research articles Journal of proteome research

Identification of Subunit−Subunit Interactions in Bacteriophage P22 Procapsids by Chemical Cross-linking and Mass Spectrometry

Sebyung Kang,[†] Adam M. Hawkridge,[§,||] Kenneth L. Johnson,[§] David C. Muddiman,[§,||] and Peter E. Prevelige, Jr.[‡]

Department of Biochemistry & Molecular Genetics and Department of Microbiology, University of Alabama at Birmingham, Birmingham, Alabama 35294, and W.M. Keck FT-ICR Mass Spectrometry Laboratory, Mayo Proteomics Research Center, Department of Biochemistry & Molecular Biology, Mayo Clinic College of Medicine, Rochester, Minnesota 55905

Received October 21, 2005

Viral capsids are dynamic structures which self-assemble and undergo a series of structural transformations to form infectious viruses. The dsDNA bacteriophage P22 is used as a model system to study the assembly and maturation of icosahedral dsDNA viruses. The P22 procapsid, which is the viral capsid precursor, is assembled from coat protein with the aid of scaffolding protein. Upon DNA packaging, the capsid lattice expands and becomes a stable virion. Chemical cross-linking analyzed by mass spectrometry was used to identify residue specific inter- and intra-subunit interactions in the P22 procapsids. All the intersubunit cross-links occurred between residues clustered in a loop region (residues 157−207) which was previously identified by mass spectrometry based on hydrogen/deuterium exchange and biochemical experiments. DSP and BS^3 which have similar distance constraints (12 Å and 11.4 Å, respectively) cross-linked the same residues between two subunits in the procapsids (K183−K183), whereas DST, a shorter cross-linker, cross-linked lysine 175 in one subunit to lysine 183 in another subunit. The replacement of threonine with a cysteine at residue 182 immediately adjacent to the K183 cross-linking site resulted in slow spontaneous disulfide bond formation in the procapsids without perturbing capsid integrity, thus suggesting flexibility within the loop region and close proximity between neighboring loop regions. To build a detailed structure model, we have predicted the secondary structure elements of the P22 coat protein, and attempted to thread the prediction onto identified helical elements of cryoEM 3D reconstruction. In this model, the loop regions where chemical cross-linkings occurred correspond to the extra density (ED) regions which protrude upward from the outside of the capsids and face one another around the symmetry axes.

Keywords: bacteriophage P22 • capsid interactions • chemical cross-linking • mass spectrometry

Introduction

Many proteins in biological systems carry out their functions as large macromolecular complexes.[1] The assembly and the structural transformations of macromolecular complexes must be tightly regulated by proper interactions between neighboring subunits.[2]

Viral capsids are one well-known example of dynamic supramolecular structures. Viral capsids self-assemble and undergo a series of controlled structural transformations along a well defined pathway to reach an infectious form. The capsids are assembled from hundreds of copies of a small number of protein subunits. Self-assembly is achieved through the use of repeated intersubunit bonding with the result generally being a symmetrical structure.[3] The symmetrical nature of the final structure simplifies analytical methods such as crystallography,[4] cryo-electron microscopy,[5,6] and mass spectrometry[7−9] making viral capsids a preferred model for studying protein−protein interactions, serial conformational transformations and protein dynamics.

The *Salmonella typhimurium* bacteriophage P22 is a well studied model system for icosahedral capsid assembly and serves as a prototype for the morphogenesis of mammalian viruses. The multistep assembly pathway of bacteriophage P22 has been well characterized both biochemically and genetically[10] (Figure 1). In the first step, the procapsid, a $T = 7$ metastable viral capsid precursor with a diameter of 580 Å, is assembled from 420 copies of the 46.6 kDa coat protein with the aid of approximately 300 copies of the 33.6 kDa scaffolding protein. Minor proteins which are required for infectivity are incorporated at this stage. The scaffolding protein is not required for coat protein folding, but plays a key role in

[†] Department of Biochemistry & Molecular Genetics, University of Alabama at Birmingham.
[‡] Department of Microbiology, University of Alabama at Birmingham.
[§] Mayo Clinic College of Medicine.
[||] Current address: W. M. Keck FT-ICR Mass Spectrometry Laboratory, Department of Chemistry, North Carolina State University, Raleigh, NC 27695.

10.1021/pr050356f CCC: $33.50
Published on Web 01/20/2006

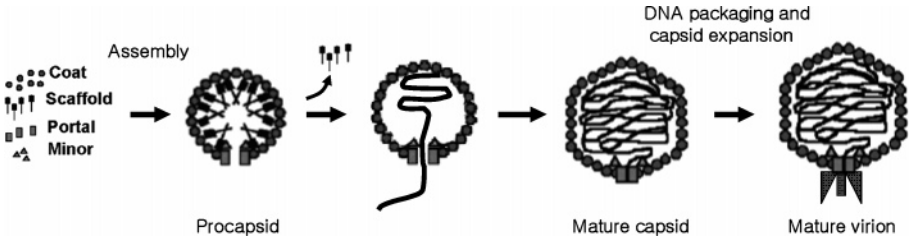

Figure 1. Assembly Pathway of Bacteriophage P22. The meta-stable procapsids are formed by co-assembly of 420 copies of coat protein monomer, 300 copies of scaffolding protein, 12 copies of portal protein and a few copies of minor proteins. As concatomeric phase DNA is packaged, procapsids are transformed into the mature capsids which are 10% larger in diameter and more stable than procapsids. Subsequently, portal ring is closed and tail attachment occurs.

assembly fidelity and activity.[11,12] The concatomeric phage DNA is packaged into the procapsid in a headful manner through the action of terminase proteins fueled by ATP hydrolysis.[10,13] Concomitant with DNA packaging, the meta-stable procapsid is transformed into the mature capsid and the scaffolding protein is released. The structural transformation results in a 10% increase in shell diameter,[6] a pronounced angular appearance and an increase in capsid stability.[14]

CryoEM based image reconstruction has provided a structure of the procapsid at ~9 Å resolution and allowed for the identification of helical elements within the subunit.[5] This analysis suggested that the structure of the subunits shared what has become known as the HK-97-fold,[5,15,16] a structure first identified in the crystal structure of the bacteriophage HK97.[4] Although a cryoEM reconstruction at this resolution can provide global structural information, it lacks sufficient resolution to directly identify the amino acids involved in subunit-subunit contacts.

Chemical cross-linking is the process of covalently joining two molecules or two regions in one molecule by the use of cross-linking reagents. Cross-linking reagents contain ends reactive for specific functional groups on proteins, nucleic acids, or drugs. Cross-linking reagents have been used to assist in determination of protein−protein interactions,[8,17,18] ligand−receptors[19,20] and 3D structures of proteins.[21,22] The use of cross-linkers with different arm lengths coupled with mass spectrometric analysis of the sites of cross-linking can provide distance constraints between reactive groups within and between protein subunits.[17,23] This information can be used to improve the resolution of molecular models of the protein complexes of interest. In this study, we have chemically cross-linked empty procapsid shells and identified inter- and intra-subunit interaction sites directly with mass spectrometry.

Material and Method

Preparation of Empty Procapsid Shells. Procapsids were prepared using *Salmonella typhimurium* strain DB7000 infected with P22 strain $2^{-}am/13^{-}am$ as described[24] and further purified by sucrose gradient centrifugation. Empty procapsid shells were prepared by repeated extraction of scaffolding protein with 0.5 M GuHCl at 4 °C.[24] Purified empty procapsid shells were stored in buffer B (50 mM Tris-HCl, 25 mM NaCl and 2 mM EDTA) at 4 °C.

Chemical Cross-Linking of the Empty Procapsid Shells. The amine-specific homobifunctional cross-linkers, dissuccinimidyl tartrate (DST), dithiobis[succinimidyl propionate] (DSP), and bis[sulfosuccinimidyl] suberate (BS^3) (Pierce Biotechnology Inc., Rockford, IL), were used to cross-link lysine residues in the empty procapsid shells according to the following protocol. The empty procapsid shells were dialyzed overnight at 4 °C against buffer containing 25 mM sodium phosphate, 25 mM NaCl, pH 7.6 and mixed with freshly prepared 0.1 M cross-linkers (1 mM final) which are dissolved in DMSO or water. Each reaction mixture was incubated at room temperature for 5 min and quenched with 60 mM Tris finally to avoid high oligomerization.

SDS-PAGE and In-Gel Digestion. Monomers, cross-linked dimers, and high oligomers were analyzed and separated by nonreducing 7.5% SDS-PAGE. Protein bands from SDS-PAGE were excised, destained, and digested with sequencing grade trypsin (Roche, Indianapolis, IN) in-gel at 37 °C overnight. Digestion was quenched with 1% formic acid and peptides were extracted by simple diffusion.[25]

Mass Spectrometry. Peptides digested in gel were directly analyzed by nanoLC-ESI time-of-flight mass spectrometry (LCT, Waters, Framingham, MA). 5 μL samples were manually injected using a gastight syringe into a 5 μL sample loop attached to a Rheodyne injection valve and tryptic peptide fragments separated on a prepacked nano MAGIC C_{18} reverse phase column (0.1 × 50 mm, 200 Å) (Michrom BioResources, Auburn, CA) using linear gradient from 5% to 95% acetonitrile containing 0.1% formic acid for 1 h at flow rate of 1 μL/min. Spectra were acquired in the range of m/z 200−1650. Mass calibration was performed with polyalanine to cover the range of m/z 200−1650.

Accurate mass measurements of tryptic peptide fragments were performed using nano-liquid chromatography (Eksigent, Livermore, CA) coupled to an 9.4 T electrospray ionization Fourier transform ion cyclotron resonance mass spectrometer (IonSpec Corp, Lake Forrest, CA) equipped with a dual electrospray source for internal calibration.[26] 10 μL samples were manually injected using a gastight syringe into a 50 μL sample loop attached to a Rheodyne injection valve. Samples were then loaded onto a 0.5 × 5 mm precolumn trap (μ-Trap Column, LC Packings, Sunnyvale, CA) packed with MAGIC (Michrom Bioresources, Auburn, CA) 5 μm, 200 Å, C_{18} and washed with a H_2O:ACN (98:2) buffer containing 0.2% formic acid for 2 min (20 μL/min). The samples were then exposed to a linear gradient from 95% A (H_2O:ACN, 98:2 w/w 0.2% formic acid) to 2% A for 45 min at a flow rate of 0.5 μL/min (B-phase = H_2O:ACN, 5:95 w/w 0.2% formic acid). Tryptic fragments were separated on a MAGIC 5 μm, 200 Å, C_{18} self-packed 10 cm, 75 μm i.d. Integrafrit column (New Objective, Woburn, MA) and electrosprayed from a 15 μm i.d. PicoTip emitter (New Objective, Woburn, MA). Internal calibration was accomplished by directly infusing 1 μM glucagon (H_2O:ACN, 50:50 w/w 0.1% formic acid) at 0.1 μL/min via an 8 μm i.d. PicoTip electrospray emitter (New Objective, Woburn, MA).

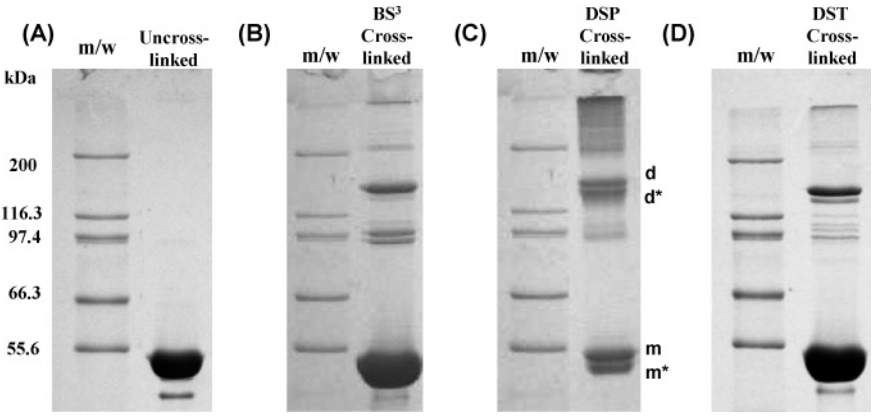

Figure 2. SDS-PAGE of Chemical Cross-linked Empty Procapsid Shells. The empty procapsid shells (20 μM coat protein) were mixed with freshly prepared cross-linkers (1 mM final) which are dissolved in DMSO or water. Each reaction mixture was incubated at room temperature for 5 min and quenched with 60 mM Tris finally to avoid high oligomerization. The cross-linking reactions were separated and analyzed on the nonreducing 7.5% SDS-PAGE. (A) Un-cross-linked empty procapsid shells. (B) BS^3 cross-linking reaction mixture. (C) DSP cross-linking reaction mixture. (D) DST cross-linking reaction mixture.

nanoLC−MS/MS measurements were performed on a QToF API−US quadrupole-time-of-flight mass spectrometer controlled by MassLynx 4.0 (Waters, Framingham, MA) interfaced to a Waters CapLC and autosampler. NanoLC separations were done on a 75 μm i.d. by 15 cm Integrafrit column (New Objective, Woburn, MA) slurry-packed with 5 μm Targa C_{18}, (Higgins Analytical, Mountain View, CA) and electrosprayed from a 15 μm i.d. PicoTip emitter (New Objective, Woburn, MA). Samples were loaded onto a 0.3 mm i.d. by 5 mm long precolumn (PepMap, Dionex, Sunnyvale, CA). A multistep gradient from 3 to 50% mobile phase B over 30 min, followed by 50−95% B over the following 6 min, was used at an approximate column flow rate of 0.3 μL/min, split from a primary pump flow of 12 μL/min. Mobile phase A was 1% acetonitrile, 1% *n*-propanol, 98% water with overall 0.2% formic acid. Mobile phase B was 80% acetonitrile, 10% 1-propanol, 10% water with overall 0.2% formic acid.

Automated MS/MS experiments used a survey scan, from m/z 400−1800, to select the two most abundant ions as MS/MS precursors, using charge and m/z-dependent selection of collision energies.

MS/MS spectra were processed using the MaxEnt 3 algorithm from MassLynx version 4.0 to transform the multiply charged fragment ion spectra to the singly charged m/z scale.

Site Directed Mutagenesis. A cysteine residue was substituted at residue 182 instead to replace a threonine using established polymerase chain reaction protocols. The pET3a based plasmids encoding genes for scaffolding and coat proteins were used as templates. The amplified DNAs were transformed into $CaCl_2$ treated competent *E. coli* strain BL21 (DE) and selected for ampicillin resistance. Candidate colonies were tested for expression and verified by DNA sequencing. Cysteine mutant (T182C) procapsids were overexpressed in *E. coli* grown in LB medium by induction with 2 mM IPTG. Cells were lysed by repeated freeze−thaw cycles, treated with lysozyme, and the supernatants were collected by centrifugation at 13,000 × g for 1 h. T182C procapsids were further purified by sucrose gradient centrifugation.[24] T182C empty procapsid shells were prepared by repeated extractions of scaffolding protein with 0.5 M GuHCl at 4 °C. Purified empty procapsid shells were stored in buffer B (50 mM Tris-HCl, 25 mM NaCl and 2 mM EDTA) at 4 °C.

Sucrose Gradient Centrifugation and Native Agarose Gel Assay. Purified wild type and T182C empty procapsid shells were fractionated by ultracentrifugation on a 5−20% sucrose gradient with SW55 rotor at 30 000 rpm for 35 min and they were analyzed by nonreducing SDS-PAGE. Wild-type and T182C empty procapsid shells were heat expanded at 65 °C for 1, 3, 7, 30 min, 1, 2, and 3 h, and 30 s, 1, 2, 3, 5, and 10 min, respectively. Heat treated samples were electrophoresed through 1.2% native agarose gels and detected with Coomassie blue as described,[27] with some modifications.

Results

To obtain relative distance constraints, three different bifunctional lysine reactive cross-linkers were used to cross-link P22 empty procapsid shells.[28] Empty procapsid shells have the same capsid structure as true procapsids[28] but lack the internal scaffolding protein thus simplifying the cross-linking patterns. Dissuccinimidyl tartrate (DST), dithiobis[succinimidyl propionate] (DSP), and bis[sulfosuccinimidyl] suberate (BS^3) have 6.4, 12, and 11.4 Å distance constraints, respectively. BS^3 is water-soluble while DST and DSP are both water-insoluble. DSP has a cleavable disulfide bond in the middle of the linker. All three cross-linkers result in the formation of an amide bond between the ϵ-amino of lysine and the cross-linker. The cross-linkers result in mass increases of 113.9953, 173.9809, and 138.0680 Da, respectively. To minimize cross-linking and thereby avoid disruption of the protein structure, the cross-linking reactions were limited to 5 min. SDS-PAGE analysis revealed that longer cross-linking times and higher cross-linker concentrations resulted in the formation of a significant fraction of high molecular weight oligomers (data not shown). The pattern of cross-linked species with DST and BS^3 cross-linkers was very similar both in apparent molecular weight and relative abundance. In contrast, the cross-linking pattern observed with DSP presented one additional band in each of the monomer and dimer positions (Figure 2).

Initial experiments involved characterizing the DST-cross-linked procapsid shells where the cross-linked procapsids were dissociated with 6 M GuHCl, the subunits were refolded by dialysis and then separated by size exclusion chromatography. This approach has been used previously in this laboratory to analyze the cross-linking of HIV capsid proteins and thus

provides information about the extent of modification to each of the species.[8] Using this approach, we were able to separate monomeric protein from cross-linked oligomers but we were unable to fractionate pure oligomers. The extent of monomer modification was determined by relative abundances of the deconvoluted mass of each species assuming modification does not affect ionization and ion transfer. The charge state distributions of each modified species were similar to each other because the number of modifications was small relative to the charge state (loss of 1 or 2 charges out of ~50). There are two possible modifications of the monomer, intramolecular cross-linking and hanging cross-linking in which one arm is cross-linked to the protein and the other is hydrolyzed by reaction with solvent. Intramolecular cross-linking results in a 113.9953 Da increase in the monoisotopic mass, whereas a hanging cross-link results in a 132.0058 Da increase in the monoisotopic mass. LC−ESI−MS analysis of the monomeric fraction of DST cross-linked empty procapsid shells showed serial increases of 132 Da suggesting approximately 70% of the monomers were unmodified, 25% were modified with a single DST cross-linker and approximately 5% were modified with two cross-linkers. (data not shown)

The products of the cross-linking reactions were fractionated using nonreducing SDS-PAGE (Figure 2). Figure 2A shows the SDS-PAGE of the un-cross-linked sample where the coat protein (46.6 kDa) and residual scaffolding protein (36.6 kDa) were observed at 52 kDa and 48 kDa respectively relative to the molecular weight standards. Following cross-linking, the dominant fraction of the coat protein migrated in the monomeric position consistent with the small extent of cross-linking. However, a new and reasonably intense band was observed to migrate at ~160 kDa along with a number of less intense bands (Figure 2B−D). Although the apparent molecular weight of 160 kDa observed for the cross-linked coat proteins is consistent with the expected molecular weight of a trimer, disulfide cross-linked coat protein dimers electrophoresed at essentially the same position (vide infra) suggesting this species is in fact a dimer that displays altered electrophoretic mobility.

The bands at positions corresponding to molecular weights of 52 kDa and 160 kDa in Figure 2B−D were excised and subjected to in-gel trypsin digestion. The monomer band at 52 kDa from the untreated empty procapsid shells (Figure 2A) was also excised and tryptically digested for comparison. All digested samples were initially analyzed by nanoLC−ESI-TOF mass spectrometry to evaluate the effectiveness of the gel extraction and determine sequence coverage of the coat protein. NanoLC−ESI-TOF results provided complete (100%) sequence coverage of the coat protein including several fragments that resulted from nonspecific cleavages (data not shown). The 160 kDa bands in Figure 2B−D were also analyzed via nanoLC−ESI-TOF and found to provide ~97% coverage for both DSP- and BS^3-cross-linked samples and complete sequence coverage for DST-cross-linked sample. The digests also included several fragments that resulted from nonspecific cleavages and unique fragments arising from cross-linkng which we discuss in the following section.

A single unique quadruply charged ion ($m/z = 496.7676$) was identified in the dimer band of BS^3 cross-linking reactions relative to the un-cross-linked control (Figure 2B). Modification of the lysine side-chain by cross-linking prevents tryptic cleavage. Therefore, to identify the cross-linked peptide the masses of combinations of two peptides (each with a single missed cleavage site), plus the cross-linker were calculated and compared to the observed masses. The best fit to the observed data was obtained by assuming symmetrical cross-linking between two identical peptides spanning residues 177−184 with the cross-link occurring through lysine 183. The calculated m/z for this combination was 496.7679 in excellent agreement (<1 ppm) with the observed monoisotopic m/z (Figure 3A, left panel). The identity of the cross-linked peptides was subsequently confirmed using data-dependent nanoLC−Q-TOF MS to obtain sequence tags (Figure 3B). Under low energy collision-induced dissociation (CID) conditions, cleavage can occur in either or both of the two cross-linked polypeptide chains resulting in serial combinations of y- and b-ions of two cross-linked peptides. After the suggestion of Schilling et al.,[29] the longer peptide is designated α and the shorter is designated β. The fragmentation products of the longer cross-linked peptide were annotated as y- and b-ions whereas the fragmentation products of the shorter cross-linked peptide were annotated as y′- and b′-ions. A symmetrical fragmentation pattern was observed (y3y′5-y7y′7) because essentially identical peptides spanning residues 177−184 were cross-linked to one another. The resulting fragment ions were assigned as in Figure 3B.

A similar strategy of comparing un-cross-linked and cross-linked digests was employed to identify potential cross-linked peptides in the DSP and DST cross-linking reactions.

In the case of the DSP cross-linked samples, both the monomer and dimer bands were split in doublets labeled m, m*, d, and d* (Figure 2C). The slight differences in gel mobility suggest that these proteins have slightly different structures even when denatured by SDS-PAGE. The tryptic digest of the more slowly migrating monomer band (labeled m) showed no ions beyond those seen in the un-cross-linked sample. In contrast, a unique quintuply charged ion of $m/z = 1154.1702$ was detected in the tryptic digest of the faster migrating m* band. Using a similar rationale, this ion was identified as containing peptides 222−248 and 299−324 in which lysine 236 of one peptide is cross-linked to lysine 310 of another (calculated $m/z = 1154.1711$) (Supporting Information Figure 1). As it was identified in a monomer band, it must arise from intramolecular cross-linking with the altered mobility presumably arising from constraints on unfolding introduced by chemical cross-linking. The identity of this peptide was also confirmed by nanoLC−Q-TOF MS (Supporting Information Figure 1). No cross-linked peptide was detected in the tryptic digests of the monomeric band from either the DST cross-linked or BS^3 cross-linked samples.

NanoLC−dualESI−FT−ICR MS analysis of the tryptic digests of the dimer doublet of the DSP cross-linked samples revealed that both bands contained the same intramolecular cross-link found in the m* band and both contained an additional quadruply charged ion of $m/z = 505.7463$ (Figure 3A, right panel). The mass increment between the ions found in the dimer of BS^3 and DSP cross-linked samples corresponded precisely to the mass difference between the two cross-linkers suggesting that the same cross-linked peptide had been produced in both cases (Figure 3A). The calculated m/z for the same cross-linked peptide (177−184:177−184) cross-linked with DSP is 505.7461 also in excellent agreement (<1 ppm) with the observed value (Figure 3A, right panel). The identity of this cross-linked peptide was also subsequently confirmed by nanoLC−Q-TOF MS. (Supporting Information Figure 2). While both the d and d* bands contained the same cross-links within the capsid protein they displayed different gel mobilities. One possible explanation is that in the faster migrating band both

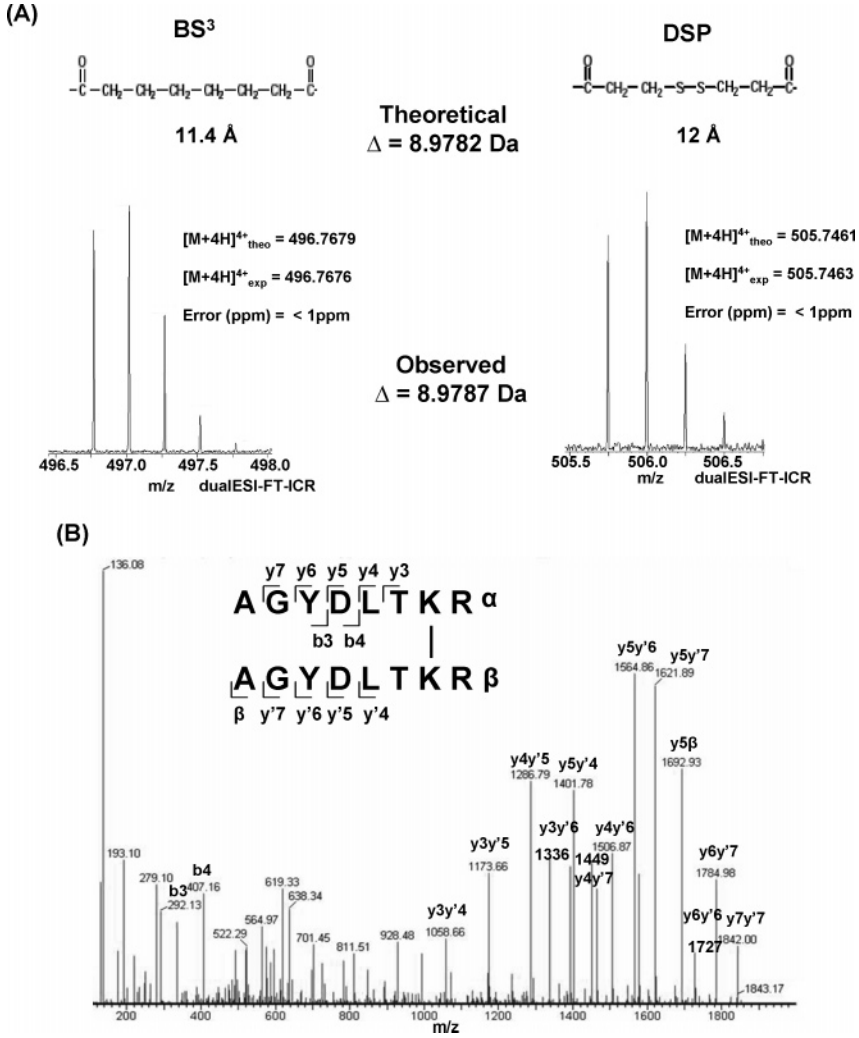

Figure 3. ESI−FT-ICR MS spectra of the BS^3 and DSP intersubunit cross-linked peptide and NanoLC−Q-TOF MS spectrum of a $[M+H]^{4+}$ ion of the BS^3 cross-linked peptide. (A) ESI−FT-ICR MS spectra of BS^3 and DSP. Monoisotopic masses of precursor ions of the BS^3 and DSP cross-linked peptides were shown at $m/z = 496.7676$ (theoretical $m/z = 496.7679$) and $m/z = 505.7463$ (theoretical $m/z = 505.7461$), respectively. The monoisotopic mass increment corresponded precisely to the mass difference between the two cross-linkers. Mass measurement errors are less than 1 ppm on the ESI−FTICR MS equipped with dual ESI source.[26] (B) NanoLC−Q-TOF MS spectrum of a $[M+H]^{4+}$ ion of the BS^3 cross-linked peptide. Nomenclature after the suggestion of Schilling et al.[29]

subunits contain the intramolecular cross-link while in the slower migrating band only one subunit contains the intramolecular cross-link. However, the peptide fragment contained within the intramolecular cross-link was also observed in the un-cross-linked form in both bands, an observation not consistent with the hypothesis that altered mobility arises from the identified intramolecular cross-link but leaving open the possibility that the altered mobility arises from heterogeneous intramolecular cross-linking below the limit of detection.

A single quadruply charged ion of $m/z = 720.07$ not seen in the digest of the un-cross-linked shells was detected in the DST cross-linked dimer band (Figure. 2D) and sequenced by nanoLC−Q-TOF MS (Supplementary Figure 3). A series of y-ions which are cross-linked to the uncleaved shorter peptide (y6β-y10β) or y′-ions (y6y′5-y9y′6) were observed. On the basis of this analysis, the cross-linked peptide was identified as containing a peptide spanning residues 162−176 cross-linked to a peptide spanning residues 177−184 with lysine 175 in one subunit cross-linked to lysine 183 in the other (Supporting Information Figure 3).

Previous mass spectrometry based hydrogen/deuterium exchange experiments as well as proteolytic digestion studies had identified the region spanning residues 157−207 as being flexible and exposed in the procapsids while becoming protected upon maturation.[7,9,30] The lysine reactive cross-linkers suggest that these residues lie within 12 Å of one another. To determine whether this region was sufficiently flexible and close enough to form zero-length cross-links between subunits, a cysteine mutation was engineered in to replace the threonine residue at position 182 (T182C) immediately adjacent to the cross-linking site (lysine 183). Site-directed mutagenesis was carried out on a pET vector based cloned genes which encode scaffolding and coat proteins. The T182C procapsids were overexpressed in *E. coli* and purified following the protocol used for the overexpression and purification of wild-type procapsids.[24] The molecular weight of the mutant monomer as determined by LC−ESI-TOF mass spectrometry was 46 624 Da in good agreement with the predicted values of 46 622.81 Da. The presence of the mutation was also verified by DNA sequencing. The purified T182C empty procapsid shells sedi-

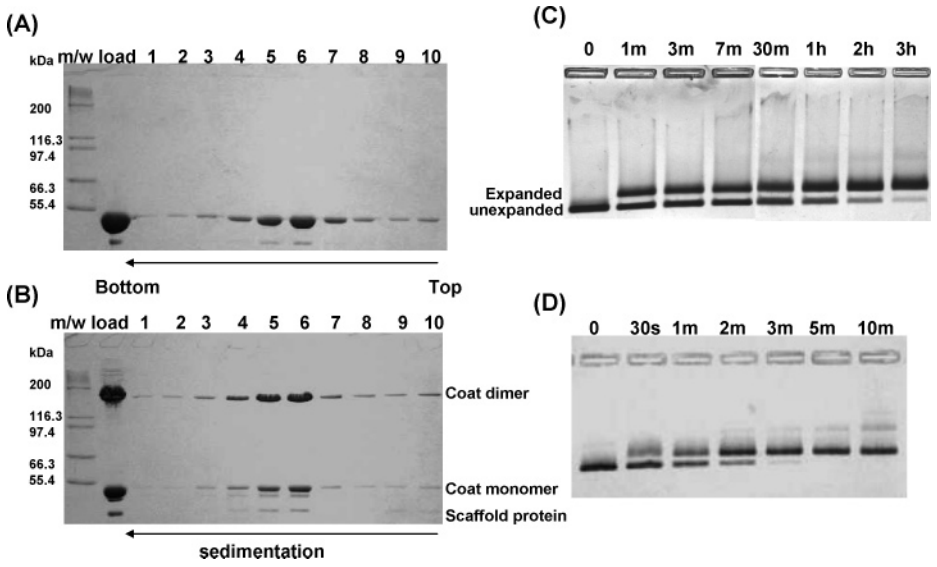

Figure 4. Sucrose gradient sedimentations and heat expansions of wild type and T182C empty procapsid shells. Purified wild type and T182C empty procapsid shells were separated on a 5−20% sucrose gradient. Fractions from the gradient were analyzed on the nonreducing 8.5% SDS-PAGE. (A) Wild-type empty procapsid shells. (B) T182C empty procapsid shells. Wild type and T182C empty procapsid shells were heat expanded at 65 °C for the indicated time periods. Heat treated samples were analyzed on the 1.2% native agarose gel and detected with Coomassie blue.[27] (C) Wild-type empty procapsid shells. (D) T182C empty procapsid shells.

mented to the same position as wild-type empty procapsid shells on a 5−20% sucrose gradient (Figure 4A,B) and displayed the same mobility on native agarose gels (data not shown). These techniques are sensitive to changes in the mass, size and structure of the capsid,[27] and therefore the results indicate that the T182C coat protein assembles into procapsid shells that are similar in form to the wild-type shells. Disulfide bonds were observed to form spontaneously within the T182C procapsid shells and did so without disturbing shell integrity (Figure 4B) and when these shells were analyzed by LC−ESI-TOF they contained a protein of mass 93 246 Da in good agreement with the value of 93 243.58 Da predicted for a disulfide cross-linked dimer. Spontaneous, dimer formation was relatively slow. After 1 day only 10% had formed disulfide bonds but the figure rose to ∼70% at 4 °C after two weeks. The particles remained unexpanded and no further disulfide formation was observed even under completely oxidizing condition (data not shown). These results suggest that the flexible loop region can form zero-length cross-links between residues of neighboring subunits in the intact procapsid lattice.

Previous experiments using both partial proteolysis and hydrogen/deuterium exchange have demonstrated that residues 157−207 of the capsid protein undergo a pronounced change in exposure upon maturation.[7,9,30] Therefore, we investigated the effect of the cysteine mutation on the kinetics of expansion. While expansion is normally induced by DNA packaging, this process can be mimicked in vitro by gentle heating or mild denaturant treaments such as SDS and urea.[31,32] Expansion is irreversible and the expanded form migrates more slowly on an agarose gel than the unexpanded form making it relatively straightforward to follow the kinetics of expansion. The thermally induced expansion of wild-type procapsids has a half time of approximately 7 min and takes upward of 2 h to reach completion (Figure 4C). While the reduced form of the T182C mutant displayed expansion kinetics that were similar to wild-type, the spontaneously oxidized form of the T182C mutant had an accelerated half time for expansion of approximately 1 min and were expanded completely within 5 min (Figure 4D). Expansion did not result in any change in the extent of cross-linking. Thus, while substitution of a cysteine residue at this location does not alter the global structure of the capsid, it has a pronounced effect on the dynamics of the capsid lattice transformation.

Discussion

Large structural rearrangements in viral capsids during their lifecycle are the norm and probably represent a compromise solution to the need for flexibility during assembly and protective function during infection.[33,34] Because all of the information required for both self-assembly and structural transformation is encoded in the protein subunits themselves, it is of fundamental interest to understand the molecular basis for these structural conversions. In the case of the dsDNA phage such as P22, the transformation is driven by the ATP dependent packaging of DNA into the preformed capsid.[10, 13]

Limited proteolysis studies have indicated that the P22 coat protein monomer consists of two approximately equal domains connected by a flexible loop region and that the loop region is susceptible to protease in the procapsids, but becomes completely resistant in the mature capsids.[30] Mass spectrometry based hydrogen/deuterium exchange experiments also showed that the loop region exchanged rapidly in the empty procapsid shells, but was well protected in the mature capsids.[7] Taken together these results suggest that the loop region is exposed to solvent and flexible in the procapsid lattice, but becomes buried and fixed in the mature capsid lattice. Thus, the rearrangement of this flexible loop region was proposed to be a key component for capsid lattice transformation. Interestingly, all the intersubunit cross-linking occurred between residues clustered in the loop region (residues 157−207). DSP and BS^3 which have similar distance constraints (12 Å and 11.4 Å, respectively) cross-linked the same residues between two subunits in the empty procapsid shells (K183−K183), whereas DST, a shorter cross-linker, cross-linked lysine 175 in one subunit to lysine 183 in another subunit.

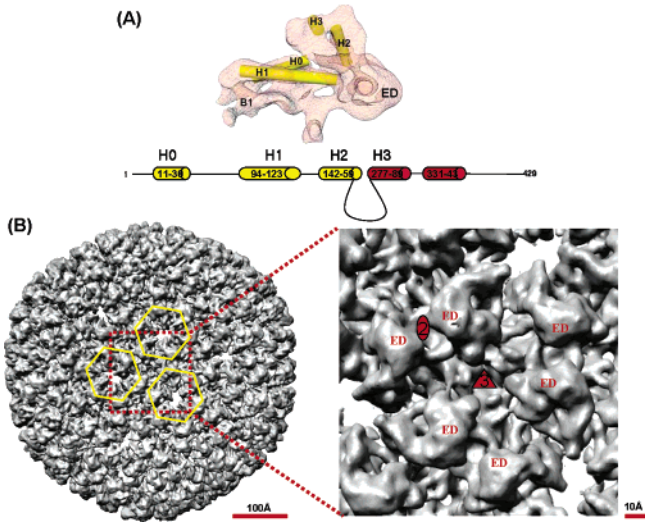

Figure 5. Outer surface representation of the P22 procapsid three-dimensional reconstruction and secondary structure prediction of P22 sequence. (A) Nonicosahedrally averaged coat protein monomer of procapsids (top)[5] and secondary structure prediction (bottom). The helices (H0−H3) are represented as cylinders. The red helices indicate the alternative possibility. Extra density region is labeled as ED. (B) Outer surface representation of the 8.5 Å resolution density map of the P22 procapsid (left) and zoomed-in view of the 3-fold rotational symmetry axis (right). Yellow hexagons represent the position of hexamers. Locations of the local 2-fold and the 3-fold rotational symmetry axis are marked.

CryoEM based three-dimensional image reconstruction has provided a structure of the procapsid at ~9 Å resolution. At this resolution, it is possible to identify α-helical elements within the structure[5] (Figure 5A, top). Although there is less than 20% sequence similarity, based on the arrangement of the helical elements, the P22 coat protein was classified as having an HK97 like fold indicating an ancestral relationship.[5] However, the P22 coat protein subunit is significantly larger than the HK97 coat protein subunit (282 vs 429 residues) suggesting that insertions have occurred. When compared to the HK97 subunit structure, P22 displays extra protein density, labeled ED, which presumably arises from the additional amino acid residues.[5]

We have performed a sequence based prediction of the secondary structure elements of P22 coat protein, and attempted to thread the prediction onto the identified helical elements in the reconstruction (Figure 5A). The predicted helices corresponded well to the helical elements identified by cryoEM (H0, H1, and H2) although there are two reasonable alternatives for helix H3 (red in prediction). According to our alignment, the cross-linking sites (K183 and K175) and the previously identified flexible loop lie in a region between helices 2 and 3. This region which is not present in the HK97 subunit presumably represents an insertion and is therefore a good candidate for the extra density (ED) region.

In the three-dimensional reconstruction, the ED regions protrude upward from the outside of the capsids and face one another around the 3-fold rotational symmetry axis (Figure 5B). The subunits approach one another closely at the pseudo 2-fold axes between hexamers. It is possible that the cross-linking observed is between either the subunits across the pseudo 2-fold or around the 3-fold axis. The difference in cross-linking pattern likely reflects the different distance constraints (DST 6.4 Å and BS^3 and DSP ~12 Å) imposed by these cross-linkers. For a 3-fold rotational symmetry axis, the distance between two identical points is expected to be longer than between any two other points whereas equivalent points on either side of a 2-fold axis can closely approach one another. For this reason, we slightly favor the model that the cross-link between ED domains is across the pseudo 2-fold axis, but either possibility or a combination of both are possible. The replacement of threonine 182 to cysteine, a position immediately adjacent to cross-linking site (K183) results in spontaneous disulfide bond formation relatively slowly (half-time, ~a week) suggesting not only moderate distant proximity but also flexible nature of the loop region which is consistent with previous hydrogen/deuterium exchange studies.[7,9]

A similar structural transformation has been observed for the bacteriophage HK97 whose atomic resolution structure is available.[4,33] In the cryoEM studies of HK97, the E-loop of HK97 which protrudes upward from the outside of the capsid in the prohead II, becomes flattened against the capsid surface in the mature capsid.[33] Interestingly, K166 is susceptible to trypsin in the prohead but not in the mature capsid, a situation similar to that observed in P22.[35] Strikingly, in HK97 this transformation is accompanied by an auto-catalytic cross-linking event in which K169 which is contained within the E-loop cross-links to N356 of a neighboring subunit resulting in the formation of chain mail like structure.[4,33] In the HK97 prohead, the knobs of the E-loop face one another around the 3-fold and protrude outward from the capsid surface.[33] The alignment between P22 mature capsid monomer and HK97 head II monomer in the cryoEM studies indicated that the E-loop region of HK97 corresponds to the region of extra density (ED) seen in P22.[5] Taken together these data suggests that the flexible loop region identified biochemically in the P22 subunit may correspond to the extra density in the cryoEM reconstructions and is therefore arranged like the E-loop in the HK97 prohead.

The functional importance of this region was evident from the effect of point mutations on the kinetics of expansion. For wild-type P22 procapsid shells expansion was complete in about 3 h at 65 °C and the kinetics of the reaction were biphasic. We could separate two populations (fast expanded capsids and unexpanded procapsids) using ion exchange column chromatography after 5 min of heating (data not shown). The unexpanded population remained unexpanded until approximately 1 h was almost completely expanded at 3 h. The origin of these two populations is still not clear. However, the formation of the intersubunit disulfide bonds in the T182C empty procapsid shells altered the heat expansion kinetics such that all the particles expanded rapidly. When the disulfide bonds were reduced with DTT or TCEP, the particles also showed biphasic expansion kinetics indicating that it was disulfide bond formation and not merely the mutation that altered the kinetics. It is possible that the intersubunit disulfide bonds form a cage-like network through the procapsid shell which serves to synchronize the structural changes in the subunits resulting in a lower activation energy for expansion and a more concerted transformation. Although not every subunit within the particle is disulfide bonded (on average 70% disulfide bond formation) each particle is expected to contain several hundred cross-linked subunits, therefore the minimum number of cross-links required to alter the expansion kinetics is unknown.

Characterizing the protein−protein interactions and conformational changes in large supramolecular complexes pre-

sents a substantial challenge to structural biology due to limitations on crystallizablity. In the case of P22 procapsids, despite significant effort to date it has been impossible to grow diffraction quality crystals. In these cases cryo-electron microscopy can yield information on the protein fold and subunit disposition but cannot identify key side chain interactions. Chemical cross-linking analyzed by mass spectrometry is a powerful approach for obtaining residue-specific information about protein−protein interactions in the supramolecular complexes such as virus particles. Putative interacting residues can then be altered through the targeted introduction of site specific mutations to validate the hypothesis.

Acknowledgment. This work was supported by NIH Grant No. GM47980 (P.E.P.) and the W.M. Keck Foundation (D.C.M.).

Supporting Information Available: Supporting Information Figures 1−3. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Alberts, B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell* **1998**, *92* (3), 291−294.
(2) Caspar, D. Movement and self-control in protein assemblies. Quasi-equivalence revisited. *Biophys. J.* **1980**, *32* (1), 103−138.
(3) Crick, F. H. C.; Watson, J. D. Structure of small viruses. *Nature* **1956**, *177* (4506), 473−475.
(4) Wikoff, W. R.; Liljas, L.; Duda, R. L.; Tsuruta, H.; Hendrix, R. W.; Johnson, J. E. Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* **2000**, *289* (5487), 2129−2133.
(5) Jiang, W.; Li, Z.; Zhang, Z.; Baker, M. L.; Prevelige, P. E., Jr.; Chiu, W. Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions. *Nat. Struct. Biol.* **2003**, *10* (2), 131−135.
(6) Prasad, B. V.; Prevelige, P. E.; Marietta, E.; Chen, R. O.; Thomas, D.; King, J.; Chiu, W. Three-dimensional transformation of capsids associated with genome packaging in a bacterial virus. *J. Mol. Biol.* **1993**, *231* (1), 65−74.
(7) Kang, S.; Prevelige, P. E., Jr. Domain study of bacteriophage p22 coat protein and characterization of the capsid lattice transformation by hydrogen/deuterium exchange. *J. Mol. Biol.* **2005**, *347* (5), 935−948.
(8) Lanman, J.; Lam, T. T.; Barnes, S.; Sakalian, M.; Emmett, M. R.; Marshall, A. G.; Prevelige, P. E., Jr. Identification of novel interactions in HIV-1 capsid protein assembly by high-resolution mass spectrometry. *J. Mol. Biol.* **2003**, *325* (4), 759−772.
(9) Tuma, R.; Coward, L. U.; Kirk, M. C.; Barnes, S.; Prevelige, P. E., Jr. Hydrogen−deuterium exchange as a probe of folding and assembly in viral capsids. *J. Mol. Biol.* **2001**, *306* (3), 389−396.
(10) King, J.; Lenk, E. V.; Botstein, D. Mechanism of head assembly and DNA encapsulation in Salmonella phage P22. II. Morphogenetic pathway. *J. Mol. Biol.* **1973**, *80* (4), 697−731.
(11) King, J.; Casjens, S. Catalytic head assembling protein in virus morphogenesis. *Nature* **1974**, *251* (5471), 112−119.
(12) Casjens, S.; King, J. P22 morphogenesis. I: Catalytic scaffolding protein in capsid assembly. *J. Supramol. Struct.* **1974**, *2* (2−4), 202−224.
(13) Botstein, D.; Waddell, C. H.; King, J. Mechanism of head assembly and DNA encapsulation in Salmonella phage p22. I. Genes, proteins, structures and DNA maturation. *J. Mol. Biol.* **1973**, *80* (4), 669−695.
(14) Galisteo, M. L.; King, J. Conformational transformations in the protein lattice of phage P22 procapsids. *Biophys. J.* **1993**, *65* (1), 227−235.
(15) Morais, M. C.; Choi, K. H.; Koti, J. S.; Chipman, P. R.; Anderson, D. L.; Rossmann, M. G. Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudoatomic structure of phi29. *Mol. Cell* **2005**, *18* (2), 149−159.
(16) Fokine, A.; Leiman, P. G.; Shneider, M. M.; Ahvazi, B.; Boeshans, K. M.; Steven, A. C.; Black, L. W.; Mesyanzhinov, V. V.; Rossmann, M. G. Structural and functional similarities between the capsid proteins of bacteriophages T4 and HK97 point to a common ancestry. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102* (20), 7163−7168.
(17) Sinz, A. Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. *J. Mass Spectrom.* **2003**, *38* (12), 1225−1237.
(18) Trester-Zedlitz, M.; Kamada, K.; Burley, S. K.; Fenyo, D.; Chait, B. T.; Muir, T. W. A modular cross-linking approach for exploring protein interactions. *J. Am. Chem. Soc.* **2003**, *125* (9), 2416−2425.
(19) Schulz, D. M.; Ihling, C.; Clore, G. M.; Sinz, A. Mapping the topology and determination of a low-resolution three-dimensional structure of the calmodulin-melittin complex by chemical cross-linking and high-resolution FTICRMS: direct demonstration of multiple binding modes. *Biochemistry* **2004**, *43* (16), 4703−4715.
(20) Cai, K.; Itoh, Y.; Khorana, H. G. Mapping of contact sites in complex formation between transducin and light-activated rhodopsin by covalent cross-linking: use of a photoactivatable reagent.-[see comment]. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (9), 4877−4882.
(21) Young, M. M.; Tang, N.; Hempel, J. C.; Oshiro, C. M.; Taylor, E. W.; Kuntz, I. D.; Gibson, B. W.; Dollinger, G. High-throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (11), 5802−5826.
(22) Back, J. W.; de Jong, L.; Muijsers, A. O.; de Koster, C. G. Chemical cross-linking and mass spectrometry for protein structural modeling. *J. Mol. Biol.* **2003**, *331* (2), 303−313.
(23) Green, N. S.; Reisler, E.; Houk, K. N. Quantitative evaluation of the lengths of homobifunctional protein cross-linking reagents used as molecular rulers. *Protein Sci.* **2001**, *10* (7), 1293−1304.
(24) Prevelige, P. E., Jr.; Thomas, D.; King, J. Scaffolding protein regulates the polymerization of P22 coat subunits into icosahedral shells in vitro. *J. Mol. Biol.* **1988**, *202* (4), 743−757.
(25) Rosenfeld, J.; Capdevielle, J.; Guillemot, J. C.; Ferrara, P. In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Anal. Biochem.* **1992**, *203* (1), 173−179.
(26) Nepomuceno, A. I.; Muddiman, D. C.; Bergen, H. R., 3rd; Craighead, J. R.; Burke, M. J.; Caskey, P. E.; Allan, J. A. Dual electrospray ionization source for confident generation of accurate mass tags using liquid chromatography Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **2003**, *75* (14), 3411−3418.
(27) Serwer, P.; Hayes, S. J.; Griess, G. A. Determination of a particle's radius by two-dimensional agarose gel electrophoresis. *Anal. Biochem.* **1986**, *152* (2), 339−345.
(28) Thuman-Commike, P. A.; Greene, B.; Malinski, J. A.; Burbea, M.; McGough, A.; Chiu, W.; Prevelige, P. E., Jr. Mechanism of scaffolding-directed virus assembly suggested by comparison of scaffolding-containing and scaffolding-lacking P22 procapsids. *Biophys. J.* **1999**, *76* (6), 3267−3277.
(29) Schilling, B.; Row, R. H.; Gibson, B. W.; Guo, X.; Young, M. M. MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically cross-linked peptides. *J. Am. Soc. Mass Spectrom.* **2003**, *14* (8), 834−850.
(30) Lanman, J.; Tuma, R.; Prevelige, P. E., Jr. Identification and characterization of the domain structure of bacteriophage P22 coat protein. *Biochemistry* **1999**, *38* (44), 14614−14623.
(31) Tuma, R.; Prevelige, P. E., Jr.; Thomas, G. J., Jr. Mechanism of capsid maturation in a double-stranded DNA virus. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95* (17), 9885−9890.
(32) Earnshaw, W.; Casjens, S.; Harrison, S. C. Assembly of the head of bacteriophage P22: X-ray diffraction from heads, proheads and related structures. *J. Mol. Biol.* **1976**, *104* (2), 387−410.
(33) Conway, J. F.; Wikoff, W. R.; Cheng, N.; Duda, R. L.; Hendrix, R. W.; Johnson, J. E.; Steven, A. C. Virus maturation involving large subunit rotations and local refolding. *Science* **2001**, *292* (5517), 744−748.
(34) Zlotnick, A.; Aldrich, R.; Johnson, J. M.; Ceres, P.; Young, M. J. Mechanism of capsid assembly for an icosahedral plant virus. *Virology* **2000**, *277* (2), 450−456.
(35) Hendrix, R. W.; Duda, R. L. Bacteriophage HK97 head assembly: a protein ballet. *Adv. Virus Res.* **1998**, *50*, 235−288.

PR050356F