In Vitro Unfolding/Refolding of Wild Type Phage P22 Scaffolding Protein Reveals Capsid-binding Domain*

Barrie Greene‡ and Jonathan King
From the Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

The scaffolding proteins of double-stranded DNA viruses are required for the polymerization of capsid subunits into properly sized closed shells but are absent from the mature virions. Phage P22 scaffolding subunits are elongated 39-kDa molecules that copolymerize with coat subunits into icosahedral precursor shells and subsequently exit from the precursor shell through channels in the procapsid lattice to participate in further rounds of polymerization and dissociation. Purified scaffolding subunits could be refolded in vitro after denaturation by high temperature or guanidine hydrochloride solutions. The lack of coincidence of fluorescence and circular dichroism signals indicated the presence of at least one partially folded intermediate, suggesting that the protein consisted of multiple domains. Proteolytic fragments containing the C terminus were competent for copolymerization with capsid subunits into procapsid shells in vitro, whereas the N terminus was not needed for this function. Proteolysis of partially denatured scaffolding subunits indicated that it was the capsid-binding C-terminal domain that unfolded at low temperatures and guanidinium concentrations. The minimal stability of the coat-binding domain may reflect its role in the conformational switching needed for icosahedral shell assembly.

Scaffolding proteins are essential for the formation of the icosahedral capsids of several classes of virus, including double-stranded DNA bacteriophage (1–3), herpesviruses (4), and adenoviruses (5). These proteins are required for assembly of the viral capsid but are not found in the mature virus after DNA packaging. In the absence of scaffolding proteins, the viral coat protein polymerizes into aberrant structures that cannot package DNA (1, 6–8).

In the presence of the 33-kDa scaffolding protein, the coat subunits of the bacteriophage P22 form an icosahedral T = 7 lattice (9, 10), which includes four distinct conformations of the coat monomers (8, 10). The scaffolding subunits are thought to be involved in the conformational switching of coat subunits occurring during polymerization at the edges of the growing shell (2, 3, 11).

Following successful procapsid assembly, on receipt of some physiological signal (probably docking of the DNA packaging complex at the portal vertex), all 200–300 scaffolding protein molecules dissociate from the procapsid lattice and recycle (12). The icosahedral coat lattice exhibits 25 Å channels at the centers of the pentamers and hexamers (9), which appear to be the site of exit for the scaffolding subunits prior to DNA packaging. Soluble subunits can reenter the procapsid through these channels in vitro (13). Upon DNA packaging the capsid lattice expands, and the channels are closed by a domain of the coat protein, which swings over to close the channels (9, 14). In vitro the scaffolding subunits can be released from procapsids by increasing temperature or by incubation with very low levels of denaturant (13, 15).

This process may involve conformational changes in the scaffolding protein as well as the coat protein. Within the assembled procapsid, 200–300 scaffolding protein molecules are relatively tightly packed with each other into an inner shell or ball (10, 16). Scaffolding subunits released from the procapsid, however, are highly soluble and do not assemble any large structures in the absence of free coat protein subunits (2, 17, 18). This presumably represents a conformational change in the scaffolding subunits back to the precursor conformation. Though direct structural evidence is limited, scaffolding proteins appear to switch back and forth between alternate conformations without phosphorylation, ATP, or other nucleotide binding steps.

Unfortunately, little is known about scaffolding protein structures. Although the structures of both phage P22 and herpesvirus procapsids containing scaffolding subunits have been determined, the scaffolding organization was not revealed in either case, suggesting that it was not icosahedral (10, 19).

Analysis of the sequences of both phage and herpesvirus scaffolding proteins suggested that they tend to be predominantly α-helical in structure (20, 21). This prediction has been confirmed for the P22 scaffolding protein by Raman spectroscopy (23) and circular dichroism (24). Because there are no leucine zipper or heptad repeats evident in the P22 scaffolding protein sequence (22), the protein is probably not a coiled-coil. Analytical ultracentrifugation or gel filtration have revealed that the scaffolding proteins of T4, A, and P22 are all highly elongated molecules (18, 24–26). Parker et al. (26) estimated dimensions for the P22 scaffolding protein of 247 Å in length by 22 Å in diameter.

Initial thermal denaturation studies of P22 procapsids by differential scanning calorimetry revealed no evidence of a cooperative melting transition for the scaffolding subunits (15). Preliminary nuclear magnetic resonance studies of a 163-amino acid C-terminal fragment, however, showed that stable elements of secondary structure were present in solution (27).

A recent study of scaffolding protein thermal denaturation followed by Raman spectroscopy revealed that the protein lost its helical secondary structure over a broad temperature range in a noncooperative manner and appeared to have little packed tertiary structure (28). The authors proposed that the scaffolding protein consisted of several loosely packed helical segments.

These segments might correspond to functional domains for
the various roles of scaffolding protein. In addition to morphogenesis, the process by which the double-stranded DNA phages ensure the incorporation of a DNA packaging portal at a unique vertex also requires scaffolding protein. Interactions between scaffolding and portal proteins have been inferred genetically (29–31) or observed directly, in the case of φ29 (32). The unassembled P22 scaffolding protein regulates its own synthesis at the translational level (33, 34) presumably by binding to its mRNA. Mutations within a specific region of the P22 scaffolding protein affect the ability of the scaffolding protein to be released from the procapsid (35), suggesting that release or sensing of DNA entry may be another scaffolding protein function. Fig. 1 summarizes the roles of the scaffolding protein in procapsid assembly and DNA packaging that have been identified in previous studies.

We were particularly interested in determining the scaffolding protein region involved in binding to the coat protein. Also of interest was the possibility that this cycle of association and dissociation was coupled to the folding and unfolding of domains of the scaffolding subunits. Accordingly, we have purified and characterized the denaturation transition of soluble wild type scaffolding subunits. In the accompanying paper (36) we describe a set of mutant scaffolding proteins which are defective in aspects of shell assembly or function.

**EXPERIMENTAL PROCEDURES**

**Chemicals**—Ultrapure GuHCl1 was purchased from Pierce. Proteases (V8, chymotrypsin-agarose, trypsin-agarose) and inhibitors were ordered from Sigma. Molecular weight standards were purchased from Bio-Rad. All other chemicals were reagent grade from common sources.

**P22 Phage Strains**—The phage strains used in the preparation of wild type phages carried the c1-7 mutation for lytic growth and two nonsense mutations, 2amH202, in one of the DNA packaging genes to block DNA packaging, and 1amH101, which delays lysis.

**Purification of Scaffolding and Coat Proteins**—Wild type procapsids were purified from cells infected with a phage mutant blocked in DNA packaging, and both scaffolding monomers and empty coat shells were separated by SDS-polyacrylamide gel electrophoresis. The peptide fragments were generated with protease V8, trypsin, or chymotrypsin as described and separated by SDS-polyacrylamide gel electrophoresis on 15% acrylamide gels.

**Thermal Denaturation**—Wild type scaffolding protein was dialyzed into 20 mM phosphate buffer, pH 7.6, with phosphoric acid. The scaffolding protein solution was dialyzed to a final concentration of 0.6 mg/ml (1.8 μM) in either phosphate buffer or phosphate buffer with GuHCl and incubated at 10 °C for 2 h to allow unfolding to occur. 2.5 μl of protease V8 was added to give a 1:10 ratio of protease to scaffolding protein. The samples were spun on a Nutator platform rocker. After 6 min, the protease sample was centrifuged to pellet the beads, and the supernatant was removed. The temperature was then increased to 330 °C with 2 min of equilibration at each point after reaching constant temperature. At the end of each experiment the sample was cooled back to 3 °C to check for recovery of the original signal.

**Circular Dichroism**—Wild type scaffolding protein was added to tubes containing varying concentrations of GuHCl in the phosphate/NaCl buffer to give a final protein concentration of 100 μg/ml (0.3 μM) and stored for at least 12 h at 4 °C. The signal from these samples remained constant up to 24 h later, demonstrating that they had reached equilibrium. The same samples were assayed by both fluorescence and CD and compared with blanks of the same GuHCl concentration. Denaturation was monitored by either fluorescence intensity at 330 nm with excitation at 280 nm or by circular dichroism ellipticity at 222 nm as described above for the thermal denaturation. Because even low concentrations of GuHCl caused unfolding of scaffolding protein as measured by CD, the CD of scaffolding protein in the absence of GuHCl was used as X0 for calculation of the CD denaturation curve. The temperature of the cuvette was maintained at 10 °C.

**Proteolytic Fragments**—Wild type scaffolding protein was dialyzed into 20 mM phosphate buffer, pH 7.6, with phosphoric acid. The scaffolding protein solution was diluted to a final concentration of 0.6 mg/ml (1.8 μM) in either phosphate buffer or phosphate buffer with GuHCl and incubated at 10 °C for 2 h to allow unfolding to occur. 2.5 μl of protease V8 was added to give a 1:10 ratio of protease to scaffolding protein. The samples were spun on a Nutator platform rocker. After 6 min, the protease sample was centrifuged to pellet the beads, and the supernatant was removed to a tube containing 3 μl of 1 mg/ml bovine pancreatic trypsin inhibitor. The chymotrypsin sample was centrifuged after 10 min, and the supernatant was added to 3 μl of 20 mg/ml N-tosyl-L-phenylalalanine chloromethyl ketone. To these digests were added either empty procapsid shells, for a final concentration of 1 mg/ml (53 nM), or coat monomers to 0.5 mg/ml (1.1 μM). These reactions were incubated at room temperature for 2 h. The samples were then centrifuged through 5–20% sucrose gradients for 35 min at 35 K. 18 fractions were collected from each gradient through a pippine at the bottom of the tube. The protein compositions of the fractions were determined by SDS-polyacrylamide gel electrophoresis on 15% acrylamide gels.

**N-terminal Sequencing**—Scaffolding protein fragments were generated with protease V8, trypsin, or chymotrypsin as described and separated by SDS-polyacrylamide gel electrophoresis. The peptide fragments were transferred to polyvinylidene difluoride membranes and analyzed by Edman degradation using the Procys protein sequencer (Beckman) at the Molecular Biology Resource Facility at the University of Michigan.

---

1 The abbreviation used is: GuHCl, guanidine hydrochloride.
Wild type scaffolding protein at 100 μg/ml (0.3 μM) was heated from 3 to 90 °C, and the unfolding process was monitored by fluorescence emission at 330 nm (●) with the excitation at 280 or by circular dichroism at 222 nm (○). The percentage folded was calculated as described under “Experimental Procedures.” The solid lines represent fits of the data to a two-state transition model.

(Immobilon polyvinylidene difluoride, Millipore) according to the method of Matsudaira (39). The transfer was performed in a Hoefer transfer unit at 100 mA for 16 h at room temperature. The membrane was stained with Coomassie Blue to reveal the location of the transferred peptide fragments. The 7 N-terminal residues of each selected band were sequenced by the MIT Biopolymers Laboratory using an Applied Biosystems model 477A Perkin-Elmer Sequencer with an on-line model 120 PTH Amino Acid Analyzer.

RESULTS

Thermal Denaturation of Wild type Scaffolding Protein—To detect possible folding domains of the scaffolding protein, the thermal denaturation of purified scaffolding protein was monitored using two signals: fluorescence at 330 nm, which measures the solvent exposure of the single tryptophan residue, amino acid 134, and circular dichroism at 222 nm, a signal of α-helical structure. The concentration of scaffolding protein was 100 μg/ml (0.3 μM), a concentration at which the protein should be almost entirely monomeric (26).

The thermal denaturation curve observed by CD did not coincide with that monitored by the fluorescence signal (Fig. 2), demonstrating the presence of at least one partially folded intermediate. The CD data were similar to that observed by Tuma et al. (28) despite the fact that the concentration of scaffolding used in their experiment was 18-fold greater than that used here. In that case, the authors fit the data as a single two-state transition. Although our CD data were also well fit as a single transition with a $T_m$ of 46 °C, the noncoincidence of the fluorescence curve (with a $T_m$ of 58 °C) reveals that there must actually be multiple steps in the denaturation process.

A portion of the scaffolding protein began to lose secondary structure at 15 °C. Approximately one-third of the scaffolding protein secondary structure is α-helical, the rest being predominantly coil and turns (28, 36). About 13% of this α-helical structure had been lost by 30 °C, the normal temperature of growth in vivo.

GuHCl-induced Reversible Denaturation of Wild type Scaffolding Protein—Low concentrations of GuHCl are sufficient to extract scaffolding protein from procapsids (13, 17). To determine whether GuHCl causes extraction by denaturing that part of the scaffolding protein required for binding to the coat shell, we investigated the effects of GuHCl on the stability of purified scaffolding protein.

Scaffolding protein was unfolded by incubation in concentrations of GuHCl from 0 to 4 M for at least 12 h, after which time the samples had reached equilibrium and no further spectral changes occurred. The experiments were carried out at 10 °C to remain below the temperature at which thermal unfolding begins.

The unfolding curves measured by CD and by fluorescence did not coincide, demonstrating once again that the unfolding process was not two-state and consistent with the presence of more than one domain. As shown in Fig. 3, the loss of secondary structure began upon the addition of minimal amounts of denaturant. The fluorescence transition did not begin until GuHCl concentrations at which about 40% of the helical secondary structure had already been lost. As the fluorescence signal is due to a single tryptophan at residue 134, one explanation is that the fluorescence signal monitors the unfolding of a particular folding domain containing the tryptophan that is more stable than other regions of the protein. Because a GuHCl concentration of only 0.5 M is sufficient to extract all the scaffolding from the coat shells (13), this could be the domain required for mediating binding of scaffolding to the coat shell. This domain presumably does not include the region surrounding tryptophan 134.

Dilution of the protein out of GuHCl led to complete recovery of the original signal, demonstrating that the GuHCl-induced unfolding was reversible. The efficient refolding of the scaffolding protein without interference from competing aggregation reactions was quite distinct from the in situ refolding of the coat protein and the tailspike proteins. For both of these proteins, off pathway association reactions competed with productive refolding (40–44).

Proteolytic Digestion of Folded and Partially Unfolded Scaffolding Protein—The lack of coincidence between the CD and fluorescence signals indicated the presence of a partially folded intermediate during both thermal and GuHCl-induced denaturation transitions. This suggests that the native state may consist of distinct folding/unfolding domains. We were particularly interested in the region of the molecule that unfolds at low GuHCl concentrations, because this might be a region required for binding to the coat shell. To determine which part of the scaffolding protein unfolded first, we looked for regions of the molecule that became more accessible to protease V8 after incubation in low concentrations of GuHCl at 10 °C. Protease V8 was chosen because it is resistant to denaturation by GuHCl (45).

The protease digestion patterns of native and partially unfolded scaffolding protein are shown in Fig. 4. Under the buffer conditions used, protease V8 cleaves after both glutamic and aspartic acid residues. As the amount of GuHCl increased, the pattern of cleavage changed, with new bands becoming more prominent. The action of the protease was also inhibited, re-
Scaffolding protein at 0.6 mg/ml was incubated with polyacrylamide gel electrophoresis. After the times indicated under “Experimental Procedures.” Empty procapsid shells were added to 200 μl of each digest for a final concentration of 1 mg/ml (53 nM). The reactions were incubated at room temperature for 2 h and then centrifuged through 5–20% sucrose gradients for 35 min at 35 K. 18 fractions were collected from each gradient through a pinhole at the bottom of the tube. The fractions were run on 15% SDS gels, and the gels were stained by Coomassie Blue.

FIG. 4. V8 proteolysis of native or partially denatured scaffolding protein. Scaffolding protein at 0.6 mg/ml was incubated with protease V8 at 10 °C in buffer or buffer with GuHCl (or NaCl) added as described under “Experimental Procedures.” The reactions were allowed to proceed for 2 h and then centrifuged through sucrose gradients to separate scaffolding protein by GuHCl.

Two assays were carried out to test for the assignment of the C terminus as the essential end for binding, we prepared proteolytic scaffolding protein fragments. These were tested for their ability to copolymerize with the coat protein into procapsid-like shells. Fragments were generated with either trypsin or chymotrypsin rather than V8, because inhibitors were available to stop the proteolysis without denaturing the scaffolding. Samples of scaffolding protein were digested for a given amount of time with each protease before addition of the appropriate inhibitor to stop the proteolysis.

Two assays were carried out to test the ability of the fragments to interact with coat subunits. In the first, the entire digest was mixed with purified empty procapsid shells to test whether the fragments could enter and stably bind to the procapsid lattice. Complete scaffolding protein molecules efficiently re-enter empty procapsids and refill the interior space (13). In the second assay the fragments were incubated with monomeric coat subunits to test for assembly of procapsids in vitro. The reactions were allowed to proceed for 2 h and then centrifuged through sucrose gradients to separate scaffolding fragments stably associated with procapsids from those remaining as monomers.

Digestion of scaffolding protein by chymotrypsin resulted in only one cleavage, generating a large fragment of 30 kDa and a small one of 10 kDa. The large chymotryptic fragment was active in the reentry assay and was recovered associated with shell structures in the middle fractions of the sucrose gradient shown in Fig. 5b. The large fragment also copolymerized with coat subunits, forming procapsid shells in vitro, as shown in Fig. 6b. The small fragment did not associate with the capsids under any conditions and remained at the top of the sucrose gradients.

Trypsin digestion of scaffolding protein resulted in a range of bands, from approximately 35 to 10 kDa. The bands from the trypsin digest migrating at 30 kDa were able to associate with the coat shells in both assays, whereas a larger band migrating at 35 kDa was not active in either assay (Figs. 5a and 6a).

The 30-kDa chymotrypsin band, and the 30- and 35-kDa trypsin bands were chosen for N-terminal sequencing. Both the 30-kDa bands had lost their N termini, from cleavages at Tyr63 and Arg65. The 35-kDa fragment had an intact N terminus and was estimated to have lost approximately 20–30 residues from the C terminus. These results demonstrated that the scaffolding protein contains multiple domains, the unfolding of P22 scaffolding protein to an apparent equilibrium by both GuHCl and heat revealed a transition that was definitely not two-state. Two signals, fluorescence and CD, failed to give coincident transitions. In both cases, the transition monitored by fluorescence lagged behind that monitored by CD. Because both transitions were monitored, the same was true for secondary structure, in this case α-helix,
capsids. Scaffolding protein at 1 mg/ml (3 μM) that had been digested by either trypsin (a) or chymotrypsin (b) as described under “Experimental Procedures.” Coat monomers were added to 200 μl of each digest to give final coat protein concentration of 0.5 mg/ml (10.5 μM). The reactions were incubated at room temperature for 2 h and then centrifuged through 5–20% sucrose gradients for 35 min at 35 K. 18 fractions were collected from each gradient through a pinhole at the bottom of the tube. The fractions were run on 15% SDS gels. The gels were stained by Coomassie Blue.

Whereas the fluorescence signal indicates the local environment of the single tryptophan residue, the fluorescence signal would be expected to drop first if scaffolding protein denaturation proceeded through a molten globule intermediate (46). That some secondary structure was lost before any change in fluorescence was detected implies that the intermediates retained some tightly packed structure. These intermediates are therefore not molten globules but molecules in which only particular domains have unfolded.

This analysis is supported by the results reported in the accompanying paper for the denaturation of the Y214W mutant protein, which has an additional tryptophan (36). This protein displayed an extra, lower temperature transition by fluorescence.

Identification of the Coat-binding Region—Because low concentrations of GuHCl are sufficient to disrupt the binding interactions between scaffolding protein and coat shells, it is reasonable to conclude that the part of the molecule unfolded at these GuHCl concentrations is the region involved in binding. Identification of a region within the C-terminal half of the protein as the binding region was confirmed by the behavior of proteolytic scaffolding protein fragments in the shell binding and assembly assays; loss of the C-terminal 20–30 residues made the scaffolding inactive, whereas loss of the 65 N-terminal residues had no effect.

Parker et al. (47) have recently reported the results of similar experiments using scaffolding protein fragments expressed from cloned genes. A scaffolding fragment lacking 140 residues from the N terminus of the P22 scaffolding protein is an unusually unstable protein that begins to lose secondary structure upon addition of the smallest amount of denaturant and at surprisingly low temperatures. The coat protein of P22 is equally unstable to denaturant (42). It may be that unassembled viral proteins must be conformationally flexible to permit conformational changes required during the assembly process. The crystal structures of the RNA plant viruses (56) as well as SV40 showed that the viral capsomeres are held together by entwined arms (57). These arms could not have been folded before assembly, because they would not have been free to make such extensive interactions. A similar flexible arm might attach the scaffolding protein to the coat protein shell. Consistent with this analysis, the C-terminal regions of the scaffolding protein, which contain the coat-binding domain, are the least stable.

In addition to low stability of the secondary structure, the scaffolding protein appears to have little packed tertiary structure, because all peptide amide groups are available to rapid deuterium exchange (28). Consistent with a large exposure of nonpolar residues, each scaffolding monomer binds 12–16 molecules of 1,1’-b(4-anilino)napthalene-5-sulfonic acid (bisANS) (23), a small dye used to probe the accessibility of hydrophobic surfaces on proteins (58).

Although these features are characteristic of a molten globule (59), the scaffolding protein does display what appears to be cooperative unfolding of distinct regions. The scaffolding protein simply may not have a hydrophobic core as for typical globular proteins. Perhaps it consists of a series of loosely interacting helices. Tight interactions of exposed helical surfaces are probably formed within the closely packed interior of the procapsid. The formation of these interactions might be a
driving force in procapsid assembly. In this sense, the final product of the scaffolding protein folding pathway is not an individual scaffolding subunit but an assembled procapsid.

It is interesting that the microtubule-binding protein tau is highly extended and contains almost no secondary structure in solution (60), leading to the suggestion that it exists in a "natively denatured" state (61). The MAP-2 protein has similar properties (62). Although scaffolding protein does have substantial secondary structure, the proteins are similar in that all can function despite lacking the characteristics of folded proteins. The scaffolding protein may represent a class of proteins that are not meant to lead an independent existence and thus do not require the same degree of structure as a typical soluble enzyme.

REFERENCES
1. Casjens, S., and Hendrix, R. (1988) in The Bacteriophages (Calendar, R., ed) Vol. 1, pp. 15–93, Plenum, New York.
2. Prevelige, P. E., Jr., Thomas, D., and King, J. (1988) J. Mol. Biol. 202, 743–757.
3. Prevelige, P. E., Jr., Thomas, D., and King, J. (1993) Biophys. J. 64, 824–835.
4.ixon, F. J. (1995) Semin. Virol. 4, 135–144.
5. Horwitz, M. S. (1991) in Fundamental Virology (Fields, B. N., and Knipe, D. M., eds) pp. 529–576, Raven Press, New York.
6. Matsuick-Kumar, L., Newcomb, W. W., Brown, J. C., McCann, P. J., Hurlbut, W., Weinheimer, S. P., and Gao, M. (1995) J. Virol. 69, 4347–4356.
7. Thomas, D. R., Newcomb, W. W., Brown, J. C., and Homa, F. L. (1995) J. Virol. 69, 3680–3703.
8. Thaman-Connmikke, P. A., Greene, B., Malinski, J., King, J., and Chiu, W. (1996) Biophys. J. 74, 559–568.
9. Prasad, B. V. V., Prevelige, P. E., Jr., Marietta, E., Chen, R. O., Thomas, D., King, J., and Chiu, W. (1996) J. Mol. Biol. 260, 85–98.
10. Caspar, E. L. (1988) Biophys. J. 52, 103–138.
11. King, J., and Casjens, S. (1974) Nature 244, 112–119.
12. Greene, B., and King, J. (1994) Virology 205, 188–197.
13. King, J., and Chiu, W. (1996) in Structural Biology of Viruses (Chiu, W., Burnett, R., and Garreau, R., eds) pp. 268–311, Oxford University Press, New York.
14. Galisteo, M. L., and King, J. (1993) Biophys. J. 65, 227–235.
15. Earnshaw, W., Casjens, S., and Harrison, S. C. (1976) J. Mol. Biol. 104, 387–410.
16. Fuller, M. T., and King, J. (1981) J. Virol. 40, 292–300.
17. Fuller, M. T., and King, J. (1982) J. Mol. Biol. 156, 633–665.
18. Earnshaw, W., Casjens, S., and Harrison, S. C. (1976) J. Mol. Biol. 104, 387–410.
19. Fuller, M. T., and King, J. (1981) J. Virol. 40, 292–300.
20. Earnshaw, W., Casjens, S., and Harrison, S. C. (1976) J. Mol. Biol. 104, 387–410.