Role of Gene 10 Protein in the Hierarchical Assembly of the Bacteriophage P22 Portal Vertex Structure†

Adam S. Olia,‡ Anshul Bhardwaj,‡ Lisa Joss,§ Sherwood Casjens,|| and Gino Cingolani*,‡

Department of Biochemistry and Molecular Biology, SUNY Upstate Medical University, 750 E. Adams Street, Syracuse, New York 13210, Department of Biochemistry, University of Utah Medical School, Salt Lake City, Utah 84112, and Department of Pathology, University of Utah School of Medicine, Salt Lake City, Utah 84112

Received January 29, 2007; Revised Manuscript Received April 25, 2007

ABSTRACT: The portal vertex structure of the phage P22 is a 2.8 MDa molecular machine that mediates attachment and injection of the viral genome into the host *Salmonella enterica* serovar Typhimurium. Five proteins form this molecular machine: the portal protein, gp1; the tail-spike, gp9; the tail-needle, gp26, and the tail accessory factors, gp4 and gp10. In order to understand the assembly of the portal vertex structure, we have isolated the gene encoding tail accessory factor gp10 and defined its structural composition and assembly within the portal vertex structure. In solution, monomeric gp10 is a β-sheet-rich protein with a stable conformational structure, which spontaneously assembles into hexamers, likely via a dimeric intermediate. This oligomerization enhances the structural stability of the protein, which then becomes competent for assembly to a preformed portal protein:gp4 complex, and acts as a structural adaptor bridging the nascent phage tail to gp26 and gp9. Notably, *in vitro* purified tail accessory factors gp4, gp10, and gp26 do not significantly interact with each other in solution, but their assembly takes place efficiently when these factors are added sequentially onto an immobilized portal protein. This suggests that the assembly of the P22 tail is a highly sequential and cooperative process, likely mediated by structural rearrangements in the assembly components. The assembled portal vertex structure represents both a membrane-binding and penetrating device as well as a plug that retains the pressurized phage DNA inside the capsid.

P22 is a double-stranded DNA (dsDNA) phage that infects *Salmonella enterica* serovar Typhimurium. The mature phage consists of an icosahedral T = 7 capsid approximately 650 Å in diameter (*1*−*6*), which is interrupted at a single vertex by a 2.8 MDa tail apparatus (*7*) (also known as portal vertex structure). The portal vertex structure is used by the phage to attach to the *Salmonella* surface, penetrate the host cell wall and inject its 43 kbps chromosome into the host (*8*). In the virion, the portal vertex structure replaces a single coat protein penton of the capsid lattice and protrudes about 320 Å outside the coat protein shell (*7*). *In vitro*, the portal vertex structure can be extracted from the mature virus using chemical treatment (*7*), and it is composed of five distinct polypeptide chains assembled in varying stoichiometries, which include 1 dodecamer of portal protein gp1 (M.W.[1] ∼83.5 kDa) (*9*), 6 trimers of tail-spike protein gp9 (M.W. ∼71.9 kDa) (*10*), 12 copies of tail accessory factor gp4 (M.W. ∼18.0 kDa) (*6*, *11*), a trimer of gp26 (M.W. ∼24.7 kDa) (*12*−14), and, as shown in this article, a hexamer of gp10 (M.W. ∼52.3 kDa).

The assembly of phage P22 (*8*) takes place via the formation of a T = 7 meta-stable intermediate known as the procapsid, which consists of a protein shell, formed by coat protein gp5, surrounding the scaffolding protein gp8 and containing a dodecameric ring of the portal protein gp1 (*15*−*17*). The 43 kbps P22 chromosome is packaged into the newly formed virion through the portal protein ring in a process that requires ATP hydrolysis and a complex of the terminase subunits gp2 and gp3 (*18*, *19*). After packaging, the portal protein channel is closed by the action of the three tail accessory factors: gp4, gp10, and gp26, which for this reason are also known as portal closure or head completion proteins (*20*). Assembly of the three tail accessory factors is essential to stabilize the virus DNA inside the capsid, which is packaged at a quasi-crystalline concentration (*6*, *21*, *22*). Mutant phages that lack functional gp4, gp10, and gp26 encapsidate viral DNA efficiently but lose their DNA quickly into the surrounding environment. Such a characteristic DNA-leakage phenotype indicates that gp4, gp10, and gp26 are not involved in DNA packaging but mediate the critical function of plugging the portal protein channel to retain the dsDNA inside the capsid (*20*, *23*).

Gp4 is the first factor to bind to the growing tail structure after DNA packaging (*20*). This was demonstrated using both a complementation assay with virus extracts (*20*) as well as, more recently, ectopically expressed purified protein (*11*). The assembly of gp4 to portal protein involves the sequential binding of 12 copies of gp4 to form a structural hub, which

† This work was supported in part by NSF Grant MCB-990526 to S.R.C.

* Corresponding author. Tel: (315) 464-8744. Fax: (315) 464-8750; E-mail: cingolag@upstate.edu.

‡ SUNY Upstate Medical University.

§ Department of Biochemistry, University of Utah Medical School.

|| Department of Pathology, University of Utah School of Medicine.

[1] Abbreviations: M.W., molecular weight; BS[3], Bis(sulfosuccinimidyl) suberate, CD, circular dichroism; ITC, isothermal titration calorimetry.

10.1021/bi700186e CCC: $37.00
Published on Web 07/10/2007

extends the portal protein channel by approximately 30 Å (*11*). It was proposed that the binding of gp4 to portal protein involved two sets of *nonequivalent* binding sites in a process that may require conformational changes in the portal protein (*11*). Attachment of gp4 to portal protein is followed by the two other tail accessory factors gp10 and gp26. Although the correct order of assembly has not been conclusively proven, two sets of indirect evidence suggested gp10 assembles first to the nascent tail, followed by gp26, and then gp9 (*24*). First, it was observed that only gp10 and gp4 are required for gp9 addition (*18*, *20*), which attaches efficiently to capsids that lack gp26 (*25*). In contrast, capsids made in the absence of gp10 do not bind tail-spikes, suggesting that gp10 and not gp4 may physically connect the tail-spike protein to the virion (*20*). Second, in the recent cryo-EM reconstruction of the entire P22 tail extracted from mature phage (*7*) and in virions (*6*), a striking elongated density emanates from the neck of the tail that is thought to be gp26 (*6*, *26*). This is consistent with the observation that recombinant gp26 expressed in bacterium forms an ~220 Å elongated triple-stranded trimeric coiled-coil structure (*12*).

In this article, we have characterized the tail accessory factor gp10, defined its structural composition, and examined its assembly into the portal vertex structure. Our data indicates that gp10 oligomerizes in solution to form a hexamer, which acts as a structural scaffold to bridge gp4 to gp9 as well as gp26. Additionally, we have shown that the P22 tail can be assembled *in vitro* using immobilized portal protein and the sequential addition of recombinant purified tail accessory factors gp4, gp10, and gp26.

MATERIALS AND METHODS

Molecular Cloning of the Gene Encoding gp10. The genes encoding the tail accessory factor gp10 were PCR amplified from P22 DNA with primers 5′-CACGATGCATATGC-CGATTCAACACTCCCC and 5′-GT AGAGAGGATC-CGTGGCCTGAATAACGACAGG, respectively (see Pedulla et al. (*27*, *28*) for the nucleotide sequence of phage P22). These DNAs were cleaved with *Nde*I and *Bam*HI and ligated into similarly cleaved plasmid expression vector pET-15b (Novagen). The gp10 gene cloned in pET-15b (named pET15b-gp10) was sequenced to confirm the fidelity of the DNA sequence.

Expression and Purification of Recombinant Proteins. Recombinant proteins were expressed in *E. coli* (strain BL21) cells (New England BioLabs, Ipswich, MA), in LB broth supplemented with 2.5 g/L glucose. After growth at 37 °C to an OD_{600} of 0.6, gp4, gp26, and gp10 expression was induced with 0.5 mM IPTG, and the culture was incubated with shaking at 22 °C for 16 h. Cells were collected and lysed by sonication in Lysis Buffer (250 mM NaCl, 20 mM Tris-HCl at pH8.0, and 3 mM β-mercaptoethanol (β-ME)) plus various protease inhibitors. The target proteins fused to an N-terminal-6×His tag were purified by metal chelating affinity chromatography using Qiagen Nickel-NTA beads. Typically, 1 L of *E. coli* yielded about 15 mg of pure gp10 and 5 mg of pure gp4 or gp26, each of which was concentrated to approximately 50 mg/mL using a Millipore-Amicon concentrator (MW cutoff 10 KDa). Freshly eluted and concentrated protein was incubated at room temperature overnight (for gp10 only) and further purified by gel filtration on a Superdex S-200 column (Amersham Biosciences) equilibrated in Gel Filtration Buffer (200 mM NaCl, 20 mM Tris-HCl at pH 8.0, 3 mM β-ME, and 0.1 mM PMSF). Calibration of the S-200 column was carried out using high molecular weight globular protein standards (BioRad, Richmond, CA). Recombinant gp1 was expressed and purified as previously described (*11*).

Cross-Linking. Chemical cross-linking experiments were carried out using the homobifunctional cross-linker Bis-(sulfosuccinimidyl) suberate (BS^3) (Pierce Reagents) using a procedure adapted from ref *29*. Gp10 from the oligomer peak of SEC was concentrated to 20 mg/mL. Concentrated gp10 in 200 mM NaCl and 20 mM Tris-HCl at pH 8.0 was diluted to 8 mg/mL in 1× PBS (20 mM sodium phosphate buffer and 170 mM NaCl at pH 8.0). The cross-linker was added at a concentration ranging from 7 to 250 μM. Incubation was allowed to progress for 60 min at room temperature, at which point the reaction was brought to 50 mM Tris-HCl at pH 8.0 to quench the reaction. After quenching for approximately 20 min, SDS sample buffer was added and the reactions analyzed on 7% SDS−PAGE with coomassie staining.

Circular Dichroism Spectroscopy. Circular dichroism (CD) experiments were performed on an Aviv 62A DS circular dichroism spectrometer equipped with a temperature-controlled cell. The wavelength dependence of ellipticity was monitored at 25 °C on a 0.25 mg/mL (4.8 μM GP10 or 0.8 μM $(GP10)_6$) protein solution in PBS from 197 to 260 nm in a 0.1 cm path length cuvette. Secondary structure contents were calculated by the *K2d* method available on DichroWeb Server (*30*). Thermal stability was determined by monitoring the change in ellipticity at 208 nm as a function of temperature. Thermal melts were performed in 1 °C increments with an equilibration time of 60 s at the desired temperature and an integration time of 15 s, using a 0.1 cm path length cuvette. The melting temperatures or midpoints of thermal unfolding curves (T_m) have been calculated.

Native Gel Electrophoresis and Pull-Down Assay. Native agarose gels were prepared as described in ref *11*. The pull-down assay was performed in PBS buffer using CNBr activated Sepharose beads (Amersham). Prior to coupling, the beads were extensively washed with 1 mM HCl. Dodecameric gp1 was covalently cross-linked to beads overnight at 4 °C in coupling buffer (0.1 M NaHCO3 at pH 8.3 and 0.5 M NaCl). The reaction was then stopped with 100 mM Tris-HCl at pH 8. Gp1-beads were washed with PBS and incubated with 100 μg His-tagged gp4. Gp4:gp1-beads were washed and then incubated with the same amounts of gp10 and gp26 added either simultaneously or sequentially. All binding reactions were performed in a total volume of 50 μL for 30 min at room temperature. Beads were then washed 3 times with 1 mL of PBS and bound proteins eluted by boiling in SDS−sample buffer and analyzed by SDS−PAGE.

Isothermal Titration Calorimetry. Isothermal titration calorimetry experiments were carried out in a VP-ITC (Microcal), and the isotherm data was analyzed using the Origin software included with the instrument. Briefly, known concentrations of ligand, gp10, was injected into a known amount of a gp4 or portal:gp4. The heat released from each injection was integrated, and the baseline manually corrected, and a buffer experiment subtracted as a blank. Following

these corrections, the data was fit to a model, which in this case was the predefined One Set of Sites model (Microcal−Origin) (defined by the equation $\Delta Q(i) = Q(i) + (dV(i)/V_0)((Q(i) + Q(i+1))/2) - Q(i-1))$. The number of binding sites (n), enthalpy (ΔH), and association constant (K_A) were extracted from the model, whereas entropy ($T\Delta S$), Gibbs free energy (ΔG), and dissociation constant (K_D) were calculated using the equations $\Delta G = \Delta H - T\Delta S$ and $K_A = 1/K_D$.

RESULTS

GP10 Forms an Oligomer in Dynamic Exchange with Monomers and Dimers. P22 gene *10*, which encodes gp10, was PCR amplified from P22 DNA and cloned into the pET15b vector (Novagen). Recombinant gp10 fused to an N-terminal 6-histidine tag (MW 54.6 kDa) was purified from soluble *E. coli* extracts by metal chelating affinity chromatography. Attempts to determine the oligomeric state of the gp10 protein in solution using sedimentation equilibrium analysis were not conclusive due to the tendency of gp10 to form large aggregates during prolonged centrifugation, complicating the analysis. But if the highest MW data was excluded, the data were compatible with a gp10 size of 300−360 kDa at 23 °C (data not shown); because the MW of His-tagged gp10 is ∼55 kDa, therefore, this data is compatible with 5, 6, or 7 subunit oligomers. On a Superdex 200 size exclusion chromatography (SEC) column, gp10 migrates differently depending on the concentration of the sample analyzed, suggesting that monomeric gp10 spontaneously oligomerizes to form larger molecular weight assemblies. To examine gp10 oligomerization in more detail, we systematically varied the concentration, temperature, and time of incubation of purified gp10 protein. Freshly purified gp10 at 2 mg/mL was divided into six aliquots of 1 mL, each containing 2 mg of pure protein. Two of these aliquots were concentrated 10-fold to a final concentration of 20 mg/mL, and other two aliquots were brought to 40 mg/mL. A fraction of each of three gp10 samples at 2, 20, and 40 mg/mL was incubated at 4 °C, whereas a second set of fractions was held at 23 °C for about 18 h. After incubation, samples were analyzed at 4 °C by SEC in buffer containing 200 mM NaCl. The gp10 samples incubated at 4 °C (Figure 1A) eluted primarily as two peaks: a predominant species with an apparent molecular mass of 55 kDa (most likely monomeric gp10, MW 54.6 kDa) and a larger species with an apparent molecular mass of about 120 kDa (consistent with dimeric gp10) (Figure 1A). Comparing the three gp10 samples in Figure 1A, the gp10 fraction concentrated to 20 mg/mL and 40 mg/mL resulted in slightly decreased monomeric gp10 and to a more pronounced dimer and higher MW species. In contrast, the samples of gp10 incubated at 23 °C displayed a clear concentration-dependent oligomerization (Figure 1B). While gp10 at 2 mg/mL eluted primarily as putative monomers and dimers, gp10 at 20 mg/mL displayed a large peak at a molecular weight corresponding to about 340 kDa, which is consistent with a hexamer (6×54.6 kDa $= 328$ kDa). Finally, gp10 at 40 mg/mL eluted in a manner similar to that of the sample at 20 mg/mL, with a lower fraction of monomers and dimers (Figure 1B). Samples from the three peaks appear identical on SDS−PAGE (data not shown) but migrated as three distinct species when analyzed by native gel electrophoresis at room temperature (Figure 1C). Interestingly, monomeric gp10 migrated as a diffuse and smeared band (Figure 1C, lane 3), likely reflecting the intrinsic flexibility of this protein that may adopt different conformations in solution and on agarose gel. Upon dimerization, gp10 formed a faster migrating species (Figure 1C, lane 2), which may result from the neutralization of positive charges at the monomer−monomer binding interface, resulting in a larger overall negative charge. Finally, oligomeric gp10 also ran as a single band that was retarded on native gel with respect to dimeric gp10 (Figure 1C, lane 1). The pattern of separation on agarose gel reflects the different size/shape and overall charge of gp10 monomer, dimer, and oligomer, which strengthens the idea that each of these three species adopts a distinct structure. Finally, gp10 concentrated to >100 mg/mL also retained a significant amount of monomer and dimer in solution, as determined by SEC analysis. Additionally, when the gp10 oligomer peak is injected onto SEC once more, the protein redistributes into the distinct populations. Taken together, this data indicates that gp10 has the propensity to form higher order assemblies, which dynamically exchange in solution with monomers and dimers, in a process that that strongly depends on gp10 concentration and temperature.

Chemical-Linking Confirms that Oligomeric gp10 Is a Hexamer. To accurately assess the oligomeric state of the gp10 oligomeric peak (pkO) seen on gel filtration (Figure 1B), we used the homobifunctional cross-linker Bis(sulfosuccinimidyl) suberate (BS^3), which is a lysine-specific cross-linking agent. BS^3 was added to the oligomeric gp10 peak from SEC in a cross-linker concentration ranging from 7 to 250 μM, and cross-linked products were analyzed on 7% SDS−PAGE followed by staining with coomassie blue. As seen in Figure 1D (lanes 4−9), five distinct bands could clearly be discerned on the gel. At higher concentrations of cross-linker (100− 250 μM) the third, fourth, and fifth bands began to decrease in intensity and populate a sixth band, migrating slightly slower (Figure 1D, lanes 10−13). Cross-linked species may not necessarily run in a manner identical to that of a single linear polypeptide of the same size; therefore, the lower three bands correspond reasonably well with the predicted sizes of a gp10 monomer, dimer, and trimer (∼ 55, 110, and 165 kDa). Likewise, the top three bands migrating above the 201 kDa molecular weight marker likely represent cross-linked populations of gp10 tetramer, pentamer, and hexamer. Even in the presence of large excesses of cross-linker (1−5 mM), no bands are apparent above the hexamer band (data not shown), indicating that the hexamer represents the actual species in solution. Interestingly, at the highest concentration of cross-linker (Figure 1D, lane 13), only three predominant bands are visible on gel, corresponding to monomeric, dimeric, and hexameric gp10 in an approximate molar ratio of 40:10:40. This agrees well with the observation that gp10 oligomerization takes place via a stably populated dimeric intermediate, which is visible on a gel filtration column (Figure 1A and B).

Gp10 Secondary Structure and Thermal Stability. To characterize the secondary structure of gp10, we recorded circular dichroism (CD) spectra of the three gp10 populations identified on gel filtration. The CD spectra of isolated $(gp10)_1$, $(gp10)_2$, and $(gp10)_6$ show a shallow minima at 220 and 208 nm (Figure 2A), consistent with a poor helical content and a large random coil/β-sheet component. To

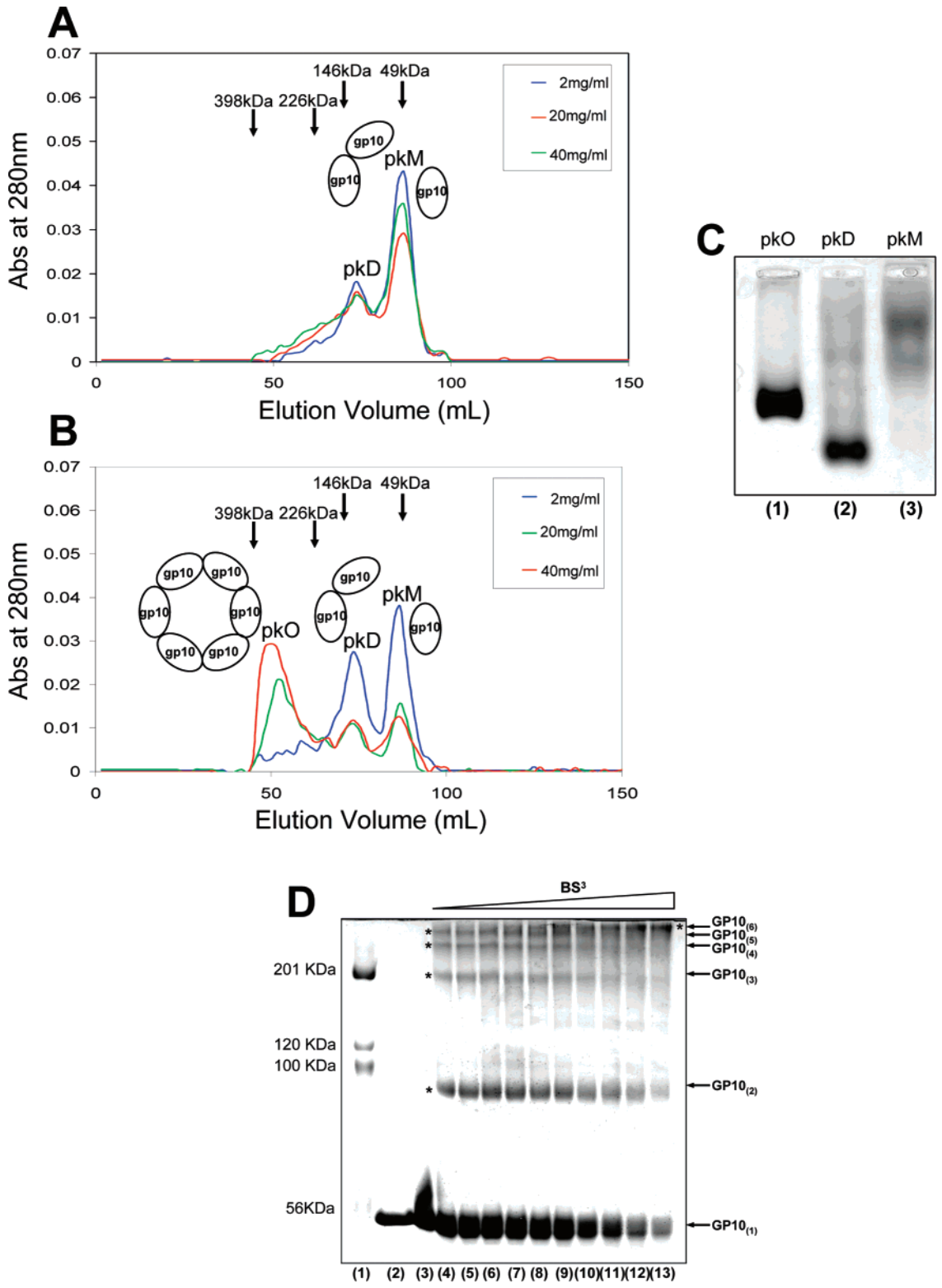

FIGURE 1: Characterization of gp10 protein. Gp10 forms an oligomer in dynamic equilibrium with monomer and dimer. (A) Gel filtration analyses of gp10 at 2 mg/mL, 20 mg/mL, and 40 mg/mL (in blue, red, and green, respectively). Gp10 samples were incubated for 18 h at 4 °C prior to loading on a Superdex 200 gel filtration column. (B) Gp10 samples were preincubated at 23 °C for 18 h to allow oligomerization and subsequently separated by Superdex 200 gel filtration. In both panels, the arrows indicate the elution volumes of the markers and their apparent MWs. The MW markers were analyzed at 4 °C using identical flow rates and ionic strength and correspond to Ovalbumin (49 kDa), Aldolase (146 kDa), Catalase (226 kDa), and Ferritin (398 kDa). On the basis of this calibration, peaks 1, 2, and 3 of gp10 (indicated as pkM, pkD, and pkO) likely correspond to monomer, dimer, and hexamer. (C) Native agarose gel of oligomer (pkO), dimer (pkD), and monomer (pkM) peaks from gel filtration. (D) Cross-linking of purified gp10 by BS^3. Purified gp10 corresponding to the high molecular weight species in pkO was incubated at 20 °C for 60 min at 8 mg/mL without (lane 2) or with increasing concentrations of homobifunctional cross-linker BS^3, in a concentration range 7−250 μM (lanes 3−13). After quenching the cross-linking with 50 mM Tris-HCl at pH 8.0, the reaction mixture was analyzed on a 7% SDS−PAGE. In lanes 3−13, the cross-linked species (indicated by an asterisk) appear as a ladder above the monomeric gp10 band. The very uppermost band (likely corresponding to a hexamer) does not appear in less than 100 μM cross-linker, at which concentration the other 5 bands diminish and begin to populate the hexamer.

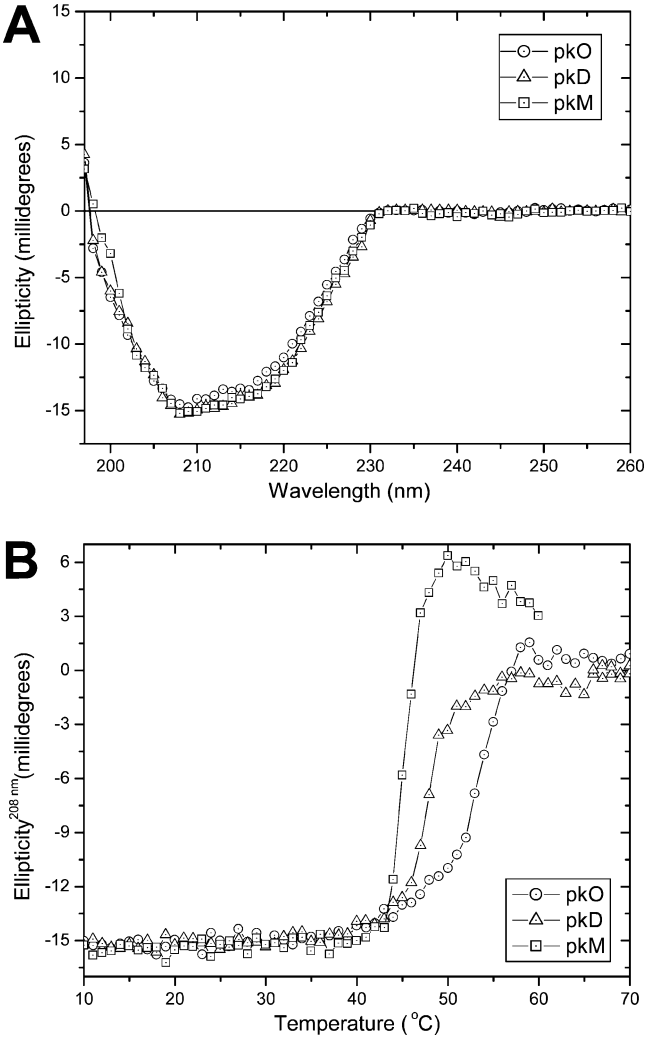

FIGURE 2: Secondary structure and stability of gp10. Far-UV CD spectra from 197 to 260 nm at 25 °C (A) and thermal unfolding curves (B) of the three populations of gp10 purified by size exclusion chromatography corresponding to monomer $(gp10)_1$, pkM (□); dimer $(gp10)_2$, pkD (△); and hexamer $(gp10)_6$, pkO (○) were recorded in 20 mM sodium phosphate buffer at pH 8 and 170 mM NaCl. Changes in the mean ellipticity at 208 nm as a function of temperature were measured at 1 °C increments at constant intervals of 120 s. The apparent T_m values for monomer $(gp10)_1$, dimer $(gp10)_2$, and hexamer $(gp10)_6$ are 45, 48, and 53 °C, respectively.

quantify the CD data, the data were analyzed using the DichroWeb server (*30*), which predicts the secondary structure to be largely random coil and to contain nearly 1/3 β-sheet.

To determine the structural stability of gp10 in the different oligomeric states, we recorded thermal denaturation curves. In this experiment, variations in ellipticity at 208 nm were measured while increasing the temperature from 10 to 70 °C. Regardless of the oligomeric states, thermal denaturation profiles for gp10 were found to be sigmoidal with significant variations in steepness (Figure 2B), pointing toward a cooperative two-state transition. All melting experiments were found to be irreversible with characteristics of heavy aggregates at increased temperatures. The apparent T_m values for $(gp10)_1$, $(gp10)_2$, and $(gp10)_6$ are 45, 48, and 53 °C, respectively. Oligomerization shifts the apparent T_m toward higher temperature, consistent with the idea that quaternary structure stabilizes the structure of gp10.

Gp4 Bridges the Interaction of Portal Protein with Tail Factor gp10. Previous work has shown that the tail accessory factor gp4 acts as a structural adaptor, bridging the P22 portal vertex to the other two tail accessory factors gp10 and gp26. Increasing biochemical (*12*) and structural (*7*) evidence indicates that the latter protein forms an elongated fiber-like needle emanating from the bottom center of the P22 tail. Consequently, gp26 interaction to the portal protein: gp4 complex is not likely to be direct but rather to be mediated by gp10. To test this hypothesis, we first tested the binding of purified oligomeric gp10 to dodecameric portal protein using native agarose gel electrophoresis. As shown in Figure 3A, no interaction was observed between these two proteins. In a second set of experiments, we examined the binding of gp10 to a preformed portal:gp4 complex. In this experiment, increasing concentrations of gp10 were titrated into a preformed gp1:gp4 complex, resulting in a slower migrating trimeric complex gp1:gp4:gp10 (Figure 3B, lane 5−14). Efficient binding of gp10 to the preformed gp1: gp4 complex was observed only using oligomeric gp10 from the high molecular weight elution peak (pkO) in Figure 1B, whereas monomeric and dimeric gp10 showed no appreciable binding (data not shown). Approximately 6 to 9 M equiv of gp10 complex were necessary to fully shift the $(portal)_{12}$: $(gp4)_{12}$ (Figure 3B, lane 10−13), which further emphasizes the idea that only oligomeric gp10 is biologically competent for assembly but leaves unresolved how many copies of gp10 attach to the portal:gp4 complex.

To more accurately access the stoichiometry of gp10 binding to the forming tail as well as to quantify the thermodynamic parameters of binding, isothermal titration calorimetry (ITC) was performed using hexameric gp10 titrated into either pure gp4 or the fully oligomeric $(portal)_{12}$: $(gp4)_{12}$ complex. Whereas no productive binding was observed mixing oligomeric gp10 with gp4 (Figure 3C), saturable release of enthalpy was measured when the portal: gp4 complex was mixed with the oligomeric gp10 (Figure 3D). ITC data fit well to a binding isotherm with n (number of binding sites) equal to ~0.5, consistent with the idea that the portal:gp4 complex exists as a dodecamer, and the oligomer of gp10 corresponds to a hexamer (Figure 3E). The dissociation constant, (K_D), ΔH, and the number of binding sites (n) obtained from the fit were $K_D = 5.03 \pm 0.52\ \mu M$, $\Delta H = -1486 \pm 58.79$ cal/mol, $n = 0.503 \pm 0.0142$. The remaining thermodynamic values were calculated from these values and determined to be $\Delta S = 19.4$ cal/K·mol and $\Delta G = 7377$cal/mol. The entropic component of the binding sizably outweighs the enthalpic component, $\Delta H = -1486$ cal/mol vs $T\Delta S = 5897$cal/mol (Figure 3F), characteristic of a hydrophobic interaction, which may help to explain the stability of this portion of the tail at high salt concentrations. In addition, the low micromolar K_D is not as strong as one would expect for a component of the P22 tail, which is an ultrastable macromolecular complex that survives prolonged heating in 2 M urea plus detergent (*7*). This indicates that binding of gp26 may act as a lock to cement the complex together.

Our data indicates that most likely a hexamer of gp10 assembles to the portal:gp4 complex of a nascent portal vertex structure. In contrast to gp4, which exists as a monomer in solution and oligomerizes upon binding to portal protein (*11*), a preformed hexamer of gp10 appears to be

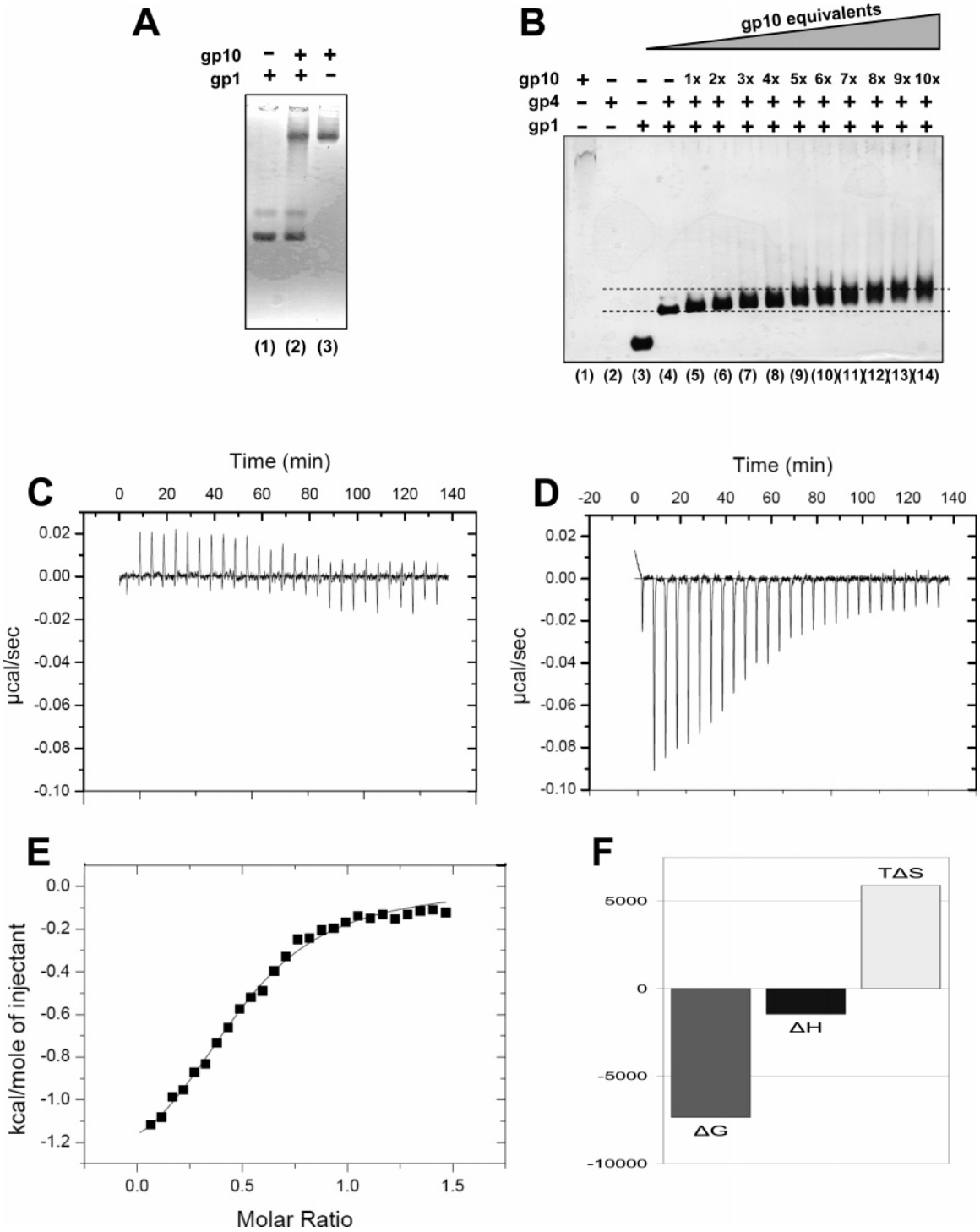

FIGURE 3: Gp10 binds to a pre-assembled portal protein:gp4 complex. (A) Oligomeric gp10 fails to interact with dodecameric portal protein on agarose gel. (B) Gp10 binds to a preformed portal ring:gp4 complex in a concentration-dependent manner. Between 6 and 9 equiv of gp10 are necessary to fully shift the gp1:gp4 complex (lanes 10−14). (C) Isothermal titration calorimetry (ITC). Raw enthalpy variations were measured by releasing injected oligomeric gp10 into gp4. The data are non-fittable to a binding isotherm; the random distribution of heat is diagnostic of unproductive binding, suggesting that isolated gp4 and oligomeric gp10 have no binding specificity in solution. (D) Raw data of ITC of GP10 injected into the portal protein:gp4 complex. The heat released shows good saturable binding. (E) Integration of injection peaks from D. The isotherm fits well to the model shown with $K_D = 5.03 \pm 0.52$ μM. The variation of enthalpy (ΔH), entropy ($T\Delta S$), and Gibbs (ΔG) energy associated to each binding event are shown in panel F. As shown, the entropic component $T\Delta S$ (5897 cal/mol) greatly outweighs the enthalpic contribution ΔH (−1486 cal/mol).

necessary and sufficient for efficient binding to the growing tail. Overall, these findings agree well with the recent three-dimensional reconstruction of P22 tail, where six lobes of density beneath the portal protein ring were assigned to a hexamer of gp10 (*6*, *7*).

Binding of Tail-Needle gp26 to the Nascent Tail Requires a Tail-Bound gp10. After determining that gp10 interaction with the portal protein is mediated by gp4 in its portal-bound oligomeric state, we focused on the interaction of gp10 with tail-needle gp26. We determined that on agarose gel the

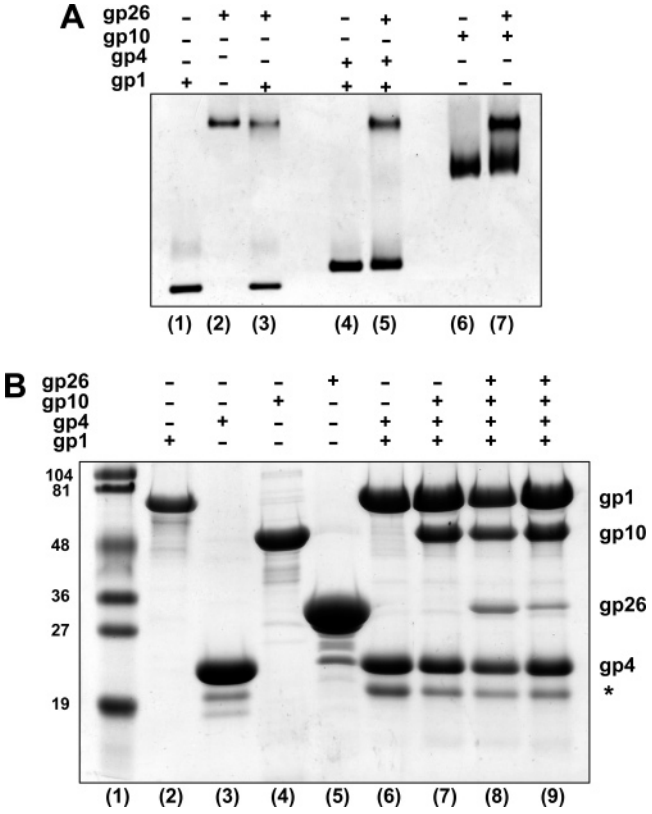

FIGURE 4: Tail accessory factor gp4, gp10, and gp26 assemble sequentially to portal protein. (A) Native agarose gel showing the failure of gp26 to bind portal protein, portal:gp4 complex, or gp10. Lanes 1, 2, and 3 show purified oligomeric portal protein, gp26, and portal protein mixed with gp26, respectively. Lanes 4 and 5 show portal:gp4 and portal:gp4 mixed with gp26, whereas lanes 6 and 7 show purified oligomeric gp10 and gp10 with gp26. As clearly seen, gp26 does not interact with any of the tested pure proteins or complexes. (B) Assembly of gp4, gp10, and gp26 *in vitro* to immobilized portal protein on CNBr beads, analyzed by SDS−PAGE. Dodecameric gp1 was covalently coupled to CNBr-beads (lane 1). The purified tail accessory factors gp4, gp10, and gp26 used in the assembly assay are shown in lanes 2, 3, and 4, respectively. Gp1-bound beads were used to selectively pull down gp4 (lane 6). A preformed gp1:gp4 complex efficiently pulls-down gp10 (lane 7). To incorporate gp26, recombinant purified gp26 was either sequentially added to a preformed gp1:gp4:gp10 complex (lane 8) or added to gp1:gp4 beads together with gp10 (lane 9). The latter resulted in a lower amount of gp26 incorporated into the complex (compare the intensity of the gp26 band in lanes 8 and 9). The minor band below gp4 (marked with an asterisk) represents an N-terminal deletion fragment of gp4, which has lost the N-terminal His-tag, and preserves full binding activity for gp1.

addition of gp26 trimers did not alter the electrophoretic mobility of oligomeric portal protein (Figure 4A, lanes 1−3), and likewise, no interaction was observed between free gp26 and the gp1:gp4 complex (Figure 4A, lanes 4−5). This suggests that the interaction of the tail needle with the portal ring:gp4 complex is indirect and likely mediated by gp10. However, by this assay oligomeric gp10 failed to bind to gp26 (Figure 4A, lanes 6−7), suggesting that only the conformation of gp10 assembled in the tail is competent for high affinity binding to tail-needle gp26. Unfortunately, when the preformed portal:gp4:gp10 complex was incubated with gp26 and analyzed by native gel electrophoresis, the resulting product was badly smeared and difficult to interpret. Therefore, we devised a pull-down assay that uses portal protein dodecameric rings covalently cross-linked to CNBr beads (see Materials and Methods). Purified tail accessory factors gp4, gp10, and gp26 were added in various combinations to beads that had covalently attached portal protein rings, and SDS−polyacrylamide gels were used to determine the composition of any formed complexes. As seen in Figure 4B, lanes 6−9, each addition of gp4 and gp10 resulted in strong and specific binding. Likewise, gp26 was pulled down by gp1:gp4 beads either by preincubating gp10 with the beads or by simultaneously adding gp10 and gp26 to the beads. Interestingly, sequential addition of gp10 and gp26 resulted in a greater amount of bound gp26 than when the two factors were added simultaneously (compare the gp26 band in Figure 4B, lane 8 and 9). Similar pull-down assays using portal protein bound to the CNBr beads and any combinations of gp10 and gp26 in the absence of gp4 clearly indicated no significant interaction, and no interaction between portal:gp4 and gp26 could be detected in the absence of gp10 (data not shown). Taken together, this data confirms that gp4 is a structural adaptor to gp10, which in turn provides an attachment site for tail-needle gp26. Because gp10 and gp26 do not bind one another outside the portal vertex context, structural and conformational changes in gp10 upon binding to the portal:gp4 complex may be necessary to expose the quaternary structure conformation competent for binding to gp26.

DISCUSSION

Molecular Characterization of gp10. The assembly of the portal vertex structure is a critical event in the morphogenesis of the phage P22 virion. It requires the ordered and sequential addition of four soluble proteins (gp4, gp10, gp26, and gp9) to capsid-bound dodecameric portal protein (gp1). In an attempt to better define the role of tail accessory factor gp10, we have characterized the structure, composition, and binding interactions of this protein in the phage tail. Interestingly, we found that gp10 has a propensity to oligomerize in solution and form hexamers $(gp10)_6$ likely via the formation of a dimeric intermediate $(gp10)_2$. The equilibrium $(gp10)_1 \leftrightarrow (gp10)_2 \leftrightarrow (gp10)_6$ is sensitive to both temperature and concentration, but even at high concentration (e.g., 100 mg/mL), a significant component of monomer and dimer is observed *in vitro*. Oligomerization of gp10 enhances the structural stability of the protein, as confirmed by thermal unfolding studies where the temperature of melting (T_m) of oligomeric gp10 was ~8 °C higher than that of the monomer (T_m ~45 vs 53 °C). Likewise, hexameric but not monomeric gp10 binds with high specificity to a preformed portal:gp4 complex, suggesting that this oligomeric conformation is biologically relevant. The stoichiometry of this interaction is such that six copies of gp10 associate with a bi-dodecameric portal:gp4 complex, implying that a symmetry transition exists at the gp4:gp10 binding interface. We hypothesize that such a symmetry transition is a direct consequence of the two *nonequivalent* sets of gp4 bound to the portal protein (*11*). As previously reported, gp4 uses the portal dodecamer to switch from a meta-stable monomer to an oligomer, likely formed by two sets of *nonequivalent* binding sites (*11*). In this article, we show that the preformed portal protein:gp4 complex becomes the attachment site for gp10, which oligomerizes into a hexamer likely as a triplet of dimers. The symmetry is then transmitted to gp26, which exists as a trimeric coiled-coil fiber. The final result of this

hierarchical assembly, the virus tail, manifests extraordinary structural stability (*7*) and represents the first weapon used by the phage to pierce the host cell wall (*31*).

Hierarchical Assembly of Tail Factors in the Portal Vertex Structure. The assembly of icosahedral viruses is strongly governed by symmetry. Consistently, our analysis of phage P22 portal vertex structure assembly provides clues that the sequential and ordered self-assembly of this molecular machine takes place in response to at least three structural and entropic determinants. First, the portal protein is locked at the 5-fold vertex of P22 icosahedral capsid, which acts as a stator preventing the portal from disassociating and equilibrating with monomeric gp1. This provides an elegant solution to the intrinsic structural polymorphism of P22 portal protein, which spontaneously assembles *in vitro* into 12- and 11-fold symmetric rings (*32*).

Second, the five proteins forming the P22 tail present distinct oligomerization properties. We distinguished at least three different oligomerization *modes* by which the P22 tail proteins can assemble to form higher order oligomers. Tail factors such as tail-spike gp9 and tail-needle gp26 always exist as trimers (*12*, *33*). This trimeric quaternary structure is necessary for the correct folding of each monomer, and their oligomerization *in vivo* likely takes place in the ribosome-bound nascent chain (*34*, *35*, *36*). In contrast, tail accessory factor gp10 and portal protein appear to oligomerize as a function of the monomer concentration and temperature (*37*), and in both cases, the correct quaternary structure (hexamer for gp10 and dodecamer for portal protein) is selected during assembly. Finally, gp4 oligomerizes only upon binding to portal protein and forms two sets of *nonequivalent* hexamers (*11*). This peculiar oligomerization is essential to link the 12-fold symmetric topology of portal protein to a hexamer of gp10.

Third, an essential determinant for the correct assembly of P22 portal vertex structure lies in the hierarchical nature of the addition of gp4, gp10, and gp26 to the portal dodecamer. In this biological *pipeline*, individual tail factors cannot interact with each other in solution, and their assembly into the tail vertex structure requires a preformed complex on the portal protein for addition to occur. Gp4 is the first protein to add to the portal protein and, in the process of binding, oligomerizes to form a dimer of hexamers (*11*). This then forms the determinant for gp10 binding. Gp10 requires gp4 in the correct oligomeric state for productive binding, which precludes gp10 and gp4 from interacting in solution. Along the same lines, gp26 shows no interaction to soluble hexameric gp10, which suggests that the tail needle recognizes the conformation adopted by gp10 onto the nascent tail, rather than the simple oligomeric state of gp10, and therefore, gp10 likely changes its conformation while assembling into the tail. This cascade of sequential association and conformational changes renders the assembly of phage P22 an irreversible downhill reaction, which takes place with extraordinary efficiency.

ACKNOWLEDGMENT

We thank Nancy Walker at SUNY Upstate Medical University for excellent technical support. We also thank Rebecca Oot for her assistance in the purification of the tail factors and Danella Winn-Stapely for construction of the histidine-tagged gp10 expression plasmid.

REFERENCES

1. Casjens, S. (1979) Molecular organization of the bacteriophage P22 coat protein shell, *J. Mol. Biol. 131*, 1−14.
2. Prasad, B. V., Prevelige, P. E., Marietta, E., Chen, R. O., Thomas, D., King, J., and Chiu, W. (1993) Three-dimensional transformation of capsids associated with genome packaging in a bacterial virus, *J. Mol. Biol. 231* , 65−74.
3. Thuman-Commike, P. A., Greene, B., Jakana, J., Prasad, B. V., King, J., Prevelige, P. E., and Chiu, W. (1996) Three-dimensional structure of scaffolding-containing phage p22 procapsids by electron cryo-microscopy, *J. Mol. Biol. 260*, 85−98.
4. Thuman-Commike, P. A., Greene, B., Malinski, J. A., Burbea, M., McGough, A., Chiu, W., and Prevelige, P. E. (1999) Mechanism of scaffolding-directed virus assembly suggested by comparison of scaffolding-containing and scaffolding-lacking P22 procapsids, *Biophys J. 76* (6), 3267−3277.
5. Thuman-Commike, P. A., Greene, B., Jakana, J., McGough, A., Chiu, W., Prevelige, P. E., and Chiu, W. (2000) Identification of additional coat-scaffolding interactions in a bacteriophage P22 mutant defective in maturation, *J. Virol. 74*, 3871−3873.
6. Lander, G. C., Tang, L., Casjens, S. R., Gilcrease, E. B., Prevelige, P., Poliakov, A., Potter, C. S., Carragher, B., and Johnson, J. E. (2006) The structure of an infectious p22 virion shows the signal for headful DNA packaging, *Science 312*, 1791−1795.
7. Tang, L., Marion, W. R., Cingolani, G., Prevelige, P. E., and Johnson, J. E. (2005) Three-dimensional structure of the bacteriophage P22 tail machine, *EMBO J. 24*, 2087−2095.
8. Casjens, S., and Weigele, P. (2005) Headful DNA Packaging by Bacteriophage P22 in Viral Genome Packaging Machines: Genetics, Structure and Mechanism, in *The Bacteriophages* (Catalano, R., Ed.) pp 457-468, Landes Publishing, Georgetown, TX.
9. Bazinet, C., Benbasat, J., King, J., Carazo, J. M., and Carrascosa, J. L. (1988) Purification and organization of the gene 1 portal protein required for phage P22 DNA packaging, *Biochemistry 27*, 1849−1856.
10. Goldenberg, D., and King, J. (1982) Trimeric intermediate in the in vivo folding and subunit assembly of the tail spike endorhamnosidase of bacteriophage P22, *Proc. Natl. Acad. Sci. U.S.A., 79*, 3403−3407.
11. Olia, A. S., Al-Bassam, J., Winn-Stapley, D. A., Joss, L., Casjens, S. R., and Cingolani, G. (2006) Binding-induced stabilization and assembly of the phage P22 tail accessory factor Gp4, *J. Mol. Biol. 363*, 558−576.
12. Andrews, D., Butler, J. S., Al-Bassam, J., Joss, L., Winn-Stapley, D. A., Casjens, S., and Cingolani, G. (2005) Bacteriophage P22 tail accessory factor GP26 is a long triple-stranded coiled-coil, *J. Biol. Chem. 280*, 5929−5933.
13. Casjens, S., and King, J. (1982) P22 morphogenesis. I: catalytic scaffolding protein in capsid assembly, *J. Supramol. Struct. 2*, 202−224.
14. Hartweig, E., Bazinet, C., and King, J. (1986) DNA injection apparatus of phage P22, *Biophys. J. 49*, 24−26.
15. King, J., Lenk, E. V., and Botstein, D. (1973) Mechanism of head assembly and DNA encapsulation in *Salmonella* phage P22. II. Morphogenetic pathway, *J. Mol. Biol. 80*, 697−731.
16. Earnshaw, W., Casjens, S., and Harrison, S. C. (1976) Assembly of the head of bacteriophage P22: X-ray diffraction from heads, proheads and related structures, *J. Mol. Biol. 104*, 387−410.
17. Thuman-Commike, P. A., Greene, B., Malinski, J. A., King, J., and Chiu, W. (1998) Role of the scaffolding protein in P22 procapsid size determination suggested by T = 4 and T = 7 procapsid structures, *Biophys. J. 74*, 559−568.
18. Botstein, D., Waddell, C. H., and King, J. (1973) Mechanism of head assembly and DNA encapsulation in *Salmonella* phage p22. I. Genes, proteins, structures and DNA maturation, *J. Mol. Biol. 80*, 669−695.
19. Poteete, A. R., and Botstein, D. (1979) Purification and properties of proteins essential to DNA encapsulation by phage P22, *Virology 95*, 565−573.
20. Strauss, H., and King, J. (1984) Steps in the stabilization of newly packaged DNA during phage P22 morphogenesis, *J. Mol. Biol. 172*, 523−543.
21. Smith, D. E., Tans, S. J., Smith, S. B., Grimes, S., Anderson, D. L., and Bustamante, C. (2001) The bacteriophage straight phi29

portal motor can package DNA against a large internal force, *Nature 413*, 748−752.
22. Earnshaw, W. C., and Casjens, S. R. (1980) DNA packaging by the double-stranded DNA bacteriophages, *Cell. 21*, 319−331.
23. Lenk, E., Casjens, S., Weeks, J., and King, J. (1975) Intracellular visualization of precursor capsids in phage P22 mutant infected cells, *Virology. 68*, 182−199.
24. Israel, V. (1978) A model for the adsorption of phage P22 to *Salmonella* typhimurium, *J. Gen. Virol. 40*, 669−673.
25. Poteete, A. R., Jarvik, V., and Botstein, D. (1979) Encapsulation of phage P22 DNA in vitro, *Virology 95*, 550−564.
26. Berget, P. B., and Poteete, A. R. (1980) Structure and functions of the bacteriophage P22 tail protein, *J. Virol. 34*, 234−243.
27. Sampson, L., and Casjens, S. (1993) Nucleotide sequence of *Salmonella* bacteriophage P22 head completion genes 10 and 26, *Nucleic. Acids. Res. 21*, 3326.
28. Pedulla, M. L., Ford, M. E., Karthikeyan, T., Houtz, J. M., Hendrix, R. W., Hatfull, G. F., Poteete, A. R., Gilcrease, E. B., Winn-Stapley, D. A., and Casjens, S. R. (2003) Corrected sequence of the bacteriophage p22 genome, *J. Bacteriol. 185*, 1475−1477.
29. McDermott, A. M., and Haslam, R. J. (1996) Chemical cross-linking of pleckstrin in human platelets: evidence for oligomerization of the protein and its dissociation by protein kinase C, *Biochem. J. 317*, 119−124.
30. Whitmore, L., and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, *Nucleic. Acids. Res. 32*, W668−W673.
31. Israel, V., Rosen, H., and Levine, M. (1972) Binding of bacteriophage P22 tail parts to cells, *J. Virol. 10*, 1152−1158.
32. Poliakov, A., van Duijn, E., Lander, G., Fu, C. Y., Johnson, J. E., Prevelige, P. E., Jr., and Heck, A. J. (2007) Macromolecular mass spectrometry and electron microscopy as complementary tools for investigation of the heterogeneity of bacteriophage portal assemblies, *J. Struct. Biol. 157*, 371−383.
33. Steinbacher, S., Seckler, R., Miller, S., Steipe, B., Huber, R., and Reinemer, P. (1994) Crystal structure of P22 tailspike protein: interdigitated subunits in a thermostable trimer, *Science 265*, 383−386.
34. Clark, P. L., and King, J. (2001) A newly synthesized, ribosome-bound polypeptide chain adopts conformations dissimilar from early in vitro refolding intermediates, *J. Biol. Chem. 276*, 25411−25420.
35. Jain, M., Evans, M. S., King, J., and Clark, P. L. (2005) Monoclonal antibody epitope mapping describes tailspike beta-helix folding and aggregation intermediates, *J. Biol. Chem. 280*, 23032−23040.
36. Bhardwaj, A., Olia, A. S., Walker-Kopp, N., and Cingolani, G. (2007) Domain organization and polarity of tail needle gp26 in the portal vertex structure of bacteriophage P22, *J. Mol. Biol.* [Online early access], DOI:10.1016/j.jmb.2007.05.051.
37. Moore, S. D., and Prevelige, P. E., Jr. (2001) Structural transformations accompanying the assembly of bacteriophage P22 portal protein rings in vitro, *J. Biol. Chem. 276*, 6779−6788.

BI700186E