ELSEVIER

Biophysical Chemistry 68 (1997) 17–31

Biophysical Chemistry

Mechanisms of virus assembly probed by Raman spectroscopy: the icosahedral bacteriophage P22

Roman Tuma, George J. Thomas Jr *

Division of Cell Biology and Biophysics, School of Biological Sciences, University of Missouri, Kansas City, MO 64110, USA

Received 19 November 1996; accepted 13 December 1996

Abstract

A microdialysis flow cell has been developed for time-resolved Raman spectroscopy of biological macromolecules and their assemblies. The flow cell permits collection of Raman spectra concurrent with the efflux of small solute molecules into a solution of macromolecules and facilitates real-time spectroscopic detection of structural transitions induced by the effluent. Additionally, the flow cell is well suited to the investigation of hydrogen–isotope exchange phenomena that can be exploited as dynamic probes of viral protein folding and solvent accessibility along the assembly pathway. Here, we describe the application of the Raman dynamic probe to the maturation of the icosahedral capsid of bacteriophage P22, a double-stranded DNA virus. The P22 virion is constructed from a capsid precursor (procapsid) consisting of 420 coat subunits (gp5) in an outer shell and a few hundred scaffolding subunits (gp8) within. Capsid maturation involves expulsion of scaffolding subunits coupled with shell expansion at the time of DNA packaging. Raman static and dynamic probes reveal that the scaffolding subunit is highly α-helical and highly thermolabile, and lacks a typical hydrophobic core. When bound within the procapsid, the α-helical fold of gp8 is thermostabilized; however, this stabilization confers no apparent protection against peptide NH → ND exchange. A molten globule model is proposed for the native scaffolding subunit that functions in procapsid assembly. Accompanying capsid expansion, a small conformational change (α-helix → β-strand) is also observed in the coat subunit. Domain movement mediated by hinge bending is proposed as the mechanism of capsid expansion. On the basis of these results, a molecular model is proposed for assembly of the P22 procapsid. © 1997 Elsevier Science B.V. © 1997 Elsevier Science B.V.

Keywords: Virus structure; Assembly; Protein dynamics; Hydrogen-isotope exchange; Raman spectroscopy; Bacteriophage P22

1. Introduction

A formidable challenge in structural biology is the elucidation of pathways of supramolecular assembly. One frequently studied example is the morphogenesis of an icosahedral virus. Typically, the pathway of viral morphogenesis involves many steps, including the folding of subunit domains that participate in protein recognition processes, the generation of transient assembly intermediates, the packaging of a viral genome and the incorporation of a host attachment apparatus [1,2]. At the center of this scheme is the formation of a viral capsid or shell from multiple copies of identical protein subunits, a process which also serves as a model for the construction of other architecturally complex cellular assemblies. Yet, the structural details and dynamic principles governing

* Corresponding author.

PII S0301-4622(97)00001-X

viral capsid assembly remain largely unknown. Raman spectroscopy provides an ideal probe for investigating the molecular mechanisms of such assembly processes. The Raman spectrum is informative of the protein secondary structures and environments of key side chains involved in capsid formation and maturation. Additionally, the Raman method is not hindered by the size of the supramolecular assembly or by light scattering artifacts that may restrict the use of other spectroscopic probes [3].

The tasks of identifying and characterizing structural intermediates along a macromolecular assembly pathway can be greatly facilitated by use of a time-resolved spectroscopic approach. Recently, a Raman microdialysis flow cell device has been designed for this purpose [4–6]. The flow cell permits collection of Raman spectra concurrent with the efflux of small solute molecules into the macromolecular solution to facilitate real-time measurement of structural transformations induced by the effluent. The flow cell is particularly well suited to investigating hydrogen–isotope exchange phenomena that accompany replacement of solvent H_2O by D_2O (or vice versa). The kinetics of such exchanges serve as dynamic probes of assembly mechanisms. Accordingly, the Raman spectrum can be employed not only as a fingerprint of protein structure and intermolecular interaction, but also as a probe of macromolecular dynamics. Applications for the measurement of proton exchange kinetics in viral protein subunits and in the packaged genomes of isometric viruses have been described [4,6,7].

To enhance the value of the Raman dynamic probe, a quantitative understanding of the operation of the microdialysis flow cell is necessary. Here, we show that a simple diffusion model satisfactorily correlates the design and performance of the device. The validity of the diffusion model is confirmed by favorable comparison of calculated and experimental efflux rates in applications to representative biological macromolecules. Understanding the principles governing operation of the microdialysis flow cell has permitted novel applications of the Raman dynamic probe, specifically to address problems of viral capsid assembly and maturation. We report results obtained recently on key assembly intermediates of the double-stranded (ds) DNA virus, P22.

In most dsDNA viruses, including bacteriophages and animal viruses, the initial product of subunit polymerization is not the mature virion capsid but a precursor capsid, or procapsid, into which the DNA is subsequently packaged. Procapsid construction requires the involvement of scaffolding subunits that reside in the procapsid interior [8,9] and direct the polymerization of the coat protein into a properly dimensioned shell [10,11]. In phage P22, the procapsid comprises 420 molecules of a 47-kDa coat protein (product of viral gene 5, or gp5) surrounding an inner core of approximately 300 molecules of the 33-kDa scaffolding protein (gp8) [12]. The P22 scaffolding subunits exit the procapsid shell intact at or near the time of DNA packaging to be recycled for further rounds of assembly. The P22 procapsid shell undergoes a concurrent transformation to the mature capsid lattice, which involves a substantial conformational change in the gp5 subunits [12,13]. This transformation results in a 10% increase in the internal volume, which is sufficient to package the P22 genome and yield a more robust particle. The procapsid-to-capsid transition is tightly controlled in vivo and is apparently triggered by the presence of the viral DNA [14,15]. We have undertaken a detailed study of the structural basis of the temperature induced transformations of P22 particles using methods of static and time-resolved Raman spectroscopy. In this paper, we characterize the structure and thermostability of the scaffolding protein, both when isolated from and when contained within procapsids. On the basis of the present results, we propose an assembly mechanism that involves coupling between release of the scaffolding protein from the procapsid and unfolding of its secondary structure. The data also identify interactions that may be important in retention of scaffolding subunits within the viral procapsid. A model for the organization of scaffolding subunits within the procapsid is proposed.

2. Materials and methods

2.1. Preparation of procapsids, empty shells and scaffolding protein

Procapsids, empty shells and scaffolding protein were prepared as described previously [16]. *Salmonella typhimurium* strain DB7136 was grown

to $\approx 4 \times 10^8$ cells ml^{-1} in superbroth and infected with P22 strain $2^-amH200/13^-amH101$ at moi = 5 and grown at 37°C for 4 h. The cells were chilled on ice, harvested by low speed centrifugation and frozen in 50 mM Tris, 20 mM $MgCl_2$. The cells were lysed by 3 freeze–thaw cycles in the presence of saturating quantities of chloroform and 1 mM PMSF (phenylmethylsulfonyl fluoride). DNAse was added to a final concentration of 100 μg ml^{-1} and digestion was continued for 30 min at 37°C. Cell debris was removed by centrifugation, and the pellet was resuspended in an equal volume of 100 mM EDTA, pH 7.6, frozen and thawed. The resuspended pellet was centrifuged at 11 000 rpm for 15 min in a Ti-45 rotor and the supernatants combined, and RNAse was added to a final concentration of 10 μg ml^{-1} and digestion was continued for 30 min at 37°C. Debris was pelleted by centrifugation at 11 000 rpm in a Ti-45 rotor for 20 min at 4°C.

The procapsids were harvested by centrifugation at 35 000 rpm in a Ti-45 rotor for 45 min at 4°C. The harvested procapsids were resuspended in 14 ml of 50 mM Tris, 25 mM NaCl, 2 mM EDTA, pH 7.6 and applied to a Sephacryl S-1000 column (volume ≈ 200 ml) equilibrated and developed with the same buffer. The procapsid containing fractions were initially located visually by turbidity and confirmed by SDS–PAGE. The procapsid containing fractions were pooled and the procapsids harvested by centrifugation and resuspended in 50 mM Tris, 25 mM NaCl, pH 7.6. The procapsids were further purified by banding on a CsCl step gradient using CsCl solutions with densities of 1.4, 1.3 and 1.2 g cm^{-3}. The procapsids banded at 1.3 g cm^{-3} and were harvested by dilution and centrifugation.

The empty shells were prepared from procapsids that had not undergone the CsCl step by repeated extraction with 0.5 M GuHCl at 4°C, followed by centrifugation. The scaffolding protein was collected from this 0.5 M GuHCl and purified as described previously [16].

2.2. Raman spectroscopy

The procapsid solution (2.5 mg ml^{-1}) was centrifuged at 95 000 rpm for 15 min in a Beckman Airfuge using rotor A-95. The resulting pellet was rinsed and resuspended in 20 μl of 50 mM sodium phosphate, 25 mM NaCl buffer, pH 7.5 to achieve a final concentration of about 70 mg ml^{-1}. A 10 μl aliquot was sealed in a Kimax #34504 glass capillary. Similarly, the scaffolding protein solution was concentrated to 50 mg ml^{-1} for Raman spectroscopy. Samples were thermostated during data collection [17]. Raman spectra were collected on a Spex model 1877 triple spectrograph (Spex Industries, Metuchen, NJ) equipped with a liquid-nitrogen cooled CCD detector (model ST130, Princeton Instruments, Princeton, NJ). Typically, 30–60 exposures of 1 min each were averaged to produce the spectra shown below.

Procedures for computation of Raman difference spectra of P22 constituents have been described [18,19]. Additionally, to compare the observed spectrum of the procapsid with the synthesized spectral sum of its constituents, Raman intensities were normalized with respect to the 621 cm^{-1} band of phenylalanine. The synthesized spectrum was constructed by taking into account both the amino acid compositions of the coat (gp5) and scaffolding (gp8) subunits and their average stoichiometric ratio (420:100) in the procapsid, as estimated by quantitative SDS–PAGE. We also employed a linear least-squares protocol [20] to minimize the spectral differences between procapsid and constituents. This procedure, which is free of potential normalization errors, was compensated for buffer contributions and baseline artifacts by including the latter in the fitting protocol. Both the intensity-normalization and least-squares procedures yielded identical results.

2.3. Raman microdialysis flow cell

A schematic diagram of the Raman microdialysis flow cell for use in the 90° laser scattering geometry is shown in Fig. 1. The sample solution is contained in a standard glass capillary of 1.0 mm internal diameter and ca. 2.0 cm length (Kimax #34504). A single fiber of SpectraPor (Spectrum, Houston, TX) hollow microdialysis tubing of 210 μm diameter and 6000 Da molecular-weight cutoff is also contained within the cell. During operation, the glass capillary is sealed at both ends with a generic rosin wax and the microdialysis tubing is attached by means of a modified pipette tip to a 10 ml syringe, which serves as the source of the dialyzing solution. Because of the low molecular-weight cutoff of the microdialysis tubing, only small molecules (effluent) are trans-

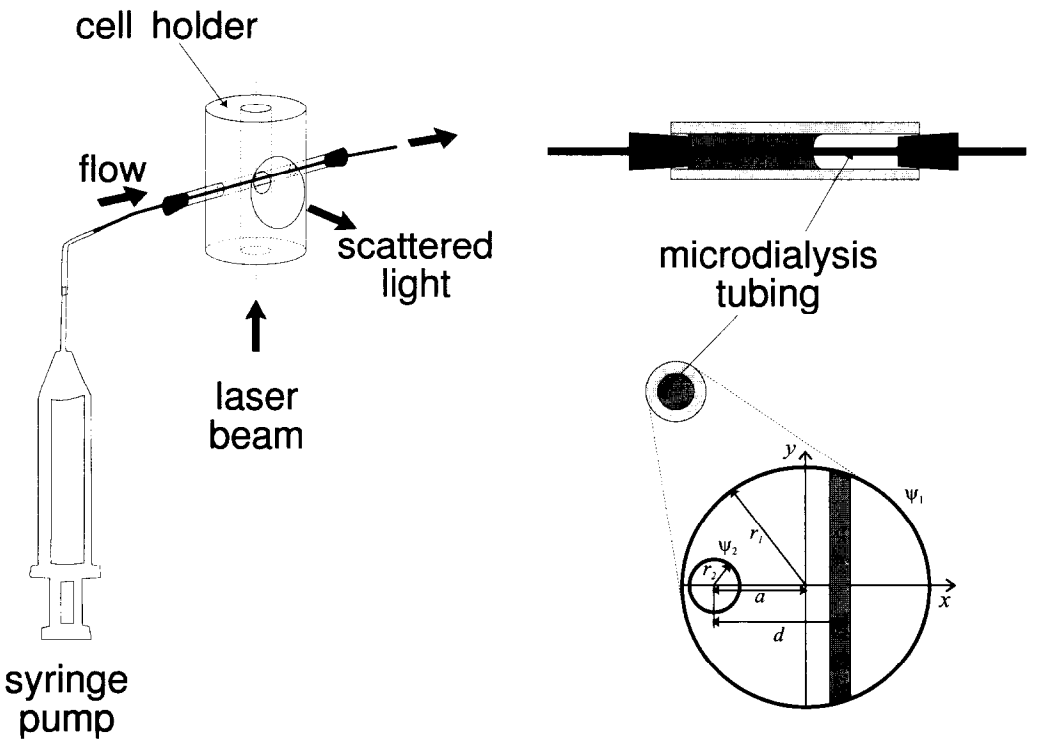

Fig. 1. (Left) Frontal view showing the laser beam path, thermostated sample cell holder and capillary sample cell in the 90° scattering geometry. (Right) Two sections through the capillary cell containing the sample solution (light stippling) and microdialysis tubing (dark stippling). In the lower right corner is shown a cross-sectional view of a chamber representing diffusion of a low-molecular-weight solute from within a small cylindrical boundary (Ψ_2, microdialysis tubing) to a larger cylindrical boundary (Ψ_1, capillary sample cell). Typically, the radii r_1 of Ψ_1 and r_2 of Ψ_2 are 0.5 and 0.106 mm, respectively. The shaded area Ω (width < 0.1 mm) is illuminated by the laser beam and provides the Raman signal. The distances between the centers of the microdialysis tubing and capillary cell and between the centers of the microdialysis tubing and laser beam are labeled a (ca. 0.37 mm) and d, respectively. The origin of the Cartesian coordinate system used in the calculations coincides with the center of the capillary.

ported across the permeable barrier. A constant flow of 2–3 ml h^{-1}, which is sufficient to maintain constant effluent concentration at the tubing boundary, is accomplished with a syringe injection pump (Harvard Instruments, South Natick, MA). The 1.0 mm capillary cell permits use of a standard Raman sample illuminator and thermostat jacket as described above.

3. Theory: numerical solution of the diffusion equation

Design of the Raman flow cell (Fig. 1) ensures that diffusion governs the transport of small molecules between the membrane boundary of the microdialysis tubing and the glass capillary containing the macromolecular solution. Thus, the solute concentration as a function of time $c(x,y,t)$ is determined by the diffusion equation (Eq. (1)), which can be solved with appropriate boundary conditions provided that the diffusion coefficient (D) is known for the solute species.

$$D\Delta c(x,y,t) = \partial c(x,y,t)/\partial t \quad (1)$$

Because the length of the capillary cylinder is much greater than its diameter, the solution of Eq. (1) is simplified to the planar scheme depicted at the lower right of Fig. 1. We define a Cartesian coordinate system with origin at the center of the capillary and x-axis bisecting the circular boundaries of both the glass capillary (Ψ_1) and microdialysis tubing

(Ψ_2). For Ψ_2, the boundary thickness is negligible in comparison to the path of diffusion (ca. 1 mm), and permeability to small molecules is very high. Accordingly, the solute concentration gradient across Ψ_2 can be ignored and the Dirichlet boundary condition (Eq. (2)) applies:

$$c(x,y,t) = c_0 : (x,y) \in \Psi_2 \quad (2)$$

where c_0 is the concentration of solute in the effluent solution. On the other hand, the glass capillary boundary Ψ_1 is impermeable, and the Neumann boundary condition (Eq. (3)) must be applied:

$$\partial c(x,y,t)/\partial n = 0 : (x,y) \in \Psi_1 \quad (3)$$

where the derivative is with respect to the direction normal to the Ψ_1 boundary.

The boundary problems represented by Eqs. (1)–(3) can be solved numerically using the direct method on a square lattice with curved boundary conditions [21].

4. Results and discussion

4.1. Experimental and theoretical determination of efflux rates for small solutes

In order to compare theoretical and experimental results, the following time-dependent diffusion processes were monitored in the microdialysis flow cell by Raman spectroscopy: (1) efflux of H_2O into a 10% (w/v) solution of polyethylene glycol (PEG) in D_2O; (2) efflux of D_2O into a 10% (w/v) solution of PEG in H_2O; (3) efflux of an aqueous solution of 0.1 M $Ca(NO_3)_2$ into pure H_2O; and (4) efflux of an aqueous solution of EGTA into pure H_2O. The results, shown graphically in Fig. 2, yielded the following first order rates for processes (1), (2), (3) and (4), respectively: $k_H = 0.48 \pm 0.01$ min^{-1}, $k_D = 0.287 \pm 0.005$ min^{-1}, $k_{Ca} = 0.10 \pm 0.01$ min^{-1}, $k_{EGTA} = 0.07 \pm 0.01$ min^{-1}. Each rate constant represents a single exponential fit to the corresponding data set in Fig. 2. Table 1 shows that the experimental rates are in excellent agreement with theoretical values calculated from available diffusion coefficients.

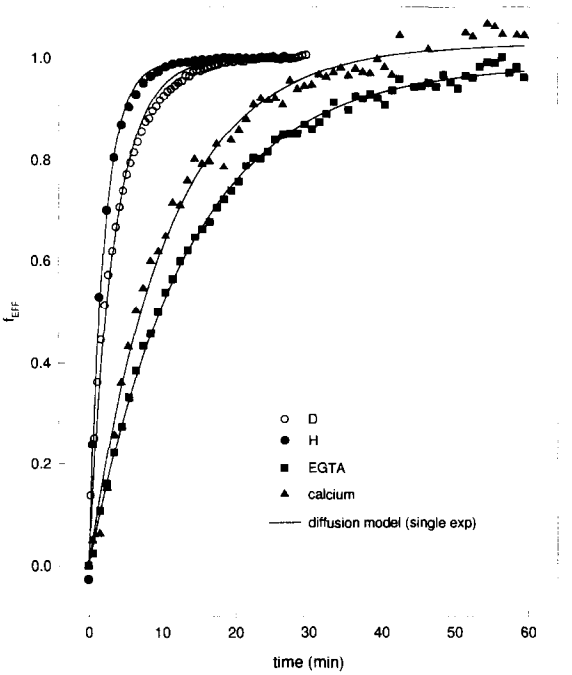

Fig. 2. Raman measurements of the efflux of low-molecular-weight species in the microdialysis flow cell of Fig. 1. Calcium nitrate concentration was determined from the intensity of the Raman band of NO_3^- at 1050 cm^{-1}, EGTA concentration from the intensity of the Raman band of CO_2^- at 1405 cm^{-1}, and protium and deuterium concentrations from the integrated intensities of their respective solvent Raman bands at 3000–3600 and 2300–2800 cm^{-1}. Intensities for $Ca(NO_3)_2$ and EGTA were normalized with respect to the H_2O band at 1635 cm^{-1}; intensities for H and D efflux were normalized using the bands at 2700–2950 cm^{-1} from 10% PEG 8000 added as an internal standard [22].

Table 1
Comparison of calculated and experimental efflux rate constants[a]

Effluent	D (cm^2 s^{-1})	k_{calc} (min^{-1})	k_{exp} (min^{-1})
H_2O	5.0×10^{-5}	0.46	0.48 ± 0.01
D_2O	3.0×10^{-5}	0.28	0.29 ± 0.01
$Ca(NO_3)_2$	1.2×10^{-5}	0.11	0.10 ± 0.01

[a] Values for the diffusion constants are at 20°C [23]. Calculated rates, k_{calc}, are from Eqs. (1)–(3), as described in the text.

4.2. Structural properties of the P22 scaffolding protein

Raman spectra of aqueous solutions of the P22 scaffolding protein (gp8) were collected over the

temperature range 5–95°C at intervals of 5 and 10°C. Representative data are shown in Fig. 3A. At 5°C, the Raman spectrum of gp8 is dominated by an intense amide I band centered near 1650 cm^{-1}, which is diagnostic of a predominantly α-helical secondary structure [25]. The conformationally informative amide III region of the spectrum (1230–1340 cm^{-1}) also exhibits bands in the interval 1275–1300 cm^{-1}, as expected for an α-helical protein. Additionally, a distinct feature centered at 935 cm^{-1}, assigned to skeletal C–C and C–N stretching vibrations, indicates that the prevailing peptide bond geometry is within the conformation characteristic of α-helix. Circular dichroism (CD) spectra, collected at 10°C also demonstrate that the dominant secondary structural motif is α-helix (Fig. 3B). These results confirm earlier preliminary findings [26].

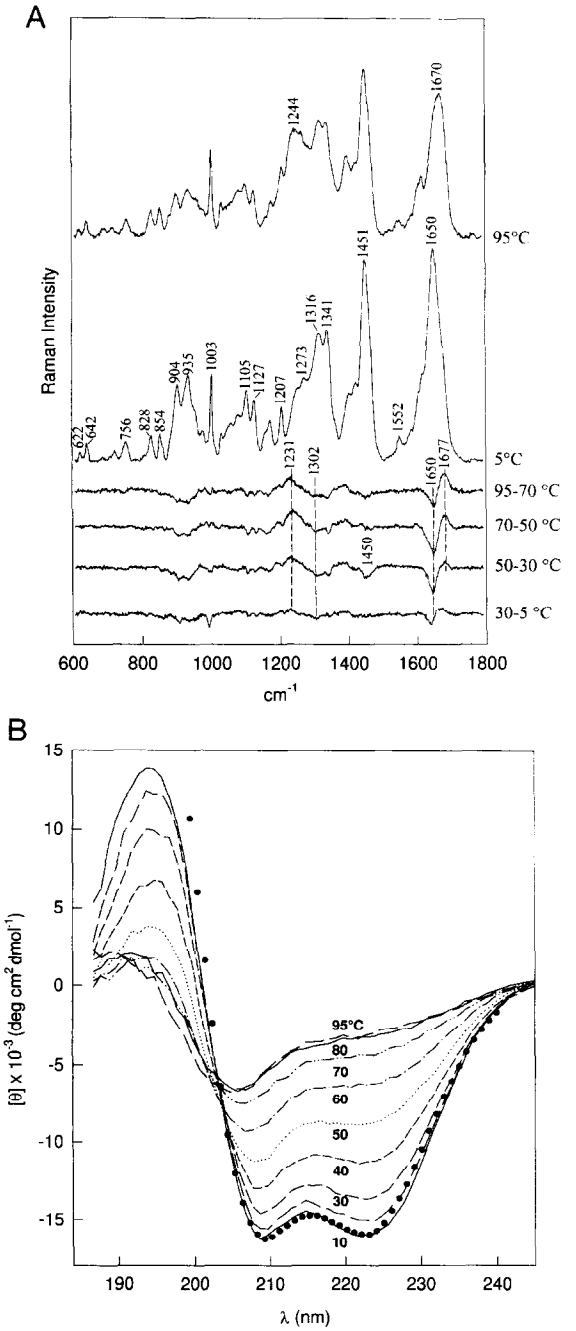

Fig. 3. (A) Raman spectra of the P22 scaffolding protein (gp8) at 5°C (second trace from top) and 95°C (top trace), and difference spectra corresponding to the temperature intervals 95–70, 70–50, 50–30 and 30–5°C, as labeled. Intensities are normalized to the spectrum at 5°C. Experimental conditions are described in the text. (B) CD spectra (185–245 nm) of the P22 scaffolding protein as a function of temperature (10–95°C). Solid symbols (●) represent the CD profile corresponding to a least-squares fit of the experimental curve to 37±3% α-helix, 33±6% β-turns and 30±6% random coil [24].

The data of Fig. 3 show that although gp8 is highly α-helical at physiological temperatures, the secondary structure is thermolabile. The structure of gp8 is observed to change gradually over the temperature range 10–95°C. The change in secondary structure is from an α-helix to a disordered chain. During the course of this noncooperative secondary structure transition, we observe little or no change in intensities of Raman marker bands assigned to specific side chains. These results indicate that, despite the definitive transformation of secondary structure, the side chain environments of subunit domains are little affected. Because more significant changes in side chain environments would accompany changes in tertiary and/or quaternary structures, we conclude that the α-helices of gp8 do not participate in extensive tertiary or quaternary structure formation. The observed denaturation of gp8 secondary structure without apparent change in higher order structure implies that the scaffolding subunit is not a typical globular protein, i.e., it appears to lack a tightly folded core.

We have analyzed gp8 thermal denaturation using

markers of α-helicity in Raman spectra (amide I band at 1650 cm^{-1}) and CD spectra (molar ellipticity at 222 nm). Apparent changes in unfolding enthalpy (ΔH_u) and entropy (ΔS_u) have been obtained using a simple two-state equilibrium model (Table 2). Although markedly different protein concentrations are employed in the two types of spectroscopy, comparable ΔH_u and ΔS_u values are obtained and no cooperativity of unfolding is evident in either case. On the other hand, a somewhat higher T_m value is observed for the Raman melting profile, which reflects greater subunit stability at the higher concentration employed in the Raman experiments. This may be due to unfavorable interactions of unfolded gp8 molecules at high concentration.

Dynamics of the native state of the scaffolding subunit have been characterized by hydrogen–deuterium exchange. The rapid exchange of gp8 peptide sites (NH → ND) has been monitored by time-resolved Raman spectroscopy at 10°C (Fig. 4A). Several distinct time-dependent features are discernible in these spectra. First, the amide III band complex associated with NH peptides (1250–1340 cm^{-1} region) is rapidly replaced by an amide III′ complex (900–1030 cm^{-1} region) associated with ND peptides, indicating rapid deuteriation of the gp8 polypeptide chain. The amide III′ intensity increase does not lag behind the efflux of D_2O, the latter measured by the growth of intensity of the D_2O band at 1206 cm^{-1}. Second, an amide II′ band, similarly diagnostic of ND peptides, is rapidly generated at 1458 cm^{-1}. (The amide II′ band overlaps another strong band near 1455 cm^{-1} that is due to CH_2 deformation vibrations). Third, a shift of the α-helix marker from 1650 cm^{-1} (amide I) to 1640 cm^{-1} (amide I′) is also apparent. The kinetics of these amide band changes indicate that most of the α-helical peptide groups are deuteriated within 30 min of exposure to D_2O and that no further deuteriation occurs after 30 min, even when the gp8 molecules are reversibly unfolded and refolded in D_2O (data not shown). These results reveal that peptide exchange of the native gp8 structure is complete after 30 min in D_2O.

Table 2
Thermodynamic parameters for thermal unfolding of P22 scaffolding protein

Method	ΔH_u (kcal mol^{-1})	ΔS_u (cal mol^{-1} K^{-1})	T_m (°C)
CD	16.0 ± 1.2	48 ± 5	50 ± 2
Raman	17.1 ± 1.7	52 ± 5	60 ± 2

The effect of peptide and side chain deuteriations on the Raman signature of gp8 can be visualized by computing a difference spectrum between the signature of fully deuteriated gp8 and its protiated counterpart (Fig. 4B). This procedure assures that spectral contributions from deuteriation-insensitive vibrations [27] are compensated. The resulting difference intensity in the 900–1030 cm^{-1} region is thus confidently assigned to amide III′ modes. Similarly, contributions from CH_2 deformation modes, which overlap the amide II′ region (1400–1500 cm^{-1}) but are not sensitive to labile proton exchange, are compensated in the resulting difference spectrum.

The difference spectrum of Fig. 4B displays several interesting features. The amide III′ region (900–1030 cm^{-1}) is dominated by the 950 cm^{-1} band, which is assigned to deuteriopeptides in α-helical secondary structure [28]. In addition, three other distinct components (at 919, 970 and 1005 cm^{-1}) are discernible within the amide III′ band complex. These are assigned to deuteriopeptides in turns (970 and 1005 cm^{-1}) and irregular or coil conformations (919 cm^{-1}) [29]. The amide III′ profile indicates that gp8 contains mostly α-helical and irregular secondary structures and β-turns. The series of troughs (at 1243, 1273, 1285 and 1341 cm^{-1}) represent amide III modes of the protiopeptides in α-helices, turns and irregular conformations. The amide II′ region also consists of multiple components at 1436, 1458 and 1470 cm^{-1}. The 1436 and 1470 cm^{-1} bands are tentatively assigned to deuteriopeptides in irregular and turn conformations [29], whereas the 1458 cm^{-1} band can be assigned to deuteriopeptides in α-helices [28,29]. The deuteriation shifts of amide I frequencies are much smaller than shifts of amide II and III modes, owing to the nature of the amide I vibration (i.e., mainly C=O stretching). Nevertheless, amide I → amide I′ shifts are apparent as a complex derivative pattern in the difference spectrum of Fig. 4B. This pattern reflects mainly the exchange of peptides in α-helical secondary structure (1655 → 1630 cm^{-1}) [26]. Deuteriation of peptides in nonhelical domains probably accounts for

the shift of intensity within the high frequency shoulder (1680 → 1666 cm^{-1}) of the central amide I peak [28,30].

The partially resolved amide III′ profile of gp8 can be decomposed into a set of four band components, as illustrated in the inset to Fig. 5. The

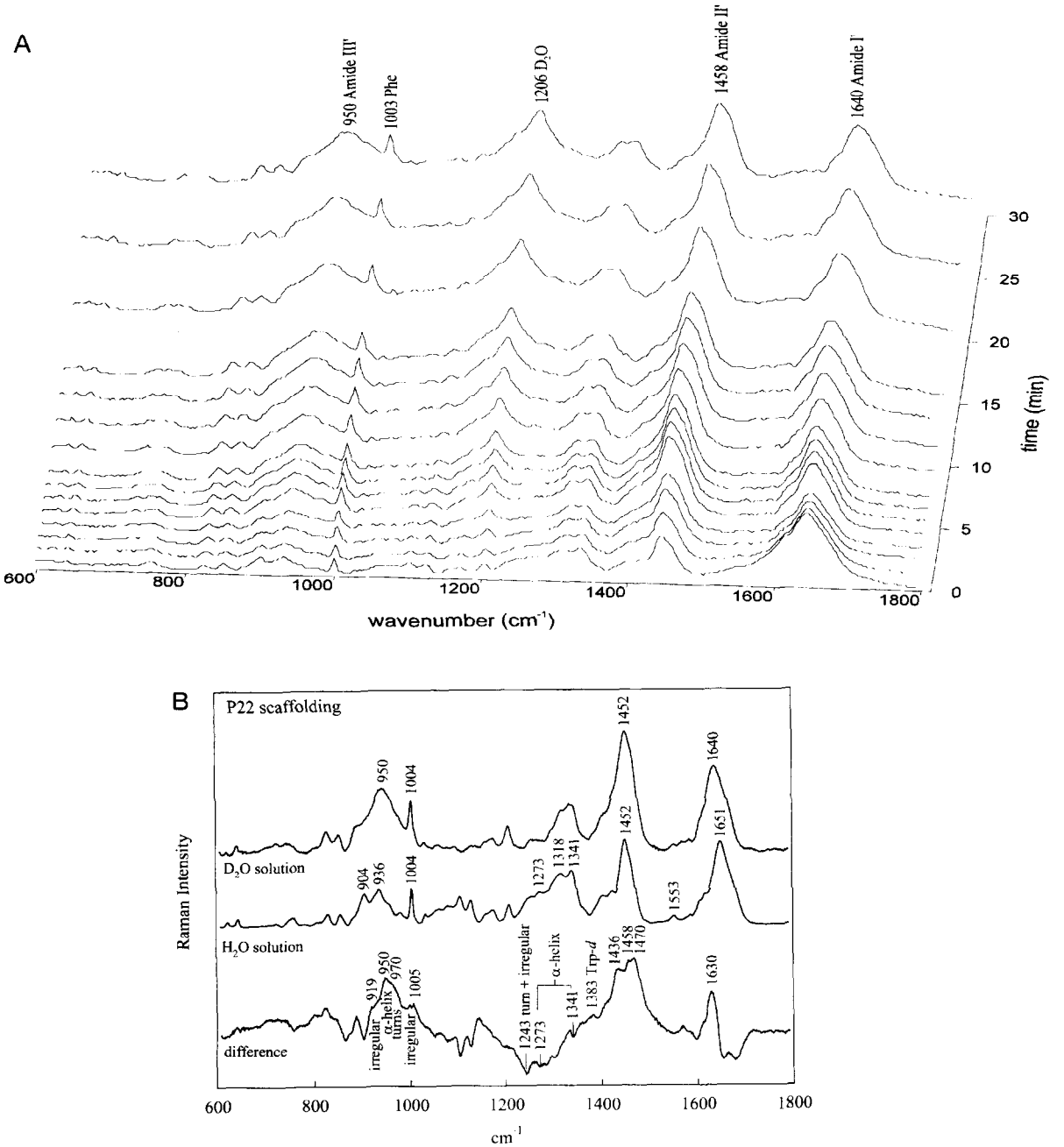

Fig. 4. (A) Time-resolved Raman spectra of the P22 scaffolding protein. Data were collected at 10°C and at 1 min intervals. Data collected at intermediate time intervals have been omitted for clarity. (B) Raman signatures of fully deuteriated gp8 (unfolded and refolded in D_2O) (top trace), fully protiated gp8 (in H_2O) (middle trace) and their difference spectrum (bottom).

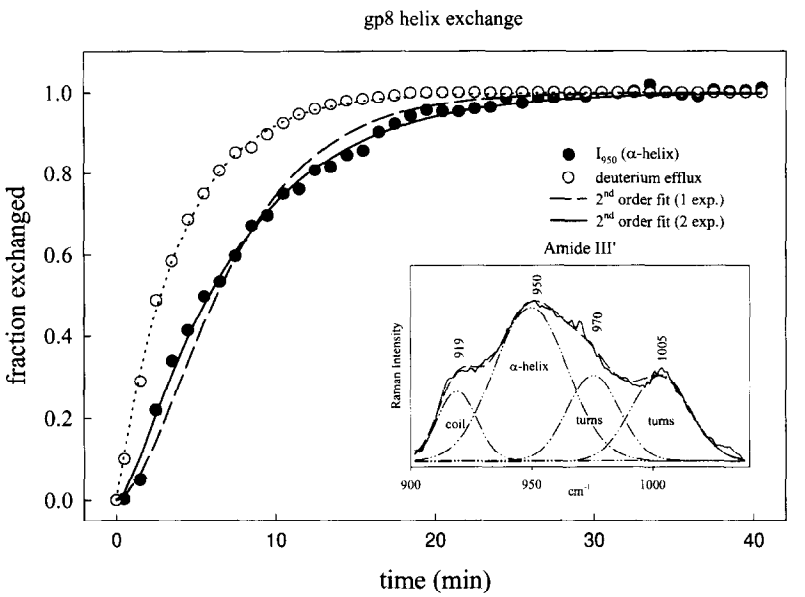

Fig. 5. Time-dependent exchange of peptide groups in helices of the P22 scaffolding protein. Progress of peptide deuteriation was calculated from data of Fig. 4 using the band decomposition shown in the inset. The inset shows the decomposition (curve fit) of the complex Raman amide III′ band profile of the fully exchanged gp8 molecule.

component bands are assigned to deuteriopeptides in different conformations, in accordance with the foregoing discussion. The time-dependent intensities of these amide III′ components may be exploited to monitor exchanges of peptide groups in different structural domains. We have monitored the two most reliably assigned band components, the Raman markers of α-helices at 950 cm^{-1} and of β-turns at 1005 cm^{-1}, to follow the exchange kinetics of peptides in these two conformations. The results for the 950 cm^{-1} band (filled circles) are plotted in Fig. 5, where comparison is made to the rate of efflux of D_2O solvent (open circles) into the sample cell. Peptides in β-turns exchange relatively rapidly, as revealed by the time-dependent growth of intensity at 1005 cm^{-1} (data not shown). This exchange approaches the limiting rate of solvent efflux. On the other hand, the exchange of α-helical domains of gp8 is characterized by considerable retardation. The data cannot be explained by a single exponential term (broken line curve of Fig. 5). However, a satisfactory fit to two exponential terms can be obtained (solid line curve), indicating two kinetically distinct populations of α-helical sites in gp8. Moreover, the two populations of helices are of roughly equal size. To compare the retardation effects, we define a protection factor, P, as the quotient of rate constants for peptide exchanges in a reference state and in the protein native state, i.e., $P \equiv k_{ex}$ (peptide reference state)$/k_{ex}$ (peptide in gp8 structural domain). The more rapidly exchanging class of gp8 α-helices ($k_{ex} = 1.3$ min^{-1}) exhibits a protection factor of 220, when compared with a reference state of 'free' peptides ($k_{ex} = 288$ min^{-1}). The more slowly exchanging gp8 α-helices ($k_{ex} = 0.16$ min^{-1}) are protected by a factor of 1800. These protection factors include protection due to secondary, tertiary and quaternary structures. In order to assess the protection conferred by tertiary and/or quaternary structure(s) alone, we also compared the exchange rates of gp8 with a reference state consisting of a small unprotected helix for which $k_{ex} = 24$ min^{-1} [31,32]. In this comparison, the more rapidly exchanging helices of gp8 are only marginally protected vis-à-vis an isolated helix ($P \approx 18$), whereas the more slowly exchanging helices of gp8 are considerably protected ($P \approx 150$). These results are summarized in Table 3.

The above protection factors are related to an apparent free energy of helix unfolding by the relation $\Delta G_{local} = -RT \ln(P)$. On the other hand, the ΔG_{helix} values, which are derived from comparison between apparent protection of gp8 helices and a model helical peptide, represent relative strength of tertiary and quaternary interactions (Table 3). Low magnitudes of the ΔG_{helix} values ($\approx$ 1.6–2.8 kcal mol^{-1}) imply that only marginally stable tertiary and quaternary structures exist for the gp8 subunit.

Several bands in the time-resolved Raman spectra of gp8 are diagnostic of exchange of side chains. For example, the early appearance of the 1380 cm^{-1} marker of the N1-deuteriated indole ring indicates unhindered exchange of the single tryptophan residue in the gp8 subunit. Similarly, the deuteriation shift of the tyrosine Fermi doublet develops early in the time-resolved spectra. Therefore, all tyrosines within the gp8 sequence are freely exchangeable.

4.3. Structural properties and thermostability of the P22 procapsid

Raman spectra of aqueous solutions of P22 procapsids were collected over the temperature interval 30–95°C at increments of 5 and 10°C. Representative data are shown in Fig. 6. The broad denaturation profile of the procapsid reflects the overlap of the denaturation profile of scaffolding protein with that of the empty procapsid shell. The temperature profile of the *W3* mode of tryptophan reveals a procapsid transition between 30 and 50°C not seen for either

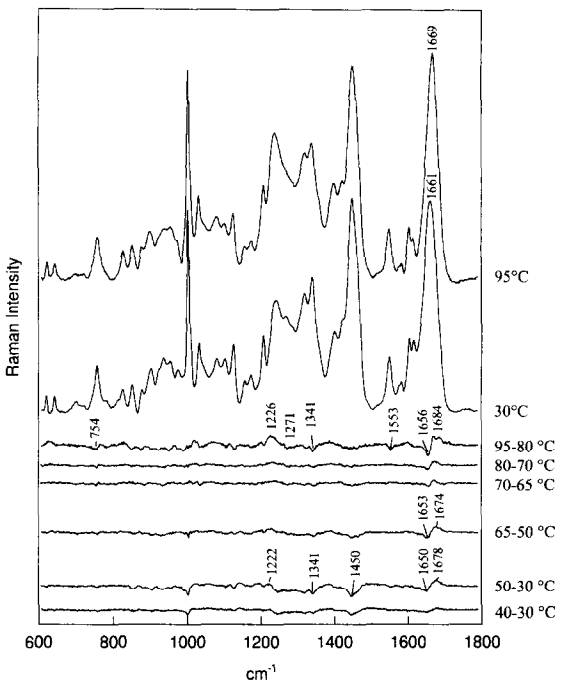

Fig. 6. Raman spectra of the P22 procapsid (gp8/gp5) at 30°C (second trace from top) and 95°C (top trace), and difference spectra corresponding to the temperature intervals 95–80, 80–70, 70–65, 65–50, 50–30 and 40–30°C, as labeled. Intensities are normalized to the spectrum at 30°C.

Table 3
Exchange rates and protection factors for α-helices of the P22 scaffolding protein

	k_{ex} (min^{-1})	P^a	P^b	ΔG^c_{local} (kcal mol^{-1})	ΔG^c_{helix} (kcal mol^{-1})
Rapid	1.3±0.5	220	18	3.0	1.6
Slow	0.16±0.01	1800	150	4.2	2.8

[a] Protection factor relative to a free peptide as the reference state, for which $k_{ex} = 288$ min^{-1} [33].
[b] Protection factor relative to a solvated helix as the reference state, for which $k_{ex} = 24$ min^{-1} (exchange rate of the central peptide in a 21-residue helical oligopeptide [32]).
[c] Free energy changes ΔG_{local} and ΔG_{helix} correspond to protections versus a free peptide and a solvated helix, respectively, as described in the text.

the scaffolding protein alone or the empty shells [34]. This transition represents indole ring reorientations for a small population of tryptophans ($\approx$ 5% of Trp residues) from $|\chi^{2,1}| \approx 110°$ to $|\chi^{2,1}| \approx 80°$. It is accompanied by a decrease in α-helicity and an increase in β-strand conformation. In contrast to the free scaffolding protein, in which the conformational changes take place gradually with increasing temperature, 80% of the conformational changes accompanying the 30–50°C transition take place between 40 and 50°C (Fig. 6). This transition is completely reversible [34].

4.4. Interaction of coat and scaffolding subunits in the P22 procapsid

If there are no changes in the structures of the coat protein lattice and scaffolding protein subunits upon exit, it should be possible to reconstruct the

spectrum of the procapsid at any given temperature by summing spectra of the empty shells (gp5) and scaffolding protein (gp8) at the same temperature, each weighted according to its mass composition in the procapsid and normalized according to its known amino acid composition, as described in Section 2.2. In fact it is not possible to adequately represent the procapsid spectrum as a sum of spectra of gp5 and gp8 [34]. This constitutes evidence that coat subunits of the empty shell and scaffolding subunits, when segregated from one another, are structurally different than when combined in the assembled procapsid. Such evidence implies interaction between coat and scaffolding subunits of the native procapsid.

The difference between the spectrum of the procapsid and the weighted sum of constituent spectra provides an interesting insight into coat and scaffolding protein interactions. At 30°C, most of the observed difference bands can be assigned to specific side chains. For example, the intensity differences at 758, 1204 and 1543 cm^{-1} are due to tryptophan and indicate perturbation of the Trp side chain upon interaction of gp8 with gp5 within the procapsid. The difference peaks observed at 1332 and 1438 cm^{-1} are assigned to methylene deformation modes of aliphatic side chains [35] and similarly reflect structural perturbations in these side chains upon gp8:gp5 interaction. All of the above noted differences occur below 50°C, when the scaffolding is retained in the procapsid, and disappear above 50°C, when the scaffolding has been released. Melting profiles for Raman markers of tryptophan (1543 cm^{-1}) and aliphatic side chains (1440 cm^{-1}) are shown in the lower panel of Fig. 7. The profiles display a rather abrupt transition near 50°C, which is consistent with the expected highly cooperative exit of scaffolding protein from the procapsid [36]. Quantitatively, however, the differences account for only a small fraction of the parent band intensities, indicating that only a limited number of side chains participate in the interactions.

The differences observed in the 1630–1700 cm^{-1} region upon heating from 30 to 50°C are assigned to the amide I mode and represent altered secondary structure of gp5 and/or gp8 upon procapsid assembly. On the basis of the positive peak located at 1652 cm^{-1}, we conclude that the average secondary structure of the procapsid assembly is more α-helical than

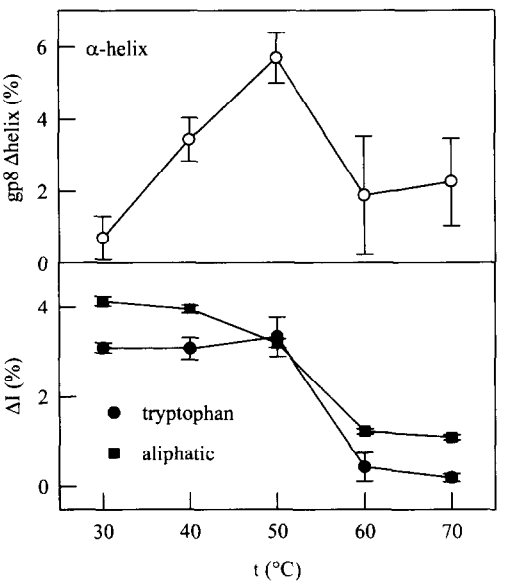

Fig. 7. (Top) Temperature dependence of the excess α-helicity of scaffolding protein in the P22 procapsid. The excess is defined as the additional α-helical secondary structure of the gp8 subunit in the procapsid compared with the extracted state, as monitored by the amide I band at 1650 cm^{-1}. A similar result is obtained for the amide III α-helix marker at 1348 cm^{-1} (not shown). (Bottom) Temperature dependence of Raman intensity differences between the P22 procapsid and the sum of its constituents for Raman markers of tryptophan (1544 cm^{-1}, filled circles) and of aliphatic side chains (1440 cm^{-1}, filled squares).

its constituents. The trough observed at 1682 cm^{-1} indicates replacement of the helical structure by β-strand upon disassembly. Based upon the α-helix to β-strand transition observed for free gp8 (see above) and the high α-helicity of native gp8, these changes occur most probably in gp8. Interestingly, the differences increase in magnitude with increasing temperature between 30 and 50°C, achieve a maximum at 50°C, then sharply attenuate with further increase of temperature between 50 and 60°C. This is illustrated in the upper panel of Fig. 7 by the temperature dependence of the excess α-helicity, as measured by the 1652 cm^{-1} difference intensity (normalized with respect to the total amount of gp8 α-helical secondary structure).

The Fig. 7 profile demonstrates that the secondary structure of gp8 is fully retained within the procapsid up to 50°C. Thus, both the secondary and tertiary structural transitions appear to be more cooperative

for the procapsid assembly than for its isolated constituents. The cooperativity of main chain unfolding is evidently linked to gp8 exit, shown previously to be a cooperative phenomenon [36]. The cooperativity in altered side chain environments, on the other hand, can be attributed to specific interactions between gp5 and gp8.

4.5. Subunit conformational changes accompanying P22 capsid maturation

The procapsid-to-capsid transformation is accompanied by numerous spectral changes, which are assigned to both main chain and side chain conformational markers [19]. Here, we limit our discussion to changes assigned to the coat protein subunit secondary structure and those attributed to reorientation of tryptophan residues. Fig. 8 compares Raman spectra of empty procapsid shells and mature (expanded) shells in the region 1500–1750 cm^{-1}. The amide I peak is centered at 1662 cm^{-1} (solid line) in the spectrum of the empty procapsid shells and at 1667 cm^{-1} in the spectrum of the mature shells (dotted line). The shift of the amide I peak generates a difference spectrum with a trough at 1672 cm^{-1} and a maximum at 1653 cm^{-1}. Quantitatively, this result can be interpreted as a conversion of about 2 to 3% of the α-helical secondary structure in subunits of the procapsid shell to the β-strand conformation in the native shell, as represented in the cartoon in Fig. 8.

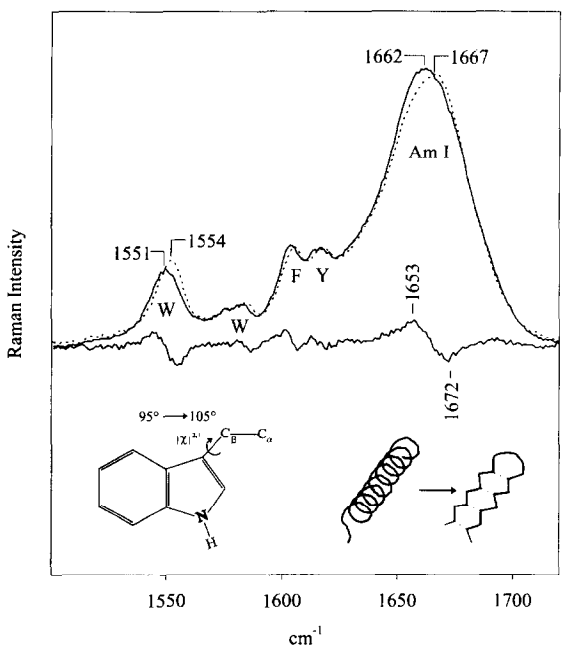

Fig. 8. Raman spectra (1500–1750 cm^{-1}) of the empty procapsid shell (solid line) and mature or expanded shell (dotted line) are shown in the top part of the figure. Amide I bands are centered at 1662 and 1667 cm^{-1}, respectively, for these two particles. The difference spectrum obtained by subtracting the latter from the former is also shown and reveals the shift of amide I intensity from 1653 to 1672 cm^{-1} accompanying expansion. This corresponds to an α-helix-to-β-strand conformational change in gp5 subunits attendant with shell maturation. A conformational change involving tryptophan side chains is also evident from the shift of the 1551 cm^{-1} band to 1554 cm^{-1}.

The spectra of procapsid shells and mature shells also exhibit differences in the position of the tryptophan *W3* band. The shift is consistent with reorientation of the average tryptophan residue, such that the torsion $|\chi^{2,1}|$ changes from 95° in the procapsid shell to 105° in the mature shell.

5. Summary and conclusions

5.1. A model for the P22 scaffolding protein

Thermal denaturation of the α-helical secondary structure of gp8 is characterized in a simple two-state approximation by an unfolding enthalpy ΔH_u of 17 ± 2 kcal mol^{-1} and an unfolding entropy ΔS_u of 50 ± 5 cal mol^{-1} K^{-1}. Within the two-state approximation, no temperature dependence of ΔH_u or ΔS_u is observed, suggesting that $\Delta C_p \approx 0$. The small value of ΔH_u and the absence of a change in heat capacity are in agreement with calorimetric experiments [36].

Despite a highly α-helical secondary structure, gp8 exhibits a broad and noncooperative melting profile, when monitored either by Raman or CD spectroscopy. The denaturation of gp8 resembles that of a polypeptide lacking the hydrophobic core typical of a globular protein [37]. The absence of a hydrophobic core is also demonstrated by the accessibility of all peptide NH groups of gp8 to relatively unhindered deuterium exchange. These findings suggest a highly flexible tertiary structure comprising multiple α-helical domains. Nevertheless, the peptide exchange measurements reveal three sub-popula-

tions of somewhat different exchange kinetics. About 50% of peptides exchange too rapidly ($k_{ex} > 2$ min^{-1}) to be resolved kinetically in the microdialysis flow cell. This population probably comprises turns, loops and irregular conformations. The remaining 50% are sub-divided roughly equally into a relatively rapidly exchanging fraction ($k_{ex} = 1.3$ min^{-1}) and a relatively slowly exchanging fraction ($k_{ex} = 0.16$ min^{-1}). Both of these peptide sub-classes are believed to consist primarily of α-helices, although they evidently differ in stability or accessibility. Free energies of local helix unfolding, which are obtained on the basis of peptide NH → ND exchange retardation (Table 3), are close in value to the free energy of global unfolding ($\Delta G_u \approx 3$ kcal mol^{-1} at 10°C). Therefore, we may conclude that the α-helical domains of gp8 exchange through a global unfolding mechanism.

The presence of secondary structure, the relative noncooperativity of unfolding, and the marginal protection against peptide NH → ND exchange all indicate a 'molten globule' state for gp8. The molten globule concept was introduced originally to describe equilibrium folding intermediates [38,39]. Here, we extend the concept to describe a mature native protein structure that functions not as an intermediate along a subunit folding pathway but as an intermediary in the assembly of a large quaternary complex, viz. a viral capsid.

Sedimentation analysis shows that gp8 possesses the compact structure of a prolate ellipsoid with an axial ratio of 9:1 [40]. The molten globule-like state proposed for gp8 is envisioned here as consisting of segmented α-helices loosely connected by flexible turns or loops. The gp8 sequence is extraordinarily rich in proline, as well as in acidic and basic residues. These features are consistent with our proposed structural model.

The molten globule-like character of the native state of gp8 is likely to play a key functional role in its exit from the procapsid. A proposed mechanism of scaffolding exit involves passage of gp8 through channels of only 25–30 Å width, which are located at the quasi-sixfold vertices of the $T = 7$ procapsid [41]. An elongated and flexible gp8 molecule, lacking a hydrophobic core, would exit through these channels more easily than would a rigid protein structure.

5.2. A procapsid assembly model

A simple model for procapsid assembly is illustrated in Fig. 9. Salient features of this model, proposed on the basis of previous and present results, are the following: (1) Assembly is initiated by recognition between the highly α-helical scaffolding (gp8) subunits and the highly β-stranded (gp5) coat subunits. (2) The gp8 subunits interact specifically with gp5 subunits to form the correct architecture of the procapsid shell. Interaction of gp8 and gp5 in the procapsid results in thermostabilization of the α-helical conformation of gp8 (Fig. 7). (3) The procapsid-bound scaffolding protein is in equilibrium with similarly folded but unbound and more thermolabile gp8 subunits. (4) The unbound gp8 subunits represent in turn an equilibrium between unfolded and incompletely folded forms. (5) The ensemble of

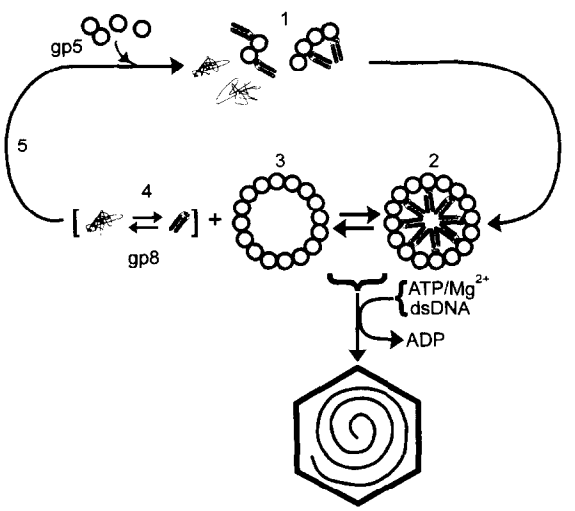

Fig. 9. A model for assembly of the P22 procapsid. Salient features include: (1) Initiation of assembly by molecular recognition involving helical scaffolding subunits (gp8) and globular coat subunits (gp5); (2) thermostabilization of gp8 secondary structure through formation of procapsids of the correct architecture, which are in equilibrium with (3) empty shells and (4) a nested equilibrium mixture of folded and incompletely folded gp8 subunits. (5) The latter are available for subsequent rounds of assembly. The in vitro transition from (2) to (3) and (4) is accomplished in vivo by ATP-dependent packaging of dsDNA, represented at the lower right. DNA packaging in vivo does not generate the empty shell (3), but a structurally related mature shell consisting of gp5 subunits [2].

folded and unfolded gp8 subunits is available for subsequent rounds of gp5 recognition and procapsid assembly.

The procapsid assembly model of Fig. 9 is consistent with a number of observations based upon physico-chemical and genetic experiments. First, the exit of a scaffolding subunit from the procapsid shell is coupled to its folding equilibrium. Thus, the in vitro denaturation of gp8 by chemical (GuHCl, urea) or physical (pressure and temperature) means is effective in driving its release from the procapsid. The in vivo counterpart of this driving force is presumably the dsDNA packaging event. The packaged genome competes for the same excluded volume and binding sites otherwise occupied by the scaffolding protein. As expected, the latter phenomenon shifts the system (gp8 subunits) toward the ensemble of folded and unfolded states of gp8, which become available for further rounds of assembly in vivo. Finally, because in vivo packaging of DNA is coupled with irreversible expansion of the procapsid shell, the exit of scaffolding protein is essentially an irreversible process.

5.3. P22 capsid maturation

The results presented here indicate that the procapsid-to-capsid transformation is accompanied by a small change in the coat subunit secondary structure and much larger changes in side chain environments. The results are consistent with domain movement mediated by hinge bending. The three-dimensional image reconstructions from electron cryomicroscopy of the procapsid and capsid shells reveal that holes located at six-fold vertices are closed upon maturation [12]. Thus, maturation requires new intersubunit contacts that result from the proposed domain movement. Similar domain movements have been proposed to explain the expansion of plant virus and bacteriophage λ capsids [42,43].

Acknowledgements

Support of this research by the US National Institutes of Health (Grant GM50776) is gratefully acknowledged.

References

[1] C.M. Teschke and J. King, Biochemistry 32 (1993) 10839.
[2] P.E. Prevelige, Jr. and J. King, Prog. Med. Virol. 40 (1993) 206.
[3] T. Miura and G.J. Thomas, Jr., in: B.B Biswas, S. Roy (Eds.) Subcellular Biochem. Vol. 24 (Plenum Press, New York, 1995).
[4] R. Tuma, J.H.K. Bamford, D.H. Bamford, M.P. Russell and G.J. Thomas, Jr., J. Mol. Biol. 257 (1996) 87.
[5] R. Tuma, J.H.K. Bamford, D.H. Bamford and G.J. Thomas, Jr., J. Mol. Biol. 257 (1996) 102.
[6] T. Li, J.E. Johnson and G.J. Thomas, Jr., Biophys. J. 65 (1993) 1963.
[7] K.E. Reilly and G.J. Thomas, Jr., J. Mol. Biol. 241 (1994) 68.
[8] J. King and S. Casjens, Nature 251 (1974) 112.
[9] S. Casjens and R. Hendrix, in: R. Calendar (Ed.), The Bacteriophages, Vol. 1, Plenum Press, New York, 1988.
[10] J. King, E.V. Lenk and D. Botstein, J. Mol. Biol. 80 (1973) 697.
[11] P. Roy and H. Murialdo, Virology 64 (1975) 247.
[12] B.V.V. Prasad, P.E. Prevelige, M.E., R.O. Chen, D. Thomas, J. King and W. Chiu, J. Mol. Biol. 231 (1993) 65.
[13] T. Dokland and H. Murialdo, J. Mol. Biol. 233 (1993) 682.
[14] W. Earnshaw and S. Casjens, Cell 21 (1980) 319.
[15] C. Bazinet and J. King, Annu. Rev. Microbiol. 39 (1985) 109.
[16] P.E. Prevelige, Jr., D. Thomas and J. King, J. Mol. Biol. 202 (1988) 743.
[17] G.J. Thomas, Jr. and J. Barylski, Appl. Spectrosc. 24 (1970) 463.
[18] P.E. Prevelige, Jr., D. Thomas, K.L. Aubrey, S.A. Towse and G.J. Thomas, Jr., Biochemistry 29 (1990) 5626.
[19] P.E. Prevelige, Jr., D. Thomas, K.L. Aubrey, S.A. Towse and G.J. Thomas, Jr., Biochemistry 32 (1993) 537.
[20] P. Pelikán, M. Ceppan and M. Liška, Applications of Numerical Methods in Molecular Spectroscopy, CRC Press, Boca Raton, 1994.
[21] M.N. Ozisik, Heat Conduction, Wiley, New York, 1980.
[22] T. Li, Investigation of molecular recognition by static and dynamic methods of Raman spectroscopy: Applications to RNA plant viruses and DNA-protein complexes. PhD Thesis, School of Biological Sciences, University of Missouri, Kansas City, 1992.
[23] R.C. Weast and S.M. Selby, CRC Handbook of Chemistry and Physics, CRC Press, Cleveland, 1966.
[24] J.T. Yang, C.W. Chuen-Shang and H.M. Martinez, Meth. Enzymol. 130 (1986).
[25] M.C. Chen and R.C. Lord, J. Am. Chem. Soc. 96 (1974) 4750.
[26] G.J. Thomas, Jr., Y. Li, M.T. Fuller and J. King, Biochemistry 21 (1982) 3866.
[27] S. Krimm, in: T.G. Spiro (Ed.), Biological Applications of Raman Spectroscopy, Vol. 1, Wiley-Interscience, New York, 1987.

[28] T.-J. Yu, J.L. Lippert and W.L. Peticolas, Biopolymers, 12 (1973) 2161.
[29] J. Sugawara, I. Harada, H. Matsuura and S.T., Biopolymers 17 (1978) 1405.
[30] J. Bandekar, Biochim. Biophys. Acta 1120 (1992) 123.
[31] C.A. Rohl, J.M. Scholtz, E.J. York, J.M. Stewart and R.L. Baldwin, Biochemistry 31 (1992) 1263.
[32] C.A. Rohl and R.L. Baldwin, Biochemistry 33 (1994) 7760.
[33] J.J. Englander, D.B. Calhoun and S.W. Englander, Anal. Biochem. 92 (1979) 517.
[34] R. Tuma, P.E. Prevelige, Jr. and G.J. Thomas, Jr., Biochemistry 35 (1996) 4619.
[35] D. Lin-Vien, N.B. Colthup, W.G. Fateley and J.G. Grasselli, The Handbook of Infrared and Raman Characteristic Frequencies of Organic Molecules, Academic Press, San Diego, 1991.
[36] M.L. Galisteo and J. King, Biophys. J. 65 (1993) 227.
[37] P.R. Privalov, Adv. Protein Chem. 33 (1979) 167.
[38] H. Christensen and R.H. Pain, in: R.H. Pain (Ed.), Mechanisms of Protein Folding, Oxford Univ. Press, New York, 1994.
[39] K.-S. Kim, J.A. Fuchs and C.K. Woodward, Biochemistry 32 (1993) 9600.
[40] M.T. Fuller and J. King, Virology 112 (1981) 529.
[41] P.A. Thuman-Commike, B. Green, J. Jakana, B.V.V. Prasad, J. King, P.E. Prevelige, Jr. and W. Chiu, J. Mol. Biol. 260 (1996) 85.
[42] I.K. Robinson and S.C. Harrison, Nature 297 (1982) 563.
[43] K. Kawaguchi, H. Noda and I. Katsura, J. Mol. Biol. 164 (1983) 573.