Role of Entropy in Protein Thermostability: Folding Kinetics of a Hyperthermophilic Cold Shock Protein at High Temperatures Using 19F NMR†

Benjamin Schuler,*†,§ Werner Kremer,§ Hans Robert Kalbitzer,§ and Rainer Jaenicke§

Institut für Biophysik und Physikalische Biochemie, Universität Regensburg, Universitätsstrasse 31, 93040 Regensburg, Germany, and Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

Received June 13, 2002; Revised Manuscript Received July 16, 2002

ABSTRACT: We used 19F NMR to extend the temperature range accessible to detailed kinetic and equilibrium studies of a hyperthermophilic protein. Employing an optimized incorporation strategy, the small cold shock protein from the bacterium Thermotoga maritima (TmCsp) was labeled with 5-fluorotryptophan. Although chaotropically induced unfolding transitions revealed a significant decrease in the stabilization free energy upon fluorine labeling, the protein’s kinetic folding mechanism is conserved. Temperature- and guanidinium chloride-dependent equilibrium unfolding transitions monitored by 19F NMR agree well with the results from optical spectroscopy, and provide a stringent test of the two-state folding character of TmCsp. Folding and unfolding rate constants at high temperatures were determined from the 19F NMR spectra close to the midpoint of thermal unfolding by global line shape analysis. In combination with results from stopped-flow experiments at lower temperatures, they show that the folding rate constant of TmCsp and its temperature dependence closely resemble those of its mesophilic homologue from Bacillus subtilis, BsCspB. However, the unfolding rate constant of TmCsp is two orders of magnitude lower over the entire temperature range that was investigated. Consequently, the difference in conformational stability between the two proteins is solely due to the unfolding rate constant over a wide temperature range. A thermodynamic analysis points to an important role of entropic factors in the stabilization of TmCsp relative to its mesophilic homologues.

Proteins from thermophilic organisms exhibit remarkable thermal stability, making life possible at temperatures above the boiling point of water. Despite two decades of active research in this field, a general concept of how this stability is achieved has remained elusive (1). The small differences that are necessary in terms of stabilization free energy appear to be implemented by a variety of strategies using combinations of virtually all known structural parameters (2, 3). Experimental analyses of the kinetic and thermodynamic basis of thermostability, which might extend our understanding beyond the static picture from sequence and structure comparisons, have often been hampered by the irreversible unfolding of thermophilic proteins (4). Recently, protein families have been identified that do not show these complications and allow a comparison of homologous proteins from organisms with different thermophilicities (5–7). To extend the temperature range and detail accessible to the investigation of the folding of the cold shock protein from the hyperthermophilic bacterium Thermotoga maritima (TmCsp, Figure 1) (5, 8, 9), we decided to use NMR. This method allows not only simultaneous monitoring of multiple nuclei in a protein individually but also the extraction of kinetic parameters from equilibrium measurements by line shape analysis (10). Due to signal overlap in proton NMR spectra, detailed analyses of this kind have so far been limited to the signals from aromatic amino acids in very small proteins (10–15). With its large fraction of aromatic amino acids, the assignment and separation of single resonances were not possible in proton spectra of TmCsp, particularly in the unfolding transition region. Thus, selective labeling of specific atoms or amino acid side chains had to be applied. Here we use 19F as a nucleus, which combines several advantages, including 100% natural abundance, a receptivity to NMR detection close to that of the proton, a large chemical shift range, and exquisite sensitivity for changes in the local van der Waals environment and electrostatic fields (16–19).

Most commonly, aromatic amino acids labeled with 19F at one of the ring positions are incorporated biosynthetically by microbial protein expression in the presence of the desired amino acid analogue. The ability of cells to synthesize that

† This work was supported by the Fonds der Chemischen Industrie and the Deutsche Forschungsgemeinschaft. B.S. was supported by an Emmy Noether-Fellowship from the Deutsche Forschungsgemeinschaft.

‡ To whom correspondence should be addressed: Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0520. Phone: (301) 402-1198. Fax: (301) 496-0825. E-mail: schuler@helix.nih.gov. Current address: Physikalische Biochemie, Universität Potsdam, 14476 Golm, Germany. Phone: +49 (331) 977 5244. Fax: +49 (331) 977 5082. E-mail: bschuler@rz.uni-potsdam.de.

§ National Institutes of Health.

§ Universität Regensburg.

‡ Abbreviations: TmCsp, cold shock protein from T. maritima; BsCspB, cold shock protein from Bacillus subtilis; BcCsp, cold shock protein from Bacillus caldolyticus; CD, circular dichroism; DSC, differential scanning calorimetry; OD555, optical density at 555 nm; IPTG, isopropyl β-d-thiogalactoside; GdmCl, guanidinium chloride; 5-F-Trp, 5-fluorotryptophan; TmCspFW, TmCsp labeled with 5-F-Trp; DSS, dimethylsulphoxide.

10.1021/bi02062931 CCC: $22.00 © 2002 American Chemical Society
Published on Web 08/17/2002
amino acid endogenously is most easily eliminated by adding the herbicide glyphosate, which inhibits the de novo synthesis of the aromatic amino acids (20, 21). However, this strategy is often hampered by incomplete incorporation (20), which introduces heterogeneity into the sample and can even obscure its folding behavior, if species with different degrees of labeling have different conformational stabilities or folding kinetics. Here, we describe a modified growth protocol, which permits complete incorporation of [19 F]Trp. We used the altered absorbance of the fluorinated tryptophan side chains to conveniently determine the level of incorporation spectroscopically. Monitoring GdmCl- and temperature-induced unfolding transitions by one-dimensional (1D) 19 F NMR spectra allowed us to stringently test the two-state folding behavior of TmCsp and measure folding and unfolding rate constants at high temperatures by complete line shape analysis. Being able to measure both thermodynamic and kinetic parameters over a wide temperature range allows a much more detailed analysis of the kinetic and thermodynamic mechanism of thermostability of this protein than previously possible. The increased thermostability of TmCsp compared to that of the highly homologous CspB from Bacillus subtilis (22) is found to be due to lower unfolding rate constants over a wide temperature range, whereas the folding rate constants and their temperature dependences are very similar for the two proteins. This leads us to conclude that entropic factors play an important role in the thermostabilization of TmCsp.

MATERIALS AND METHODS

Ultrapure GdmCl was from ICN (Cleveland, OH). All other chemicals were obtained from Sigma-Aldrich or Merck at analytical grade. The herbicide Roundup (Monsanto) was used as a cost-saving source of glyphosate. Chromatography material was from Pharmacia-Amersham, and YM3 ultrafiltration membranes from Amicon were used for the concentration of protein solutions. Refractive index measurements were used to determine GdmCl concentrations (23). Protein concentrations were determined by absorbance spectroscopy (24, 25) using a molar extinction coefficient of 1.28 × 10³ M⁻¹ cm⁻¹ for unlabeled TmCsp (8) and molar extinction coefficients for labeled TmCsp as described below. All measurements were performed in buffer containing 50 mM sodium phosphate, 20 mM NaCl, 0.2 mM EDTA, and 1 mM DTE (pH 6.5) unless stated otherwise. SigmaPlot (Jandel) was used for regression analysis.

Mutagenesis, Incorporation of Fluorine, and Protein Purification. One of the major problems with the incorporation of some of the fluorinated amino acids into proteins is their toxicity to bacterial cells (21). Inclusion of ~10–20% of the unlabeled analogue in the growth medium has been suggested (20), which increases the viability of the cells. However, if the labeling changes the stability or folding mechanism of the protein, the resulting incomplete incorporation of the fluorine label leads to a heterogeneous sample, which is prohibitive for an analysis of protein folding. Therefore, we grew bacterial cells in the presence of glyphosate with a limiting concentration of unlabeled Trp in the medium to allow for the buildup of a biosynthetic machinery unaffected by the toxic effects of the fluorine-labeled amino acid. When the cell culture started to reach the stationary phase due to a limiting Trp concentration, 5-F-Trp was added and protein expression was induced. If a sufficiently tightly regulated expression system is used, this procedure ensures high viability of the cells and complete incorporation of the respective amino acid.

The cold shock protein from T. maritima (TmCsp) was expressed in Escherichia coli BL21(DE3) using a pET21a vector as described previously (8). Single tryptophan mutants were cloned using the Stratagene Quickchange site-directed mutagenesis kit and the primers 5′-CATATGAGAGGAAAGGTTAAAGTGACCTTTGATATAAGGATCAGG-3′ and 5′-CCGTAGCCCTTCTTGGAATCGAAGTACTTAACCTTTCTTCATATG-3′ to substitute a tyrosine residue for...
Trp7 and introduce a diagnostic Scal restriction site (mutant W7Y) and the primers 5′-GGAGACGTGTTCTGACAC-TACTCAGCAATTTGGAATTGGAAGTTTCAAAAAC-3′ and 5′-GTITGTTGAAACCCTTTGCAAATTGCGTAGTGTG-TACGAAACGTCTCC-3′ to replace Trp29 with tyrosine and introduce a diagnostic MunI restriction site (mutant W29Y).

The unlabeled protein was expressed as described previously (8). For the introduction of fluorotryptophan, minimal medium (NMM) was used (26) with all amino acids except for Trp at a final concentration of 50 mg/L. Trp was added to a final concentration of only 5 mg/L. All bacterial growth was under the selective pressure of 50 mg/L glycine at a final concentration of 1 g/L. Medium was prepared by weighing out the dry substance in triplicate and dissolving it in water or an aqueous 6 M GdmCl solution at neutral pH. The concentration of GdmCl did not affect the absorbance at wavelengths between 250 and 320 nm, and using the total number of Phe, Tyr, and unlabeled Trp is shown as a dashed line.

Figure 2: Absorbance of 5-F-Trp in comparison to that of unlabeled Trp and deconvolution of fluorine-labeled TmCsp using the absorbance spectra of isolated aromatic amino acids in 6 M GdmCl. (a) Absorbance spectra of Trp and 5-F-Trp in 6 M GdmCl showing the strong bathochromic effect caused by the fluorine atom. (b) The absorbance spectrum of TmCsp labeled with 5-F-Trp (TmCspFW), denatured in 6 M GdmCl (●), can be well deconvolved using a sum of the absorbance spectra of Phe, Tyr, and 5-F-Trp (solid line). The optimum deconvolution using the spectra of Phe, Tyr, and unlabeled Trp is shown as a dashed line.

Table 1: Extinction Coefficients* of 5-F-Trp

<table>
<thead>
<tr>
<th></th>
<th>262 nm</th>
<th>271 nm</th>
<th>280 nm</th>
<th>290 nm</th>
<th>300 nm</th>
</tr>
</thead>
<tbody>
<tr>
<td>water</td>
<td>(4.0 ± 0.2) × 10^3</td>
<td>(5.4 ± 0.2) × 10^3</td>
<td>(5.7 ± 0.3) × 10^3</td>
<td>(5.1 ± 0.2) × 10^3</td>
<td>(2.05 ± 0.09) × 10^3</td>
</tr>
<tr>
<td>6 M GdmCl</td>
<td>(3.76 ± 0.05) × 10^3</td>
<td>(5.31 ± 0.06) × 10^3</td>
<td>(5.75 ± 0.07) × 10^3</td>
<td>(5.33 ± 0.06) × 10^3</td>
<td>(2.61 ± 0.03) × 10^3</td>
</tr>
</tbody>
</table>

* Extinction coefficients are given in M⁻¹ cm⁻¹.

where ε_{native} is the extinction coefficient of the native protein and ε_{GdmCl} is the extinction coefficient of the denatured protein calculated from the sum of the contributing amino acids in 6 M GdmCl at the respective wavelength. A_{native} and A_{GdmCl} are the absorbances of the native and denatured protein measured with protein solutions diluted from the same stock solution to identical final protein concentrations into buffer and GdmCl solutions, respectively.

The absorbance spectrum of TmCspFW was analyzed assuming that only Phe, Tyr, Trp, and 5-F-Trp contribute significantly to the absorbance at wavelengths between 250 and 320 nm, and using the total number of Phe, Tyr, and...
where GdmCl dependence of folding kinetics was fit using the averaged and analyzed as single-exponential functions. The GdmCl concentrations. Five to fifteen kinetic traces were solutions of varying concentrations to give the desired final GdmCl was diluted 11-fold with aqueous buffer or GdmCl containing 50 mM sodium phosphate, 20 mM NaCl, 0.2 mM EDTA, and 1 mM DTE (pH 6.5).

Trp residues present in the polypeptide as a constraint. The experimental data can be described well as a sum of these amino acid spectra, in contrast to the best fit composite spectrum assuming only incorporation of unmodified aromatic side chains (Figure 2b). The optimum fit was obtained by assuming the complete absence of unlabeled Trp. This result was confirmed by quantitative NMR. As expected, two resonances of equal intensity at −44.2 and −47.6 ppm were observed in the 19F NMR spectra of TmCspFW (Figures 5 and 6 and Table 2), compared to a resonance at −47.0 ppm for pure 5-F-Trp (Table 2). Integration of these lines and comparison with the 19F NMR spectra of free 5-F-Trp at a known concentration under identical conditions gave an incorporation yield of 99 ± 3%. We also analyzed the samples with electron spray ionization mass spectrometry (ESI-MS). The upper bound for the unlabeled protein resulting from these measurements was −0.5%, and that for the singly labeled protein was 2%, corroborating virtually complete incorporation. Our analysis resulted in extinction coefficients of 1.30 × 10^4 M⁻¹ cm⁻¹ for TmCspFW at 280 nm, using the values for unlabeled amino acids given by Pace et al. (25).

\[ \lambda = k_{f,0} e^{m_{c,D}} + k_{u,0} e^{m_{u,D}} \]

where \( \lambda \) is the measured rate constant, \( k_{f,0} \) and \( k_{u,0} \) are the folding and unfolding rate constants, respectively, at a denaturant concentration of 0 M, and \( c_D \) is the denaturant concentration

\[ m_f = \frac{\partial \ln k_f}{\partial c_D} \]

and

\[ m_u = \frac{\partial \ln k_u}{\partial c_D} \]

For a protein folding according to a two-state reaction

\[ D \rightleftharpoons N \]

this results in a linear dependence of folding and unfolding rate constants on denaturant concentration if \( \lambda \) is plotted on a logarithmic scale.

**Differential Scanning Calorimetry.** Samples at a protein concentration of 1 mg/mL were dialyzed against buffer containing 50 mM sodium phosphate, 20 mM NaCl, 0.2 mM EDTA, and 1 mM DTE (pH 6.5) for at least 12 h; then D2O was added to a final concentration of 8%, and the sample and some reference dialysis buffer were filtered and degassed. Measurements were performed in a CSC 6100 II calorimeter (Calorimetry Sciences Corp., Provo, UT) at a scan rate of 1 K/min. Independent samples were tested for reversibility of the unfolding transition by repeated scanning. Data were analyzed using the deconvolution software CpCalc supplied by CSC. The partial specific volume of the protein was calculated from its amino acid composition to be 0.744 cm³/g, and the exact concentration was determined by absorbance.

**NMR Spectroscopy.** All samples were dialyzed against buffer containing 50 mM sodium phosphate, 20 mM NaCl, 0.2 mM EDTA, and 1 mM DTE (pH 6.5), and then NaDSS, D2O were added to final concentrations of 5 mM, 0.1 mM, and 8%, respectively. A sample volume of 500 μL was used, and sample tubes were sealed by melting for high-temperature experiments. 19F NMR spectra were obtained at 470 MHz in a Bruker DMX 500 spectrometer using an appropriately tuned 5 mm hydrogen probe. Ninety-degree flip-angle pulses were found to be ~11.5 μs at 300 K for all samples. For the quantitative measurements, a pulse angle of 90° and a repetition time of 1 s were used. Under these conditions, no saturation of 19F signals was observed. No pH correction was made for the D2O content; no proton decoupling was used, and spectra were processed with 1 Hz of line broadening, which does not affect line shape significantly at the line widths observed in our 19F spectra. Indirect referencing via 1H NMR spectra and DSS (2,2-dimethyl-2-silapentane-5-sulfonic acid) was used. For this purpose, 1H spectra of the samples were measured at 500 MHz under identical conditions just before 19F data acquisition, and the absolute 1H frequency of the DSS signal was multiplied by a scaling factor of 0.940 866 982 to obtain the reference frequency of trifluoroacetic acid (TFA) for 19F spectra (30). Temperature calibration was performed using the known temperature dependence of the difference of the two ethylene glycol 1H NMR resonances (31). Integration
of NMR signals was performed using PeakFit (Jandel) and Excel (Microsoft).

Thermal unfolding of thermostable proteins can be problematic due to the irreversible degradative reactions taking place at high temperatures (32), which can only be neglected if incubation at these conditions is limited to a short period of time. Due to the low signal-to-noise ratio of NMR experiments and the broad lines of the $^{19}$F-labeled protein, spectra in the transition region of thermal unfolding of TmCspFW had to be acquired for at least 1 h. Thus, even the use of fresh samples for each temperature could not completely prevent the slow appearance of small amounts of chemically modified polypeptides during data acquisition. From experiments with samples incubated at high temperatures for up to 10 h, we estimate the error introduced into our data to be less than 10%.

**Line Shape Analysis.** Folding and unfolding rate constants in the transition region of thermal melts were estimated from the line broadening in $^{19}$F NMR spectra by nonlinear regression, using the solution for the kinetic equations describing the exchange between two states and the Bloch equations describing the spin–relaxation relaxation of a nucleus (10, 33). The frequency dependence of the $^{19}$F NMR line spectrum $I(\nu)$ was simulated by

$$I(\nu) = \frac{P[1 + \tau\left(\frac{P_D}{T_{2N}} + \frac{P_N}{T_{2D}}\right)] + QR}{P^2 + R^2}$$

where $C_0$ is a scaling factor that is proportional to the protein concentration. $T_{2N}$ and $T_{2D}$ are the apparent transverse relaxation times in the absence of exchange broadening of states N and D, respectively, which were determined by linear extrapolation of the pre- and posttransition baseline values into the transition region, and $p_N$ and $p_D$ are the fractional populations of the native and denatured states, respectively. The terms $P$, $Q$, and $R$ are given by

$$P = \frac{1}{T_{2N}T_{2D}} - 4\pi^2\Delta\nu^2 + \pi^2(\partial\nu)^2 + \frac{P_N}{T_{2N}} + \frac{P_D}{T_{2D}}$$

$$Q = \tau[2\pi\Delta\nu - \pi\partial\nu(p_N - p_D)]$$

$$R = 2\pi\Delta\nu\left[1 + \tau\left(\frac{1}{T_{2N}} + \frac{1}{T_{2D}}\right)\right] + \pi\partial\nu(\frac{1}{T_{2N}} + \frac{1}{T_{2D}}) + \pi\partial\nu(p_N - p_D)$$

with

$$\partial\nu = v_N - v_D$$

and

$$\Delta\nu = \frac{v_N + v_D}{2} - \nu$$

where $v_N$ and $v_D$ are the resonance frequencies of a nucleus in the native and denatured states, respectively, which were estimated by extrapolation of the pre- and posttransition baseline values into the transition region using a quadratic function. With the folding and unfolding rate constants $k_i$ and $k_u$, respectively, the exchange correlation time $\tau$ is defined as

$$\tau = \frac{1}{k_i + k_u}$$

The relative populations of states D and N follow as

$$p_N = k_i\tau \quad \text{and} \quad p_D = k_u\tau$$

A sum of two functions $I(\nu)$ was used for a global fit of the doubly labeled protein, using independent values for $T_{2N}$, $T_{2D}$, $v_N$, and $v_D$ and identical values for $p_N$, $p_D$, $k_i$, $k_u$, and $C_0$. Standard deviations for the folding and unfolding rate constants were estimated from fits using independent values for $k_i$ and $k_u$ for the two lines in exchange. The indirect fluor–proton spin–spin couplings to the protons in positions 4 (H3), 6 (H5), and 7 (H7) in 5-fluorotryptophan (F35) are 9.9 ± 0.1 Hz for $J(F^3^H^3^N)$ and $J(F^7^H^7^N)$ and 4.68 ± 0.07 Hz for $J(F^3^H^7^N)$ (measured in aqueous solution at 293 K and pH 6.5). They are not resolved in the protein and were neglected in the simulations since the inhomogeneous line broadening caused by these couplings is small compared to the homogeneous line widths.

**Analysis of the Temperature Dependence of Folding and Unfolding Rate Constants.** Assuming that the transition state ensembles for folding and unfolding are energetically similar and that the reaction is microscopically reversible, we used transition state theory to analyze the folding kinetics of TmCspFW (34, 35). Transition state theory relates a reaction rate constant $k$ to the Gibbs free energy of the transition state $\Delta G^\ddagger$ relative to the initial state:

$$k = \frac{k_B T}{h} \exp\left(-\frac{\Delta G^\ddagger}{RT}\right)$$

where $\epsilon$ is the transmission coefficient, $k_B$ and $h$ are the Boltzmann and Planck constants, respectively, $R$ is the gas constant, and $T$ is the absolute temperature. In analogy to the equilibrium Gibbs free energy $\Delta G$, the temperature dependence of $\Delta G^\ddagger$ is given by

$$\Delta G^\ddagger = \Delta H^\ddagger - T\Delta S^\ddagger$$

where $\Delta H^\ddagger$ is the enthalpy of activation, $\Delta S^\ddagger$ is the entropy of activation, $\Delta c_{px}^\ddagger$ is the change in heat capacity between the final and the initial state, which is assumed to be independent of temperature in this derivation, and $T_0$ is the reference temperature, 298 K in our case. For the folding and unfolding rate constants $k_i$ and $k_u$, respectively, eq 2 and 3 yield the Eyring equations

$$k_i = \frac{k_B}{h} \exp\left(-\frac{\Delta G^\ddagger_{i}(T_0) - \Delta G^\ddagger_{px}(T - T_0)}{RT}\right)$$

and

$$k_u = \frac{k_B}{h} \exp\left(-\frac{\Delta G^\ddagger_{u}(T_0) - \Delta G^\ddagger_{px}(T - T_0)}{RT}\right)$$
1D 1H NMR spectra exhibited the usual dispersion and peak
Trp, the structural integrity of the protein had to be verified.
DSC (Figure 3a). In spectropolarimetric measurements, 360
the ratio of calorimetric to van’t Hoff unfolding enthalpies
proximation of the DSC data gave good fits (Figure 3a), and
the same samples to 375 K. A theoretical two-state ap-
unfolding was found to be reversible by repeated heating of
baselines in CD experiments (22). To be able to compare our data directly to those
of Schindler and Schmid (22), we used a transmission
coefficient of 1 in our analysis, which allows changes in
activation entropies with temperature to be determined,
although it clearly overestimates the absolute value of the
preexponential factor (36) and does not take into account
the temperature dependence of solvent viscosity.

RESULTS
Effects of Fluorine Labeling on Csp Structure and Ther-
mostability. To be able to assess the validity of any
conclusions drawn from folding analyses of TmCspFW, in
which both Trp residues of TmCsp are replaced with 5-F-
Trp, the structural integrity of the protein had to be verified.
1D 1H NMR spectra exhibited the usual dispersion and peak
distribution of native TmCsp. Moreover, the similarity of far-
UV CD spectra indicated virtually identical secondary
structure content of labeled and unlabeled TmCsp (data not
shown), as expected for fluorine incorporation (17). Effects
on thermostability were investigated by CD and differential
scanning calorimetry (DSC). The melting temperature for
TmCspFW was 359 K, compared to 361 K for TmCsp using
DSC (Figure 3a). In spectropolarimetric measurements, 360
K was observed as the transition midpoint for TmCspFW
from the fit to a two-state model, compared to 364 K for
the unlabeled protein (Figure 3b). Slight differences between
the two methods are probably due to the high melting point
of the protein and the resulting uncertainty of posttransition
baselines in CD experiments (8). Temperature-induced
unfolding was found to be reversible by repeated heating of
the same samples to 375 K. A theoretical two-state ap-
proach of the DSC data gave good fits (Figure 3a), and
the ratio of calorimetric to van’t Hoff unfolding enthalpies
was close to unity for all samples as expected for a two-
state transition (37), similar to what was found for the
unlabeled protein (38). Despite slight effects on the melting
temperature, fluorine labeling of TmCsp therefore conserved
both the thermostability and the two-state characteristics of
the protein.

Effect of Fluorine Labeling on the Folding of TmCspFW in GdmCl. Unlike thermal unfolding transitions, GdmCl-
induced unfolding transitions of TmCspFW exhibited a clear
shift of the midpoint, c1/2, to lower GdmCl concentrations
by ~1 M and a slight decrease in cooperativity, m, compared
to that of unlabeled TmCsp (Figure 4). This corresponds to
a decrease in the Gibbs free energy of stabilization extrapo-
lated to 0 M GdmCl, ΔG0°, by ~10 kJ/mol, to 14 ± 2 kJ/
mol (5, 8), a destabilization much greater than that typically
reported for fluorine-labeled proteins (17–19). This decrease
in stability also manifested itself in the dependence of the
folding kinetics on GdmCl concentration (Figure 4a): the
extrapolated folding rate constant in the absence of denatur-
ant, k0 = 470 s⁻¹, and its linear change with GdmCl
concentration, m0, were virtually identical to those of
unlabeled TmCsp, but the unfolding branch of the chevron
plot was shifted to higher unfolding rate constants with a
k0 of 0.47 s⁻¹ (Figure 4a). Both ΔG0° and its GdmCl
dependence calculated from kinetic and equilibrium data
were in good agreement. Identical values were obtained for
the apparent folding rate constant λ for unfolding and
refolding experiments at the same concentration of denatur-
ants, and the analysis of the amplitudes of the monophasic

FIGURE 3: Thermal denaturation of fluorine-labeled (●, Tm = 359
K) and unlabeled (○, Tm = 361 K) TmCsp. (a) Temperature
dependence of the partial molar heat capacity measured by
differential scanning calorimetry. Only every 12th data point is
given. The solid lines show theoretical two-state approximations.
(b) Temperature dependence of the circular dichroism signal at 208
nm for fluorine-labeled (●, Tm = 360 K) and unlabeled (○, Tm = 364 K) TmCsp. Only every second data point is given. The solid
lines show fits to a two-state model.

FIGURE 4: GdmCl dependence of the folding kinetics and equili-
rium unfolding of labeled TmCsp at 298 K. (a) Observed rate constant λ for folding and unfolding as a function of GdmCl
concentration. A fit of the data to a two-state model (see Materials
and Methods) is shown by the solid line. The dashed lines show
the extrapolated GdmCl dependence of the microscopic folding and
unfolding rate constants. The dotted lines represent the fits to
the corresponding data for the highly homologous unlabeled cold shock
proteins from B. subtilis (BsCspB), B. caldolyticus (BcCsp), and
T. maritima (TmCsp) shown for comparison (parameters taken from
ref 5). (b) GdmCl-induced equilibrium unfolding transition measured
by fluorescence [unfolding (●) and refolding (○)] and peak
integration from 1D 19F NMR spectra [signal from [19F]Trp29 (△)
and [19F]Trp7 (▲)]. The fraction of native protein obtained from
the fit to a two-state model is plotted, and the regression is indicated
by the solid line. For comparison, the corresponding two-state fit
for unlabeled TmCsp is included (dashed line; parameters taken from
ref 5).
unfolding and refolding kinetics of TmCspFW (39) provided no evidence for rapidly formed folding intermediates (data not shown). Despite the significant effect of fluorine incorporation on its stability, these observations strongly support the assumption that the folding of TmCspFW is a reversible two-state reaction, and that the folding mechanism of TmCsp is conserved.

**Investigation of Two-State Folding by 19F NMR.** To be able to assign the fluorine resonances of Trp7 and Trp29 in the native and denatured states of the protein, single-tryptophan mutants (W7Y and W29Y) were created by site-directed mutagenesis and labeled with 5-F-Trp. Under native conditions, the mutants exhibited 19F NMR spectra with single peaks at −44.2 ppm for W29Y and −47.6 ppm for W7Y (data not shown), allowing unequivocal assignment of the resonances of the doubly labeled protein. Moreover, identical peak positions in TmCspFW and the single-tryptophan mutants indicate a lack of direct interactions between the two tryptophan side chains, which makes them independent probes for monitoring the folding reaction and allows an even more stringent test of the two-state model than with optical methods. A GdmCl-induced unfolding transition of TmCspFW monitored by 19F NMR is shown in Figure 5. Above 1 M GdmCl, the intensity of the signals from W7 and W29 decreased synchronously with the simultaneous appearance of a single resonance for the denatured protein close to the signal of free 5-F-Trp (Table 2). This suggests the absence of secondary or tertiary structure in the unfolded state. There was hardly any dependence of the chemical shifts of all three peaks on GdmCl concentration, which means that there are no detectable changes in the electronic vicinity of the fluorine atoms, and therefore that the conformational ensembles of both the native and the denatured state do not undergo significant changes in protein structure or dynamics with changing GdmCl concentration. Calculating the fraction of native protein from the peak integrals yielded unfolding transitions that agree well with the transition obtained from fluorescence measurements (Figure 4b), which is compelling evidence for the good agreement of the equilibrium folding of TmCspFW with a two-state model.

**Temperature Dependence of Folding by 19F NMR and Stopped Flow.** The Tm of 356 K determined from the 19F NMR spectra taken between 300 and 392 K (Figure 6) is close to the values obtained from CD and calorimetry (Figure 3), but several differences compared to the GdmCl-induced unfolding transition are obvious. First of all, there were both a pronounced temperature dependence of the chemical shifts and a gradual narrowing of the lines in the pretransition region. The increasing thermal motion in the native protein (40) can probably account for these effects without having to invoke major structural changes or folding pretransitions. Above 370 K, two sharp lines were observed in the denatured state (Figure 6) that were assigned from the single-tryptophan mutants under identical conditions (data not shown). The chemical shift difference between the lines was approximately constant over the entire posttransition region, indicating stable residual structure in the thermally unfolded state. No indication for secondary or tertiary structure was found in GdmCl-denatured TmCspFW (see above, Figure 5). This discrepancy is in agreement with the dependence of residual structure on unfolding conditions in the denatured states of other proteins (41–43), and can probably be accounted for by the shielding of electrostatic interactions at high concentrations of GdmCl.

Around the folding temperature, chemical exchange between folded and unfolded states leads to obvious line broadening, especially for the W7 resonance (Figure 6). This can be exploited to obtain folding and unfolding rate constants at high temperatures, using the functions describing NMR spectra of systems undergoing chemical exchange (10, 33; see Materials and Methods). The method is limited to conditions significantly populating both the folded and the denatured state and yields greater uncertainty in the values determined than fluorescence stopped-flow techniques; however, it extends the measurable range of rate constants beyond that of conventional techniques (10–12, 15), and it facilitates high-temperature measurements. Global line shape analysis of the spectra in the temperature range between 347 and 362 K assuming a two-state reaction yielded good fits (Figure

**FIGURE 5:** 19F NMR spectra of 5-F-Trp-labeled TmCsp as a function of GdmCl concentration measured at 298 K. Data were taken at a 19F frequency of 470 MHz and indirectly referenced relative to TFA using 1H spectra recorded under identical conditions. The vertical displacement of the spectra is proportional to the GdmCl concentration in the sample, which is indicated to the left of each spectrum. The resonances were assigned using site-directed mutagenesis.
Unfolding rate constants that increased from \(~40\) s\(^{-1}\) at 347 K to \(~300\) s\(^{-1}\) at 362 K were obtained, along with less temperature-dependent folding rate constants of \(~200\) s\(^{-1}\) (Figure 7).

Together with the values from the stopped-flow experiments between 290 and 318 K, the unfolding rate constant increased by 3 orders of magnitude over a temperature range of 72 K (Figure 7). The unfolding rate constant of TmCspFW is almost independent of GdmCl concentration (Figure 4), which implies a transition state that is very natively-like in its interactions with the denaturant, a property that is closely linked to the heat capacity change, because both values depend on the difference in solvent accessible surface area upon unfolding (44). Correspondingly, the curvature of the temperature dependence of the unfolding rate constant of TmCspFW is also very small (Figure 7), yielding a \(\Delta c_{p,u}\) close to zero (Table 3). The folding rate constant is much less temperature-dependent than the unfolding rate constant in the range investigated here. As the transition state and denatured state are very dissimilar in their interactions with the solvent (Figure 4; see above), the heat capacity difference between these two states is large. Using eq 4a, we obtained a value of \(-2.6 \pm 0.2\) kJ mol\(^{-1}\) K\(^{-1}\), with an activation enthalpy for folding \(\Delta H_p^*\) of 51 \(\pm 2\) kJ mol\(^{-1}\) at 298 K. Due to the large negative \(\Delta c_{p,u}\), \(\Delta H_p^*\) decreases with temperature. At low temperatures it is positive but becomes zero at \(\sim318\) K, and at higher temperatures, it is negative, implying that the barrier to refolding is of both enthalpic and entropic origin at low, but of entirely entropic origin at high, temperatures. This non-Arrhenius behavior of protein folding rates is well-established (45, 46). A very similar temperature dependence was found for the folding kinetics of BsCspB (22; Figure 7).

### Table 3: Comparison of Thermodynamic and Kinetic Parameters for TmCspFW and BsCspB at 298 K

<table>
<thead>
<tr>
<th></th>
<th>TmCsp</th>
<th>BsCspB</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta G_f^\circ) (equilibrium) (kJ mol(^{-1}))</td>
<td>14 (\pm) 2</td>
<td>8.9</td>
</tr>
<tr>
<td>(m) (equilibrium) (kJ mol(^{-1}) M(^{-1}))</td>
<td>5.6 (\pm) 0.2</td>
<td></td>
</tr>
<tr>
<td>(k_f^0) (s(^{-1}))</td>
<td>470</td>
<td></td>
</tr>
<tr>
<td>(k_u^0) (s(^{-1}))</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>(\Delta G_f^\circ) (kinetic) (kJ mol(^{-1}))</td>
<td>17 (\pm) 2</td>
<td></td>
</tr>
<tr>
<td>(R T m_f) (kJ mol(^{-1}) M(^{-1}))</td>
<td>51 (\pm) 2</td>
<td>32 (\pm) 2</td>
</tr>
<tr>
<td>(R T m_u) (kJ mol(^{-1}) M(^{-1}))</td>
<td>0.40 (\pm) 0.02</td>
<td></td>
</tr>
<tr>
<td>(\Delta H_f^\circ) (kJ mol(^{-1}))</td>
<td>51 (\pm) 2</td>
<td>32 (\pm) 2</td>
</tr>
<tr>
<td>(\Delta H_f^\circ) (kJ mol(^{-1}) K(^{-1}))</td>
<td>72 (\pm) 2</td>
<td>96 (\pm) 3</td>
</tr>
<tr>
<td>(\Delta c_{p,u}) (kJ mol(^{-1}) K(^{-1}))</td>
<td>0.2 (\pm) 0.2</td>
<td>0.3 (\pm) 0.4</td>
</tr>
</tbody>
</table>

* Data taken from ref 22. * Obtained from GdmCl-induced equilibrium unfolding transitions (Figure 4b). * Obtained from the GdmCl dependence of the folding and unfolding kinetics (Figure 4a). * Assumed to be independent of temperature.
As opposed to the strikingly similar folding rates of TmCspFW and BsCspB, the unfolding rate constants are almost 2 orders of magnitude higher for BsCspB over the whole temperature range that was investigated (Figure 7). From the work of Perl et al. (5), it has been known that at 298 K, the differences in stability between BsCspB, Csp from Bacillus caldolyticus (BcCsp), and TmCsp are entirely due to different unfolding rate constants, whereas the folding rate constants and the dependences of both folding and unfolding rate constants on GdmCl concentration are virtually identical for the three proteins. Our results suggest that the similarities in the folding kinetics and mechanism found for the mesophilic and hyperthermophilic cold shock proteins at ambient temperature extend to their temperature dependence: the lower unfolding rate constants of TmCsp can also account for the different melting temperatures of BsCspB and TmCsp. In other words, the mechanism of stabilization of the thermophilic versus the mesophilic protein is conserved from room temperature to the physiological temperature of Thermotoga beyond 80 °C.

**DISCUSSION**

The conservation of the temperature dependences of folding and unfolding in this family of small cold shock proteins has implications for the structural interpretation of their different conformational stabilities. In this case, their different stabilities are completely due to different unfolding rate constants, but what is the origin of the large increase in the activation barrier for unfolding? One possible explanation would be a stabilization of the native state by means of surface mutations leading to a larger number of enthalpic interactions that form only after the transition state (5). Assuming that the energetics of the unfolded states of the three Csp variants are very similar, one would then expect increasing activation enthalpies of unfolding with increasing thermostability. However, as shown in Figure 8, the activation enthalpy of unfolding and even more the equilibrium unfolding enthalpy at 298 K are significantly lower for the thermophilic protein, which shows that enthalpic contributions can obviously not explain its increased stability. The activation enthalpy corresponds to the slope of the Arrhenius plot (the graph of ln k_u as a function of 1/T, Figure 7), and to a first approximation, the activation entropy is proportional to the intercept with the ordinate. The large virtually parallel shift of the unfolding rate constants in Figure 7 therefore points to entropic factors dominating the effects of amino acid exchanges that lead to thermostability. Even if ΔH⁺_u is varied between 0 and 400 J mol⁻¹ K⁻¹, it is obvious from an analysis using eq 4b that the difference between the activation free energies of unfolding of TmCspFW and BsCspB is to a large part due to a difference in the activation entropies of unfolding.

Assigning these entropic factors to contributions from either solvent entropy or protein conformational entropy in experimental studies is complicated by the fact that solvent and protein form one thermodynamic system. Here, the only hint comes from the solvent dependence of the activation entropy. For BsCspB, it was found that, compared with the folding activation entropy, the unfolding activation entropy hardly changed with urea concentration (22). This argues that most of the entropy difference between the transition state and the folded state is not due to different solvent interactions, but rather due to a reduction of the degrees of freedom accessible to the protein, e.g., by tighter core packing or reduced surface side chain flexibility. Therefore, it appears to be probable that changes in conformational entropy on the native side of the barrier can be modulated more easily than solvent entropy to achieve a decrease in free energy of the folded protein. This illustrates a situation where the mechanism of thermostabilization is inaccessible to approaches using the comparison of static structures, one of the classical strategies in the field (2), but rather requires kinetic and thermodynamic or dynamic structural data.

A detailed structural interpretation of our finding does not appear to be feasible on the basis of our current molecular understanding of protein stability. Molecular dynamics simulations, currently the only method that can address all components of the molecular partition function of a protein individually for all states, are not yet able to give reliable quantitative estimates of their relative contributions. However, it has become clear over the past two decades that, despite the high packing densities of proteins, residual conformational entropy in the folded state can by no means be neglected and may even surpass the entropy change upon folding (47). More recent simulations indicate that previously unexpected entropic terms can significantly contribute to the delicate free energy balance of protein stability. Vibrational flexibility, for instance, can increase upon secondary structure formation (48), especially in the case of β-structure (49). The only currently available experimental methods with the potential capability of addressing some of these issues are nuclear spin relaxation rate measurements (50, 51; for a review, see ref 52). In a recent study, Stone and co-workers (53) found a striking correlation between the stability and backbone flexibility of several protein G B1 domain mutants, suggesting that increased backbone conformational entropy is a means of protein stabilization. Similar analyses might be able to clarify whether related mechanisms are utilized to achieve the remarkable stability of proteins from thermophilic organisms.

**ACKNOWLEDGMENT**

We thank G. Hummer, T. Kiefhaber, and T. Oas for discussion, F. X. Schmid, D. Perl, and W. A. Eaton for access...
REFERENCES
