In Vitro Folding of Phage P22 Coat Protein with Amino Acid Substitutions That Confer *in Vivo* Temperature Sensitivity†

Carolyn M. Teschke* and Jonathan King

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received September 26, 1994; Revised Manuscript Received November 30, 1994®

ABSTRACT: The coat protein that forms the icosahedral shell of phage P22 can be efficiently refolded *in vitro* [Teschke, C. M., & King, J. (1993) *Biochemistry 32,* 10839−10847]. Temperature-sensitive mutants of coat protein interfere with folding or assembly *in vivo* [Gordon, C. L., & King, J. (1993) *J. Biol. Chem. 268,* 9358−9368]. The folding of a set of phage P22 coat proteins carrying the temperature-sensitive for folding *(tsf)* substitutions W48Q, A108V, G232D, T294I, and F353L has been investigated *in vitro.* Denatured *tsf* polypeptides were able to fold into soluble species with high efficiency. The efficiency of folding of the wild-type (WT) and mutant polypeptides at different temperatures showed sharp transitions where aggregation predominated over folding. The refolding of the *tsf* mutant proteins did not show an obvious thermal defect in yield. The *tsf* polypeptides folded through the long-lived kinetic intermediate previously described for WT coat protein with similar relaxation times. The folding kinetics of the *tsf* polypeptides in bisANS, a hydrophobic fluorescent dye, were also similar to those of the WT protein. However, the folded *tsf* proteins showed decreased secondary structure compared to WT coat protein. Analysis of the folded state by native gel electrophoresis revealed that the *tsf* coat proteins folded into dimers and trimers, not monomers. The dimer and trimer species were incompetent for assembly. Once formed, dimers and trimers showed no propensity toward aggregation. The folding pathway of the mutant polypeptides must be quite similar to the WT pathway, but at some step inappropriate intersubunit interactions occur due to the amino acid substitutions, trapping the subunits from assembly.

The amino acid sequence of a polypeptide codes for both its native structure and folding pathway (Anfinsen, 1973). Investigations of *in vitro* refolding pathways have concentrated on small highly soluble enzymes and proteins such as lysozyme, ribonuclease, barnase, dihydrofolate reductase, staphlyoccocal nuclease, and BPTI (Kim & Baldwin, 1982, 1990; Matthews, 1993; Creighton, 1990). Reports of the refolding of structural proteins have been less frequent. In fact, the structural proteins, actin, tubulin and collagen, have not been refolded back to their native states *in vitro.* Actin and tubulin require a special chaperone (Gao *et al.,* 1992; Yaffe *et al.,* 1992), and collagen chains are unable to reach the correct register in the absence of their registration peptides, instead forming gelatin (Harrington & Rao, 1970; Engel & Prockop, 1991).

The folding of the structural proteins of viral capsids poses additional problems; subunits in different positions of the icosahedral shell must take different conformations (Rossman & Erickson, 1985; Valegard *et al.,* 1990; Golmohammadi *et al.,* 1993). Seven different conformations are required to form the topologically closed $T = 7$ icosahedral shells of viruses such as cauliflower mosaic virus or phage P22 (Cheng *et al.,* 1992; Prasad *et al.,* 1993). In P22, one conformation is present in the capsid pentamers, and the other six are found in the capsid hexamers (Prasad *et al.,* 1993). The hexamers show a pronounced distortion in symmetry about the quasi 6-fold axis. The coat protein of P22 takes these different conformations without covalent modification.

The P22 coat protein is one of the few structural proteins whose refolding kinetics have been described *in vitro* (Teschke & King, 1993). Since monomeric coat protein subunits can assemble, with the assistance of scaffolding protein, into correctly dimensioned closed icosahedral shells, they either fold into a precursor conformation that necessitates switching into one of the quasiequivalent conformations during polymerization, or the subunits partition between the seven necessary conformers in solution. Coat subunit folding and virion assembly can be uncoupled *in vitro* in phage P22 (Prevelige *et al.,* 1988, 1993). Thus, the effects of amino acid substitutions on the folding or assembly pathways can clearly be delineated.

Mutagenesis has been extensively used to probe the amino acid sequences important for protein folding. The supposition is that amino acids are involved in determining both the final structure of proteins and conformation of folding intermediates (Lecomte & Matthews, 1993). Amino acid substitutions allow the investigation of the role of a specific amino acid in the folding or structure of the protein through analysis of folding and unfolding kinetics as well as equilibrium studies (Beasty *et al.,* 1987). The approach that we have taken to identify amino acids important for protein folding or subunit assembly is to select conditional mutants resulting in defects in the folding or assembly pathways only under the restrictive conditions. Such temperature-sensitive for folding mutations *(tsf)* have been systematically studied in the thermostable tailspike of P22 (Goldenberg & King, 1981; Goldenberg *et al.,* 1983; Haase-Pettingell & King, 1988; Sturtevant *et al.,* 1989; Mitraki *et al.,* 1993). Mutations of this nature have the advantage of locating residues likely

† Supported by NIH Grant GM17980.

‡ Current address: Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT 06269-3125. E-mail address: teschke@uconnvm.uconn.edu.

® Abstract published in *Advance ACS Abstracts,* January 15, 1995.

0006-2960/95/0434-6815$09.00/0

to affect the folding pathway rather than the final structure of the protein.

Amino acid substitutions at 17 different sites in P22 coat protein have been identified that render production of viable phage temperature-sensitive (Gordon & King, 1993; Gordon & King, 1994). The *tsf* mutants of coat protein sharply reduce the yield of soluble protein *in vivo*; the newly synthesized coat polypeptides aggregate to form inclusion bodies prior to reaching the assembly-competent state. *In vitro*, the mutations do not destabilize the folded subunit since differential scanning calorimetry, and pressure and urea denaturation of the mutant proteins show them to be as stable as the WT protein (Fogel *et al.*, 1995; Galisteo *et al.*, 1995; C. Teschke, unpublished results). These observations suggested that *in vivo* the *tsf* amino acid substitutions of coat protein acted by destabilizing a folding intermediate resulting in their aggregation into inclusion bodies, similar to the *tsf* phenotype of the P22 tailspike protein (Haase-Pettingell & King, 1988; Mitraki *et al.*, 1991). Remarkably, *in vivo* all 17 *tsf* mutants are rescued by overproduction of the chaperones GroEL and GroES (Gordon *et al.*, 1994).

Since *in vivo* inclusion bodies frequently form from incorrect interactions of folding intermediates and GroEL is thought to bind folding intermediates, it seemed a reasonable hypothesis that the *tsf* amino acid substitutions would have an effect on the folding pathway of P22 coat protein (Mitraki & King, 1989; Hendrick & Hartl, 1993). Here we present evidence that the amino acid substitutions indeed affect the folding of coat polypeptides *in vitro* leading the mutant polypeptide chains to be trapped in stable, but assembly-incompetent, dimers and trimers.

MATERIALS AND METHODS

Chemicals. Ultra pure GuHCl[1] was purchased from Pierce and ultra pure urea from Schwartz-Mann ICN. BisANS was purchased from Molecular Probes. DTSSP was purchased from Pierce. All other chemicals were reagent grade purchased from common sources.

Buffer. The buffer used in all studies, except for cross-linking and circular dichroism (CD), was 50 mM Tris, 25 mM NaCl, and 2 mM ethylene diamine tetraacetic acid, pH 7.6. For cross-linking, the buffer was the same except that HEPES was used in place of Tris. For CD the buffer was 50 mM Tris and 25 mM NaCl, pH 7.6.

Concentration of Coat Protein Shells. The concentration of the coat protein shells was determined by dissociating and denaturing the shells in 6 M GuHCl and determining the absorbance at 280 nm. All extinction coefficients were determined by the method of Johnson (1988). The extinction coefficient used for WT and all the mutants except for W48Q was 0.957 mL mg^{-1} cm^{-1}. The extinction coefficient of W48Q was determined to be 0.830 mL mg^{-1} cm^{-1}. The shells were diluted so that the storage concentration was 8 mg/mL.

Fluorescence Spectra. Fluorescence spectra of the WT and mutant shells, unfolded polypeptides, refolded proteins, and the kinetic intermediate were taken with a Hitachi F4500 spectrofluorometer. The procedure for folding of coat polypeptides is described below. The spectrum of the kinetic intermediate was taken by beginning the emission scan immediately after initiation of refolding (Teschke & King, 1993). The spectrum of the folded coat proteins was taken 1 h after dilution from denaturant. Settings for tryptophan fluorescence were as follows: excitation wavelength, 280 nm; excitation bandpass, 2 nm; emission bandpass, 5 nm; photomultiplier voltage, 700 V; response, 0.5 s.

Refolding of Coat Polypeptides. Coat protein shells were dissociated and denatured at 2 mg/mL in either 4.5 M GuHCl or 6.75 M urea at room temperature for a least 30 min. Refolding was initiated by rapidly diluting the GuHCl or urea 200-fold so that final the protein concentration was 10 μg/mL. The refolding was monitored by either the intrinsic fluorescence of the six tryptophans of coat protein or by the extrinsic fluorescence of the dye, bisANS. The relaxation times of refolding monitored by tryptophan fluorescence were determined using the formula for a first order decay using two exponentials. The data were fit using the computer program Kaleidagraph (Abelbeck software). The settings for tryptophan fluorescence were as follows: excitation wavelength, 280 nm; excitation bandpass, 2 nm; emission wavelength, 340 nm; emission bandpass, 5 nm; photomultiplier voltage, 700 V; response, 0.5 s. The settings for bisANS fluorescence were as follows: excitation wavelength, 390 nm; emission wavelength, 495 nm. The remaining settings were the same. The concentration of bisANS used was 2 μM (Teschke & King, 1993).

Aggregation Experiments. Shells were denatured in GuHCl or urea at 2 mg/mL as described above. The denatured protein, 10 μL, was placed in the bottom of a siliconized microfuge tube and refolding initiated by the addition of buffer, held at the experimental temperature, so that the final concentration of protein was 100 μg/mL. The refolding protein was incubated for 1 h and moved to ice, and an aliquot was centrifuged in a TLA-100 rotor in a Beckman TL100 ultracentrifuge at 100 000 rpm for 15 min at 4 °C. This protocol was designed to pellet anything sedimenting faster than 28S. The supernatant was removed and diluted with 3× SDS sample buffer. The pellet was suspended in 1× sample buffer so that the volume was equal to that of the diluted supernatant. Equal volumes of the pellet and supernatant samples were run on 10% SDS−polyacrylamide gels, visualized by silver stain (Rabilloud *et al.*, 1988) and quantified using a LKB densitometer. The percent soluble is the fraction of the total protein that was found in the supernatant.

Circular Dichroism. WT and mutant coat proteins were prepared for CD by dilution 20-fold from 6.75 M urea so that the final protein concentration was 100 μg/mL and incubated overnight at 4 °C. CD spectra were taken with an AVIV 60DS spectrophotometer. The CD spectra, the average of nine repetitions, were taken in a 0.5 cm cell with an averaging time of 2 s, a step size of 1 nm, and the bandwidth set to 1.5 nm. The protein concentration was determined by the absorbance at 280 nm using the extinction coefficients described above.

Native Gel Electrophoresis. Coat protein samples, 2 mg/mL, were refolded from 6.75 M urea to a final protein concentration of 100 μg/mL at either 4 or 20 °C as described above. Aliquots were mixed with 3× native gel sample buffer and run on a native gel that was 4.3% pH 8.3 stacking gel/7.5% pH 9.5 separation gel as described in Andrews (1986). The gel was run at 4 °C at 10 mA constant current

[1] Abbreviations: GuHCl, guanidine hydrochloride; bisANS, 1,1′-bi(4-anilinonaphthalene-5-sulfonic acid); SDS, sodium dodecyl sulfate; DTSSP, 3,3′-dithiobis(sulfosuccimidylpropionate); βME, 2-mercaptoethanol.

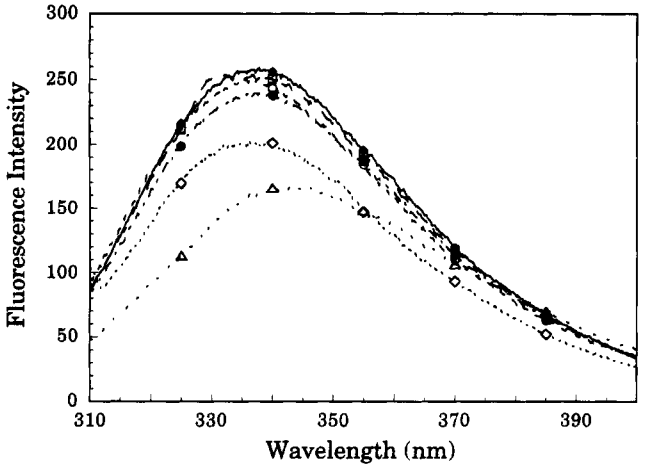


FIGURE 1: Fluorescence emission spectra of WT and mutant empty procapsid shells. The spectra of WT (◆), W48Q (◇), A108V (○), G232D (□), T294I (●), and F353L (△) excited at 280 nm. The protein concentration was 10 μg/mL.

and visualized by silver stain (Rabilloud *et al.*, 1988).

Cross-Linking of Coat Proteins. Coat protein samples, 2 mg/mL, were refolded from 6.75 M urea so that the final protein concentration was 100 μg/mL at 20 °C for 15−25 min. The samples were moved to ice, and DTSSP was added to a final concentration of 100 μM. The cross-linking reaction was quenched after 1 h on ice by the addition of 3× native gel or 3× SDS gel sample buffer with or without 150 mM βME. The samples run on SDS gels were boiled for 5 min.

Assembly Assay. WT or T294I coat protein samples, 3 mL final volume, were prepared by diluting the proteins unfolded in 6.75 M urea 20-fold to 100 μg/mL at 4 °C overnight as described above. The refolded proteins were concentrated using a Centricon-10 (Amicon) concentration device with or without the addition of 300 μg of scaffolding protein. Two milliliters of buffer was added after the first concentration step to decrease the residual urea concentration and the solution concentrated again. The final volumes of the proteins were adjusted to be equal at ~100 μL. Uranyl acetate negatively stained grids were made of 10 μL of each sample doped with T4 phage for focusing. A Jeol 1200C electron microscope was used to view the negatively stained grids.

Aggregation Propensity of Folded Coat Proteins. WT and T294I were unfolded in 4 M GuHCl at 2 mg/mL for 30 min and refolded by dilution 20-fold with cold buffer. The samples were incubated at 4 °C overnight. An aliquot of each refolded protein was held on ice for the duration of the experiment and then diluted with native gel sample buffer. The rest of the refolded protein was moved to a tube that had been prewarmed at 28 °C in a water bath. Aliquots were

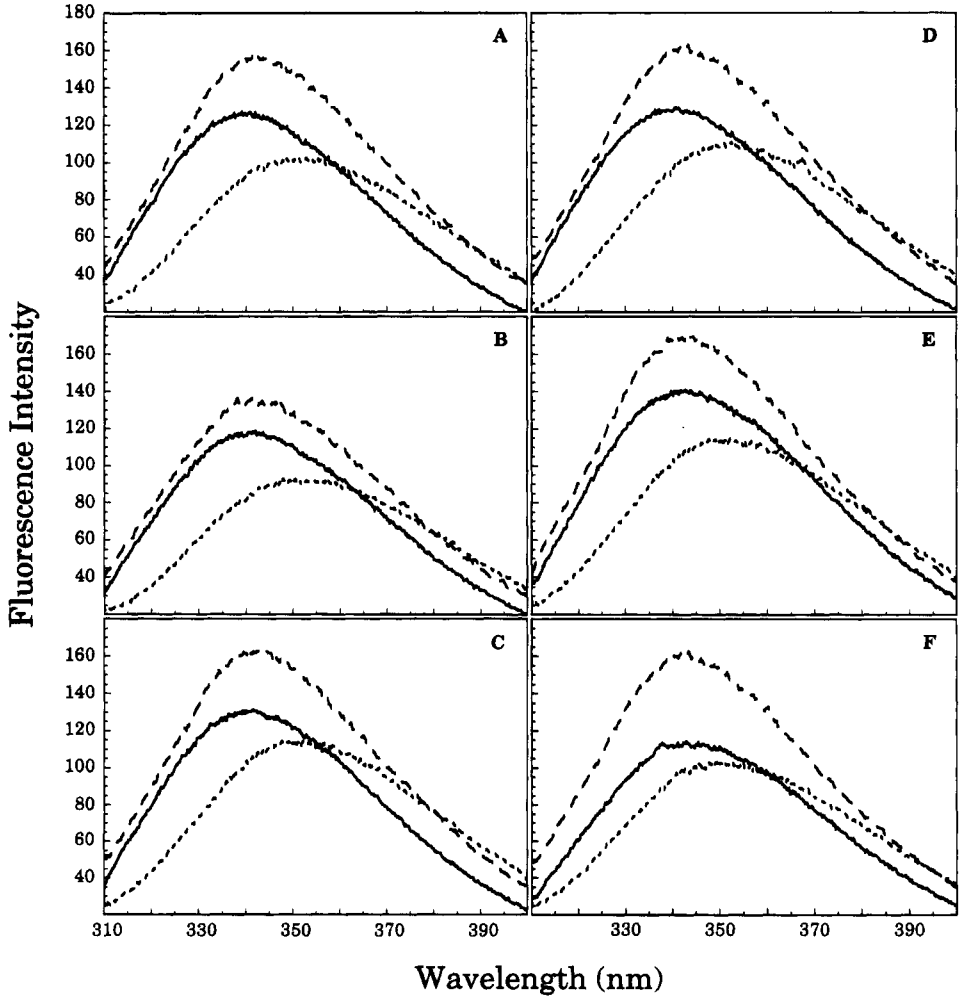


FIGURE 2: Fluorescence emission spectra of WT and mutant unfolded coat protein in 4.5 M GuHCl (···), the kinetic intermediate (- - -), and folded coat protein (—). The panels are as follows: (A) WT; (B) W48Q; (C) A108V; (D) G232D; (E) T294I; (F), F353L. The emission scan of the kinetic intermediate was initiated within 7 s after dilution and took 22 s to complete. The spectra of the folded proteins were taken 1 h after initiation of refolding. The protein concentration was 10 μg/mL.

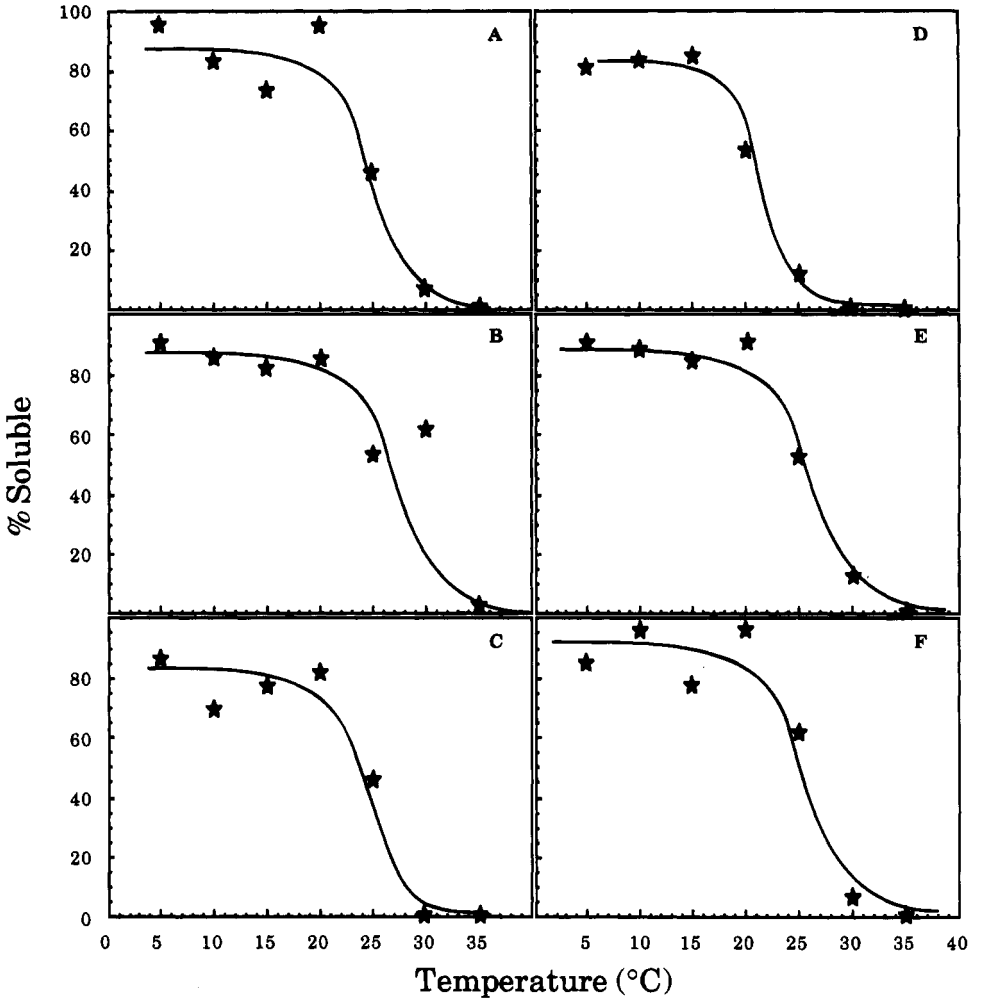


FIGURE 3: Efficiency of folding at different temperatures when diluting 20-fold from 4.5 M GuHCl. Soluble protein was separated from the aggregates by ultracentrifugation as described in the Materials and Methods. The final protein concentration was 100 μg/mL. The panels are as follows: (A) WT; (B) W48Q; (C) A108V; (D) G232D; (E) T294I; (F) F353L.

Table 1: Midpoint in the Efficiency of Folding of WT and Mutant Coat Proteins by Rapid Dilution from Urea or GuHCl[a]

	midpoint (°C)			midpoint (°C)	
	GuHCl	urea		GuHCl	urea
WT	25	36	G232D	21.5	43
W48Q	27	41	T294I	26	41
A108V	24.5	41	F353L	25.5	44

[a] The midpoints of the curves shown in Figures 3 and 5 were determined by (percent soluble at low temperature−% soluble at high temperature)/2. This midpoint percentage was used to determine the midpoint temperature.

taken periodically after the temperature shift and then diluted with native gel sample buffer on ice. A 7.5% native gel was run as described above and silver stained.

RESULTS

The *tsf* coat mutants chosen for study were W48Q, A108V, G232D, T294I, and F353L. These mutants are spread throughout the 430 amino acid sequence of coat protein and are representative of the *tsf* amino acid substitutions. To understand the effect of the amino acid substitutions on the folding of coat polypeptide, *in vitro* experiments were carried out as described in Teschke and King (1993). Procapsids were prepared by infecting cells at permissive temperatures with strains blocked in DNA packaging (Prevelige *et al.*, 1988). The yields of mutant procapsids were somewhat lower than those of the wild-type infections due to some loss to inclusion bodies. The scaffolding subunits were released from the procapsids by treatment with 0.5 M GuHCl (Fuller & King, 1981). The resulting purified empty shells served as the storage form of the mutant proteins.

Refolding of Mutant Coat Proteins. WT coat polypeptides have been successfully refolded with high yield from denaturant *in vitro* by both dialyzing and by rapid dilution of the denaturant (Prevelige *et al.*, 1988; Teschke & King, 1993). Dissociated and unfolded coat polypeptide chains were obtained by incubation of empty procapsid shells with 4.5 M GuHCl. Refolding was initiated by rapid dilution that decreased the concentration of GuHCl by 200-fold. WT coat polypeptides fold through a long-lived kinetic intermediate with a relaxation time of ~300 s at 25 °C, long enough to obtain fluorescence emission spectra of the intermediate (Teschke & King, 1993). The kinetic intermediate had characteristics suggesting that it had incomplete tertiary interactions, a high hydrophobicity as compared to the folded subunit, but completed secondary structure (Teschke & King, 1993).

To determine whether the *tsf* mutant coat polypeptides could be refolded *in vitro*, similar protocols were followed but at 20 °C. Shell stocks were diluted so that the concentration was 8 mg/mL as described in the Materials and Methods. Fluorescence spectra of empty shells were similar to the WT having emission maxima at 337 nm, the

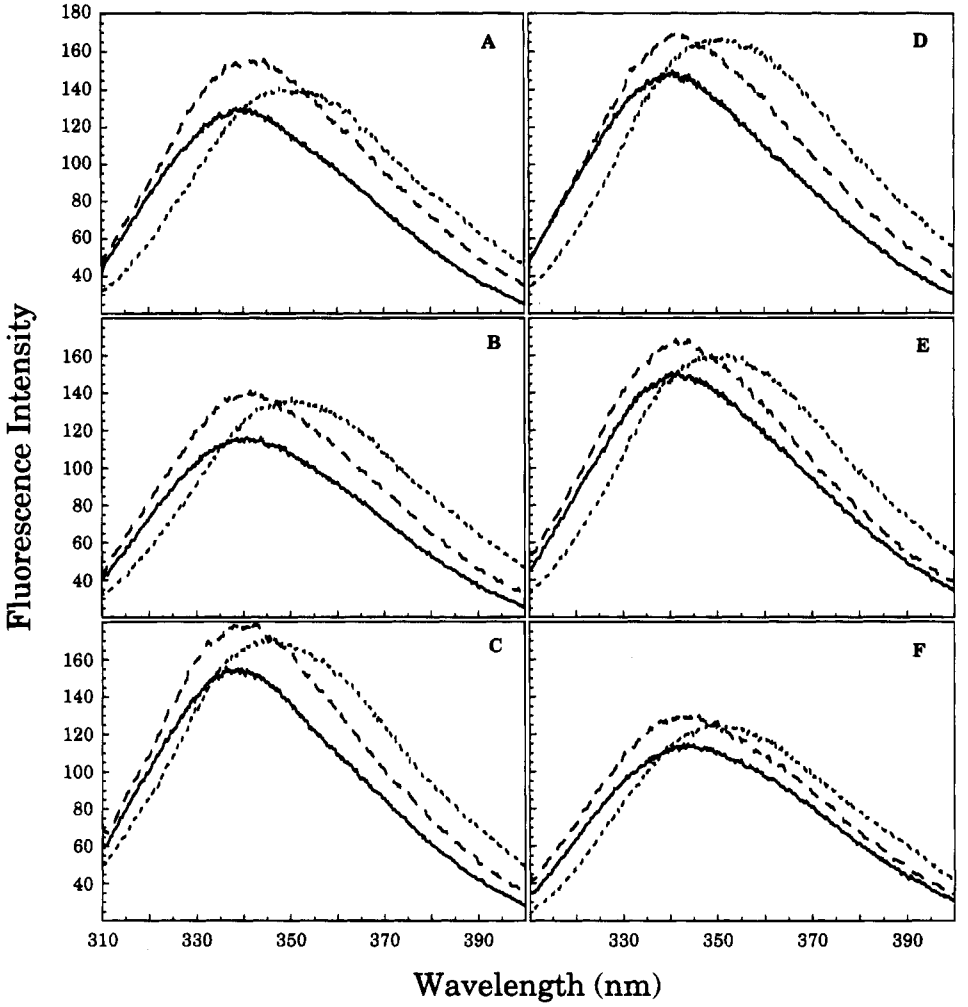


FIGURE 4: Fluorescence emission spectra of WT and mutant unfolded coat protein in 6.75 M urea (---), the kinetic intermediate (- - -) and folded coat protein (—). The panels are as follows: (A) WT; (B) W48Q; (C) A108V; (D) G232D; (E) T294I; (F) F353L. The emission scan of the kinetic intermediate was initiated within 7 s after dilution and took 22 s to complete. The spectra of the folded proteins were taken an hour after initiation of refolding. The protein concentration was 10 μg/mL.

exception being F353L which dissociates quite quickly upon storage (Figure 1). W48Q showed a decrease in intensity as expected for a protein with one less tryptophan. Investigations of F353L by differential scanning calorimetry have shown that shells of F353L are drastically less stable than WT or other *tsf* shells (M. Galisteo, personal communication). Furthermore, the dissociated form of F353L has been determined to be trimeric (P. Prevelige, personal communication).

The fluorescence spectrum of each mutant polypeptide in 4.5 M GuHCl showed an emission maximum centered around 352 nm as seen with the WT coat polypeptide (Figure 2). Upon refolding, none of the mutant proteins showed visible aggregates, allowing fluorescence spectra to be recorded. All the mutant polypeptides showed a shift to lower wavelength of ~340 nm as expected if folding occurred (Figure 2). W48Q again showed lower fluorescence intensity as expected. The spectra were similar to the WT spectrum. The spectra of the kinetic intermediate were obtained immediately after rapid dilution from GuHCl. Data collection took 22 s. Again the *tsf* mutant spectra were quite similar to the WT spectrum with an emission maxima of ~343 nm (Figure 2). These results are consistent with those previously reported for the refolding of WT coat protein (Teschke & King, 1993).

In vivo the mutant polypeptides fail to fold productively and accumulate in inclusion bodies at restrictive temperatures. Therefore, we examined the yield of soluble protein at different temperatures. To separate the aggregates from soluble subunits, the solutions were centrifuged so that particles sedimenting faster than 28S (~10 subunits) would pellet. The refolding of WT coat polypeptides gave high yields at temperatures from 5 to 20 °C. At higher temperatures, the yield of soluble protein was reduced and the aggregation pathway dominated over productive folding. The sharp transition in the temperature dependence of productive folding had a midpoint at 25 °C. The mutants showed similar overall behavior; efficient refolding from 5 to 20 °C, with a sharp transition having a midpoint from 21 to 27 °C (Figure 3 and Table 1). This result was unexpected as the *tsf* mutants in tailspike showed reduced efficiency of refolding *in vitro* in comparison to WT tailspike (Mitraki *et al.*, 1993).

One possibility for the invariance in the efficiency of refolding of the *tsf* coat polypeptides as compared to the WT protein was that the critical thermolabile intermediate in all of the proteins was destabilized by the residual GuHCl, thus masking the effect of the amino acid substitutions. Consequently, the refolding from urea was examined. WT and mutant shells were dissociated and denatured in 6.75 M urea so that the final concentration of coat polypeptides was 2 mg/mL. This concentration of urea was determined to be sufficient to dissociate and denature WT shells (Foguel *et al.*, 1995). To initiate refolding, the urea was rapidly diluted

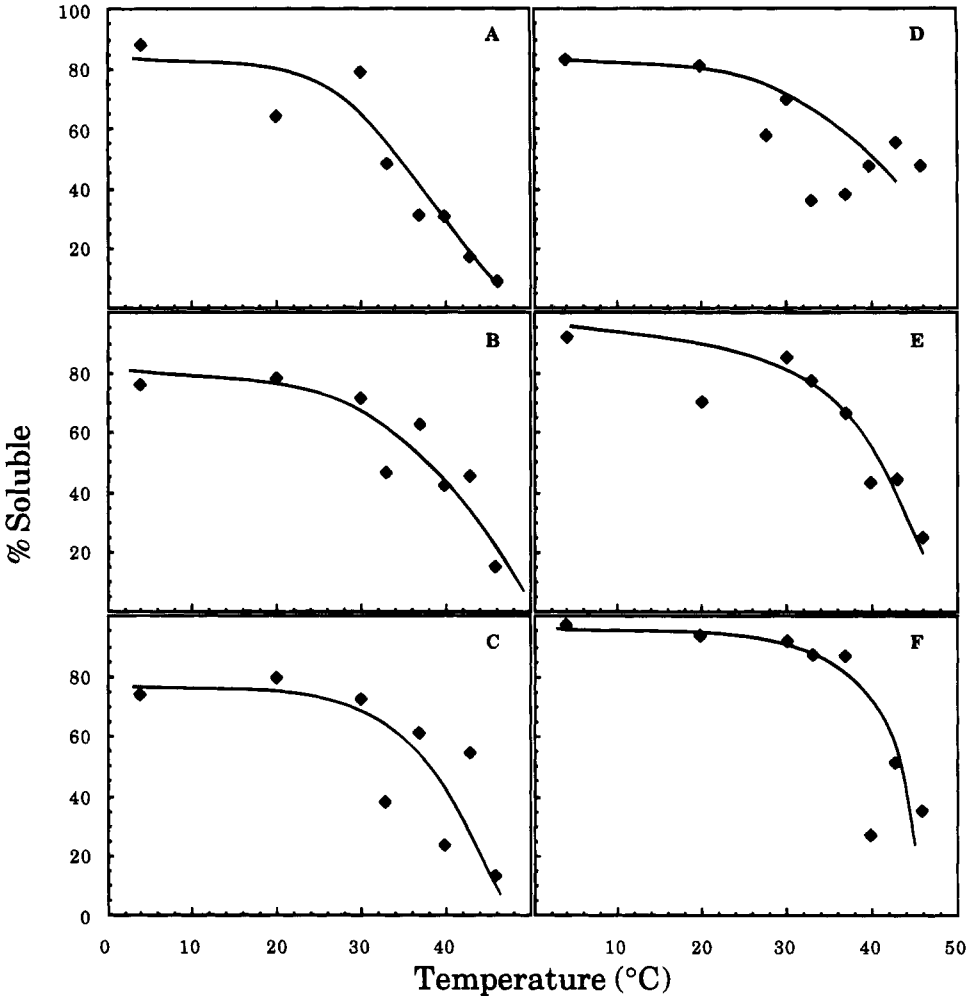


FIGURE 5: Efficiency of folding at different temperatures when diluting 20-fold from 6.75 M urea. Soluble protein was separated from the aggregates by ultracentrifugation as described in the Materials and Methods. The panels are as follows: (A) WT; (B) W48Q; (C) A108V; (D) G232D; (E) T294I; (F) F353L. The final protein concentration was 100 μg/mL.

200-fold so that the residual urea concentration was 34 mM. The spectrum of the unfolded WT protein in urea showed the same emission maxima as in GuHCl (Figure 4). The increase in intensity of the unfolded polypeptides in urea was due to the effect of urea on the fluorescence of coat protein tryptophans (data not shown). The fluorescence emission spectra of the unfolded, folded, and kinetic intermediate of WT and mutant proteins were again quite similar to each other. These data suggest that the environment of the tryptophans was not substantially affected by the *tsf* amino acid substitutions, whether the polypeptides were folded from urea or GuHCl.

The propensity of the WT and mutant coat polypeptides to aggregate upon refolding from urea was examined as described above for the refolding from GuHCl. The temperature range where the yield of soluble WT protein was high increased to 36 °C when folding from urea (Figure 5 and Table 1). The recoveries of all the mutants also increased considerably as compared to the refolding from GuHCl. The recoveries of soluble mutant proteins again showed no correlation with the phenotype observed *in vivo*; in fact, the midpoints of the mutants were higher than the WT. Urea was used as the denaturant for the majority of the remaining studies since the temperature range of efficient folding is closer to that seen *in vivo*.

Kinetics of Folding. To discern the effect of the amino acid substitutions on the kinetics of folding, the relaxation times of folding after rapidly diluting form urea were determined for the WT and mutant polypeptides in triplicate (Figure 6 and Table 2). The folding of WT coat polypeptides from urea was similar to that seen from GuHCl except that the data were not fit well by a single-exponential first-order decay, rather two exponentials were needed for a good fit of the data. The mutant polypeptides folded with the same general characteristics as the WT protein and the relaxation times were the same, within experimental error. The exception was T294I whose fast relaxation was slightly outside the range of rates of the refolding of WT coat protein. Thus, the *tsf* amino acid substitutions did not affect the kinetics of folding *in vitro*.

Previously it was shown that the hydrophobic dye, bisANS, bound to the kinetic intermediate (10 bisANS/intermediate) in the folding of WT coat protein. As folding proceeded, the number of molecules of bisANS bound to the polypeptide decreased to 1 for each subunit (Teschke & King, 1993). Thus, we determined if the *tsf* amino acid substitutions affected the kinetics of folding in bisANS (Figure 7). In each case, the kinetic intermediate of the mutant polypeptides bound more bisANS than the folded state. The kinetics were not substantially different than WT coat protein. Thus, the *tsf* proteins fold through a kinetic intermediate with the same features as the WT intermediate.

Conformation of the Folded WT and Mutant Coat Proteins. The secondary structure of the mutants was examined by

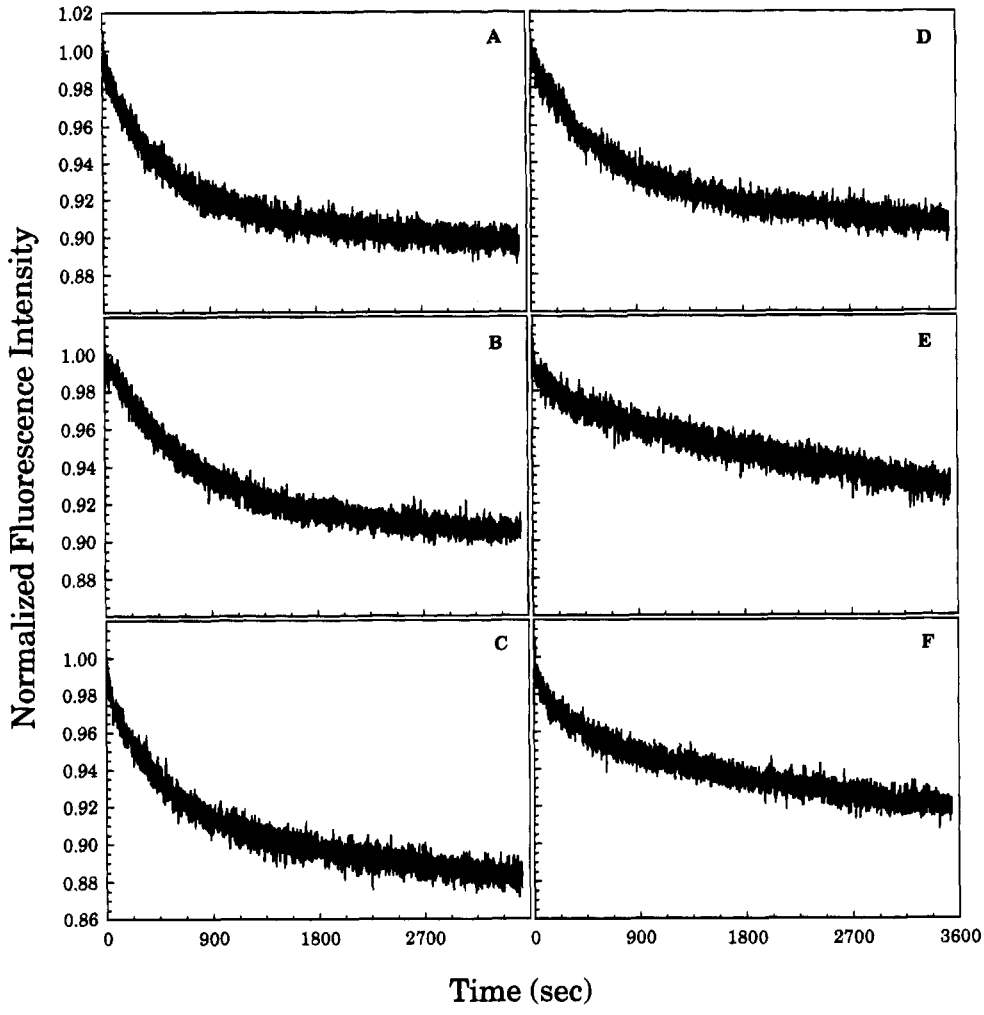


FIGURE 6: Kinetics of refolding followed by the intrinsic fluorescence of coat protein tryptophans after rapidly diluting 200-fold from 6.75 M urea. The excitation wavelength was 280 nm and the emission at 340 nm was observed. The data are normalized so that the point of highest intensity is 1.00. The panels are as follows: (A) WT; (B) W48Q; (C) A108V; (D) G232D; (E) T294I; (F) F353L. The final protein concentration was 10 μg/mL.

Table 2: Relaxation Times of Folding of the Mutant Proteins from Urea at 20 °C[a]

coat protein	τ_1 (s)	τ_2 (s)
WT	322 ± 141	2027 ± 1124
W48Q	445 ± 174	2240 ± 588
A108V	257 ± 139	1780 ±1086
G232D	356 ± 210	3495 ± 1975
T294I	89 ± 48	3518 ± 1306
F353L	236 ± 127	2436 ± 788

[a] The relaxation times of folding were determined by fitting the data to a first-order decay with two exponentials.

circular dichroism after refolding by rapid dilution to ascertain the effect of the *tsf* amino acid substitutions on the folded conformation (Figure 8). The spectrum of refolded WT showed peaks at 209 and at 222 nm, as seen previously when the protein was refolded by dialysis, and had a negative ellipticity of −6500 deg cm^2 dmol^{-1} at 209 nm and −5600 at 222 nm (Teschke *et al.*, 1993). The negative ellipticity here is lower than previously reported for WT coat protein and may be due to the effect of the residual urea on the secondary structure. In each case, the refolded mutant proteins showed a decrease in secondary structure reflected by a decrease in negative ellipticity at both 209 and 222 nm, the decrease at 222 nm being larger. Therefore, the mutant proteins fold into soluble species that have decreased secondary structure but essentially unaltered tryptophan fluorescence, suggesting that the hydrophobic cores of the proteins are not affected by the amino acid substitutions. This remains to be investigated.

The state of association of the folded, soluble WT and *tsf* coat subunits was determined by nondenaturing polyacrylamide gels (Figure 9). The proteins were refolded at 100 μg/mL at either 4 or 20 °C for 15−25 min and shifted to ice before the native gel was run. WT coat protein folded at 4 °C showed a prominent band that was taken to be the monomer species as well as small amounts of slower migrating bands that were tentatively identified as dimer and trimer bands. When folded at 20 °C, the WT protein showed monomer, dimers, trimers, and small amounts of higher order multimers. All of the refolded mutant proteins exhibited the distinct absence of significant levels of monomer. Each had prominent dimer and trimer bands at 4 °C and some higher order multimers when refolded at 20 °C. If the protein concentration was decreased 10-fold to 10 μg/mL, the mutant proteins still formed dimers and trimers and not monomers demonstrating that the dimers and trimers interact tightly (data not shown). The dimers of the proteins do not migrate to the same position on the gel suggesting that there is either a difference in their conformation or an altered equilibrium between the monomer and dimer species; those with the equilibrium shifted toward the monomer species would appear to have faster migrating dimers.

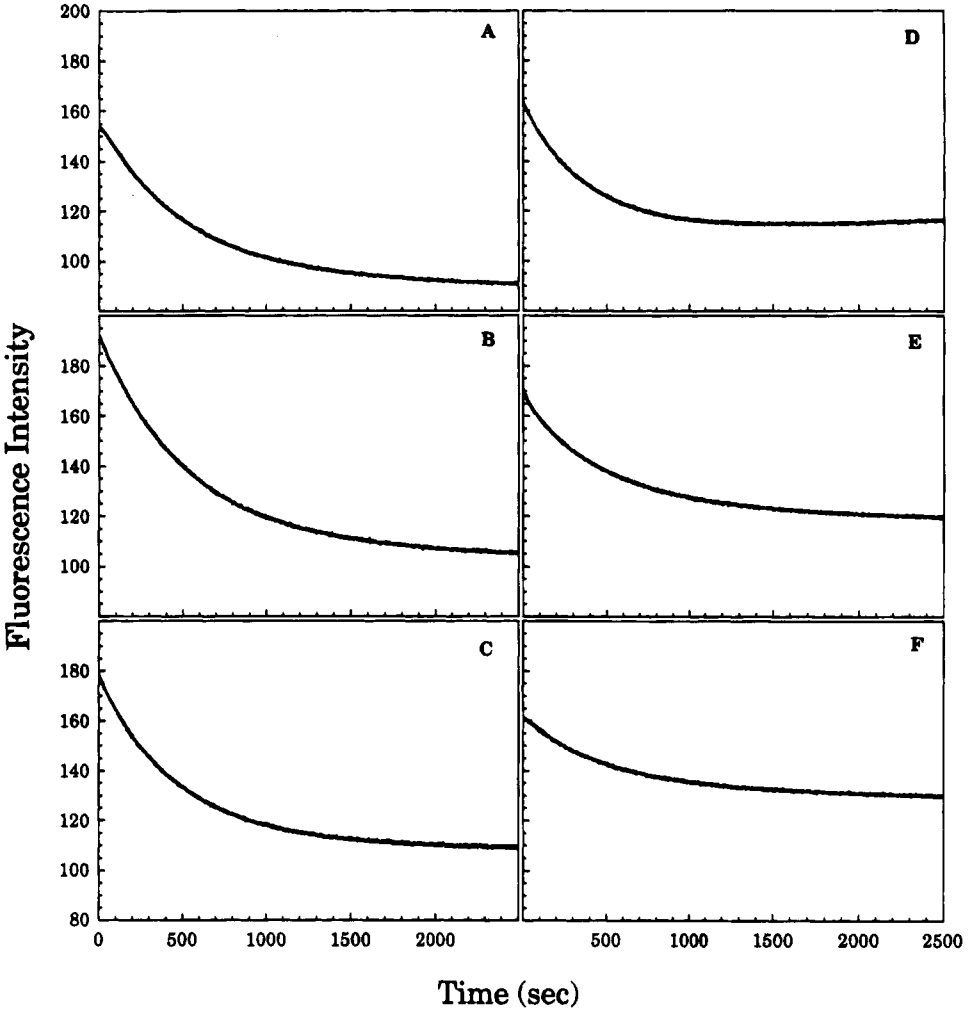


FIGURE 7: Kinetics of refolding followed by the extrinsic fluorescence of bisANS after rapidly diluting 200-fold from 6.75 M urea. The excitation wavelength was 390 nm and the emission observed at 495 nm. The panels are as follows: (A) WT; (B) W48Q; (C) A108V; (D) G232D; (E) T294I; (F) F353L. The protein concentration was 10 μg/mL, and the concentration of bisANS was 2 μM.

Chemical cross-linking with DTSSP, an amino-reactive cross-linker, was used to demonstrate that the bands tentatively identified as dimers and trimers actually corresponded to those species (Figure 10A). WT and mutant coat proteins (100 μg/mL, 2 μM) were refolded from urea at 20 °C for 15−25 min and transferred to ice as described in the Materials and Methods. The cross-linking reagent was added so that its concentration was 50-fold over the coat protein concentration to encourage intramultimer cross-linking. Cross-linking was stopped by the addition of SDS gel sample buffer or native gel sample buffer which quenched the reaction due to the presence of Tris buffer. Lanes on the SDS gel where cross-linker was added to the samples showed new bands that were calculated to be from 92 to 106 kDa (2.0−2.3 monomers), 123 to 149 kDa (2.7−3.3 monomers), and 181.5 to 200 kDa (4−4.4 monomers) corresponding to the molecular mass expected for dimers, trimers, and tetramers. Bands corresponding to pentamers were also visible on the gel. The cross-linker, DTSSP, can be reduced by the addition of βME. When the cross-linked samples were boiled in the presence of βME, the bands that corresponded to dimers, trimers, and tetramers returned to the position of monomer on the SDS gel, demonstrating that the oligomeric species were the result of the cross-linking. Samples that were folded but not cross-linked were applied to the SDS gel without the addition of βME to the sample buffer to determine if the dimers were the result of disulfide bond formation between the single cysteine residue at 405 of coat protein (Eppler *et al.*, 1991). Only F353L showed a substantial level of disulfide-linked dimers. It is not clear why this substitution should have this effect. The cysteine is closest to this substitution, and perhaps there is local destabilization of this region of the protein such that the cysteine becomes accessible during folding.

Folded samples that were cross-linked or not treated with cross-linker were also run on a native gel (Figure 10B). The amounts of dimers and trimers were not affected by the cross-linking in that cross-linking did not cause the dimers and trimers to become larger oligomers. However, new bands in the region of WT monomer were observed. These new bands in the position of monomer could be due to either charge changes brought about by the cross-linking of lysines or by capturing transient conformations. Distinguishing between these possibilities will require further investigation. Some of the mutant proteins (especially G232D) also gave rise to these new bands, but only of a subset of those observed in the WT lane. Interestingly, the mobility of the dimer bands became more uniform with the cross-linking. This suggests that the reason for the difference in mobility observed on the native gel was due to an equilibrium between monomer and dimer species. Also, the native gel demonstrated that the change of buffer from Tris to HEPES had no deleterious effect on the folding of the proteins; in fact, use of HEPES buffer seemed to improve folding. G232D

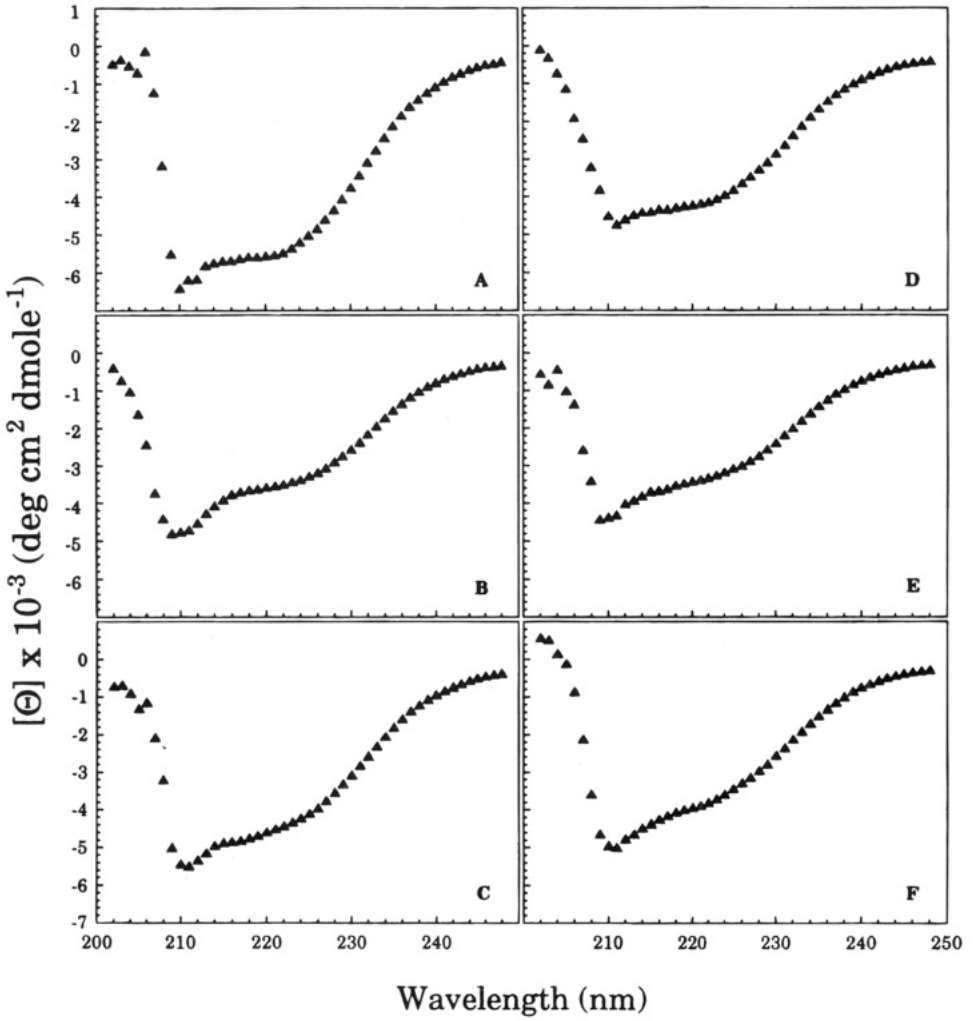


FIGURE 8: Secondary structure of refolded WT and mutant coat proteins monitored by circular dichroism. The experiment was done as described in the Materials and Methods. The panels are as follows: (A) WT; (B) W48Q; (C) A108V; (D) G232D; (E) T294I; (F) F353L. The protein concentration was 100 μg/mL.

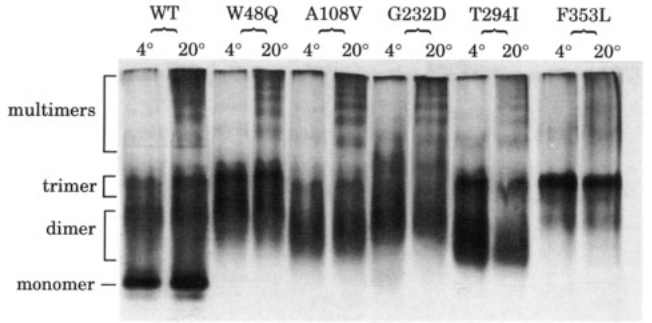


FIGURE 9: Native gel electrophoresis of refolded WT and mutant coat proteins. The species of coat protein and the temperature of folding are indicated above the gel. The position of the monomer, dimer, trimer and higher oligomers are indicated on the left side of the gel. The protein concentration in the folding reaction was 100 μg/mL. Each lane was loaded with 10 μg of protein. The gel was silver stained.

appeared to form some monomer and the all of the mutants formed decreased amounts of trimer and more dimer (Figure 10B).

Assembly Competence of the Refolded WT and tsf Mutant Proteins. To assess if the refolded mutant coat proteins were biologically active, their competence for assembly into procapsids was tested. Since assembly into procapsids is a nucleation limited reaction needing at least 0.3 mg/mL of coat protein, it was necessary to concentrate the proteins using a Centricon 10 filtration device (Prevelige *et al.*, 1988; Teschke *et al.*, 1993). If the concentration is done in the absence of scaffolding protein, spirals of coat protein are made instead of closed shells. Such spirals structures are also formed *in vivo* in scaffolding amber infections (Earnshaw & King, 1978). If the concentration step is done in the presence of scaffolding protein, then closed procapsids are formed. Refolded WT and T294I were concentrated from 0.1 mg/mL in the absence and presence of scaffolding protein (Figure 11). When WT coat protein was concentrated in the absence of scaffolding protein, many spirals were seen in the negatively stained electronmicrograph. When WT coat protein was concentrated in the presence of scaffolding protein, closed procapsids were seen in the micrograph. When the same concentration procedure was done with refolded T294I, no structures were seen in either the presence or absence of scaffolding protein, even though the protein concentration reached during concentration was higher than with WT coat protein. The few shells observed in the micrographs of T294I were residual undissociated shells from the denaturation and are also observed in micrographs of the proteins before concentration (data not shown). These data indicated that the *tsf* mutant subunits are unable to assemble, presumably because they are dimers and trimers rather than monomers.

Propensity of Folded WT and Mutant Proteins to Aggregate. One explanation for the existence of the dimers

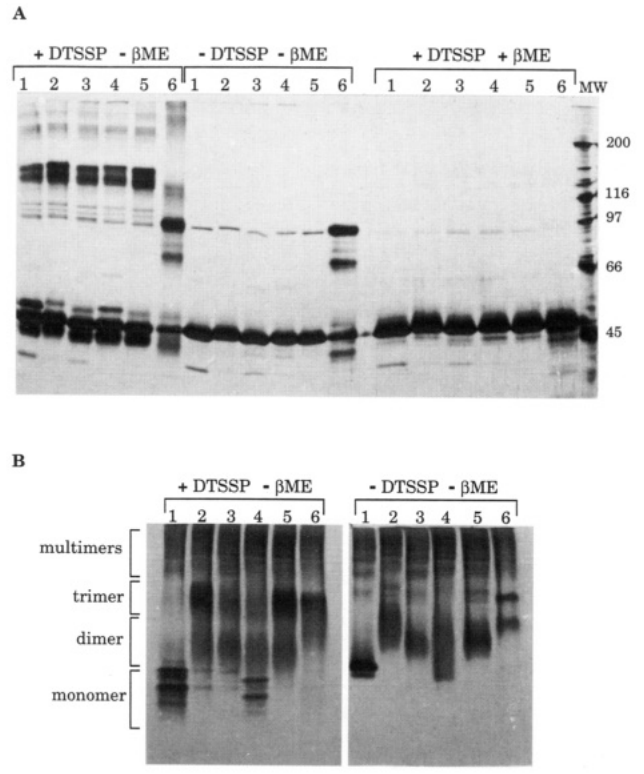


FIGURE 10: Analysis of the refolded multimeric species by cross-linking. Panel A shows samples of refolded proteins run on a 7.5% SDS gel. The lanes are as follows: (1), WT; (2) W48Q; (3) A108V; (4) G232D; (5) T294I; (6) F353L. The treatment of the samples is indicated above the gel. The concentrations of protein, cross-linking reagent, and βME used are outlined in the Materials and Methods. The molecular weight (MW) of the SDS gel markers in kDa is given on the right side of panel A. Panel B shows the samples run on a 7.5% native gel. The lanes are as follows: (1) WT; (2) W48Q; (3) A108V; (4) G232D; (5) T294I; (6) F353L. The treatment of the refolded proteins is indicated above the gel and the positions of monomer, dimer, trimer, and oligomers are indicated on the right side of panel B. The protein concentration in the folding reactions was 100 μg/mL. Each lane was loaded with 10 μg of protein. The gels were silver stained.

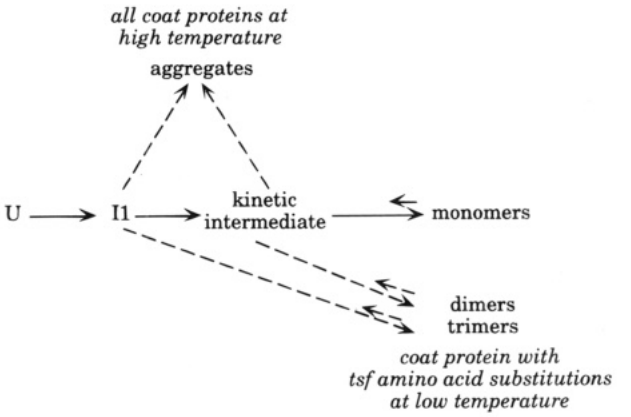
and trimers in the refolding reactions was that they were simply an early species on the aggregation pathway. To determine if this was the case, a temperature shift experiment was done as described in the Materials and Methods. GuHCl was used as a denaturant in this experiment since the transition between soluble species and aggregates (Figure 3) is below the melting temperature (~39 °C) of the WT and mutant proteins (Galisteo *et al.*, 1995). Thus, we could determine if the dimers and trimers had a propensity to aggregate independent of melting of the proteins. WT and T294I that had been refolded from GuHCl overnight at 100 μg/mL were shifted to 28 °C, the approximate midpoint of the productive folding transitions. Samples were taken with time and run on a native gel (Figure 12). Dimers and trimers of T294I showed no change in gel pattern with time at 28 °C. WT monomers also did not aggregate with time. These data indicated that the dimers and trimers were not just species on the aggregation pathway and that aggregates of coat protein come from folding intermediates rather than from these folded species.

DISCUSSION

Purified WT coat protein and coat proteins with *tsf* amino acid substitutions folded efficiently from the denatured state into soluble species *in vitro*. The folding pathways of the WT and mutant proteins were quite similar but the end results were different; WT coat polypeptide chains folded into assembly-competent monomers whereas the mutant polypeptides folded into dimers and trimers that were assembly-incompetent. The relaxation times of folding for the WT and mutant proteins from urea were within experimental error of each other, as were the tryptophan emission spectra. Folding in bisANS also appeared the same; the kinetic intermediates bound more bisANS than the folded proteins. The folded state of the *tsf* mutant polypeptides was less structured by CD compared to WT coat protein. The refolding of the mutant proteins had a similar temperature dependence of aggregation as the WT protein. This was seen whether refolding from either GuHCl or urea. Refolding from urea increased the range where folding was efficient compared to refolding from GuHCl, probably due to GuHCl destabilization of the intermediate that undergoes aggregation. Once folded, mutant dimers and trimers were not prone to further temperature-induced aggregation.

The in Vitro Defect in the Folding of the tsf Coat Proteins. The sharp transition seen in the folding efficiency with temperature suggests that two processes are occurring: folding into soluble species, whether they be monomer, dimer, or trimer, at low temperatures and aggregation at higher temperatures. This indicates that there is an intermediate that exhibits a thermal destabilization. Whether this intermediate is the same one that leads to monomers in the WT protein or dimers and trimers with mutant proteins is not clear. Since we observe dimers and trimers in low abundance in refolded WT protein, the path leading to dimers and trimers must exist in the folding of WT coat protein and have a thermodynamic barrier low enough to occasionally allow folding intermediates to cross. This implies that the *tsf* amino acid substitutions may decrease this barrier and since the refolded mutant subunits are slightly more stable than WT monomers suggests that the folding path is under thermodynamic control.

We believe that these data are consistent with the model presented below.



In vitro, the folding of WT coat protein from the unfolded state (U) has been shown to go through a least two intermediates: the long-lived kinetic intermediate, and at least one intermediate (I1) prior to the kinetic intermediate since the secondary structure was already established in the kinetic intermediate (Teschke & King, 1993). Here we have demonstrated that aggregation of coat polypeptides must occur through interactions of a thermally sensitive folding intermediate. We are not yet certain when dimer and trimer interactions occur during folding.

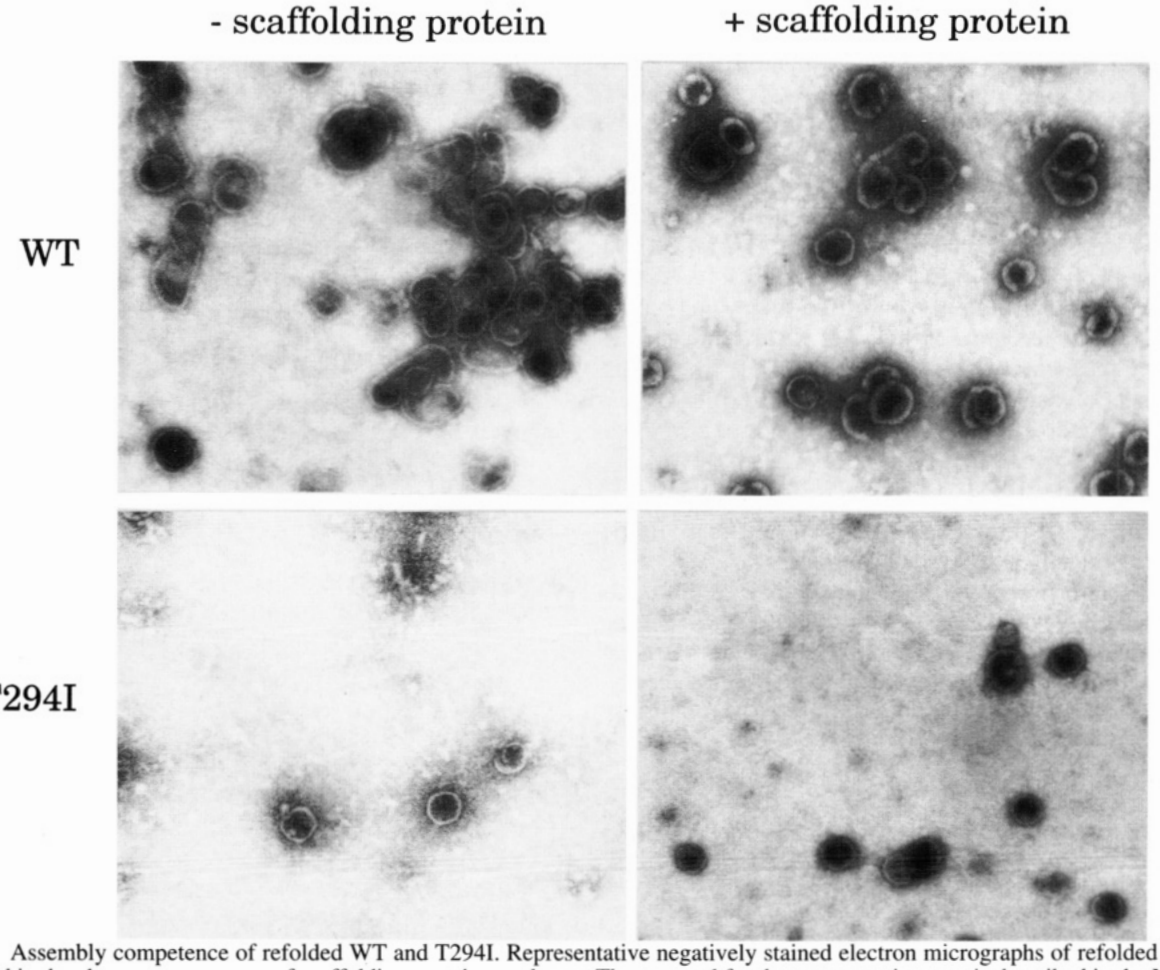


FIGURE 11: Assembly competence of refolded WT and T294I. Representative negatively stained electron micrographs of refolded proteins concentrated in the absence or presence of scaffolding protein are shown. The protocol for the concentration step is described in the Materials and Methods. The final concentration of WT coat protein was ~3 mg/mL and T294I was ~5 mg/mL. T294I was diluted after concentration to have the same concentration as WT coat protein.

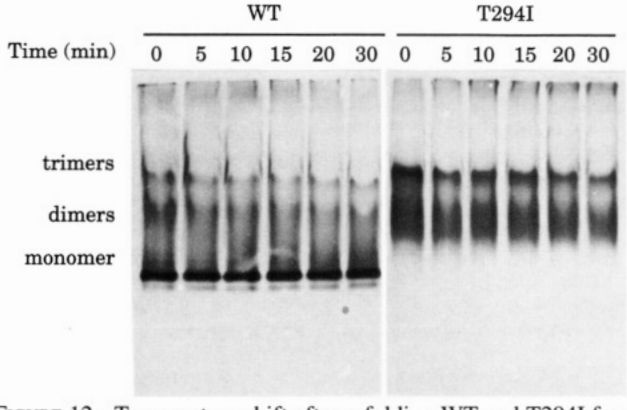


FIGURE 12: Temperature shift after refolding WT and T294I from GuHCl to 28 °C visualized by native gel electrophoresis. WT and T294I were refolded at 4 °C overnight at 100 μg/mL as summarized in the Materials and Methods. The proteins were shifted to 28 °C, and samples taken at the times indicated. The position of monomer, dimer, and trimer bands is shown on the right side of the gel. Each lane was loaded with 10 μg of protein, and the gel was silver stained.

Bennett *et al.* (1994) have recently described a phenomenon they called "domain swapping" in the dimer form of diphtheria toxin. Essentially, one domain of each toxin monomer replaces the identical domain in the other monomer of the dimer. The dimer formation occurs during conditions that might partially open the monomer so that the swapping domain does not tightly interact with the rest of the monomer. They also observe higher order oligomers, including trimers, tetramers, and pentamers, and suggest that the proteins are making linear or cyclized oligomers by interacting with two neighboring monomers. Perhaps this sort of domain swapping is what is occurring during folding of the *tsf* coat proteins.

Bacterial luciferase is a heterodimer comprised of an α and β subunit. If the purified subunits are refolded together, then active $\alpha\beta$ heterodimers are formed (Ziegler *et al.*, 1993). If the β subunit is folded in the absence of the α subunit, then $\beta 2$ homodimers form that are more stable than the heterodimer (Sinclair *et al.*, 1994). This is an example of how proteins can attain an alternative native state and, in the case of bacterial luciferase, in the absence of amino acid alterations.

The Potential Role of GroEL in Folding Coat Polypeptides in Vivo. *In vitro*, the folding of the *tsf* mutant coat proteins is always nonproductive in that they are assembly-incompetent independent of the temperature of folding. *In vivo*, the coat protein mutants are assembly-competent until the restrictive temperature (Gordon & King, 1993). Overproduction of GroEL and GroES rescue the folding of the *tsf* mutants at the restrictive temperature (Gordon *et al.*, 1994). One possibility for the difference between the temperature dependence of proper folding between *in vivo* and *in vitro* is that *in vivo* the mutant polypeptides require GroEL and GroES to fold at all temperatures and therefore bypass the formation of dimers and trimers seen *in vitro*. At the *ts* temperature, the folding would become too defective for the chromosomal level of GroEL and GroES to be sufficient for

proper folding, leading to inclusion bodies. Overproduction of GroEL and GroES at this temperature would rescue the folding. WT P22 grows on host cells with GroEL or GroES point mutations suggesting that the folding pathway of WT coat protein may not require GroEL and GroES (C. Teschke, unpublished results).

Current models for the function of GroEL and GroES suggest that GroEL recognizes early folding intermediates (Martin *et al.*, 1994; Todd *et al.*, 1994). Therefore, we believe it likely that GroEL recognizes such a folding intermediate of coat protein. It will be interesting to see which species are recognized by GroEL *in vitro* and whether assembly-competent species can be obtained by folding mutant coat polypeptides *in vitro* in the presence of GroEL and GroES.

ACKNOWLEDGMENT

We thank Peter Prevelige, Anna Mitraki, and Barrie Greene for critically reading the manuscript. We appreciate the generous gift from Peter Prevelige of some of the shells used in this study. We are grateful to Maria Luisa Galisteo for communicating unpublished results and to Margaret Speed for finding the best native gel to use for this study.

REFERENCES

Andrews, A. T. (1964) in *Electrophoresis: Theory, techniques and Biochemical and Clinical Application* (Peacocke, A. R., & Harrington, W. F., Eds.) pp 5−78, Oxford University Press, New York.

Anfinsen, C. B. (1973) *Science 181*, 223−230.

Beasty, A. M., Hurle, M., Manz, J. T., Stackhouse, T., & Matthews, C. R. (1987) *Protein Eng. 1*, 91−102.

Bennet, M. J., Choe, S., & Eisenberg, D. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 3127−3131.

Cheng, R. H., Olson, N. H., & Baker, T. S. (1992) *Virology 186*, 655−668.

Creighton, T. E. (1990) *Biochem. J. 270*, 1−16.

Earnshaw, W., & King, J. (1978) *J. Mol. Biol. l26*, 721-747.

Engel, J., & Prockop, D. J. (1991) *Annu. Rev. Biophys. Biophys. Chem. 20*, 137−152.

Eppler, K., Wykoff, E., Goates, J., Parr, R., & Casjens, S. (1991) *Virology 183*, 519−538.

Fogel, D., Teschke, C. M., Prevelige, P. E., Jr., & Silva, J. L. (1995) *Biochemistry 34*, 1120−1126.

Fuller, M. T., & King, J. (1981) *Virology 112*, 529−547.

Galisteo, M. L., Gordon, C. L., & King, J. (1995) *J. Biol. Chem.* (in press).

Gao, Y., Thomas, J. O., Chow, R. L., Lee, G. H., & Cowan, N. J. (1992) *Cell 69*,1043−1050.

Goldenberg, D., & King, J. (1981) *J. Mol. Biol. 145*, 633−651.

Goldenberg, D. P., Smith, D. H., & King, J. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 7060−7064.

Golmohammadi, R., Valegard, K., Fridborg, K., & Liljas, L. (1993) *J. Mol. Biol. 234*, 620−639.

Gordon, C. L., & King, J. (1993) *J. Biol. Chem. 268*, 9358−9368.

Gordon, C. L., & King, J. (1994) *Genetics 136*, 427−438.

Gordon, C. L., Sather, S. K., Casjens, S., & King, J. (1994) *J. Biol. Chem. 269*, 27941−27951.

Haase-Pettingell, C., & King J., (1988) *J. Biol. Chem. 263*, 4977−4983.

Harrington, W. F., & Rao, N. V. (1970) *Biochemistry 9*, 3714−3724.

Hendrick, J. P., & Hartl, F.-U. (1993) *Annu. Rev. Biochem. 62*, 349−384.

Johnson, W. C. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 145−166.

Kim, P. S., & Baldwin, R. L. (1982) *Annu. Rev. Biochem. 51*, 459−489.

Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.

Lecomte, J. T. J., & Matthews, C. R. (1993) *Protein Eng. 6*, 1−10.

Liddington, R. C., Yan, Y., Moulsi, J., Sahli, R., Benjamin, T. L., & Harrison, S. C. (1992) *Science 354*, 278−284.

Martin, J., Mayhew, M., Langer, T., & Hartl, F.-U. (1993) *Nature 366*, 228−233.

Matthews, C. R. (1993) *Annu. Rev. Biochem. 62*, 653−683.

Mitraki, A., & King, J. (1989) *Bio/Technology 7*, 690−697.

Mitraki, A., Haase-Pettingell, C. & King, J. (1991) in: *Protein Refolding* (Georgiou, G., & de Bernardez, E., Eds.) American Chemical Society, Washington, D.C.

Mitraki, A., Danner, M., King, J., & Seckler, R. (1993) *J. Biol. Chem. 268*, 20071−20075.

Prasad, B. V. V., Prevelige P. E., Marieta, E., Chen, R. O., Thomas, D., King, J., & Chiu, W. (1993) *J. Mol. Biol. 231*, 65−74.

Prevelige, P., Thomas, D., & King, J. (1988) *J. Mol. Biol. 202*, 743−757.

Prevelige, P. E., Thomas D., & King, J. (1993) *Biophys. J. 64*, 824−835.

Rabilloud, T., Carpentier, G., & Tarroux, P. (1988) *Electrophoresis 9*, 288−291.

Rossman, M. G., & Erickson, J. W. (1985) in *Virus Structure and Assembly* (Casjens, S., Ed.) pp 29−74, Jones and Bartlett Publishers, Inc., Boston, MA.

Sinclair, J. F., Ziegler, M. M., & Baldwin, T. O. (1994) *Struct. Biol. 1*, 320−326.

Sturtevant, J., Yu, M.-h., Haase-Pettingell, C., & King, J. (1989) *J. Biol. Chem. 264*, 10693−10698.

Teschke, C. M., & King, J. (1993) *Biochemistry 32*, 10839−10847.

Teschke, C. M., King, J., & Prevelige, P. E., Jr. (1993) *Biochemstry 32*, 10658−10665.

Todd, M. J., Viitanen, P. V. & Lorimer, G. H. (1994) *Science 265*, 659−666.

Valegard, K., Liljas, L., Fridborg, K., & Unge, T. (1990) *Nature 345*, 36−41.

Yaffe, M. B., Farr, G. W., Miklos, D., Horwich,A. L., Sternlicht, M. L., & Sternlicht, H. (1992) Nature 358, 245−252.

Ziegler, M. M., Goldberg, M.-E., Chaffotte, A. F., & Baldwin, T. O. (1993) *J. Biol. Chem. 268*, 10760−10765.

BI942265U