 1–3) and especially prominent in cleavage of expanded TBSV, can then be understood as attack at the C terminus of the R domain. Degradation of 29K to 19.5K plus 9.5K in urea (Figures 5 and 6) can safely be taken, from the sizes of the fragments, to correspond to cleavage at the interdomain hinge. These assignments are summarized in Table I.

The structure of southern bean mosaic virus (SBMV) closely resembles that of TCV or TBSV (Abad-Zapatero et al., 1980): its subunit has a basic R domain, a flexible connector that forms a scaffold of ordered C-subunit arms, and an S domain of nearly identical fold. (SBMV completely lacks a P domain, however.) Subunits of expanded SBMV can be cleaved by trypsin at several sites near the N terminus (Tremaine & Ronald, 1978) in a pattern strongly reminiscent of the TCV cleavages shown in Figure 4. Accessibility of the N-terminal part of the molecule, normally buried in the native virus, thus appears to be a common feature of the expanded forms.

The susceptibility of various segments of the TCV subunit to proteolytic attack can readily be understood in terms of the known conformations of the subunit in the virus. Arms of A and B subunits are disordered; those of C subunits can fold only when subunits associate to form a shell. We therefore

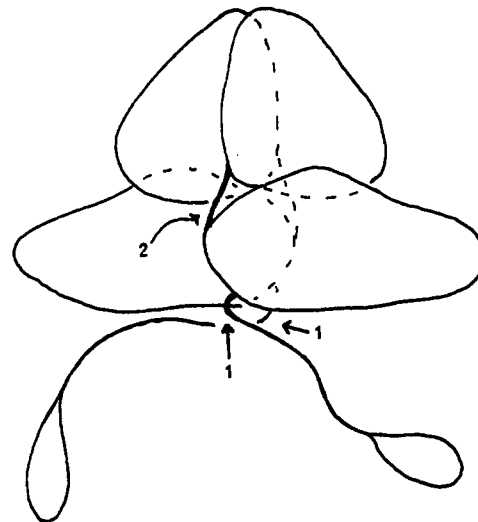


FIGURE 8: Schematic view of TCV dimer in solution. Arrows labeled 1 indicate positions of cleavage to form 29K species, still dimeric in solution. The arrow labeled 2 shows a hinge cleavage, which occurs only in urea.

expect this part of the molecule to be readily degraded, since it is not likely to be protected by a compact fold. By contrast, the interdomain hinge is buried at a 2-fold contact, and dimeric association of the 29K fragment should lead to significant protection (Figure 8).

**State of Association.** It was anticipated from the structure of TBSV and TCV that the subunit might dimerize readily under conditions that do not favor further association. The P domains have an extensive face for hydrophobic interaction with polar groups at the periphery forming H bonds and salt bridges (Harrison, 1980). The results here show that TCV subunits do indeed form "P-clamped" dimers in solution (Figure 8). Removal of the arm and R domain (production of the 29K fragment) does not affect this association, whereas dissection at the P/S hinge appears to lead to the monomeric 19.5K fragment, as expected from the modest extent of dimer contact in the S domain. The 9.5K fragment, corresponding to the P domain, was too unstable to isolate effectively. The effect of 2 M urea might indeed be to destabilize this domain selectively, since the hinge region is relatively protected by the dimer contact as it exists in the virus.

**Assembly Intermediates.** We have at present only rather sketchy information concerning the pathway of TCV assembly. Virus-like particles can be reassembled from protein and RNA under conditions where protein alone forms no larger structures (D. Kimmelman, P. G. Stockley, and S. C. Harrison, unpublished results). The design of the TCV particle is such that the inner "scaffold" of C-subunit arms precisely determines its size (Harrison, 1980). Formation of this scaffold requires two classes of interaction: contact around the 3-folds and contacts across dyads. The  $\beta$  annulus establishes contacts among three C-position subunits, and P-domain pairing can strongly enforce the dyad interaction. A P-clamped subunit dimer is thus a plausible intermediate in viral assembly. The SBMV structure (Abad-Zapatero et al., 1980) shows that the P domain is not an essential feature, however. Other 180-subunit structures, such as brome mosaic virus and cowpea chlorotic mottle virus, also appear to assemble from subunit dimers (Adolph & Butler, 1974; Pfeiffer & Hirth, 1974). If a scaffold of N-terminal arms determines the  $T = 3$  icosahedral lattice in these viruses as it does in TCV, similar intermediates may reflect similar architectural features of the particles.

## Acknowledgments

We thank James Hogle for TCV and Robert Sauer, Peter Stockley, and Judy White for thoughtful advice.

## References

- Abad-Zapatero, C., Abdel-Meguid, S. S., Johnson, J. E., Leslie, H. G. W., Rayment, J., Rossmann, M. G., Suck, D., & Tsukihara, T. (1980) *Nature (London)* 286, 33-39.
- Adolph, K., & Butler, P. J. G. (1974) *J. Mol. Biol.* 88, 327-341.
- Davies, G. E., & Stark, G. R. (1970) *Proc. Natl. Acad. Sci. U.S.A.* 66, 651-656.
- Harrison, S. C. (1980) *Biophys. J.* 32, 139.
- Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K., & Bricogne, G. (1978) *Nature (London)* 276, 368.
- Laemmli, U. K. (1970) *Nature (London)* 227, 680-685.
- Leberman, R. (1966) *Symp. Soc. Gen. Microbiol.* 18, 183-204.
- Pfeiffer, P., & Hirth, L. (1974) *Virology* 58, 362-368.
- Swank, R. T., & Munkres, K. D. (1971) *Anal. Biochem.* 39, 462-477.
- Tremaine, J. H., & Ronald, W. P. (1978) *Virology* 91, 164-172.
- Ziegler, A., Harrison, S. C., & Leberman, R. (1974) *Virology* 59, 509.

## Structural Studies of P22 Phage, Precursor Particles, and Proteins by Laser Raman Spectroscopy<sup>†</sup>

George J. Thomas, Jr.,\* Yinglin Li, Margaret T. Fuller, and Jonathan King

**ABSTRACT:** For the study of the protein-protein and protein-nucleic acid interactions in the assembly of virus particles, laser Raman spectra have been obtained in H<sub>2</sub>O and D<sub>2</sub>O solutions and as a function of temperature for the following *Salmonella* phage P22 components: mature phage particles, isolated mature phage DNA, mature protein shells empty of DNA, precursor protein shells (procapsids), and purified coat, scaffolding and tail-spike proteins. The spectra confirm that the condensed DNA within the phage capsid assumes the B-form secondary structure similar to aqueous DNA and reveal no evidence of specific molecular interactions between subgroups of DNA and protein subunits of the phage capsid. No differences were detected in the highly irregular secondary

structure of the major capsid protein in mature capsids, empty capsids (lacking DNA), procapsids, and empty procapsids (lacking scaffolding protein). Features of both primary and secondary structures of the viral scaffolding and tail-spike proteins are also revealed by the spectra. Differences in thermal stability of tyrosyl side-chain interactions were observed between scaffolding protein extracted from the procapsid and within the procapsid. These differences correspond to different hydrogen bonding configurations of *p*-hydroxy-phenyl groups and provide indirect evidence for the participation of the scaffolding proteins in specific macromolecular interactions within the procapsid.

In the maturation of viruses and other complex particles, numerous transformations occur within the organized structures themselves. For the assembly of double-stranded DNA (dsDNA) phages, a doubly shelled procapsid is first assembled with coat protein on the outside and scaffolding protein on the inside (Casjens & King, 1975; Murialdo & Becker, 1978; Earnshaw et al., 1979). In packaging dsDNA into this structure, the interior scaffolding is lost while the outer coat protein shell expands and undergoes transformation to the mature capsid structure (Earnshaw & Casjens, 1980). Although the DNA is tightly and stably packed within the coat protein shell, it must be able to exit efficiently from the capsid during the DNA injection process (Laemmli et al., 1974). We have used laser Raman spectroscopy to monitor the states of the macromolecules participating in these processes.

Phage P22 is a temperate phage of *Salmonella typhimurium*. It contains dsDNA of molecular weight  $2.7 \times 10^7$  packed in concentric coils within the isometric capsid of di-

ameter of approximately 560 Å (Earnshaw et al., 1976; Earnshaw & Harrison, 1977). In the assembly of the P22 procapsid, of diameter 510 Å, about 200 molecules of the gene 8 scaffolding protein (42 000 dalton) copolymerize with about 420 molecules of the gene 5 coat protein (55 000 dalton) (Fuller & King, 1982). Procapsids also contain a relatively small number of minor proteins (King et al., 1973). The scaffolding protein molecules exit the capsid without proteolytic processing during the packaging of DNA and are recycled for further rounds of capsid assembly (King & Casjens, 1974). The particle which results from DNA packaging is unstable until the "neck" structure has been completed by addition of three more proteins (Botstein et al., 1973; Poteete & King, 1977). If one of these is removed by mutation, the capsids lose their DNA spontaneously and in so doing provide a source of empty "mature" capsids.

Procapsids and empty capsids can be isolated from mutant-infected cells and purified in large quantities (Casjens & King, 1975; Earnshaw et al., 1976; Fuller & King, 1981). Empty procapsids can be prepared by removal of the scaffolding protein in vitro (Fuller & King, 1982). The coat (gp5) and scaffolding (gp8) proteins of P22 procapsids have been purified in biologically active forms (Fuller & King, 1981, 1982), as has the tail-spike (gp9) protein (Berget & Poteete, 1980) of mature phage. We have attempted to ascertain

<sup>†</sup> From the Department of Chemistry, Southeastern Massachusetts University, North Dartmouth, Massachusetts 02747 (G.J.T. and Y.L.), and the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 (M.T.F. and J.K.). Received December 8, 1981. This is part X in the series Studies of Virus Structure by Laser Raman Spectroscopy. Supported by National Institutes of Health Grants AI 11855 (G.J.T.) and GM 17980 (J.K.).