Hendrickson, W. A., & Konnert, J. H. (1979) in *Biomolecular Structure, Conformation, Function, and Evolution* (Srinivasan, R., Ed.) Vol. 1, pp 43-57, Pergamon, New York.
Henry, E. R., Sommer, J. H., Hofrichter, J., & Eaton, W. A. (1983) *J. Mol. Biol. 166*, 443-451.
Johnson, K. A., Olson, J. S., & Phillips, G. N. (1989) *J. Mol. Biol. 207*, 459-463.
Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268-272.
Jongeward, K. A., Magde, D., Taube, D. J., Marsters, J. C., Traylor, T. G., & Sharma, V. S. (1988) *J. Am. Chem. Soc. 110*, 380-387.
Kabsch, W. (1988) *J. Appl. Crystallogr. 21*, 67-71.
Kottalam, J., & Case, D. A. (1988) *J. Am. Chem. Soc. 110*, 7690-7697.
Kuriyan, J., Wilz, S., Karplus, M., & Petsko, G. A. (1986) *J. Mol. Biol. 192*, 133-154.
Mathews, A. J., Rohlfs, R. J., Olson, J. S., Tame, J., Renaud, J.-P., & Nagai, K. (1989) *J. Biol. Chem. 264*, 16573-16583.
Mims, M. P., Porras, A. G., Olson, J. S., Noble, R. W., & Peterson, J. A. (1983) *J. Biol. Chem. 258*, 14219-14232.
Nagai, K., & Thøgersen, H.-C. (1984) *Nature 309*, 810-812.
Perutz, M. F., & Matthews, F. S. (1966) *J. Mol. Biol. 21*, 199-202.
Petrich, J. W., Poyart, C., & Martin, J. L. (1988) *Biochemistry 27*, 4049-4060.
Phillips, S. E. V. (1980) *J. Mol. Biol. 142*, 531-554.
Phillips, S. E. V. (1981) *The X-ray Structure of Deoxy-Mb (pH 8.5) at 1.4-Å Resolution*, Brookhaven Data Bank.
Phillips, S. E. V., & Schoenborn, B. P. (1981) *Nature 292*, 81-82.
Remaut, E., Stanssens, P., & Fiers, W. (1981) *Gene 15*, 81-93.
Ringe, D., Petsko, G. A., Kerr, D. E., & Ortiz de Montellano, P. R. (1984) *Biochemistry 23*, 2-4.
Rohlfs, R. J., Olson, J. S., & Gibson, Q. H. (1988) *J. Biol. Chem. 263*, 1803-1813.
Rohlfs, R. J., Mathews, A. J., Carver, T. E., Olson, J. S., Springer, B. A., Egeberg, K. D., & Sligar, S. G. (1990) *J. Biol. Chem. 265*, 3168-3176.
Shanaan, B. (1983) *J. Mol. Biol. 171*, 31-59.
Smerdon, S. J., Oldfield, T. J., Dodson, E. J., Dodson, G. G., Hubbard, R. E., & Wilkinson, A. J. (1990) *Acta Crystallogr. B46*, 370-377.
Springer, B. A., & Sligar, S. G. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8961-8965.
Springer, B. A., Egeberg, K. D., Sligar, S. G., Rohlfs, R. J., Mathews, A. J., & Olson, J. S. (1989) *J. Biol. Chem. 264*, 3057-3060.
Traylor, T. G., Koga, N., & Deardruff, L. A. (1985) *J. Am. Chem. Soc. 107*, 6504-6510.
Varadarajan, R., Szabo, A., & Boxer, S. G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 5681-5684.
Zoller, M. J., & Smith, M. (1983) *DNA 3*, 479-488.

# Thermal Unfolding Pathway for the Thermostable P22 Tailspike Endorhamnosidase†

Bao-lu Chen and Jonathan King*
*Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*
*Received August 23, 1990; Revised Manuscript Received April 1, 1991*

ABSTRACT: The conditions in which protein stability is biologically or industrially relevant frequently differ from those in which reversible denaturation is studied. The trimeric tailspike endorhamnosidase of phage P22 is a viral structural protein which exhibits high stability to heat, proteases, and detergents under a range of environmental conditions. Its intracellular folding pathway includes monomeric and trimeric folding intermediates and has been the subject of detailed genetic analysis. To understand the basis of tailspike thermostability, we have examined the kinetics of thermal and detergent unfolding. During thermal unfolding of the tailspike, a metastable unfolding intermediate accumulates which can be trapped in the cold or in the presence of SDS. This species is still trimeric, but has lost the ability to bind to virus capsids and, unlike the native trimer, is partially susceptible to protease digestion. Its N-terminal regions, containing about 110 residues, are unfolded whereas the central regions and the C-termini of the polypeptide chains are still in the folded state. Thus, the initiation step in thermal denaturation is the unfolding of the N-termini, but melting of the intermediate represents a second kinetic barrier in the denaturation process. This two-step unfolding is unusually slow at elevated temperature; for instance, in 2% SDS at 65 °C, the unfolding rate constant is $1.1 \times 10^{-3}$ $s^{-1}$ for the transition from the native to the unfolding intermediate and $4.0 \times 10^{-5}$ $s^{-1}$ for the transition from the intermediate to the unfolded chains. The sequential unfolding pathway explains the insensitivity of the apparent $T_m$ to the presence of temperature-sensitive folding mutations [Sturtevant, J. M., Yu, M.-H., Haase-Pettingell, C., & King, J. (1989) *J. Biol. Chem. 264*, 10693-10698] which are located in the central region of the chain. The metastable unfolding intermediate has not been detected in the forward folding pathway occurring at lower temperatures. The early stage of the high-temperature thermal unfolding pathway is not the reverse of the late stage of the low-temperature folding pathway.

Efforts to understand the stability of the native states of proteins have concentrated on small globular proteins which display reversible unfolding/refolding transitions in denaturant solutions (Schellman, 1987; Goldenberg, 1988; Alber, 1989; Pace, 1990). Under more realistic conditions, for example, under physiological conditions, denaturation is often rapidly followed by aggregation and is not reversible (Wetzel et al., 1990; Volkin & Klibanov, 1989; Mitraki & King, 1989).

† This work was supported by NIH Grant GM17980 and NSF Engineering Research Center Initiative Grant ECD 8803014 to the MIT Biotechnology Process Engineering Center.
* To whom correspondence should be addressed.

0006-2960/91/0430-6260$02.50/0 

Denaturation of larger oligomeric proteins shows additional complexity (Jaenicke, 1987; Edwards et al., 1990). In oligomeric complexes, contributions to stability from the subunit interfaces may be equal or greater than the contributions of the intrachain interactions. We have been interested in the thermostability of complex structural proteins, in particular the thermostable P22 tailspike endorhamnosidase, the host cell recognition and attachment organelle. Unlike intracellular enzymes, the stabilities of such proteins presumably represent the outcome of selection for stability in the external environment.

The tailspike is a structural protein of *Salmonella* phage P22 encoded by gene *9* (Botstein et al., 1973). The mature tailspike protein contains 3 identical polypeptide chains of 666 residues (Sauer et al., 1982; Goldenberg et al., 1982). Six trimeric tailspikes assemble onto the neck of the capsid to form the cell attachment organelle (Israel et al., 1967). The mature tailspike has an endorhamnosidase activity which hydrolyzes the O-antigen polysaccharide of the host cell *Salmonella* (Iwashita & Kanegasaki, 1976).

The intracellular folding pathway of the tailspike protein has been characterized by trapping in vivo folding intermediates in the cold (Goldenberg & King, 1982) and by the use of temperature-sensitive folding (*tsf*) mutants which block certain steps in the folding pathway (Goldenberg et al., 1983a; Yu & King, 1984). The newly synthesized chain first folds to partially folded single-chain intermediates (Goldenberg et al., 1983a,b; Haase-Pettingell & King, 1988), and these partially folded monomers than associate into a trimeric intermediate called the protrimer (Goldenberg & King, 1982). Finally, the protrimer folds to the native form.

After denaturation by urea, the tailspike chains can be refolded back into the native thermostable trimer in vitro (Seckler et al., 1989). The in vitro refolding reaction appears to proceed through intermediates similar to those in the intracellular folding pathway (R. Seckler and A. Fuchs, personal communication).

Although the native protein is resistant to heat denaturation up to temperatures above 80 °C (Sturtevant et al., 1989), the folding process is very sensitive to high temperature (Goldenberg & King, 1981). Maturation of the wild-type protein at the 39 °C restrictive temperature rather than at the 30 °C permissive temperature lowers the yield of correctly folded and assembled protein to 25%, although the biosynthesis of the polypeptide chain is the same at both temperatures (Goldenberg et al., 1982).

Gene *9* encoding the tailspike has been the subject of extensive genetic investigations (Smith et al., 1980; Berget & Poteete, 1980; Fane & King, 1987, 1991; Villafane & King, 1988; Berget & Chidambaram, 1989; Schwarz & Berget, 1989a,b; Maurides et al., 1990). There are over 30 sites of temperature-sensitive folding mutations clustered in the central region of the tailspike polypeptide chain (Smith et al., 1980; Villafane & King, 1988). When released from the ribosome at the restrictive temperature, these *tsf* mutations block the folding before the formation of the protrimer (Goldenberg et al., 1983a). However, when these mutant polypeptide chains mature at 30 °C, they form stable and functional proteins and retain their native chain conformation at elevated temperatures. With some exception (Sargent et al., 1988), *tsf* mutant tailspikes have melting temperatures very similar to wild type (Goldenberg & King, 1981; Sturtevant et al., 1989; Thomas et al., 1990). This is in contrast to the effect of thermolabile temperature-sensitive mutations, for example, in T4 lysozyme, which lower the melting temperature of the native state (Alber, 1989). Mutations affecting the functions and stability of the mature tailspike protein have been isolated and characterized by Schwarz and Berget (1989a,b; Maurides et al., 1990).

The previous studies of tailspike thermostability measured the loss of the native activity or structure (Goldenberg & King, 1981; Sargent et al., 1988; Sturtevant et al., 1989; Thomas et al., 1990). In all cases, thermal denaturation in physiological buffers was not reversible on cooling. In the Raman spectroscopy (Sargent et al., 1988) and calorimetry studies (Sturtevant et al., 1989), the irreversibility was associated with aggregation of the denatured chains. Sargent et al. (1988) studied the denatured aggregated product directly by Raman spectroscopy and found that it contained significant $\beta$-sheet structure.

To further elucidate the nature of the highly thermostable wild-type and *tsf* mutant tailspike proteins, we have initiated a kinetic analysis of the thermal unfolding processes. The results suggest that tailspikes undergo a sequential unfolding transition in which the initiation step is the denaturation of a particular region of the native protein, while the loss of reversibility follows from the further melting of a metastable unfolding intermediate.

## Materials and Methods

*Proteins, Reagents, and Media.* Tailspike protein was purified as described previously by King and Yu (1986). Trypsin (EC 3.4.21.4) from bovine pancreas (TPCK treated), protease from *Staphylococcus aureus* strain V8 type XVII-B (V8 protease), phosphorylase *b* (EC 2.4.1.1) from rabbit muscle, and L-cysteine amino acids were purchased from Sigma Chemical Co. All other chemicals used were either electrophoresis pure for SDS–PAGE or reagent grade.

LB broth, LB bottom agar, and soft agar were used for routine propagation of host cells and phage (Fane & King, 1991). 1XM9 solution contained 0.6% $Na_2HPO_4$, 0.3% $KH_2PO_4$, 0.05% NaCl, and 0.1% $NH_4Cl$. Dilution fluid contained 0.1% tryptone and 0.7% NaCl.

*Phage and Bacteria Strains.* All bacterial strains were derivatives of *Salmonella typhimurium* LT2. The suppressor minus host DB7136 (*leu A414-am, his C525-am*) and its suppressor plus derivative DB7155 (*leu A414-am, his C525-am, supE*) have been described elsewhere (Winston et al., 1979).

Phage capsids were prepared by infecting suppressor minus host cells DB7136 with phage carrying a mutation in the *c1* gene (*c1-7*) to ensure entry into the lytic cycle, an amber mutation in gene *13* (*amH101*) to prevent spontaneous lysis of the infected cells, and an amber *E1017* mutation in gene *9* (Berget & Poteete, 1980). Phage was stored in 1XM9 solution containing 2 mM $MgSO_4$.

*Thermal Denaturation Procedures.* Protein samples for thermal denaturation were usually prepared in 50 mM Tris/1.7 mM 2-mercaptoethanol with or without 2% SDS. Buffer pH was adjusted with 1 N HCl and measured by an Orion Research Model 601A/digital Ionalyzer with a Ross pH electrode.

Thermal denaturation was carried out in a water bath with samples arranged next to a thermometer with 0.2 °C accuracy. Following incubation at high temperature, samples were withdrawn and mixed with an equal volume of SDS sample buffer. The SDS sample buffer contained 2% SDS, 62.5 mM Tris, 2.14 mM 2-mercaptoethanol, 10% glycerol, and 0.012% Bromophenol blue dye and was adjusted to pH 7. This mixture was placed on ice to quench the unfolding reaction and later analyzed by SDS–PAGE with a 7.5% acrylamide slab gel. Electrophoresis was carried out at 150-V constant voltage for

about 3 h at room temperature. Then the gels were stained with Coomassie blue. The bands were scanned to quantitate their intensities. The percentages of the intensities among the N, I, and M species were used to describe the kinetics of these three species upon unfolding. The kinetic rate constants were calculated according to eq 2–4 under Results by an IBM PC computer with the Lotus *123* program.

*Gel Electrophoresis and Scanning.* SDS–polyacrylamide gel electrophoresis (SDS–PAGE) was carried out according to Laemmli (Laemmli, 1970; King & Laemmli, 1971), but with a slab gel of 0.75-mm thickness. The gels were stained with Coomassie blue and scanned by an LKB 2202 Ultroscan densitometer equipped with an LKB 2220 recording integrator for obtaining the band intensity of interest.

2-D gel electrophoresis was used to trace the composition of the limited proteolysis product. Protease-treated samples were applied onto 7.5% acrylamide cylindrical rod SDS gels (3 mm in diameter) as the first dimension. After the electrophoresis was completed, gels were heated in a boiling water bath for 3 min to dissociate native tailspikes and the limited proteolytic products. Then these gels were laid onto the top of 10% acrylamide slab SDS gels (0.75 mm in thickness) as the second dimension, the electrophoresis in the second dimension was carried out.

*Sucrose Gradient Centrifugation.* Sucrose gradients (5–20%) were made up in the buffer used for thermal denaturation studies. Centrifugation was performed with a SW50.1 rotor in a Beckman L8-55M ultracentrifuge. After the sedimentation, the liquid columns were fractionated from a pinhole at the bottom of the tube, and a total of 20 fractions were collected for each gradient. These fractions were analyzed by SDS–PAGE with 7.5% acrylamide gels.

*Cysteine Determination.* Three samples with tailspike protein concentration at about 1.0 mg/mL were made in 40 mM $Na_2HPO_4$, 1 mM EDTA, and 2% SDS at pH 8. Protein concentration was measured by the OD at 278 nm with a Gilford Response spectrophotometer using the molar extinction coefficient $7.3 \times 10^4$ $mol^{-1}$ $cm^{-1}$ L (Sauer et al., 1982). After thermal denaturation, these samples were immediately cooled down to room temperature. The relative percentages of the N, I, and M species in these samples were analyzed by SDS–PAGE. The reactive cysteine residues in these samples were detected by the Ellman's reaction (Ellman, 1959). These samples were mixed with 4 mg/mL 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB) made in 0.1 M $Na_2HPO_4$ in a ratio of 10:1 (v/v) in a 1-cm path-length cuvette. The absorbance at 412 nm of these reaction mixtures was compared with that of freshly prepared free L-cysteine solutions of known concentration to calculate the number of the reactive cysteines.

*Limited Digestion by Trypsin and V8 Protease.* Fresh trypsin solutions were prepared in the experimental buffer just before each proteolytic experiment; 1 mg/mL V8 protease was prepared in the experimental buffer and stored at 4 °C.

About 1 mg/mL tailspike was prepared in 50 mM Tris/1.7 mM 2-mercaptoethanol at pH 9.5. Thermal denaturation was done at 80 °C to generate the N, I, and M species. The proteolytic reactions were carried out by incubating these samples with the proteases at room temperature. The resulting products were examined by SDS–PAGE.

In order to find the N-terminal sequence of the limited proteolytic products of the intermediate species, the SDS–polyacrylamide gels containing trypsin and V8 protease cleavage products were first electroblotted onto poly(vinylidene difluoride) membranes (PVDF Immobilon Transfer membranes from Millipore) (Matsudaira, 1987). Then the bands of interest were cut out and subjected to automated Edman degradation analysis.

*Tailspike Binding to Phage Capsids.* Native tailspikes bind noncovalently but irreversibly to phage capsids (Israel et al., 1967). For the tailing reaction, 0.2 mL of the diluted tailspike samples and 0.2 mL of $2 \times 10^9$ capsids/mL capsid solution were mixed, and the mixture was incubated at 30 °C for 30 min to allow the reaction to complete. Then the reaction mixture was diluted and plated on the suppressor plus strain DB7155 at 30 °C. The plaque forming units (PFU) on the plates were counted to determine the number of infectious phage formed.

To assay the unfolding intermediate, 0.4 mg/mL tailspike solutions in 50 mM Tris buffer at pH between 7 and 7.5 were prepared with or without 2% SDS. Thermal treatment of these samples was done either at 80 °C for the samples without SDS or at 70 °C for samples with 2% SDS. The reactions were stopped by chilling on ice. A dilution to 0.1 mg/mL tailspike was first made with the buffer for the thermal denaturation. Then subsequent serial 3-fold dilutions of these samples were carried out. These dilutions were later mixed with the capsid solution for the tailing reaction.

## Results

*Identification of the Unfolding Intermediate.* The tailspike is resistant not only to heat but also to denaturation by detergent and to proteolysis (Goldenberg et al., 1982). The native tailspike protein binds few SDS molecules and thus migrates anomalously slowly during SDS gel electrophoresis (Goldenberg et al., 1982). If the native tailspike is dissociated into its constituent polypeptide chains by heating above its melting temperature, it binds SDS molecules similarly to other denatured polypeptide chains and migrates as expected for its molecular weight (Reynolds & Tanford, 1970; Goldenberg et al., 1982).

In a preliminary investigation of the thermostability of the tailspike protein, Myeong-Hee Yu and J. King (personal communication) observed that during incubation of the native tailspike above 75 °C, a species was generated with a mobility in SDS gels between the native and denatured species. To study the unfolding process more carefully, thermal unfolding was carried out by heating tailspike samples in Tris buffer at 80 °C. At various times, samples were withdrawn and mixed with SDS sample buffer on ice to quench the reaction. These samples were analyzed by SDS–PAGE without further heating of the sample. Under these conditions, native tailspikes remain native (Goldenberg et al., 1982).

A typical thermal unfolding experiment is shown in Figure 1. The sample containing the starting material before heating displayed a single band with the mobility of native tailspike. As the incubation proceeded, the concentration of the native species, "N", decreased and a new species, "I", was observed migrating slightly faster than the native species. This mobility would be expected of a species with higher charge than the native. At an incubation time of about 50 min, the I species became dominant and a small amount of a species "M" corresponding to the SDS–polypeptide chain complex appeared, as well as aggregated chains which failed to enter the gel. This suggested that upon heating at 80 °C, the native form was slowly converted to the I species first, and then the I species dissociated to monomeric species (possibly still containing partially folded structures; see Discussion) which aggregated. The ratios of the N to I and to M species changed very slowly during storage in the cold.

The increased mobility of the intermediate species detected during gel electrophoresis would be expected for a trimeric

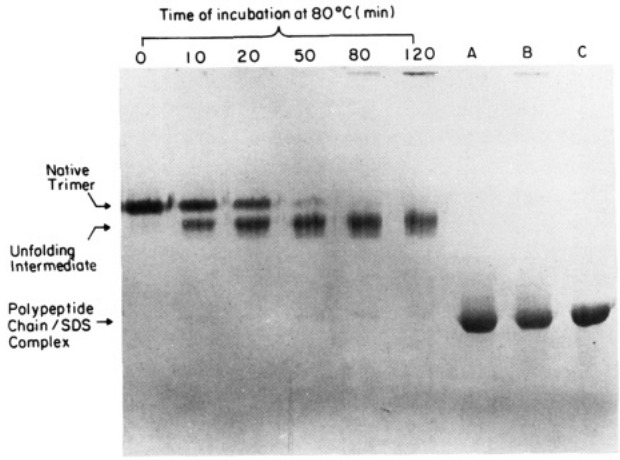


FIGURE 1: SDS–PAGE analysis of the thermal unfolding of tailspike protein. Tailspike (0.4 mg/mL) in Tris buffer at pH 8 was incubated at 80 °C. Samples were withdrawn at times 0, 10, 20, 50, 80, and 120 min, and the unfolding reaction was quenched by mixing it with SDS sample buffer on ice. These samples were electrophoresed through a SDS–polyacrylamide gel to analyze the thermal unfolding products. Lanes A and B show the $t$ = 0- and 120-min samples after 3-min incubation at 100 °C, respectively. Lane C also shows the $t$ = 120-min sample after 3-min incubation at 100 °C but in 0.36 M 2-mercaptoethanol.

species binding more SDS than the fully native state. Stabilization of an unfolding/refolding intermediate by detergent has been described for the prokaryotic enzyme rhodanese in which exposed hydrophobic interdomain surfaces in the intermediate interact with lauryl maltoside (Tandon & Horowitz, 1986; Horowitz & Criscimagna, 1986). However, since SDS was not present during the sample heating, the tailspike intermediate must be relatively long-lived even in the absence of detergent.

If these species represented different conformations of the tailspike chain, rather than cleavage or other covalent modification products, heating to boiling should convert them to homogeneous tailspike chain/SDS complexes. The $t$ = 0- and 120-min samples were heated at 100 °C for 3 min and electrophoresed through lanes A and B, respectively. All N and I species were denatured to M species. The aggregated material in the $t$ = 120 min sample required higher concentrations of 2-mercaptoethanol as shown in lane C, after which almost all the aggregates were converted to the unfolded monomeric chains.

If I was an unfolding intermediate, it might be expected to convert back to the native tailspike. To determine if the I species could refold back to native, a sample was first incubated at 80 °C for 50 min to generate the I species and then was directly transferred into a room temperature water bath and incubated. The refolding was followed at times 1, 4, 9, and 24 h by taking aliquots of the sample for SDS–PAGE analysis. The results are shown in Figure 2. At the beginning of refolding, the sample contained about 15% N species, 80% I species, and 5% M species. After 1 day, about half of the I species was converted back to the native while the M species showed little change. To determine if the refolded species is nativelike, we treated the refolded sample with trypsin in a 1:1 (w/w) ratio of trypsin to tailspike for 30 min at room temperature. The refolded form showed the same resistance to the protease as the intact native protein (lane B in Figure 2). The I species was partially sensitive, yielding a band migrating slightly faster (lanes A and B in Figure 2). This band is designated as "I1" and will be discussed later. Besides the N and I species, the other species, including the aggregates, were cleaved completely by trypsin.

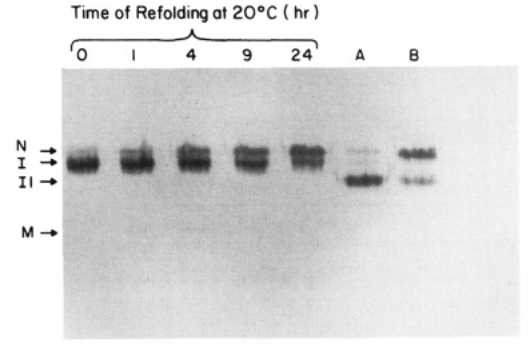


FIGURE 2: Refolding of tailspike protein at room temperature. A tailspike sample previously heated to 80 °C for 50 min was placed in a room temperature water bath for refolding. Samples were withdrawn after 0, 1, 4, 9, and 24 h and mixed with SDS sample buffer to quench the refolding reaction. Then these samples were subjected to SDS–PAGE analysis. The last two lanes (A and B) show the $t$ = 0- and 24-h refolding samples after treatment with trypsin, respectively.

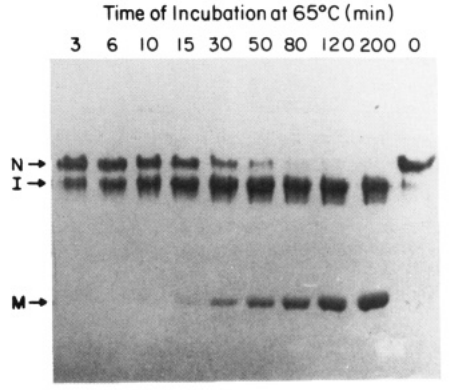


FIGURE 3: Thermal denaturation of tailspike protein in the presence of SDS. Tailspike (0.4 mg/mL) in 2% SDS at pH 8 was incubated at 65 °C. Samples were taken at the times indicated and mixed with SDS sample buffer on ice to quench the reaction. These samples were subjected to SDS–PAGE analysis.

The high temperature required for the thermal denaturation posed the possibility of covalent damage or modification of the side groups to the polypeptide chain (Volkin & Klibanov, 1989). In addition, the aggregation reaction obscured the conversion of the N to I and to M species. To avoid these problems, the quantitative kinetic measurements of the unfolding reaction were carried out in the presence of SDS which lowered the melting temperature as well as accelerated the unfolding reaction.

Samples containing 0.4 mg/mL tailspike protein were prepared in 50 mM Tris and 2% SDS at neutral pH. Figure 3 shows the time course of thermal unfolding of tailspike at 65 °C. The concentration of the native protein decreased rapidly and disappeared at about 60 min. Meanwhile, the concentration of the intermediate species initially increased and reached a maximum at about 50 min and then decreased slowly. The M species appeared at early time and increased till the termination of the experiment, as the end product. It was found that incubation of the samples at temperatures between 60 to 75 °C quantitatively transformed the native protein to the intermediate species and then to the M species with half-times ranging from minutes to hours. No aggregation was observed during denaturation.

Since the kinetic assay was done by first transferring the thermal denaturation samples to the cold and then analyzing them by SDS–PAGE, the effect of SDS on these thermal unfolding products in the cold was examined. A tailspike

Table I: Tailing Assay of the Thermal Unfolding Products of the Tailspike Protein at 30 °C

| sample | time (min) at 70 °C | % N | % I | % M | $SP/SP_0$[a] | % capsid binding activity |
|---|---|---|---|---|---|---|
| A | 0 | 100 | | | 1 | 100 |
| B | 3 | 44 | 56 | | 2.3 | 44 |
| C | 15 | | 94 | 6 | 90 | 1 |
| D | 120 | | 73 | 27 | ~2350[b] | ~0.04[b] |

[a] SP stands for the saturation point, and $SP_0$ is the SP for sample A. [b] These are estimated from extrapolation of the data.

protein sample made in 2% SDS was stored in refrigerator. An aliquot was withdrawn every day and analyzed by SDS–PAGE. It was observed that about half of the native was converted to the intermediate species after 10 days, whereas the conversion from the intermediate to the unfolded monomer was undetected. Thus, the intermediate species generated at high temperature is stable for long periods in 2% SDS in the cold.

The unfolding process can be described by a three-component sequential reaction mechanism:

$$N \xrightarrow{k_1} I \xrightarrow{k_2} M \quad (1)$$

Here, N, I, and M stand for the native, intermediate, and unfolded monomeric species, respectively. $k_1$ and $k_2$ are the two rate constants for the transitions from N to I and from I to M, respectively. The change in the concentration of each species of this irreversible process with time follows

$$c(N) = N_0 \exp(-k_1 t) \quad (2)$$

$$c(I) = N_0 k_1/(k_2 - k_1)[\exp(-k_1 t) - \exp(-k_2 t)] + I_0 \exp(-k_2 t) \quad (3)$$

$$c(M) = N_0 + I_0 - c(N) - c(I) \quad (4)$$

Here $c(N)$, $c(I)$, and $c(M)$ are the relative protein subunit concentrations of N, I, and M species, respectively, at time $t$. $N_0$ and $I_0$ are the initial concentrations of N and I, respectively. We have estimated the concentration of each species by scanning the Coomassie blue stained SDS–polyacrylamide gels and calculated the two rate constants by fitting the kinetic curves according to the above relations. The calculated rate constants for the kinetic data in Figure 3 are $1.1 \times 10^{-3}$ s$^{-1}$ for $k_1$ and $4.0 \times 10^{-5}$ s$^{-1}$ for $k_2$, respectively. Simulated kinetic curves based on these two rate constants are drawn in Figure 4 as the continuous lines through the data points. These two rate constants were not affected by doubling the protein concentration in the sample.

*Characterizing the Thermal Unfolding Intermediate. (A) Attachment to the Capsid (Tailing Ability).* During phage morphogenesis, tailspike binds to the virion as the last structural protein to form a baseplate structure which serves as the cell attachment apparatus. The binding of the tailspike protein to the capsid proceeds efficiently in vitro and was found to be an irreversible and noncovalent process (Israel et al., 1967). To determine if partial unfolding affects this biological function, we have examined the capsid binding activity of the thermal unfolding intermediate.

Four identical tailspike protein samples were made in 50 mM Tris, pH 7.5, and 2% SDS for the tailing assay. These samples were subjected to thermal unfolding at 70 °C for various amount of time to generate different ratios among the N, I, and M species (Table I). The tailing reaction was done at 30 °C by mixing these samples with the purified phage capsids as described under Materials and Methods. The results are presented in Figure 5 as log–log plots for the values of the plaque forming unit (PFU) versus the total tailspike concentration. Sample A, which contained all native protein, showed

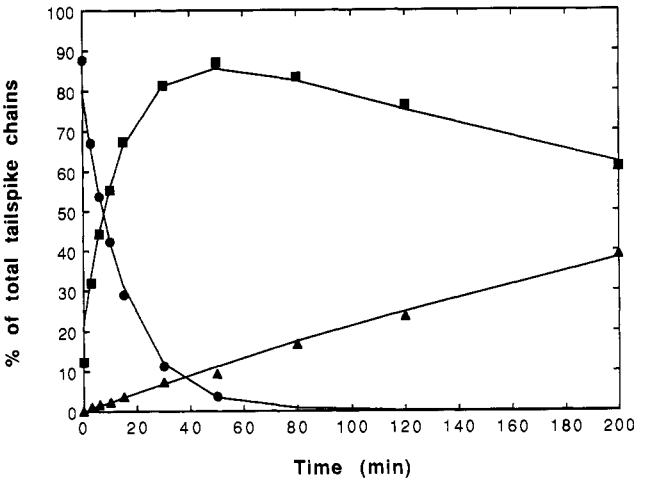


FIGURE 4: Kinetics of the thermal unfolding of tailspike protein in 2% SDS at 65 °C. The data points are the relative intensity of the N (●), I (■), and M (▲) stained bands on the SDS–polyacrylamide gel shown in Figure 3. The continuous lines were calculated according to eq 2–4 with $k_1 = 1.1 \times 10^{-3}$ s$^{-1}$ and $k_2 = 4.0 \times 10^{-5}$ s$^{-1}$.

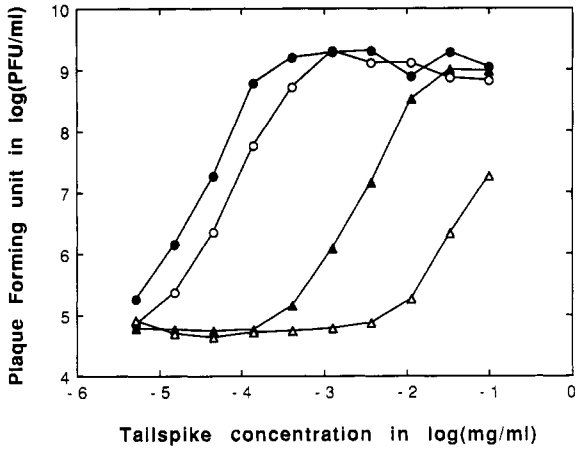


FIGURE 5: Tailing assay of thermal unfolding samples. Four tailspike samples, A (●), B (○), C (▲), and D (△), were subjected to thermal unfolding at 70 °C. The relative concentrations of the N, I, and M species for these samples are listed in Table I. Serial 3-fold dilutions of these samples were incubated with phage capsids at a concentration of $1 \times 10^9$ capsids/mL at 30 °C for 30 min. After the incubation, the reaction mixture was diluted and plated to determine the titer of viable phages formed in the reaction.

normal capsid attachment in the thermal denaturation buffer in the presence of SDS as reported previously (Israel et al., 1967; Berget & Poteete, 1980). The tailing curve for sample A in Figure 5 displays the typical two-phase curve with the plateau determined by the input concentration of phage capsids, and the saturation point (SP) determined by the concentration of active tailspikes. At tailspike concentrations above the saturation point, there is no detectable increase in phage titer. At tailspike concentrations below the saturation point, the phage titer increases with the cube of the tailspike concentration, since a minimum of three tailspikes are needed to form an infectious phage (Israel et al., 1967).

Samples B–D contained mixtures of the N, I, and M species. The basic features of the tailing curves for these samples are not changed. The slopes of all these curves below their saturation points are close to 3, and the final plateaus reached are similar to that for sample A. The only difference in these tailing curves is that their saturation points are shifted to higher concentrations of protein. This indicates a lowering in the capsid binding activity in these samples. The estimated saturation points and the capsid binding activities relative to sample A for these samples are presented in Table I. The capsid binding activity of each sample correlated with the remaining native species in the sample. Thus, the thermal unfolding intermediate is defective in binding to the phage capsid.

There is a possibility that the intermediate species binds to the capsid but makes noninfectious phage due to a defect in that function. If the I species is able to bind to the capsid but is defective in cell attachment and injection, the binding of the I species would lower the population of the infectious phage and decrease the final plateau PFU values for these samples. Since this was not observed in this case, we concluded that the defect is in the binding reaction.

If the I species is defective in capsid binding, added exogenous native tailspike will bind to the capsid and increase the PFU values. Excess native tailspike was added into the above four samples and plated. Sample A displayed no increase in the PFU values for the tailing reaction mixture with tailspike concentration above its saturation point ($SP_0$). However, samples B–D, which contained different concentrations of the I species, showed increases in the PFU values for the tailing reaction mixture with tailspike concentrations between $SP_0$ and their own SP values. Therefore, the intermediate species is unable to attach to the phage capsid.

*(B) Sedimentation Behavior.* If the I species is a partially unfolded intermediate as implied from its altered mobility on the SDS gel, its sedimentation rate should be different from the native form as well as from the totally unfolded monomeric species. Three tailspike protein samples made in 50 mM Tris buffer, 0.71 M 2-mercaptoethanol, and 2% SDS were subjected to different heating treatments to generate samples containing greater than 90% of the N, I, or M species. Then each sample was mixed with protein phosphorylase *b* ($M_r$ 97 000) as a centrifugation marker before being loaded on a 5–20% sucrose gradient containing 2% SDS. The centrifugation was carried out at 45 000 rpm in an SW50.1 rotor for 16 h at 4 °C. After the centrifugation, the gradients were fractionated and analyzed by SDS–PAGE.

The results of sedimentation of these samples are presented in Figure 6. The leading peak (see panel A) represents native tailspike ($M_r$ 216 000, $s_{20,w}$ = 9.3 S; Berget & Poteete, 1980). The M species ($M_r$ 72 000) (see panel C) sedimented slowly, consistent with its unfolded structure. The intermediate species (see panel B) sedimented slightly behind the native and far from the monomeric species. This indicates that the intermediate is a nativelike species and likely a trimer.

*(C) Exposure of the Cysteine Residues.* The tailspike protein contains eight cysteines in each of its polypeptide chains. They are all buried inside when the protein is folded, and no disulfides are found in this protein (Sauer et al., 1982). The eight cysteines are located at positions 169, 267, 287, 290, 458, 496, 613, and 635. These residues span almost the entire central region (residues 150–500 also called the "*tsf* region" for clustering of the *tsf* mutant sites) and most of the C-terminal region (residues 500–666). If thermal unfolding results in exposure of any of these cysteines, the chemical reactivity

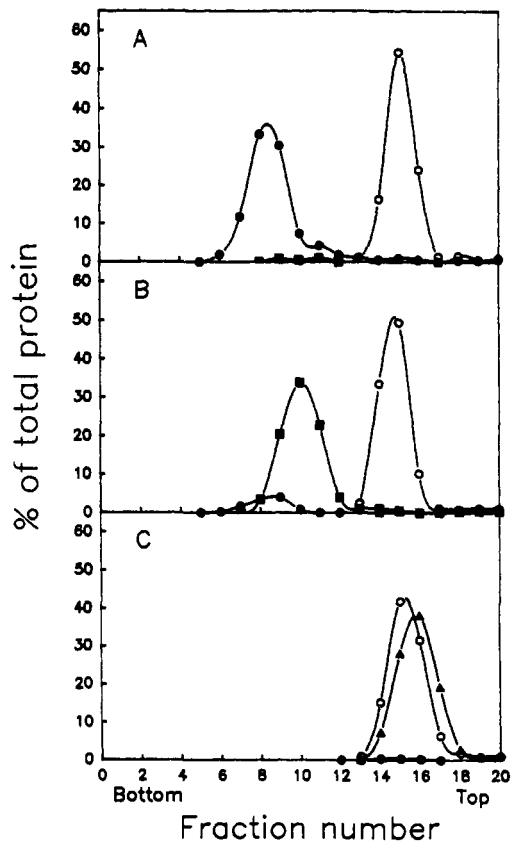


FIGURE 6: Sucrose gradient sedimentation. Three tailspike samples, each containing more than 90% of one of the three species N (A), I (B), or M (C), were mixed individually with the marker protein phosphorylase *b* after thermal denaturation. Two hundred microliters of each sample containing 4.5 mg/mL tailspike and 1 mg/mL phosphorylase *b* was applied to a 4.5-mL 5–20% sucrose gradient. After centrifugation, 20 fractions were collected for each gradient and analyzed by SDS–PAGE. The relative intensity among the three species, N (●), I (■), and M (▲), as well as the marker protein (○) is plotted against the fraction number.

of these residues will be greatly enhanced. Therefore, these cysteines make good structural probes for detecting the unfolded regions in the polypeptide chain of the intermediate.

Exposed cysteine residues were detected by using Ellman's reaction as described under Materials and Methods. There was no exposure of any cysteines after the native protein was incubated with DTNB at room temperature for 1 h. The intermediate species showed less than 0.5 reactive cysteine residue per polypeptide chain, and the monomeric species showed about 8 reactive cysteine residues per polypeptide chain. Thus, at the intermediate stage, all the cysteines are still buried inside the protein. The regions of the polypeptide chain containing these eight cysteines (residues 169–635) apparently remain in the conformation similar to the native form.

*(D) Susceptibility to Proteases.* Trypsin and V8 protease were used to examine the proteolytic susceptibility of the intermediate species. In the proteolytic experiments, we have varied the ratio of protease to tailspike from 1:100 to 1:1 by weight and the digestion time from 1 min to 1 day. Under these conditions, the native tailspike was resistant to cleavage. The unfolded monomeric form was completely cleaved, leaving no bands on the SDS gel. In contrast, the thermal unfolding intermediate was susceptible to limited proteolysis. Figure 7 displays the kinetics of proteolysis of the three tailspike species by trypsin at a ratio of 1:10 (w/w) of trypsin to tailspike. The starting sample shown on the first lane to the left contained about 20% native, 70% intermediate, and 10% unfolded

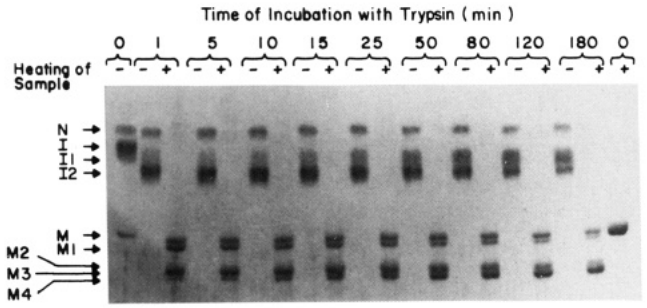


FIGURE 7: Kinetics of the limited digestion of the thermal unfolding intermediate species by trypsin. Thermal denaturation samples were treated with the protease at room temperature. Samples were withdrawn at various times and mixed with the SDS sample buffer. One portion was loaded onto the SDS–polyacrylamide gel directly (labeled as "–" on top of the gel), and the other was incubated at 100 °C for 3 min to dissociate the trimers before the loading (labeled as "+" on top of the gel). The first lane (far left) shows the control sample containing about 20% N, 70% I, and 10% M species without protease. The last lane shows the control sample but after 3-min incubation at 100 °C. Other lanes show the products of the limited proteolysis with or without 3-min incubation at 100 °C as indicated.

tailspike. After incubation at 100 °C, both N and I species were converted to unfolded chain as shown in the last lane.

The limited proteolysis of the intermediate species by trypsin resulted in two bands on the SDS gel. These two bands, designated as "I1" and "I2", migrated between the I and M species. Their slow mobilities indicate that they are still partially nativelike and trimeric species.

Incubation at 100 °C caused these two trimeric species to melt into four bands migrating slightly below the tailspike monomer. These four bands were named "M1", "M2", "M3", and "M4" according to their sequence toward the anode. Since incubation at 100 °C will dissociate the native trimer to the unfolded monomer, presumably M1, M2, M3, and M4 are unfolded polypeptide chains as well. The M2 and M3 bands are not well resolved in Figure 7 but were easily detected if a 1:100 (w/w) ratio of trypsin to tailspike was used for the digestion. M2 is probably a transient species observed only at a fairly low ratio of trypsin to tailspike.

The kinetic results of the proteolysis shown in Figure 7 are not sufficient for the establishment of trimer–monomer relationships. The true relationship between the trimeric species I1 and I2, and their dissociated monomeric species, was readily confirmed after using 2-D gel electrophoresis. The I1 and I2 species were first separated by electrophoresis in a SDS gel in the first dimension and dissociated in situ by placing the gel in boiling water for the second dimension.

In the second-dimension gel, the I1 species yielded two bands which corresponded to species M2 and M4. The I2 species yielded three bands with mobilities of species M1, M3, and M4. Thus, I1 is composed of M2 and M4 species, and I2 is composed of M1, M3, and M4 species.

Although species I1 moves slower than species I2 on the SDS gel, this does not mean that I1 has a higher molecular weight. The kinetics of the proteolysis indicate that species I2 is actually the precursor of species I1 (see Figure 7). Presumably, I2 is not extensively cleaved, contains substantial unfolded structure, and binds a larger number of SDS molecules. By contrast, I1 is more cleaved, contains less unfolded structure, and binds fewer SDS molecules. Thus, I2 migrates more rapidly than I1 on the SDS gels despite its greater molecular weight.

If the ratio of trypsin to tailspike was increased to 1:1 by weight, all the I band or I2 band was converted to I1 after 30 min at room temperature (see also lane A in Figure 2). After incubation at 100 °C, I1 yielded only M4. Both I1 and M4 bands remained constant after further proteolysis for another 2.5 h. Thus, M4 is the final cleavage fragment among all these progressively digested products. The molecular weight estimated from its mobility on an SDS–polyacrylamide gel for the M4 species is about 60 K, indicating a loss of about 110 residues. In order to produce a 60-kDa fragment from a 72-kDa polypeptide, all the cleavage sites must be near the chain ends.

Treatment of the intermediate species by V8 protease produced one band on the SDS gel as the limited proteolytic cleavage product (data not shown). This species, designated as "I3", had mobility between those of species I1 and I2. Upon incubation at 100 °C, the I3 trimer was converted to two major monomer bands above 60 kDa and some minor bands with lower molecular weights.

The N-terminal amino acid sequencing results showed that the I3 band is mainly composed of two polypeptide chains starting from Ser-19 and Asn-52, respectively, and the M4 band is composed of a unique polypeptide chain starting from Tyr-108. Therefore, V8 protease cleaves the I species at the carboxyl-terminal side of either Glu-18 or Glu-51 and gives two monomeric bands, and trypsin cleaves the I species finally at the carboxyl-terminal side of Lys-107. Thus, the cleavage sites for both trypsin and V8 protease are located in the N-terminal region of the 666 amino acid polypeptide chain.

## Discussion

Most of the recent detailed studies on protein stability have been carried out with small globular proteins (Schellman, 1987; Goldenberg, 1988; Alber, 1989; Pace, 1990). Proteins such as the tailspike involved in cell surface recognition and adhesion generally have a different character. The tailspike is an elongated asymmetrical structure, with one end presumably specialized for cell surface recognition and the other for binding to phage heads. Preliminary X-ray diffraction results from tailspike crystals (T. Alber, personal communication) indicate that the chains are arranged symmetrically in the molecule. It seems most likely that each subunit extends the length of the tailspike, as in influenza hemagglutinin (Wilson et al., 1981). In such a structure, the contact interfaces between the subunits would be quite large. Though it is clear from the Raman spectra that the protein is dominated by $\beta$-sheet structure, we do not know the relationship of the sheets to the subunit interfaces (Sargent et al., 1988; Thomas et al., 1990).

*Tailspike Unfolds via a Sequential Pathway.* Upon being heated at high temperature, the N-termini region of the tailspike chains comprising about the first 110 residues slowly unfolds, while the central and C-terminal parts of the chain remain in the native trimeric conformation. In a slower step, this intermediate species, "I", then unfolds further. During thermal denaturation in the absence of detergent, unfolding was accompanied by aggregation, accounting for the loss of reversibility (Sargent et al., 1988). Though direct evidence is lacking, it seems likely that the aggregation proceeds from partially folded monomeric species, which we represent as "M*" (Jaenicke, 1987; Mitraki & King, 1989).

We detected the trimeric unfolding intermediate by its increased mobility in SDS gels as compared to the native state, which does not form a polypeptide chain/SDS complex. The increased mobility of the intermediate in SDS gels is presumably due to SDS binding to the unfolded regions, increasing the charge on the complex.

The unfolding intermediate can also be trapped in the absence of detergent, by low temperature. Under the buffer and pH conditions of our experiment, I refolded back to native with

a half-time of 1 day, and no aggregation from I was observed. Refolding from the monomeric chains was not observed under this condition.

In the absence of SDS, the reaction scheme can be described as

$$\underset{\text{native trimer}}{N} \rightleftharpoons \underset{\text{partially unfolded trimer}}{[I]} \rightarrow \underset{\text{partially folded monomers}}{3[M^*]} \rightarrow \text{aggregates}$$

In the presence of detergent, the SDS molecules presumably block the refolding by tightly binding to the trimeric intermediate species I and to partially unfolded monomeric chains, so that thermal unfolding in detergent is effectively irreversible despite the absence of aggregation. The SDS binding shifts the equilibrium toward the unfolding side, lowering the apparent $T_m$. The aggregation step is blocked probably due to the binding of SDS to the partially folded monomeric species M*. This SDS binding resembles the detergent binding to unfolding intermediates of rhodanese described by Tandon and Horowitz (1986). In their model, the detergent binds to exposed hydrophobic surfaces normally buried between domain interfaces, preventing aggregation.

The unfolding of the intermediate to the unfolded monomer was very slow in the presence of SDS at 4 °C. Thus, in the cold, the intermediate species is kinetically stabilized with respect to both native and fully unfolded chains. Sargent et al. (1988) have reported evidence of a conformational change in the tailspike below 20 °C from their Raman spectrum analysis. The dramatic change at the amide I band may be a reflection of the secondary structural change in protein at low temperatures.

Therefore, SDS plays at least two roles: (1) the binding of SDS stabilizes the intermediate and the unfolded monomer chains, lowering the $T_m$ and blocking refolding; (2) the binding of SDS solubilizes the unfolded chains, inhibiting aggregation.

Thermal denaturation of the tailspike protein in the presence of SDS follows

$$\underset{\text{native trimer}}{N} \rightarrow \underset{\text{partially unfolded trimer}}{I{:}SDS} \rightarrow \underset{\text{unfolded monomer/SDS complex}}{3M{:}SDS}$$

The unfolding rate constants for the above two transitions determined in 2% SDS at 65 °C (Figure 4) were quite slow. This indicates that the energy barriers for unfolding of this protein are high even at elevated temperature. Although the initiation step for denaturation is the unfolding of the N-termini, the rate-limiting step for the whole unfolding process is the unfolding of the central region and C-termini of the polypeptide chains.

Some proteins with independent domains such as the λ repressor (Pabo et al., 1979) display a sequential thermal unfolding pathway in which the less stable domain melts out independently of the more stable domain. However, differential scanning calorimetry of the tailspike reveals only a single melting transition (Sturtevant et al., 1989). Furthermore, the structure of the tailspike suggests that all regions of the chain are trimeric. Though the N-terminal region is differentiated from the rest of the protein in the unfolding pathway, it does not appear to correspond to a truly independent domain (Wetlaufer, 1973).

*Comparison of Kinetic and Calorimetric Approaches.* In these experiments, we observed denaturation of tailspikes at 80 °C, considerably lower than the 88 °C $T_m$ reported by Sturtevant et al. (1989) using differential scanning calorimetry. The scanning calorimetry data were obtained for protein samples at concentrations of 2.6–2.8 mg/mL in 20 mM phosphate buffer at pH 7.4 with a scanning rate of 1 K $min^{-1}$ (Sturtevant et al., 1989). The thermal melting process was not reversible. Subsequent experiments by J. Sturtevant and A. Mitraki (personal communication) have revealed that the $T_m$ decreases with decreasing scanning rate; for example, with the scanning rate lowered to 0.125 K $min^{-1}$, the apparent $T_m$ decreased to 83.3 °C. Given the slow rate of unfolding that we have observed, the dependence of $T_m$ on scanning rate is expected and accounts for the differences in the measurements of melting temperature. This has recently been carefully studied with phosphoglycerate kinase (Galisteo et al., 1991).

It seems likely that the single transition observed by calorimetry represents both kinetic steps (the native to the intermediate and then to the unfolded monomers) observed here. Given the continual increase in temperature during calorimetry experiments, the unfolding reaction would be accelerated such that the population of molecules would have proceeded past the second barrier as the higher temperatures were reached.

*N-Terminus Is Involved in Head Binding.* The unfolding intermediate I was not able to bind to phage capsids, suggesting that the N-terminal region of the molecule binds to the capsid. Recently, Schwarz and Berget (1989a,b) have isolated a tailspike mutant, *hmH3034*, which is defective in binding to the capsid. The amino acid substitution occurs in the N-terminus, at Asp-100. Substitution of Asn resulted in a 100-fold decrease in head binding activity. More recently, Mauridies, Schwarz, and Berget (1990) reported the isolation of a second site suppressor which corrects the Asp-100 → Asn defect at Arg-13. The substitution pattern suggested that inter- or intrachain Arg/Asp salt bridges stabilized the N-terminal region of the tailspike.

*Effects of Mutations on Tailspike Stability.* The Asp-100 → Asn substitution decribed above also increased the electrophoretic mobility of the mutant native tailspike in SDS gels, but not in native gels (Schwarz & Berget, 1989b). If the alteration in mobility was due to the charge change, the mobility should have been altered in native gels (Yu & King, 1988). The mobility of the Asp-100 → Asn mutant protein in SDS gels was similar to the unfolding intermediate described here. It seems likely that the Asp-100 → Asn substitution destabilizes the N-terminal region of the tailspike, generating by mutation the partially unfolded intermediate that we have generated by thermal denaturation.

An unexpected aspect of the majority of gene *9* temperature-sensitive folding mutations is that their native proteins are as thermostable as the wild-type protein (Goldenberg & King, 1981; Sturtevant et al., 1989; Thomas et al., 1990) while their folding is vulnerable to small temperature changes. This is now understandable. Unfolding of the tailspike protein at high temperature is kinetically controlled by two high-energy barriers. Given the sequential character of the unfolding process, the apparent $T_m$ is determined by the initiation step, the melting of the N-termini of the polypeptide chains. None of these *tsf* mutations falls in the N-terminal region (Yu & King, 1984; Villafane & King, 1988). This accounts for their small effect on the initiation step and therefore on the apparent $T_m$. This first barrier apparently acts as a kinetic clamp on the unfolding of the mutant proteins.

We have carried out preliminary experiments on the thermal unfolding of proteins carrying *tsf* substitutions. Each of 10 *tsf* mutants spanning residues 162–435 had a large effect on $k_2$, the unfolding rate constant for the transition from the intermediate to the fully unfolded chains, but little effect on $k_1$, the unfolding rate constant for the transition from the native to the intermediate (B.-L. Chen, M.-H. Yu, & J. King,

Table II: Comparison between the in Vivo Late Folding Intermediate (Protrimer) and the Thermal Unfolding Intermediate (I) of Tailspike Protein

| property | protrimer (in vivo) | I (in vitro) |
|---|---|---|
| conformational state | trimer | trimer |
| folding to native | yes | yes |
| resistance to SDS | denatured monomer | nativelike trimer |
| resistance to protease | sensitive | central region and C-terminus resistant |

unpublished data). Given their location in the central region of the chain, it is not surprising that they affected the second step but not the first one.

*Unfolding Pathways of Oligomeric Proteins.* Careful examination of the unfolding of a variety of water-soluble oligomeric enzymes reveals a coupled multiple-step transition in which the oligomer dissociates to structured but inactive monomers and these further unfold to denatured chains [for a review, see Jaenicke (1987)]. For example, with β-galactosidase, the active tetramer transforms to an inactive tetramer which subsequently dissociates into monomers which aggregate (Edwards et al., 1990). The conformational transition that triggers the dissociation in the oligomeric state appears to be quite subtle.

The thermal unfolding of the trimeric tailspike protein reported here showed a variation to this general mechanism. Instead of dissociating the subunits first, only part of the structure in the polypeptide chain dissociates while the rest of the chain is still in a compact form and interacts strongly with other subunits. In globular proteins, the subunit–subunit interactions are through limited patches of subunit surface, with very little chain interpenetration. On the other hand, a very different situation holds for molecules like collagen or myosin, in which the chains are completely interwound so that dissociation cannot occur without substantial denaturation. We suspect that the tailspike may be intermediate; for example, some of the extensive β-sheet structure in the native state (Sargent et al., 1988) is probably formed from strands originating in different chains. Such intertwining between the polypeptide chains of different subunits has been observed in the *Escherichia coli* tryptophan repressor. This dimeric protein has its helices in one polypeptide chain intertwined with the corresponding helices in the other polypeptide chain (Schevitz et al., 1985) and also displays an unusually high thermal stability (Bae et al., 1988).

Intertwining of subunits or their chains not only would explain aspects of the thermostability but also would explain why in the folding pathway association of partially folded species occurs (Goldenberg et al., 1983b).

*In Vitro Thermal Unfolding versus in Vivo Folding Mechanisms.* In the forward folding pathway, both in vivo off the ribosome and in vitro from urea-denatured chains (Seckler et al., 1989), there are not intermediates stable in SDS, other than the fully native species. The late folding intermediate protrimer in which the chains are associated but not fully folded and would seem to be related to the unfolding intermediate (Goldenberg & King, 1982). Both the protrimer and the thermal unfolding intermediate are trimeric species and are closest to native in the pathways. However, the protrimer is completely sensitive to SDS denaturation and protease digestion (Goldenberg & King, 1982) whereas the thermal unfolding intermediate is not (Table II). Thus, the thermal unfolding intermediate is a more nativelike species with more folded structures than the protrimer.

In our in vitro thermal unfolding investigation, none of the previously reported in vivo folding intermediates were detectable, because all of these species are sensitive to SDS denaturation and therefore would be unfolded during SDS–PAGE analysis. The thermal unfolding intermediate has a longer lifetime and higher resistance to both SDS and proteases compared to the in vivo folding intermediates. However, the fact that the in vitro thermal unfolding intermediate has not been observed during the folding process suggests that it is absent from the in vivo folding pathway. Therefore, we conclude that the early stage of the thermal unfolding pathway is not the reverse of the late stage of the forward folding pathway.

## References


Alber, T. (1989) *Annu. Rev. Biochem. 58*, 765–798.

Bae, S.-J., Chou, W.-Y., Matthews, K., & Sturtevant, J. M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6731–6732.

Berget, P. B., & Poteete, A. R. (1980) *J. Virol. 34*, 234–243.

Berget, P. B., & Chidambaram, M. (1989) *Genetics 121*, 13–28.

Botstein, D., Waddell, C. H., & King, J. (1973) *J. Mol. Biol. 80*, 669–695.

Edwards, R. A., Jacobson, A. L., & Huber, R. E. (1990) *Biochemistry 29*, 11001–11008.

Ellman, G. L. (1959) *Arch. Biochem. Biophys. 82*, 70–77.

Fane, B., & King, J. (1987) *Genetics 117*, 157–171.

Fane, B., & King, J. (1991) *Genetics 127*, 263–277.

Galisteo, M. L., Mateo, P. L., & Sanchez-Ruiz, J. M. (1991) *Biochemistry 30*, 2061–2066.

Goldenberg, D. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 481–507.

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

Goldenberg, D., & King, J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 3403–3407.

Goldenberg, D., Berget, P., & King, J. (1982) *J. Biol. Chem. 257*, 7864–7871.

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

Goldenberg, D. P., Smith, D. H., & King, J. (1983b) *Biopolymers 22*, 125–129.

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

Horowitz, P., & Criscimagna, N. L. (1986) *J. Biol. Chem. 261*, 15652–15658.

Israel., J. V., Anderson, T. F., & Levine, M. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 284–291.

Iwashita, S., & Kanegasaki, S. (1976) *Eur. J. Biochem. 65*, 87–94.

Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol. 49*, 117–237.

King, J., & Laemmli, U. K. (1971) *J. Mol. Biol. 62*, 465–477.

King, J., & Yu, M.-H. (1986) *Methods Enzymol. 131*, 250–266.

Laemmli, U. K. (1970) *Nature 227*, 680–685.

Matsudaira, P. (1987) *J. Biol. Chem. 262*, 10035–10038.

Maurides, P., Schwarz, J. J., & Berget, P. B. (1990) *Genetics 125*, 673–681.

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

Pabo, C. O., Sauer, R. T., Sturtevant, J. M., & Ptashne, M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 1608–1612.

Pace, C. N. (1990) *Trends Biochem. Sci. (Pers. Ed.) 15*, 14–17.

Reynolds, J., & Tanford, C. (1970) *J. Biol. Chem. 245*, 5161–5165.

Sargent, O., Benevides, J. M., Yu, M.-H., King, J., & Thomas, G. J., Jr. (1988) *J. Mol. Biol. 199*, 491–502.

Sauer, R. T., Krovatin, W., Poteete, A. R., & Berget, P. B. (1982) *Biochemistry 21*, 5811–5815.

Schellman, J. A. (1987) *Annu. Rev. Biophys. Chem. 16*, 115–137.
Schevitz, R. W., Otwinowski, Z., Joachimiak, A., Lawson, C. L., & Sigler, P. B. (1985) *Nature 317*, 782–786.
Schwarz, J., & Berget, P. B. (1989a) *Genetics 121*, 635–649.
Schwarz, J., & Berget, P. B. (1989b) *J. Biol. Chem. 264*, 20112–20119.
Seckler, R., Fuchs, A., King, J., & Jaenicke, R. (1989) *J. Biol. Chem. 264*, 11750–11753.
Smith, D. H., Berget, P. B., & King, J. (1980) *Genetics 96*, 331–352.
Sturtevant, J. M., Yu, M.-H., Haase-Pettingell, C., & King, J. (1989) *J. Biol. Chem. 264*, 10693–10698.
Tandon, S., & Horowitz, P. (1986) *J. Biol. Chem. 261*, 15615–15618.
Thomas, G. J., Becka, R., Sargent, D., Yu, M.-H., & King, J. (1990) *Biochemistry 29*, 4181–4187.
Villafane, R., & King, J. (1988) *J. Mol. Biol. 204*, 607–619.
Volkin, D. B., & Klibanov, A. M. (1989) in *Protein Function: a Practical Approach* (Creighton, T. E., Ed.) pp 1–24, IRL Press, Oxford.
Wetlaufer, D. B. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 697–701.
Wetzel, R., Perry, L. J., Mulkerrin, M. G., & Randall, L. M. (1990) in *Protein Design and the Development of New Therapeutics and Vaccines*, Smith Kline & French Laboratories Research Symposia Series (Hook, J. B., & Poste, G., Eds.) pp 79–115, Plenum Press, New York.
Wilson, I. A., Skehel, J. J., & Wiley, D. C. (1981) *Nature 289*, 366–373.
Winston, F. D., Botstein, D., & Miller, J. H. (1979) *J. Bacteriol. 137*, 433–439.
Yu, M.-H., & King, J. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6584–6588.
Yu, M.-H., & King, J. (1988) *J. Biol. Chem. 263*, 1424–1431.

# Thermodynamics of Antiparallel Hairpin–Double Helix Equilibria in DNA Oligonucleotides from Equilibrium Ultracentrifugation

Philip D. Ross,*,‡ Frank B. Howard,‡ and Marc S. Lewis§

*Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, and Biomedical Engineering and Instrumentation Program, National Center for Research Resources, National Institutes of Health, Bethesda, Maryland 20892*

*Received December 6, 1990; Revised Manuscript Received March 11, 1991*

ABSTRACT: Five highly palindromic DNA dodecamers, four of which may form G-A or I-A purine–purine mispairs at either the 5,8 or 6,7 positions, have been studied at sedimentation equilibrium in the analytical ultracentrifuge. Each DNA oligonucleotide forms an equilibrium mixture of ordered antiparallel hairpin and double-stranded helical structures in solutions of 0.1 or 0.5 M NaCl between 5 and 40 °C. The dimeric duplex is favored by conditions of high salt and low temperature. The monomer–dimer equilibrium constants vary from $5 \times 10^6$ to $5 \times 10^3$ and are unique for each DNA dodecamer. Analysis of the temperature dependence of the equilibrium constants shows that the double helix to hairpin conversion is driven by a positive entropy change and is associated with an endothermic enthalpy change. The mispair substitutions at the 5,8 positions and the IA(6,7) mispair have the greatest tendency toward hairpin formation and exhibit significantly larger entropy changes than the nonmispaired dGGTACGCGTACC parent sequence and the thermodynamically similar GA(6,7) DNA. The consequences of such hairpin–double helix equilibria must be considered in the interpretation of other kinds of experiments carried out on oligonucleotides at different concentrations.

In recent years the production of DNA oligomers of defined nucleotide sequence has become commonplace due to advances in the methods of chemical synthesis and the availability of automated machines for the preparation of such molecules. The exact sequences of binding sites for biologically relevant proteins such as repressors and restriction endonucleases may be readily reproduced, and the effect of variation of sequence upon these and other interactions such as drug binding may be examined. Recent X-ray studies of crystals and two-dimensional nuclear magnetic resonance of oligonucleotides in solution have revealed a wide variety of structures that indicate considerable backbone flexibility and local sequence-dependent variation in DNA structure. Many of the properties of oligonucleotides such as their conformation, their mechanism of melting, and their interactions with ionic components in solution are highly chain-length dependent since oligonucleotides occupy the chain-length-sensitive transition region between small molecules and infinitely long-chain polyelectrolytes. It is clear that, on account of both their importance and their complexity, oligonucleotides should be examined from as many points of view as possible.

A spectroscopic and calorimetric study of five highly palindromic DNA dodecamers, four of which were designed with the potential of forming purine–purine mispairs in a duplex helical structure (F. B. Howard, C.-q. Chen, P. D. Ross, and H. T. Miles, manuscript in preparation) has recently been carried out. Based upon the conventional assumption of a two-state process, completely unreasonable values were derived for the thermodynamic parameters associated with the thermally induced melting transition. Since it has been known

* To whom correspondence should be addressed.
‡ National Institute of Diabetes and Digestive and Kidney Diseases.
§ National Center for Research Resources.