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†

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_2O and D_2O 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*-hydroxyphenyl 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 2.7×10^7 packed in concentric coils within the isometric capsid of diameter 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

† 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.).

0006-2960/82/0421-3866$01.25/0

Table I: Identification and Composition of P22 Phage, Precursor Particles, and Protein Subunits[a]

name	description
P22 phage	the mature phage contains genome, coat proteins, tail proteins, and minor proteins
1⁻ procapsid	the procapsid is a precursor particle (diameter ~510 Å) assembled by the association of 420 coat protein subunits (gp5) and 200 scaffolding protein subunits (gp8); the minor structural protein gp1 is absent from 1⁻ procapsids
1⁻ empty procapsid	the empty procapsid is the outer protein shell (gp5) of the capsid; it is obtained by treatment of the procapsid with 1.0 M Gdn·HCl which causes the release of the scaffolding protein
10⁻ empty capsid	the empty capsid is purified from 10⁻ infected cells and has packaged and then lost its DNA; it contains 420 copies of gp5 and some minor proteins
gp5 (major coat protein)	the outer shell of the procapsid is assembled from the gp5 subunit; it is also the major component of the phage capsid
gp8 (scaffolding protein)	the scaffolding protein participates in the assembly of the capsid; it is the subunit from which the inner shell of the procapsid is assembled
gp9 (scaffolding protein)	the tail-spike of the phage is constructed from the gp9 subunit; the tail-spike is believed to consist of six trimers of gp9; it manifests enzymatic activity and participates in a specific assembly reaction during phage morphogenesis
P22 DNA	the DNA molecule is linear and double stranded with a molecular weight of approximately 2.7×10^7

[a]See Fuller & King (1981, 1982) and references cited therein.

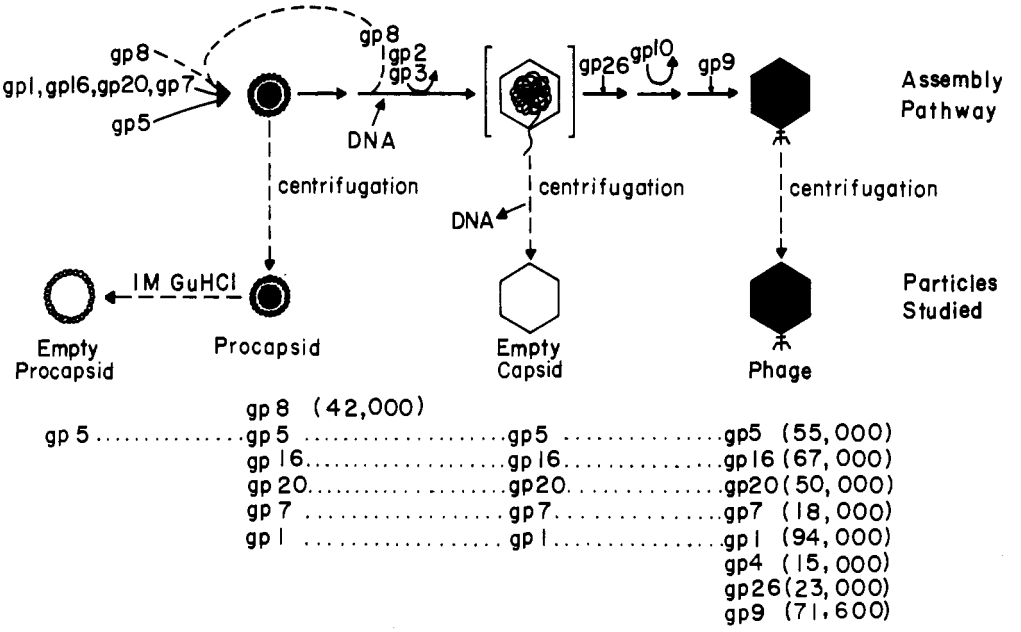

FIGURE 1: Pathway for the morphogenesis of bacteriophage P22, showing the sequence of gene product interaction in P22 assembly. Adapted from Fuller & King (1980).

macromolecular secondary structures and detect protein–protein and protein–DNA interactions by obtaining Raman spectra of these species.

Raman spectroscopy is a light-scattering technique in which the frequencies of the scattered radiation are determined by the molecular vibrations of the scattering particles. Since the intensity of Raman scattering is dependent upon the change in polarizability with vibration, Raman spectral bands (or "lines") of high intensity are generally obtained from bond stretching vibrations of multiply-bonded groups (e.g., C=O, C=N, C=C, etc.) and of heavy atoms (e.g., P–O, S–S, C–C, etc.). Certain symmetrical stretching modes of aromatic rings, including those of phenylalanine, tryptophan, tyrosine, and the DNA bases, also lead to intense Raman scattering. The intensities and frequencies of the Raman lines can be related to the interactions and structures of the molecular subgroups involved and in favorable cases have been used to measure protein and nucleic acid secondary structures (Thomas & Kyogoku, 1977). These results have permitted the use of Raman spectroscopy for the study of structures of both RNA and DNA viruses (Thomas, 1976a,b).

Two preliminary reports of the Raman spectroscopy of P22 have appeared recently (Fish et al., 1980; Li et al., 1981). These investigations demonstrate the capability for obtaining Raman spectra of high quality on P22 and its precursor particles and also indicate the sensitivity of the Raman spectrum to changes in protein structure with temperature and environment of protein subunits. The present work extends the preliminary studies and reevaluates the interpretation of the earlier Raman data in light of the more complete data now available. The results of these physical probes extend the initial interpretation of P22 organization obtained from biochemical and genetic investigations.

Experimental Procedures

Materials. The specific samples used in this work and their descriptions are given in Table I. The functions and compositions of these particles are further illustrated in Figure 1. P22 components (Table I) were derived from mutant-infected cultures and were isolated by density gradient or gel filtration methods and then purified on a DEAE column or by sucrose gradient ultracentrifugation (Fuller & King, 1981, 1982). The

scaffolding protein subunits were isolated from solutions of partially dissociated procapsids, which were treated with 1 M Gdn·HCl.[1] Samples were stored initially in a buffer of 50 mM Tris, 25 mM NaCl, 2 mM EDTA, 3 mM β-mercaptoethanol, and 1% glycerol at pH 7.6 and were subsequently transferred to the solvent media indicated below by dialysis or pelleting procedures.

Deuterated solvents (D_2O, 99.7%; DCl, 20% in D_2O; and NaOH, 40% in D_2O) were obtained from Aldrich Chemical Co. Tris buffer (Trizma base; Sigma Chemical Co.) and all other reagent grade materials were used without further purification.

Sample Handling for Raman Spectroscopy. In order to ensure a low level of background scattering in Raman spectra, to remove sulfhydryl group containing reagents, and to ensure uniformity in the solution compositions, all viral and protein samples were further treated as follows:

Heavy particles (phage, capsids, and procapsids) were pelleted by centrifugation at 25 000 rpm (51000*g*) for 2–3 h in an IEC Model 50 ultracentrifuge equipped with a Model A-321 rotor. The pellets were routinely redissolved in 10 mM Tris buffer, pH 7.6 ± 0.2, and were again pelleted by ultracentrifugation. Occasionally when debris remained, a second or third repelleting cycle was required. The homogeneous pellets finally obtained (approximately 70 mg of protein/mL) were loaded directly into Raman cells (Kimax 34507 capillaries).

Solutions of protein subunits were concentrated by vacuum ultrafiltration (Millipore, Immersible CX) to give a concentration in the range 70–100 mg/mL. The concentrate was diluted 20-fold in 10 mM Tris, pH 7.6 ± 0.2, and subsequently ultrafiltered to yield a final concentration of 70–100 mg/mL. The concentrate was loaded directly into a Raman cell.

For obtaining Raman spectra of samples in corresponding D_2O solutions, similar procedures were followed, except that D_2O replaced H_2O in all buffer preparations. The pD value of deuterated media was computed from the formula pD = pH + 0.4.

Whenever adjustment of the pH or removal of β-mercaptoethanol was required, the protein or virus solution was dialyzed twice against 500 volumes of 10 mM Tris at the desired pH for 36 h.

All samples investigated were sufficiently concentrated to yield Raman spectra of excellent signal-to-noise ratio. A comparison with protein and nucleoprotein samples investigated prviously (Thomas, 1976a) indicates that virus, precursor particles, and proteins were present at 70–100 mg/mL total concentration. The actual concentrations in the Raman cells, as estimated from UV absorption intensities of 1000-fold dilutions of the Raman samples, were consistent with total protein concentration in the range 70–100 mg/mL.

Instrumentation. The Raman system consists of an argon ion laser excitation source (Coherent Radiation Model CR-2) and a Spex Ramalog spectrometer described previously (Thomas, 1976b). The temperature of the sample cell is precisely controlled (±0.5 °C) by a thermostated brass jacket (Thomas & Barylski, 1970).

The monochromator was calibrated with emission lines of mercury and the frequency scale was adjusted prior to each day's work to fix the symmetrical stretching mode of CCl_4 at 459.5 cm^{-1}. All Raman frequencies are believed accurate to within ±2 cm^{-1}. Spectra were excited with approximately 300 mW of the 514.5-nm line of argon. The error limits (±) reported on Raman intensities are average deviations from three or more spectral scans of the same sample(s).

[1] Abbreviations: Gdn, guanidine; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; def, deformation.

Results and Discussion

Table I lists the particles and proteins which were isolated in sufficient concentration and purity to obtain Raman spectra. The origin of a number of these particles is shown in Figure 1. The pure precursor coat protein from procapsids was not available at the time of these studies. All spectra of this protein (gp5) were obtained from either empty procapsids or empty mature capsids.

(A) Raman Spectra of Bacteriophage P22 and Related Particles. The Raman spectra of H_2O solutions of P22 and several related particles are shown in Figure 2. Corresponding spectra for D_2O solutions are shown in Figure 3. The salient features of the spectra and their interpretations in terms of molecular composition and conformation are discussed next.

(1) Scaffolding Protein (gp8). The scaffolding protein displays an intense Raman line at 938 cm^{-1} with a weak shoulder at 949 cm^{-1}, a weak and broad amide III line near 1270 cm^{-1}, and a strong amide I line at 1650 cm^{-1}, all due to the protein backbone. These features are characteristic of α-helical proteins and clearly indicate that scaffolding protein in this isolated state contains a predominantly α-helical secondary structure. The additional weak amide III Raman scattering near 1248 cm^{-1} shows also that irregular or disordered chain conformations are present to only a minor extent (Lord, 1977; Frushour & Koenig, 1975).

Also prominent in the Raman spectrum of scaffolding protein are lines due to tyrosine near 857 and 830 cm^{-1}. According to Siamwiza et al. (1975) the intensity ratio of these lines ($R_{Tyr} = I_{857}/I_{830}$) is sensitive to the hydrogen-bonding environment of the *p*-hydroxyl group of tyrosine. Three types of hydrogen bonding have been recognized and correlated with R_{Tyr}: (i) If the tyrosyl –OH group is the acceptor of a strong hydrogen bond from a positive group such as $-NH_3^+$, then R_{tyr} = 2.5. (ii) With moderately strong hydrogen bonding of the –OH group as both donor and acceptor, such as occurs with H_2O solvent molecules, then R_{Tyr} = 1.2. (iii) If the –OH group of Tyr is the donor of a strong hydrogen bond to a negative acceptor, for example, to a $-CO_2^-$ group, then R_{Tyr} = 0.3 (Siamwiza et al., 1975). From Figure 2, we find R_{Tyr} = 1.2 ± 0.2, which suggests either that the –OH groups of the 11 Tyr residues of scaffolding protein participate in moderate hydrogen bonding with H_2O or equivalent groups or that comparable numbers of tyrosyl –OH groups are distributed among states of the types i, ii, and iii described above so that on average $R_{Tyr} \simeq 1.2$.

Other prominent lines in the spectrum of the H_2O solution of scaffolding protein occur at 758 (Trp), 1004 (Phe), 1208 (Phe + Tyr), and 1450 cm^{-1} (CH_2 def).[1] The weak Raman line at approximately 521 cm^{-1} (Figure 2), although within the region of disulfide stretching of cystine bridges (Lord & Yu, 1970), is not assigned to Cys because no companion C–S stretching frequency can be detected near 650–700 cm^{-1}. A band near 520–525 cm^{-1} is also found in α-helical coat proteins of filamentous phages Pf1, fd, Xf, and Pf3, which lack cystine (Thomas & Murphy, 1975; Thomas et al., 1981). We propose that this frequency is an amide mode probably characteristic of α-helical structure. Supporting evidence for this assignment comes from the occurrence of a strong line at 523 cm^{-1} in Raman spectra of α-helical poly(L-alanine) and poly(D-alanine) but not in nonhelical poly(DL-alanine) (Itoh et al., 1974; Fasman et al., 1978).

In searching the –SH stretching region (2500–2600 cm^{-1})

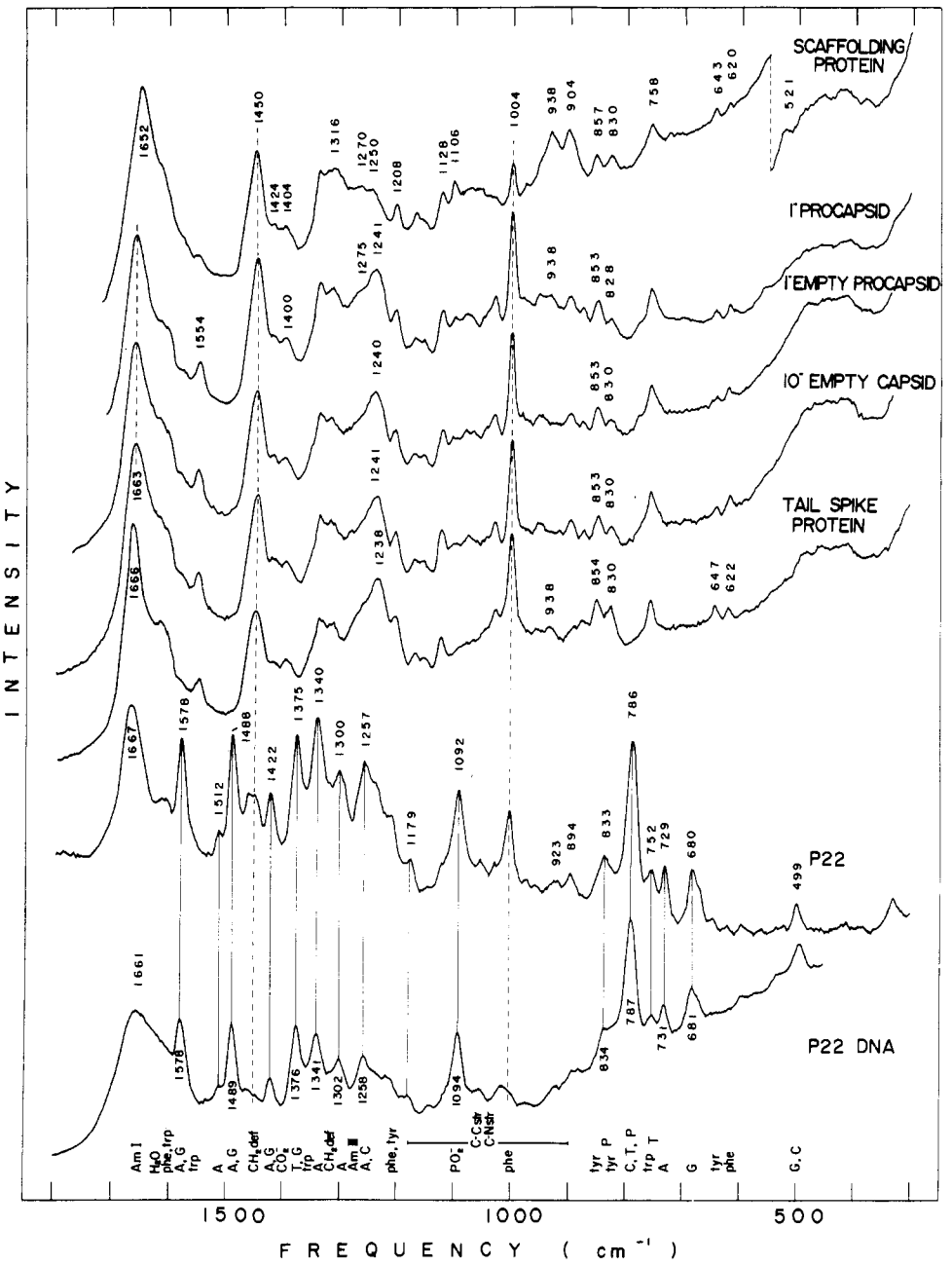

FIGURE 2: (From top to bottom) Raman spectra of H_2O solutions of scaffolding protein, 1^- procapsid, 1^- empty procapsid, 10^- empty capsid, and tail-spike protein (each at 7% w/v), P22 phage, and P22 DNA (each at 4% w/v). All samples are at 32 °C in 10 mM Tris buffer, pH 7.6. Instrument conditions: Excitation wavelength (λ) = 514.5 nm, radiant power (P) = 300 mW, slit width ($\Delta\sigma$) = 10 cm^{-1}, scan rate (r) = 50 cm^{-1} min^{-1}, and rise time (τ) = 3 s. Frequency scale alignment is illustrated by the broken vertical lines through the 1004- and 1450-cm^{-1} bands of each protein spectrum. A broken vertical line also illustrates the identical amide I frequency (1663 cm^{-1}) in capsid and procapsid which differs from the amide I frequency of scaffolding protein (1652 cm^{-1}), tail-spike protein (1666 cm^{-1}), or phage (1667 cm^{-1}). Solid vertical lines are used to connect all intense Raman lines of the phage to their counterparts in P22 DNA. Assignments for many of the prominent bands are given along the abscissa.

and –SD stretching region (1800–1900 cm^{-1}) of H_2O and D_2O solution spectra, respectively, we found no Raman lines that would indicate the presence of cysteinyl groups in the scaffolding protein. If such groups are present, they are below the minimum level detectable by Raman spectroscopy, i.e., less than one Cys per several hundred amino acids of the scaffolding protein. We note that the single cysteine group in TMV coat protein (158 residues per subunit) is readily detected by the Raman method (G. J. Thomas, Jr., unpublished results). Therefore, there can be at most one cysteine among the 351 residues of the scaffolding protein that could go undetected in the Raman spectrum.

The strong line at 904 cm^{-1} is probably due to C_α–C_β stretching of amino acid aliphatic side groups. A prominent line at this frequency occurs in Raman spectra of polyalanine (Frushour & Koenig, 1974; Itoh et al., 1974; Fasman et al., 1978), and therefore alanyl residues may be a major contributor.

When a highly α-helical protein such as the scaffolding protein is dissolved in D_2O solution, it is likely that deuteration of peptidyl –NH–groups is not instantaneously complete (Blout et al., 1961). Accordingly, the Raman spectrum in D_2O solution (Figure 3) exhibits several relatively broad Raman lines as might be anticipated for incompletely deuterated molecules. The spectrum of gp8 shows no further progress toward exchange even after several days in D_2O at room temperature. Nevertheless, many of the lines in the D_2O solution spectrum can be assigned to vibrational modes of specific subgroups.

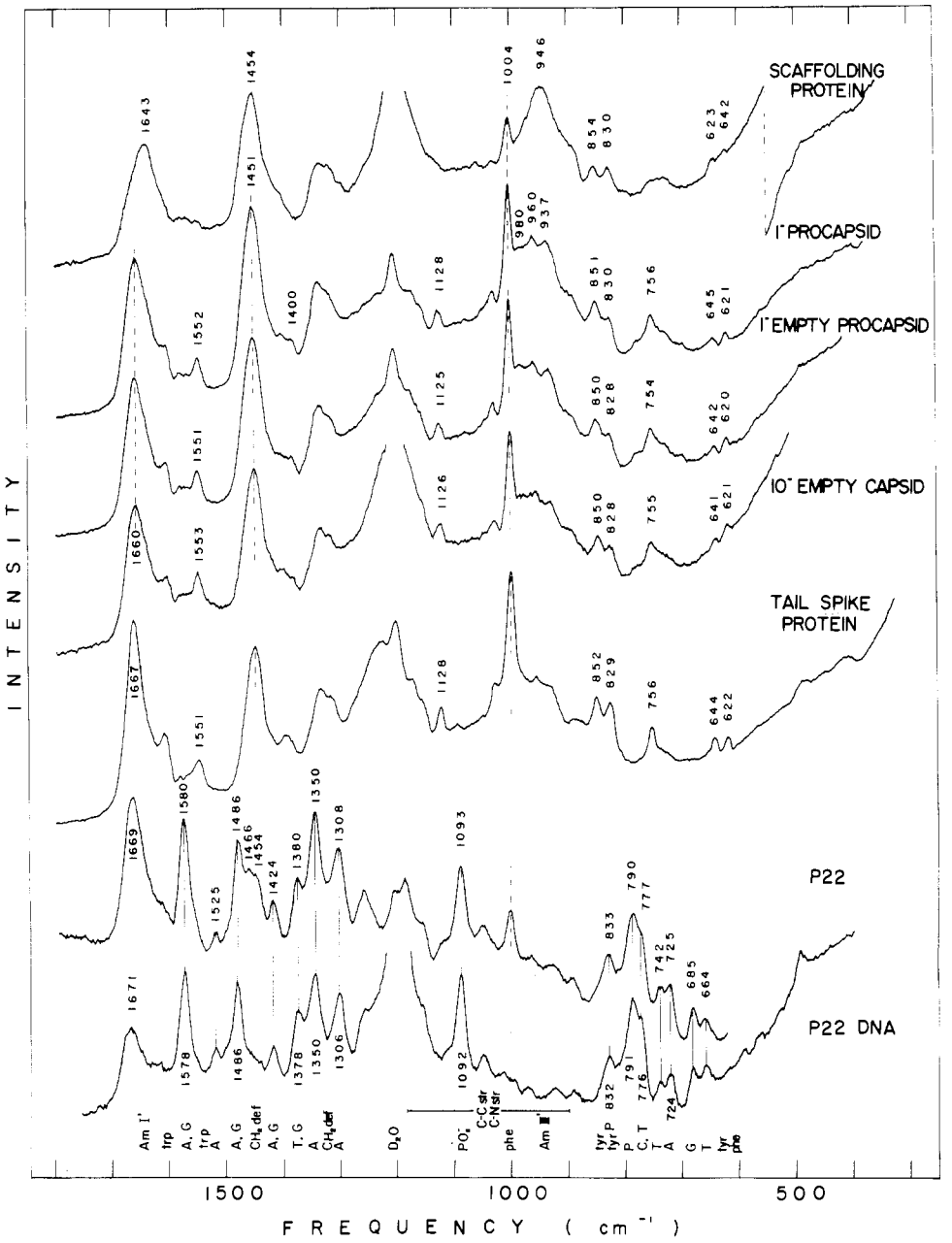

FIGURE 3: Raman spectra of D_2O solutions of phage P22 and related particles as described in Figure 2. Instrument conditions and notation are as in Figure 2.

Thus, the lines centered at 946 and 1643 cm^{-1} are assigned mainly to amide III′ and amide I′ modes, respectively. Again these are the frequencies expected for a protein that is rich in α-helical secondary structure (Frushour & Koenig, 1974; Lord, 1977; Thomas & Kyogoku, 1977). Other Raman lines of the D_2O solution spectrum which are easily assigned are as follows: 1454 (CH_2 def), 1004 (Phe), and 854 and 830 cm^{-1} (Tyr doublet). A very weak line from Trp residues occurs near 750 cm^{-1}.

Because the Raman lines of Phe, Tyr, and Trp residues are so much weaker in spectra of scaffolding protein than are the corresponding lines in spectra of capsids and procapsids (Figures 2 and 3), it is evident that the scaffolding protein is deficient in the aromatic amino acids Phe, Tyr, and Trp as compared with the major coat protein.

(*2*) *I⁻ Procapsid* (*gp8* + *gp5*). The I⁻ procapsid (Table I) displays a strong and broad amide III line near 1241 cm^{-1} with a prominent shoulder near 1275 cm^{-1}. The amide I band appears at 1663 cm^{-1}. These features indicate that the protein subunits of the I⁻ procapsid are largely nonhelical. Only the amide III shoulder at 1275 cm^{-1} can be attributed to α helix. Thus the 420 subunits of major coat protein and 200 subunits of scaffolding protein which collectively make up each procapsid are much less α helical than is purified scaffolding protein in aqueous solution.

For the I⁻ procapsid, the tyrosine doublet is located at 853 and 828 cm^{-1} with an intensity ratio $R_{Tyr} = 1.7 \pm 0.2$, which indicates that many tyrosines of the collective coat and scaffolding subunits in the procapsid exist in state i (defined in section A, part 1). In other words, for most tyrosine residues in the procapsid the –OH groups act as the acceptor of a strong hydrogen bond from a positive donor group (Siamwiza et al., 1975).

The greater relative intensities of Raman lines of Tyr (853 and 828 cm^{-1}), Trp (758 cm^{-1}), and Phe (1004 cm^{-1}) show that the average protein of the procapsid, and therefore the coat protein, is much richer in aromatic amino acid residues than scaffolding protein. Nevertheless, the absence of a distinct

Table II: Molecular Composition of Phage P22 and Precursor Particles[a]

subunit (mol wt)	1⁻ empty procapsid (wt %)	1⁻ procapsid (wt %)	10⁻ empty capsid (wt %)	P22 phage (wt %)
gp5 (55 000)	420 (100%)	420 (69.8%)	420 (90.4%)	420 (42.2%)
gp1 (94 000)			9 (3.3%)	9 (1.5%)
gp7 (18 000)		10 (0.5%)	10 (0.7%)	10 (0.3%)
gp16 (67 000)		10 (2.0%)	10 (2.6%)	10 (1.2%)
gp20 (50 000)		15 (2.3%)	15 (2.9%)	15 (1.4%)
gp4 (15 000)				10 (0.3%)
gp26 (23 000)				30 (1.3%)
gp9 trimer (215 000)[b]				6 (2.4%)
gp8 (42 000)		200 (25.4%)		
DNA (2.7×10^7)				1 (49.4%)

[a] From Fuller & King (1980, 1981, 1982). The table gives the number of copies and the percent by weight (in parentheses) of the given gene product in each particle listed. [b] R. T. Sauer, W. Krovatin, A. R. Poteete, and P. B. Berget, personal communication.

peak at 1361 cm^{-1} that could be assigned to Trp residues in hydrophobic environments suggests that most Trp residues in the procapsid are exposed to the aqueous environment (Yu, 1974; Kitagawa et al., 1979). The weak line near 938 cm^{-1} is assigned to α-helical conformers of the procapsid, in particular to the scaffolding molecules of the inner protein shell.

The Raman spectrum of the procapsid in D_2O solution (Figure 3) displays three intense bands in the amide III′ region, at 937 (α helix), 960 (assignment uncertain but probably due to irregular conformations), and approximately 980 cm^{-1} (β sheet) (Williams et al., 1980; Thomas et al., 1981).

The 1⁻ procapsid displays a Raman line of medium intensity at 1400 cm^{-1} in aqueous solution (Figure 2) which is nearly unaffected by deuteration (Figure 3). Therefore, this line is very likely due to the carboxylate ($-CO_2^-$) symmetric stretching mode of Glu and Asp residues. A line of similar behavior is observed in scaffolding protein alone (Figures 2 and 3). We note that amino acid analyses (Fuller & King, 1981, 1982) indicate 37 Asx and 51 Glx in scaffolding protein, and 52 Asx and 36 Glx in coat protein. Thus both proteins are relatively rich in the requisite acidic amino acids, although a breakdown into free acid and amidated forms of Asx and Glx is not available.

(3) 1⁻ Empty Procapsid (gp5). Raman spectra of the 1⁻ empty procapsid in H_2O and D_2O solutions are very much like Raman spectra of the corresponding solutions of 1⁻ procapsid. There are, however, a few important differences between the Raman spectra of the procapsid and empty procapsid, as next described.

First, the procapsid displays a Raman line at 938 cm^{-1}, but no such line is found in the empty procapsid. Second, the Raman spectrum of the procapsid contains a broad but discernible shoulder at approximately 1275 cm^{-1}, while diminished intensity at this frequency is evident in the spectrum of the empty procapsid. Third, the Raman lines due to aromatic amino acid side chains, viz., 1004 (Phe), 1010 (Trp), 1208 (Phe and Tyr), 1554 (Trp), 1586 (Trp), 1607 (Phe and Tyr) and 1619 cm^{-1} (Trp), are weaker in the procapsid than in the empty procapsid, when normalized to the intensity of the 1450 cm^{-1} Raman line in each respective spectrum. [These intensity differentials are not all evident from simple visual examination of the data of Figure 2. The rationale for use of the 1450 cm^{-1} line as the intensity standard has been given by Fasman et al. (1978).]

The first two differences noted above indicate that the procapsid is richer in α-helical structure than is the empty procapsid. The last difference noted above implies that the average protein subunit of the procapsid (approximately 200 gp8 + 420 gp5 subunits) contains a lower ratio of aromatic to aliphatic amino acid residues than the average protein

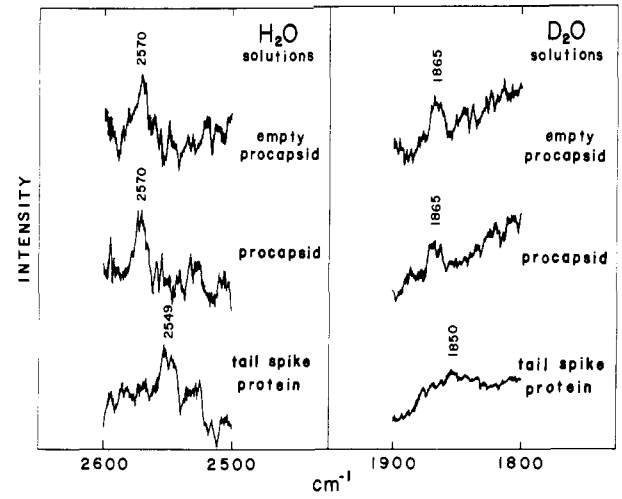

FIGURE 4: SH and SD stretching frequencies from Raman spectra of H_2O (left) and D_2O (right) solutions, respectively, of empty procapsids (gp5), procapsids (gp5 + gp8), and tail-spike protein (gp9). Sample conditions are as in Figure 2. Instrument conditions are as in Figure 2 except that $r = 25$ cm^{-1} min^{-1}, $\tau = 10$ s, and amplification is 3 times that of Figure 2.

subunit of the empty procapsid (approximately 420 gp5 subunits; Table I). In other words, all of the Raman data consistently indicate that the coat protein is richer in aromatic amino acids than is the scaffolding protein, in accord with preliminary amino acid composition analyses (Fuller & King, 1981).

The intensity ratio R_{Tyr} of the tyrosine doublet in the Raman spectrum of the empty procapsid is, within experimental error, the same as that of the procapsid. Another interesting feature is that both procapsid and empty procapsid have –SH and –SD vibrational Raman lines at 2570 and 1865 cm^{-1}, respectively (Figure 4), revealing cysteinyl groups in similar states in the two particles. However, the –SH (or –SD) line intensity is stronger in empty procapsids than in procapsids, confirming that cysteinyl groups are prevalent in coat protein but not in scaffolding protein.

(4) 10⁻ Empty Capsid (gp5 + Minor Proteins). The Raman spectra of 10⁻ empty capsids in H_2O (Figure 2) and D_2O (Figure 3) solutions are virtually identical with Raman spectra of corresponding solutions of 1⁻ empty procapsids at 32 °C. The near identical Raman scattering frequencies and intensities for the two kinds of particles have the following significance. First, with regard to *conformation insensitive* bands, it confirms the very similar average protein compositions (Table II) of empty capsid and empty procapsid. Second, with regard to *conformation sensitive* bands, it demonstrates that gp5 subunits have virtually the same secondary structures in both

empty procapsids and empty capsids. Therefore, despite possible differences in overall dimensions or shapes of the two kinds of particles (Fuller & King, 1980, 1981, 1982) and despite their different functional roles, the protein subunits manifest no significant differences in secondary structure that can be detected by Raman spectroscopy.

It is important to stress here the limit of precision of the Raman measurements. Although no absolute limit can be calculated, we estimate this limit as well within 10% on the basis of the fact that secondary structure changes of 10% or more are readily detected by changes in amide I and amide III intensities when the signal-to-noise ratio is 20:1 or greater (Thomas et al., 1981; Thomas & Day, 1981). We note also that the total weight percent composition of minor proteins in 10⁻ empty capsids is 9.5% (Table II), which is just below the limit of Raman detectability by the above criterion.

(*5*) *Tail-Spike Protein* (*gp9*). The Raman spectrum of the tail-spike protein in H_2O solution (Figure 2) shows a number of features distinguishing it from spectra of both the major coat protein and the scaffolding protein. These are as follows: (a) low intensity in the 1050–1120-cm^{-1} interval, in particular the absence of clearly resolved peaks at 1106 and 1083 cm^{-1}; (b) high intensity of bands of aromatic residues relative to the 1450 cm^{-1} band; (c) little or no Raman scattering on the high frequency side of the 758-cm^{-1} (Trp) band; (d) a reversal of intensities of the pair of lines at 647 and 622 cm^{-1} vis-à-vis their counterparts in spectra of capsids; (e) a moderately intense Raman line at 938-cm^{-1} intermediate in intensity between the cases of scaffolding protein and capsids; (f) a value of $R_{Tyr} = 1.2 \pm 0.2$, considerably smaller than in capsids but similar to the value in scaffolding protein; (g) a very strong amide III band at 1238 cm^{-1} with a shoulder at approximately 1275 cm^{-1}; (h) a strong and fairly sharp amide I band at 1666 cm^{-1}.

The Raman spectrum of tail-spike protein in D_2O solution shows some remarkable differences when compared with D_2O solution spectra of scaffolding protein and capsids (Figure 3). Foremost is the resistance of amide bands to deuterium exchange. Thus, the spectrum of "deuterated" tail-spike protein (Figure 3) displays the same amide I frequency as the H_2O solution spectrum (Figure 2). Also, the tail-spike protein shows relatively little intensity in the amide III′ region (900–1000 cm^{-1}; Figure 3) as compared, for example, with scaffolding protein. Obviously very little, if any, deuterium exchange of peptidyl –NH– groups occurs upon simply dissolving the tail-spike protein in D_2O. Moreover, samples of gp9 which remained in D_2O for 6 months at 10 °C or which were heated to 80 °C for several hours likewise showed no further progress toward deuterium exchange of peptidyl –NH– groups. Finally, heating gp9 for prolonged periods (up to 70 °C for 72 h) also had no significant impact upon deuterium exchange.

What is the secondary structure of the tail-spike protein which is so highly resistant to deuterium exchange? The appearance of amide I at 1666 cm^{-1} and amide III at 1238 cm^{-1} as strong and fairly sharp bands shows that the predominant secondary structure in the tail-spike protein is β-sheet structure (Chen & Lord, 1974; Yu et al., 1973). The weak shoulder at 1275 cm^{-1} in the amide III region and the skeletal mode at 938 cm^{-1} indicate also the presence of some α helix in gp9. However, it is clear from the sharpness of amide I at 1666 cm^{-1} and from the absence of a discernible shoulder near 1650 cm^{-1} that there is substantially more β-sheet structure than α-helix structure present in this protein.

Since deuterium exchange of the native structure of tail-spike protein is not feasible, the method of Lippert et al. (1976) cannot be used to estimate quantitatively the proportions of α, β, and disordered structures in gp9. Similarly, the method Pezolet et al. (1976) can be expected to provide an underestimate of β-sheet structure in the tail-spike protein because the amide III band is relatively broad when compared with model proteins examined by this method. Thus, from the data of Figure 2 and the procedure described by Pezolet et al. (1976), we obtain a lower limit of 34% as the fraction of β-sheet structure in tail-spike protein. We recognize this figure at best as an underestimate. Nevertheless, 34% is considerably higher than the estimates of β-sheet structure in scaffolding protein and coat protein by the same method. Further discussion is given in section B, below, and Table III.

Table III: Estimates of P22 Protein Secondary Structures from Raman Data[a]

species	α helix (%)	β sheet (%)	irregular structure (%)
10⁻ empty capsid (gp5)	20	25 (25)	55
1⁻ empty procapsid (gp5)	19	26 (25)	55
1⁻ procapsid (gp5 + gp8)	23	29 (25)	48
scaffolding protein (gp8)	*b*	(19)	
tail-spike protein (gp9)		(>34)	

[a] Figures given in parentheses are computed by the method of Pezolet et al. (1976). Other figures are computed by the method of Lippert et al. (1976). Uncertainties are estimated to be ±10% of each figure cited. [b] No quantitative estimate can be made from the available data, but comparison with other proteins suggests gp8 is predominantly α helical. See text.

Sulfhydryl groups of cysteine residues are revealed in the tail-spike protein by a Raman line at 2549 cm^{-1} (Figure 4). The relative intensity of this line indicates that the Cys content is comparable to that of the major coat protein; however, the lower –SH frequency suggests a different environment or different bonding characteristics than in the coat protein. The shift of the Cys band to 1850 cm^{-1} in D_2O solution indicates that the sulfhydryl groups, unlike amide groups, are readily accessible to solvent molecules.

(*6*) *P22 DNA*. The DNA molecule of the mature P22 phage particle is linear and double stranded with a molecular weight of approximately 2.7×10^7 (Rhodes et al., 1968). The P22 DNA investigated was high molecular weight material isolated by phenol extraction of purified phage particles (Tye et al., 1974).

The P22 DNA is similar in base composition and secondary structure to calf thymus DNA, which has been extensively studied by a variety of techniques, including Raman spectroscopy (Erfurth et al., 1972; Erfurth & Peticolas, 1975). Therefore, as expected, the Raman spectrum of P22 DNA (Figure 2) is similar to that of aqueous calf thymus DNA which assumes the B-form geometry (Arnott, 1975).

Some of the more prominent features of the Raman spectrum of P22 DNA are the following:

(a) The band at 1094 cm^{-1} in the H_2O solution spectrum and the corresponding band at 1092 cm^{-1} in the D_2O solution spectrum (Figures 2 and 3) are both assigned to the OPO^- symmetrical stretching mode, the intensities of which do not change significantly with backbone structure. The OPO^- band may therefore be used as an intensity standard for comparison with other bands of DNA in the Raman spectrum.

(b) The broad band at 834 cm^{-1} and the intense band at 787 cm^{-1} are typical of B-DNA. The former is due to a vibration of the deoxyribose phosphate backbone (coupled O–P and C–O stretching of the 5′-C–O–P–O–C-3′ network). The latter is believed due to an overlapping ring vibration of cystosine (C) and a mode of the DNA backbone (Lord &

Thomas, 1967; Thomas, 1970; Small & Peticolas, 1971). In D_2O, the former band appears at 832 cm^{-1}, while the latter is apparently split into two bands at 776 (C) and 791 cm^{-1} (backbone).

(c) The Raman bands at 731, 1302, and 1341 cm^{-1} are due to ring stretching modes of adenine (A). In D_2O, these bands appear at 724, 1306, and 1350 cm^{-1}, respectively. The Raman bands at 681 and 1376 cm^{-1} are due to ring stretching modes of guanine (G). In D_2O, these bands appear at 684 and 1378 cm^{-1}, respectively. The bands at 1489 and 1578 cm^{-1} are due to both A and G and are little affected by deuteration.

(d) The Raman band at 752 cm^{-1} is due to a ring stretching mode of T. In D_2O, this band appears at 741 cm^{-1}. T also contributes to the 1376-cm^{-1} band noted above.

The DNA assignments are based primarily on the Raman studies of nucleotides (Lord & Thomas, 1967) and polynucleotides (Small & Peticolas, 1971; Lafleur et al., 1972) which permit identification of the contributions of the DNA bases and backbone to the Raman spectrum of DNA. A similar approach has been followed for chicken erythrocyte DNA in nucleosome particles (Thomas et al., 1977).

(7) Mature Phage Particles (gp5 + gp9 + Minor Proteins + DNA). The Raman spectrum of P22 phage in either H_2O (Figure 2) or D_2O (Figure 3) solution is dominated by the Raman lines of DNA inside the phage capsid. The DNA contributions to the spectrum of the phage are shown by the connecting solid lines in Figures 2 and 3. The origins of these Raman frequencies have been discussed in the preceding section.

The prominent Raman lines at 833 and 786 cm^{-1} (Figure 2) and their intensities relative to the prominent 1092-cm^{-1} band show conclusively that P22 DNA within the phage capsid adopts the conventional B-DNA secondary structure. This conclusion is confirmed also by the Raman data for D_2O solutions (Figure 3).

Most of the protein contributions to the Raman spectrum of P22 are either masked or overlapped by the DNA contributions, except for the lines at 1004 (Phe) and 1450 cm^{-1} (CH_2 def), which have been discussed earlier. There are no significant changes in the Raman spectrum of P22 phage, as compared with the separate spectra of DNA and empty capsids, which would suggest significant changes in their respective structures in the assembled virion. Furthermore, we find no evidence of specific protein (gp5 or otherwise) interaction with nucleic acid such as is evident in Raman spectra of DNA–polylysine (Prescott et al., 1976), DNA–polyarginine (Thomas et al., 1980), ribosomes (Thomas et al., 1980), and tobacco mosaic virus (Fish et al., 1981).

(B) Estimation of Protein Secondary Structures. Several methods have been described to estimate protein secondary structure from Raman spectra (Lippert et al., 1976; Pezolet et al., 1976; Lord, 1977; Williams et al., 1980; Williams & Dunker, 1981). The methods of Lippert et al. (1976) and Williams et al. (1980) require deuterium exchange of amide –NH– groups and are therefore inapplicable to tail-spike protein which does not exchange and poorly suited to scaffolding protein which exchanges incompletely. The method of Pezolet et al. (1976) requires knowledge of the average number of $-CH_2-$ groups per residue of the protein and is limited to estimation of β-sheet content. The method of Williams & Dunker (1981) requires deconvolution of the amide I band in terms of a set of hypothetical standard spectra which are unavailable for this work. All of these methods utilize certain general correlations of Raman amide band intensities and frequencies with peptide chain conformations, which are summarized as follows:

type of structure	frequency interval (cm^{-1}) and intensity of amide band	
	amide I	amide III
α helix	1645–1655 strong	1270–1310 weak
β sheet	1665–1672 strong	1230–1243 strong
irregular structure	1660–1670 strong, broad	1243–1270 medium, broad

We applied the method of Lippert et al. (1976) to procapsids (gp5 + gp8), empty procapsids (gp5), and empty capsids (gp5), with the understanding that incomplete exchange of scaffolding protein (gp8) will reduce the reliability of results obtained from procapsids. We also applied the method of Pezolet et al. (1976) to aqueous solutions of scaffolding protein and tail-spike protein. The results are summarized in Table III.

In computing the average number of $-CH_2-$ groups per amino acid residue of tail-spike protein, the amino acid composition was obtained from the recently determined sequence of the DNA gene 9 (R. T. Sauer, W. Krovatin, A. R. Poteete, and P. B. Berget, personal communication), which is in substantial agreement with the data of Berget &Poteete (1980), except for Phe residues (4.0 mol %) that were underestimated (2.4%) in the latter study. This is in accord with the observed Raman intensity at 1004 cm^{-1}, which suggests a much higher mole percentage of Phe in gp9 than 2.4%, particularly when compared with spectra of scaffolding protein, capsids, procapsids (Figures 2 and 3), and other proteins of known Phe content (Thomas & Kyogoku, 1977; Frushour & Koenig, 1975).

(C) Temperature Dependence of Raman Spectra. A number of DNA and RNA viruses are known to undergo structural transitions promoted by changes in temperature that are conveniently studied by spectroscopic techniques. Such alteration of virus structure could be due to rearrangement in the organization of protein subunits, to differences in interaction between coat protein and packaged nucleic acid, or to changes in intramolecular conformations. For example, Raman spectra show that the coat proteins of filamentous bacteriophages Pf3 and Xf undergo large-scale transitions from α helix to β sheet without apparent dissociation from encapsulated DNA as the solution temperature is increased above 30 °C (Thomas et al., 1981; Thomas & Day, 1981). Similarly, fiber X-ray diffraction data show a contraction in the length of the filamentous phage Pf1, also without dissassembly of the virion, as the temperature of the fiber is decreased below 8 °C (Nave et al., 1979). Changes in protein and RNA secondary structures with temperature are also apparent in the tobacco mosaic virus (G. J. Thomas, Jr., unpublished results). On the other hand, the RNA bacteriophages R17 and MS2 are relatively resistant to structural change with temperature (Thomas et al., 1976).

We have investigated the temperature-dependence of Raman spectra of P22 mature phage, capsids, and procapsids, as well as available isolated proteins, in order to determine whether changes in structure or subunit interaction occur selectively in one or another assembly state. Since the Raman spectra are sensitive to *changes* in protein structure and interactions, the temperature-dependent spectra are a potential source of information on whether putative interactions between coat and scaffolding proteins are disrupted as procapsids are heated. Also, differences in the thermal behavior of coat or scaffolding protein when isolated from procapsids vis-à-vis assembled procapsids would provide indirect evidence for specific interactions between gp5 and gp8.

(1) I⁻ Empty Procapsid. Raman spectra of empty procapsids undergo rather slight changes with temperature, no-

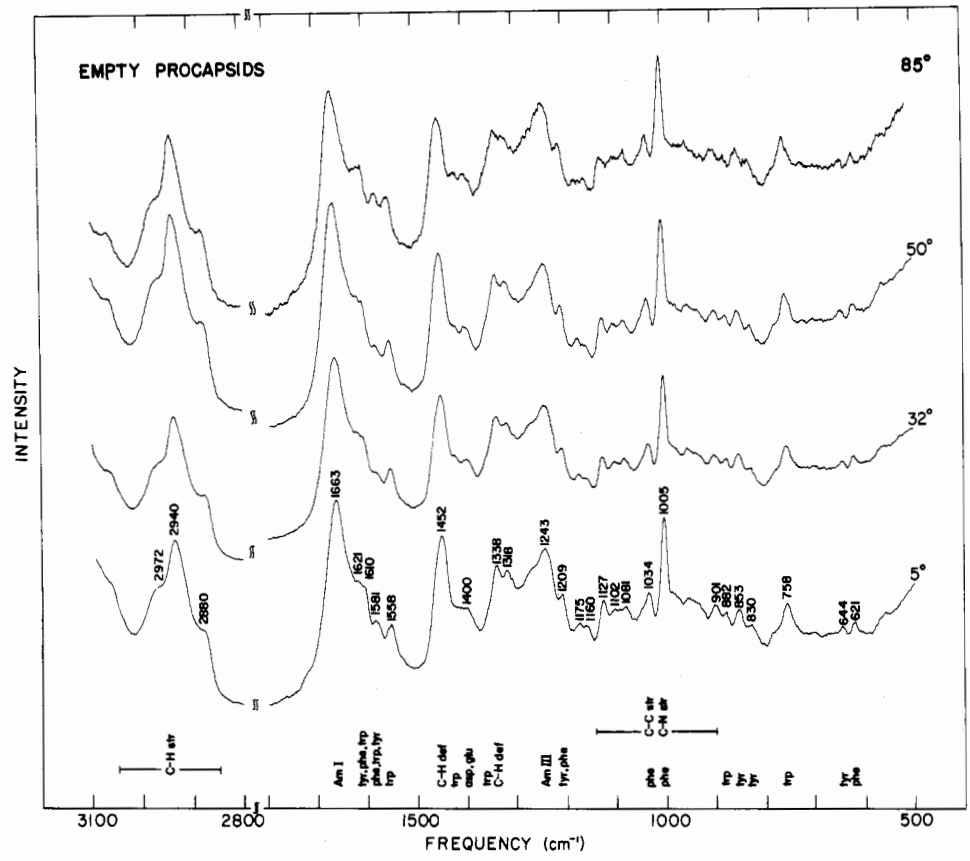

FIGURE 5: Raman spectra of H_2O solutions of empty procapsids (approximately 10% w/v in 10 mM Tris, pH 7.5) at 5, 32, 50, and 85 °C. Other conditions are as in Figure 2.

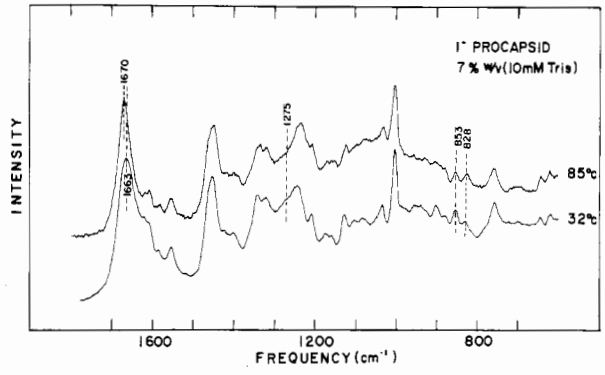

FIGURE 6: Raman spectra of H_2O solutions of procapsids at 32 and 85 °C. Broken vertical lines illustrate some of the significant changes in Raman intensities which result from heating. Other conditions are as in Figure 2.

tably above 50 °C, as shown in Figure 5. The amide I mode, for example, shifts only from 1663 (at 5, 32, and 50 °C) to 1667 cm^{-1} (at 85 °C). Similarly, the shoulder (ca. 1275 cm^{-1}) to the amide III band diminishes only slightly in intensity above 50 °C. Figure 5 also shows a very small decrease in the ratio of intensities of the lines of the tyrosine doublet (I_{853}/I_{830}). These features indicate at most a marginal change in the strength of hydrogen bonding of hydroxyl groups of tyrosine residues and possibly a very small decrease in the amount of α-helical secondary structure of gp5 with increasing temperature. The small changes in the spectrum above 50 °C are reversed upon cooling the sample back to lower temperatures.

(2) I⁻ Procapsid. The temperature dependence of the Raman spectrum of procapsids contrasts sharply with the empty procapsids, as evident in Figure 6. Between 32 and 85 °C we find the following reversible changes: (a) the amide I peak shifts from 1663 to 1670 cm^{-1} and the band width

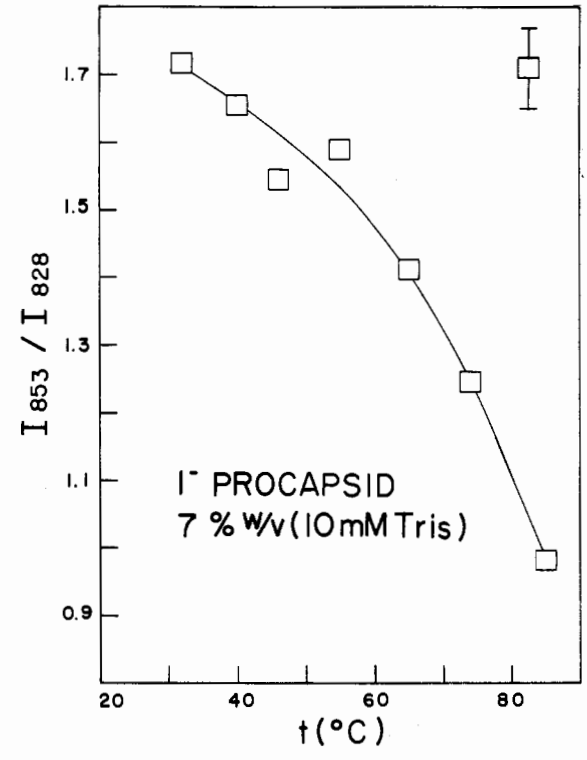

FIGURE 7: Plot of Raman intensity ratio $R_{Tyr} = I_{853}/I_{828}$ vs. temperature over the range 0–85 °C, from spectra of I^- procapsids (7% w/v in 10 mM Tris, pH 7.6).

diminishes significantly; (b) the amide III shoulder at 1275 cm^{-1} diminishes greatly in intensity; and (c) the intensity ratio I_{853}/I_{828} decreases substantially (Figure 7). These results show that the protein structures in the procapsid (i.e., gp5 + gp8)

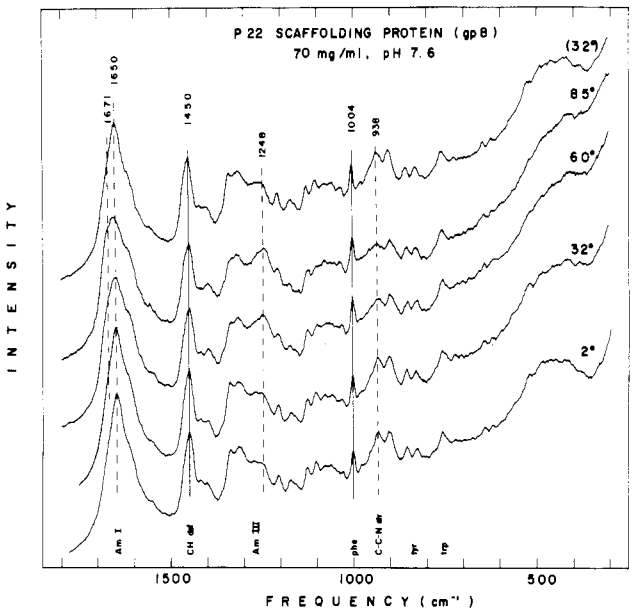

FIGURE 8: Raman spectra of H_2O solutions of scaffolding protein at 2, 32, 60, and 85 °C. Solid vertical lines connect the 1004- and 1450-cm^{-1} bands which are invariant to temperature. Broken vertical lines illustrate some of the significant changes in Raman intensities which result from heating. The top spectrum, recorded after the heated sample was cooled back to 32 °C, illustrates reversibility. Other conditions are as in Figure 2.

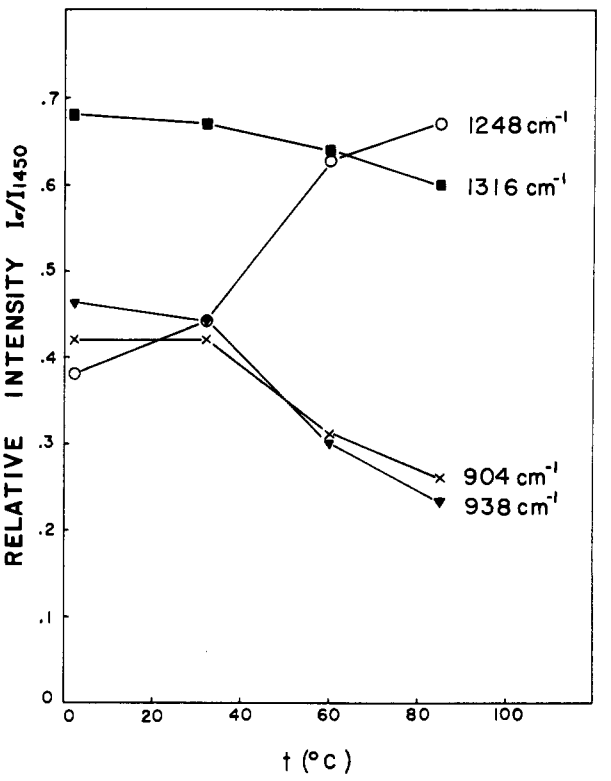

FIGURE 9: Plots of Raman intensity ratios I_σ/I_{1450} vs. temperature over the range 2–85 °C, from spectra of scaffolding protein (7% w/v in 10 mM Tris, pH 7.6). Assignments are 904 (CC stretching of side chains), 938 (α-helical backbone), 1248 (irregular backbone, amide III), and 1316 cm^{-1} (overlapping side-chain modes and α-helical backbone, amide III).

are more sensitive to temperature change than is the structure of gp5 in the empty procapsid. Moreover, the Raman data show that the overall change in the "average" molecule of the procapsid is from a more α-helical to a less α-helical secondary structure. Attendant with this change in protein conformation is a net transfer of tyrosyl –OH functions from states of hydrogen bonding with donor groups to hydrogen bonding with acceptor groups. Comparison of the empty procapsid with the procapsid suggests that the scaffolding protein molecules of the latter are responsible for the temperature-sensitive features of the Raman spectrum. It is therefore of interest to investigate the temperature dependence of the Raman spectrum of scaffolding protein directly, as discussed next.

(*3*) *Scaffolding Protein.* Raman spectra of scaffolding protein in H_2O, recorded at 2, 32, 60, and 85 °C and then after cooling back to 32 °C, are shown in Figure 8. It is evident from Figure 8 that the spectrum of gp8 is highly sensitive to the temperature. The most pronounced changes in the Raman spectrum with heating to 85 °C are the intensity decreases in bands centered at 904, 938, and 1316 cm^{-1} and the intensity increase in the band centered at 1248 cm^{-1}. These results are depicted in Figure 9. In accordance with the band assignments discussed earlier (Section A, part 1, and Figure 2), the data show a loss of α-helical secondary structure and a corresponding increase of irregular structure in gp8 with increasing temperature. The structure change is gradual and appears to be reversible since the original spectrum is recovered after returning the heated sample to 32 °C. These results on gp8 provide strong evidence that the temperature sensitivity of the Raman spectrum of the procapsid is due to the inner shell of scaffolding protein.

In addition to the clear-cut intensity changes plotted in Figure 9, a number of other noticeable changes occur in the Raman spectrum of gp8 upon elevating the temperature. The amide I band, originally centered at 1650 cm^{-1} (32 °C), develops a prominent shoulder at 1671 cm^{-1} (85 °C). This is consistent with the conversion of α helices to other (presumably disordered) chain configurations. Also with increasing temperature, Raman intensity decreases occur at 521, 1173, 1342, and 1424 cm^{-1} and increases occur in lines of aromatic side chains above 1200 and above 1600 cm^{-1}. We note, however, no significant change in the tyrosine doublet (I_{857}/I_{830}) with temperature. This is surprising in view of the results obtained on the 1^- procapsid (Figure 7, above).

Conclusions

P22 and Precursor Particles: Structure and Stability. The Raman spectrum of P22 phage is dominated by Raman lines of the encapsulated P22 DNA molecule. These lines exhibit frequencies and intensities which are identical with those of DNA extracted from the phage capsid and indicate the B-form secondary structure of P22 DNA in both the extracted and packaged states. Therefore, any tertiary structure or supercoiling of the DNA molecule which may exist within the phage capsid does not alter the normal B-DNA secondary structure which prevails for the protein-free P22 DNA molecule in aqueous solution. Since the base composition of P22 DNA is similar to that of calf thymus DNA or chicken erythrocyte DNA, its Raman spectrum is also similar to spectra of these other DNAs in the B-form secondary structure (Erfurth et al., 1972; Prescott et al., 1976; Thomas et al., 1977).

The Raman spectrum of P22 also reveals lines of the phage structural proteins. However, these lines are of feeble intensity in comparison to the lines of DNA, even though the phage particle contains approximately 50.6 wt % protein. This is consistent with the well-known fact that proteins are typically weak Raman scatterers in comparison to nucleic acids (Thomas, 1976a). In the spectrum of P22 phage all of the Raman lines not attributable to DNA are assignable to the major coat protein, gp5, which constitutes 42.2 wt % of the virion. The frequencies and intensities of the Raman lines of

Table IV: Interpretation of the Tyrosine Doublet in Raman Spectra of P22 Proteins

particle (subunit)	tyrosine doublet ratio I_{850}/I_{830}[a]	temp dependence	interpretation[c]
10⁻ empty capsid (gp5)	1.7	none	Tyr OH groups are net acceptors of H bonds at all temperatures
1⁻ empty procapsid (gp5)	1.7	none	Tyr OH groups are net acceptors of H bonds at all temperatures
1⁻ procapsid (gp5 + gp8)	1.6	decreases to 1.0 at 85 °C[b]	Tyr OH groups are net acceptors of H bonds at low temperature and net donors of H bonds at high temperature
scaffolding protein (gp8)	1.2	none	Tyr OH groups are both donors and acceptors of H bonds at all temperatures
tail-spike protein (gp9)	1.2	none	Tyr OH groups are both donors and acceptors of H bonds at all temperatures

[a] Values cited are from spectra at 32 °C and are accurate to ±10%. [b] See Figures 6 and 7. [c] See Siamwiza et al. (1975).

gp5 subunits in the mature phage particle are in accord with those of gp5 subunits in the derived particles that lack both DNA and substantial amounts of other proteins, viz., phage empty capsids and empty procapsids.

Raman spectra of empty procapsids and empty capsids are virtually identical with one another, indicating first that the Raman spectra of both particles are dominated by gp5 and second that the secondary structure of gp5 is essentially the same in both particles. Since the empty procapsids were prepared by treatment of procapsids with 1.0 M Gdn·HCl in vitro, and since the empty capsids were prepared in vivo as a consequence of a defect in the assembly pathway, it is of interest to note that two such radically different procedures for producing empty shells of gp5 yield nevertheless the same subunit structure insofar as can be detected by Raman spectroscopy. The results of biochemical studies which indicate that empty procapsids and empty capsids have no major differences in morphology or dimension (Fuller & King, 1980) are thus in accord with the present results which indicate no major structural differences at the molecular level. The recent suggestion (Fuller & King, 1982) that Gdn·HCl-treated empty procapsids are morphologically different than empty capsids is not inconsistent with the Raman data provided that any differences in shape or dimension are not due to major differences in protein secondary structures in the two kinds of shells.

Raman spectra of procapsids manifest significant differences from spectra of either empty procapsids or empty capsids. These differences must be due directly or indirectly to the presence of scaffolding protein, gp8, which constitutes 25.4 wt % of the procapsid. In comparing Raman spectra of procapsids with spectra of either empty procapsids or empty capsids, we assume that coat protein molecules make the same contribution to each spectrum. This assumption is supported by the high stability of coat protein structure as evidenced by the invariance of its Raman spectrum to Gdn·HCl treatment, heat treatment, and so forth, as discussed earlier. With this assumption, the Raman spectra of procapsids allow the conclusion that scaffolding protein molecules within the procapsid are considerably richer in α-helical secondary structure than are the gp5 molecules of the outer protein coat of the procapsid. As noted above, the scaffolding protein is also shown by Raman spectroscopy to be much poorer than coat protein in aromatic amino acid residues.

An alternative explanation of the present findings is that scaffold (inner) and coat (outer) shells of the procapsid interact with one another in such a way that the coat protein molecules also undergo a transition to a more α-helical state in the procapsid. However, we consider this alternative unlikely in view of the aforementioned stability of the coat protein and in view also of the Raman data obtained on scaffolding protein molecules isolated from procapsids: The Raman spectra of purified scaffolding protein show conclusively that this protein assumes a much more highly α-helical secondary structure than the coat protein in aqueous solution. The most straightforward interpretation of the Raman results on procapsids is, therefore, that the scaffolding protein retains its highly α-helical conformation in the procapsid, while the coat protein retains its characteristic irregular conformation. We propose, then, no significant changes in the secondary structures of either coat or scaffolding proteins attendant with their "association" in the procapsid nor in the coat protein in its different states of aggregation (phage, empty capsid, and empty procapsid).

Earlier studies (Fish et al., 1980; Li et al., 1981) have shown the Raman spectrum of P22 to be invariant to temperature over the range 0–85 °C, confirming the thermal stability of the packaged DNA molecule and phage capsid. Moreover, the rather small changes which occur in the Raman spectra of empty capsids (Li et al., 1981) and empty procapsids (Figure 5) over the same temperature interval indicate that the secondary structure of the coat protein molecules in these assembled shells is not highly vulnerable to increasing temperature. Therefore, the very substantial changes in the Raman spectrum of the procapsid which result from increasing the temperature (Figure 6 and 7) can be attributed to the inner shell of scaffolding subunits (gp8). The thermal instability of the gp8 secondary structure is confirmed also by Raman data of purified preparations of gp8 (Figures 8 and 9). All of the Raman spectral changes of gp8, either within the procapsid or in the purified aqueous state, point to a substantial loss of α-helical secondary structure with increasing temperature.

On the other hand, the marked sensitivity to temperature of tyrosine hydrogen-bonding interactions in procapsids (Figure 7) is surprisingly not replicated in either empty procapsids (Figure 5) or in isolated gp8 (Figure 8). The data on tyrosine group Raman bands and their interpretation are summarized in Table IV. We conclude from these results the following: In purified aqueous gp8, and at all temperatures (0–85 °C), comparable numbers of tyrosyl OH groups act as donors and acceptors of hydrogen bonds. This is as expected if tyrosyl OH groups are generally accessible to solvent molecules. In empty shells of coat protein, at all temperatures, more tyrosyl OH groups act as acceptors than as donors of H bonds, and therefore not all tyrosyl OH groups can be accessible to solvent molecules. This is consistent with the thermal stability of the coat protein. Finally, in procapsids at lower temperatures, tyrosyl OH groups appear to be well represented as a simple sum of scaffolding and coat protein groups; however, the tyrosine hydrogen-bonding characteristics change gradually with increasing temperature, so that at 85 °C the *average* tyrosyl OH group of the procapsid is a net donor of a hydrogen bond to a negative acceptor group, contrary to what is expected

for a sum of gp8 and gp5. Evidently the structure of gp8 within the procapsid does not respond to increasing temperature in the same way as the structure of gp8 extracted from the procapsid, at least with regard to the capacity of tyrosyl OH groups to enter into hydrogen-bonding interactions with neighboring protein subgroups and solvent molecules. Thus, in the procapsid, specific tyrosine side-chian interactions among scaffolding subunits or between scaffolding and coat protein subunits are indicated by the Raman spectra. Nevertheless, we find no evidence that the overall main-chain conformation (secondary structure) of either scaffolding or coat protein molecules is altered significantly by any associative interactions between or among subunits of the capsid.

Structural Characteristics of the Viral Proteins. Of the viral proteins examined here, the tail-spike protein is the most stable (Goldenberg & King, 1981). The Raman spectrum gives evidence of substantial β-sheet structure. A lower limit of 34% β-sheet structure is computed by the method of Pezolet et al. (1976), but gp9 almost certainly contains more than 34% β sheet by virtue of the fact that amide I and amide III frequencies and intensities are close to those observed for the β-polypeptide poly(L-lysine) (Yu et al., 1973) and for the β-sheet structures of filamentous phage proteins (Thomas et al., 1981; Thomas & Day, 1981). Unfortunately, a more reliable quantitative estimate cannot be obtained from the present Raman data because of the unavailability of spectra of deuterated and denatured forms of gp9.

The fact that the tail-spike maintained in D_2O solution for 6 months at 10 °C, or heated in D_2O for 72 h at 70 °C, failed to show any appreciable deuterium exchange of amide NH groups is evidence of an extremely rigid secondary structure. Efforts to denature gp9 by heating to 85 °C (the highest temperature at which satisfactory Raman signal could be monitored) also proved unsuccessful. Evidently the function of the P22 tail requires a highly stable and rigid protein subunit. We envision the tail fiber as a pleated sheet structure in accord with the Raman data of gp9.

In the case of the procapsid (gp5 + gp8), the method of Lippert et al. (1976) yields an estimate of 23% α-helical secondary structure. However, we believe this also to be at best a lower limit in view of the obviously incomplete deuteration of gp8 in D_2O solution. Incomplete deuteration would have the consequence of underestimating α-helix and overestimating irregular structures in the Lippert et al. (1976) procedure. Comparison with the very highly α-helical proteins of filamentous DNA bacteriophages (Thomas & Murphy, 1975; Thomas & Day, 1981) indeed suggests far more than 23% α-helix in the procapsid. Therefore the present estimate (Table III) should be regarded only as a lower limit.

Amino acid analyses of coat, scaffolding, and tail-spike proteins have provided no information on the presence or absence of cysteine and tryptophan residues (Fuller & King, 1981; Berget and Poteete, 1980). However, such information can be obtained from Raman spectra since characteristic Raman frequencies are expected from Trp (ca. 758 cm^{-1}) and Cys (ca. 2550–2575 cm^{-1}) residues. The present study thus demonstrates the presence of Trp and Cys residues in coat and tail-spike proteins. [Cys and Trp are confirmed in gp9 by the gene sequence (R. T. Sauer, W. Krovatin, A. R. Poteete, and P. B. Berget, personal communication).] Trp residues are also shown to be present in scaffolding protein, but no Cys residues can be detected in gp8 by Raman spectroscopy.

The relative abundance of Trp residues in coat (gp5), scaffolding (gp8), and tail-spike (gp9) proteins can be estimated from the Raman intensity ratios I_{758}/I_{1450} in spectra of Figure 2. We find I_{758}/I_{1450} to be in the ratio 1.5:1.0:1.7 for gp5, gp8, and gp9, respectively. Therefore, gp9 is richest and gp8 is poorest in Trp. Moreover, the low intensity of the Trp 870-cm^{-1} line (Kitagawa et al., 1979) and the diffuseness of the Trp 1350-cm^{-1} line (Yu, 1974) indicate that Trp residues in all three proteins are generally exposed to hydrophilic environments, probably solvent molecules.

A similar Raman estimate of the relative abundance of Phe residues in gp5, gp8, and gp9 by use of the intensity quotient I_{1004}/I_{1450} yields the ratio 3.0:1.0:3.3, respectively. Again, the gp9 subunit is richest in Phe while gp8 is poorest. The same conclusion follows from consideration of the intensity ratio I_{624}/I_{1450}, although this is not evident from simple inspection of the low gain spectra of Figure 2. We note also that the Raman result of more Phe in gp5 than in gp8 is in agreement with the amino acid analyses of gp5 and gp8 (Fuller & King, 1981).

Finally, we note that the hypothesis of a hinge connecting two domains of the gp5 coat protein subunit (Casjens, 1979), similar to that demonstrated for the tomato bushy stunt virus (Harrison et al., 1978; Winkler et al., 1977), remains viable for P22. We have demonstrated directly that no major difference in secondary structure of gp5 occurs between empty capsid and empty procapsid, or indirectly between procapsid and empty procapsid. However, if the in vitro emptying of procapsids with 1 M Gdn·HCl produces the same 10% (diameter) expansion of the outer protein shell that is observed upon normal release of the scaffolding protein (Steven et al., 1976), and if such shell expansion is accompanied by flexing a relatively small segment of the gp5 peptide chain, it would not be readily detected with the precision of the present Raman experiments. Therefore, while the present results rule out major differences in average subunit conformation among all shells containing gp5, conformational changes affecting less than 10% of the protein backbone are clearly consistent with our data.

Acknowledgments

The expert technical assistance of Betty Prescott (Southeastern Massachusetts University) and Donna Smith (Massachusetts Institute of Technology) is gratefully acknowledged.

References

Arnott, S. (1975) *Nucleic Acids Res. 2*, 1493–1502.
Berget, P. B., & Poteete, A. (1980) *J. Virol. 34*, 234–243.
Blout, E. R., Deloae, C., & Asadourian, A. (1961) *J. Am. Chem. Soc. 83*, 1895–1900.
Botstein, D., Waddell, C., & King, J. (1973) *J. Mol. Biol. 80*, 669–695.
Casjens, S. (1979) *J. Mol. Biol. 131*, 1–19.
Casjens, S., & King, J. (1975) *Annu. Rev. Biochem. 44*, 555–611.
Chen, M. C., & Lord, R. C. (1974) *J. Am. Chem. Soc. 96*, 4750–4752.
Earnshaw, W., & Harrison, S. C. (1977) *Nature (London) 286*, 598–602.
Earnshaw, W., & Casjens, S. (1980) *Cell (Cambridge, Mass.) 21*, 319–331.
Earnshaw, W., Casjens, S., & Harrison, S. C. (1976) *J. Mol. Biol. 104*, 387–410.
Earnshaw, W., Hendrix, R., & King, J. (1979) *J. Mol. Biol. 134*, 575–594.
Erfurth, S. C., & Peticolas, W. L. (1975) *Biopolymers 14*, 247–264.
Erfurth, S. C., Kiser, E. J., & Peticolas, W. L. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 938–941.

Fasman, G. D., Itoh, K., Liu, C. S., & Lord, R. C. (1978) *Biopolymers 17*, 125–143.
Fish, S. R., Thomas, G. J., Jr., Fuller, M., & King, J. (1980) *Biophys. J. 32*, 234–237.
Fish, S. R., Hartman, K. A., Stubbs, G. J., & Thomas, G. J., Jr. (1981) *Biochemistry 20*, 7449–7457.
Frushour, B. G., & Koenig, J. L. (1974) *Biopolymers 13*, 455–474.
Frushour, B. G., & Koenig, J. L. (1975) *Adv. Infrared Raman Spectrosc. 1*, 35–97.
Fuller, M., & King, J. (1980) *Biophys. J. 10*, 381–401.
Fuller, M., & King, J. (1981) *Virology 112*, 529–547.
Fuller, M., & King, J. (1982) *J. Mol. Biol.* (in press).
Goldenberg, D. P., & King, J. (1981) *J. Mol. Biol. 145*, 633–651.
Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K., & Bricogne, G. (1978) *Nature (London) 276*, 368–373.
Itoh, K., Hinomoto, T., & Shimanouchi, T. (1974) *Biopolymers 13*, 307–312.
King, J., & Casjens, S. (1974) *Nature (London) 251*, 112–119.
King, J., Lenk, E., & Botstein, D. (1973) *J. Mol. Biol. 80*, 697–731.
Kitagawa, T., Azuma, T., & Hamaguchi, K. (1979) *Biopolymers 18*, 451–465.
Laemmli, U. K., Paulson, J., & Hitchins, V. (1974) *J. Supramol. Struct. 2*, 276–301.
Lafleur, L., Rice, J., & Thomas, G. J., Jr. (1972) *Biopolymers 11*, 2423–2437.
Li, Y., Thomas, G. J., Jr., Fuller, M., & King, J. (1981) in *Bacteriophage Assembly, Progress in Clinical and Biological Research* (Dubow, M., Ed.) Vol. 64, pp 271–283, Liss, New York.
Lippert, J. L. Tyminski, D., & Desmeules, P. J. (1976) *J. Am. Chem. Soc. 98*, 7075–7080.
Lord, R. C. (1977) *Appl. Spectrosc. 31*, 187–194.
Lord, R. C., & Thomas, G. J., Jr. (1967) *Spectrochim. Acta 23A*, 2551–2591.
Lord, R. C., & Yu, N.-T. (1970) *J. Mol. Biol. 50*, 509–524.
Murialdo, H., & Becker, A. (1978) *Microbiol. Rev. 42*, 529–576.
Nave, C., Fowler, A. G., Malsey, S., Marvin, D. A., Siegrist, H., & Wachtel, E. J. (1979) *Nature (London) 281*, 232–234.
Pezolet, M., Gosselin, M. P., & Coulombe, L. (1976) *Biochim. Biophys. Acta 453*, 502–512.
Poteete, A., & King, J. (1977) *Virology 76*, 725–739.
Prescott, B., Chou, C. H., & Thomas, G. J., Jr. (1976) *J. Phys. Chem. 80*, 1164–1171.
Rhodes, M., MacHattie, L. A., & Thomas, C. A., Jr. (1968) *J. Mol. Biol. 37*, 21–40.
Siamwiza, M. N., Lord, R. C., Chen, M. C., Takamatsu, T., Harada, I., Matsuura, H., & Shimanouchi, T. (1975) *Biochemistry 14*, 4870–4876.
Small, E. W., & Peticolas, W. L. (1971) *Biopolymers 10*, 1377–1416.
Steven, A., Couture, E., Aebi, U., & Showe, M. (1976) *J. Mol. Biol. 106*, 187–221.
Thomas, G. J., Jr. (1970) *Biochim. Biophys. Acta 213*, 417–423.
Thomas, G. J., Jr. (1976a) *Appl. Spectrosc. 30*, 483–494.
Thomas, G. J., Jr. (1976b) *Spex Speaker 21*, 1–11.
Thomas, G. J., Jr., & Barylski, J. R. (1970) *Appl. Spectrosc. 24*, 463–464.
Thomas, G. J., Jr., & Murphy, P. (1975) *Science (Washington, D.C.)* 1205–1207.
Thomas, G. J., Jr., & Kyogoku, Y. (1977) *Pract. Spectrosc. 1C*, 717–872.
Thomas, G. J., Jr., & Day, L. A. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2962–2966.
Thomas, G. J., Jr., Prescott, B., McDonald-Ordzie, P. E., & Hartman, K. A. (1976) *J. Mol. Biol. 102*, 103–124.
Thomas, G. J., Jr., Prescott, B., & Olins, D. E. (1977) *Science (Washington, D.C.) 197*, 385–388.
Thomas, G. J., Jr., Prescott, B., & Hamilton, M. G. (1980) *Biochemistry 19*, 3604–3613.
Thomas, G. J., Jr., Prescott, B., Day, L. A., & Boyle, P. D. (1981) in *Bacteriophage Assembly, Progress in Clinical and Biological Research* (Dubow, M., Ed.) Vol. 64, pp 271–283, Liss, New York.
Tye, B., Huberman, J., & Botstein, D. (1974) *J. Mol. Biol. 85*, 501–532.
Williams, R. W., & Dunker, A. K. (1981) *J. Mol. Biol. 152*, 783–813.
Williams, R. W., Cutrera, T., Dunker, A. K., & Peticolas, W. L. (1980) *FEBS Lett. 115*, 306–308.
Winkler, F. K., Schutt, C. E., Harrison, S. C., & Bricogne, G. (1977) *Nature (London) 265*, 509–513.
Yu, N.-T. (1974) *J. Am. Chem. Soc. 96*, 4664–4668.
Yu, T. J., Lippert, J. L., & Peticolas, W. L. (1973) *Biopolymers 12*, 2161–2176.