Thermodynamic Basis for the Increased Thermostability of CheY from the Hyperthermophile *Thermotoga maritima*

William A. Deutschman and F. W. Dahlquist*

Institute of Molecular Biology, Department of Chemistry, University of Oregon, Eugene, Oregon 97403

Received April 3, 2001; Revised Manuscript Received July 17, 2001

ABSTRACT: The CheY protein isolated from the hyperthermophile *Thermotoga maritima* is much more resistant to thermally induced unfolding than is its counterpart from the mesophile *Bacillus subtilis*. To determine the basis of this increased thermostability, the temperature dependence of the free energy of unfolding was determined for these CheY homologues using denaturant-induced unfolding experiments. This allowed comparison of *T. maritima* CheY with *B. subtilis* CheY and determination of the thermodynamic qualities responsible for the enhanced thermostability of *T. maritima* CheY. The stability of the thermophilic CheY protein is a direct result of the increased enthalpy contribution at the temperature of zero entropy, T_s, and the decreased heat capacity change upon unfolding, resulting in a decreased dependence of the free energy of unfolding on temperature. It was found that neither purely entropic nor purely enthalpic contributions alone (as reflected by T_s) were sufficient to account for the increase in stability.

Comparative analysis of proteins from thermophilic and mesophilic organisms offers a unique model system in which to study the determinants of thermostability. Examination of homologous protein pairs on both sequence and structural levels has led to the formation of many hypotheses to explain enhanced thermostability of proteins from thermophilic origin. The majority of the studies have focused on the impact of individual amino acids on the structure of the protein and their effects on features such as salt-bridges (*1*, *2*), surface-to-volume ratio (*3*), surface loop size, packing density (*4*), differing amino acid distributions (*5*, *6*), or electrostatic contributions (*7*, *8*). Although it is clear that these factors may play a stabilizing role in specific examples, none could be considered paradigms for the increased stability of most proteins from thermophiles. In fact, it has been more recently proposed that the higher thermostability of thermophilic proteins may be due to small optimizations of the proteins as a whole, rather than as a result of large changes in structural features. These optimizations, although acting at a local level on the structure, likely have their impact on global thermodynamic parameters such as the changes in enthalpy, entropy, and heat capacity upon protein unfolding.

Despite the dramatic increase in the number of crystal structures and mutational studies of thermophilic proteins in recent years, experiments that thoroughly examine the folding thermodynamics of thermostable proteins in comparison to their mesophilic homologues are relatively few. In this paper we examine the global thermodynamic parameters for two CheY homologues, one from the thermophile *Thermotoga maritima* (TmY) and the other from the mesophile *Bacillus subtilis* (BsY).

CheY is a small single-domain protein with an α/β parallel fold and is the response regulator in the bacterial chemotaxis pathway (*9*). Figure 1A shows that both BsY and TmY contain 120 amino acids and are very related at the sequence level (71% identical, 79% similar). Despite this significant degree of homology, TmY is substantially more stable than BsY. Figure 2 shows a plot of the fraction of unfolded protein as a function of temperature. In the absence of denaturant TmY has no observable unfolded population below 95 °C, whereas BsY is 50% unfolded at 55 °C. Even in the presence of 1.5 M GdmCl,[1] the midpoint of the unfolding transition for TmY is nearly 20 °C higher than that of BsY in the absence of denaturant. The change in the free energy of unfolding for the two is quite large, with TmY having a free energy of unfolding at 25 °C, ΔG_{298}, of ∼9.5 kcal/mol and BsY having a ΔG_{298} of ∼3.2 kcal/mol. Due to the high level of similarity in primary structure and highly conserved tertiary structure within the CheY family (*9*), it is unlikely that there are systematic changes between the two structures which account for this stability difference. The structure of TmY has been solved to high resolution and was found to be very similar to that of CheY from *Escherichia coli* (backbone RMSD 1.87 Å) (*10*), although their primary sequences

* Author to whom correspondence should be addressed [telephone (541) 346-5882; fax (541) 346-5891; e-mail fwd@nmr.uoregon.edu].

[1] Abbreviations: BsY, *Bacillus subtilis* CheY; TmY, *Thermotoga maritima* CheY; EcY *Escherichia coli* CheY; T_s, temperature of maximal stability; ΔH_s, change in molar enthalpy measured at T_s; $\Delta H'$, change in molar enthalpy at a specific reference temperature T_0; $\Delta S'$, change in molar entropy at a specific reference temperature T_0; ΔC_p, change in heat capacity upon protein unfolding; TRIS, tris(hydroxymethyl)aminomethane; CD, circular dichroism; T_m, thermal melting temperature; T4L, bacteriophage T4 lysozyme; RNase H*, ribonuclease HI; GdmCl, guanidinium chloride; ΔG_{298}, free energy of folding at 298 K; m_g, denaturant-dependent slope of ΔG; a, temperature-dependent slope of m_g; ϵ, observed CD signal; ϵ_f, CD signal of the fully folded form of the protein; ϵ_u, CD signal of the fully unfolded form of the protein; $s_{f,t}$, temperature-dependent slope of the folded baseline of the chemical unfolding experiments; $s_{u,t}$, temperature-dependent slope of the unfolded baseline of the chemical unfolding experiments; $s_{f,d}$, denaturant-dependent slope of the folded baseline; $s_{u,d}$ denaturant-dependent slope of the unfolded baseline.

10.1021/bi010665t CCC: $20.00
Published on Web 10/05/2001

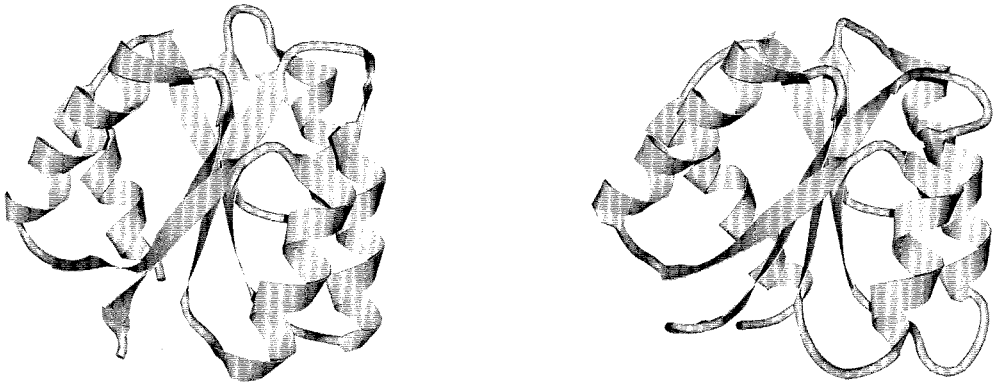
A

```
TmY MGKRVLIVDD AAFMRMMLKD IITKAGYEVA GEATNGREAV EKYKELKPDI   50
BsY MAHRILIVDD AAFMRMMIKD ILVKNGFEVV AEAENGAQAV EKYKEHSPDL   50

VTMDITMPEM NGIDAIKEIM KIDPNAKIIV CSAMGQQAMV IEAIKAGAKD  100
VTMDITMPEM DGITALKEIK QIDAQARIIM CSAMGQQSMV IDAIQAGAKD  100

FIVKPFQPSR VVEALNKVSK  120
FIVKPFQADR VLEAINKTLN  120
```

B



T. maritima CheY *E. coli* CheY

FIGURE 1: (A) Sequence alignment of CheY from *T. maritima* (top) and *B. subtilis* (bottom). Identical residues are shown in black type. Those that differ in the two proteins are shown in gray type. (B) Ribbon diagrams of CheY from *T. maritima* (left) and *E. coli* (right).

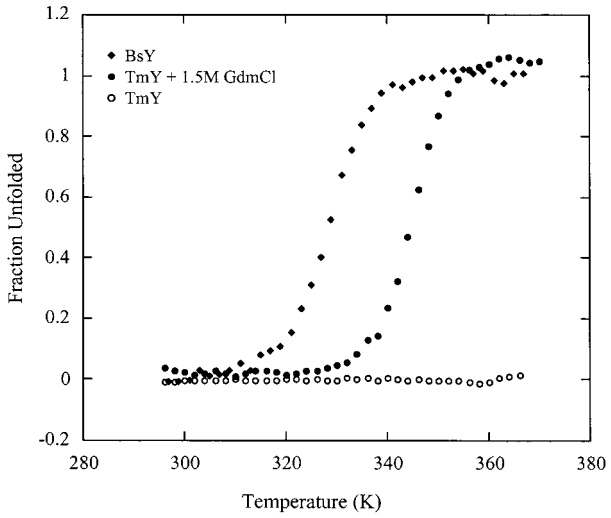


FIGURE 2: Thermal unfolding of BsY (♦), TmY (○), and TmY in 1.5 M guanidine hydrochloride (●) plotted as fraction unfolded versus temperature. The T_m for BsY is 328 K (55 °C) and for TmY in 1.5 M guanidine hydrochloride is 345 K (72 °C). Without guanidine hydrochloride, TmY remains folded at 100 °C.

are far less related (approximately 30% identical, 60% similar). Figure 1B shows a ribbon diagram of the structures of TmY and EcY. The structural differences in these two homologues are minor and are primarily located in the loop regions of the molecules. These regions tend to be less important to the stability than the core region, which is structurally highly conserved in the CheY family.

To fully understand the basis of the increased thermostability of proteins of thermophilic origin such as TmY, we have determined the contributions of enthalpy at the temperature of maximum stability (T_s) and the change in heat capacity upon protein unfolding (ΔC_p) make to the stability of the thermophilic protein TmY and two mesophilic proteins, BsY and lysozyme from the bacteriophage T4 (T4L). We chose T4 lysozyme because it has a similar value of ΔG_{298} yet undergoes thermally induced unfolding at much lower temperatures than TmY. The contributions of $\Delta H'$, T_s, and ΔC_p to the thermal stability of each protein and the implications of the thermodynamic differences between these proteins are discussed.

We find that changes in both $\Delta H'$ and ΔC_p are important contributors to the thermal stability of TmY as compared to BsY. A similar decrease in ΔC_p was reported for the mesophilic and thermophilic forms of RNase H* (*16*), suggesting that reducing ΔC_p may be an important strategy for producing heat stable proteins.

THERMODYNAMIC BACKGROUND

For a two-state N→U unfolding reaction in which ΔC_p is independent of temperature, the equations governing the dependence of ΔG, ΔH, and ΔS with respect to temperature are well established. Here we define ΔG, ΔH, and ΔS as the changes in molar free energy, molar enthalpy, and molar entropy at any temperature. Then

$$\Delta G = \Delta H - T\Delta S \quad (1)$$

$$\Delta H = \Delta H' + \Delta C_p(T - T_0) \quad (2)$$

$$\Delta S = \Delta S' + \Delta C_p \ln(T/T_0) \quad (3)$$

$$\Delta G = \Delta H' + \Delta C_p(T - T_0) - T[\Delta S' + \Delta C_p \ln(T/T_0)] \quad (4)$$

where $\Delta H'$ and $\Delta S'$ are the values of ΔH and ΔS at the standard temperature T_0.

Given that the temperature of maximal stability, T_s, is the temperature at which the net entropy of folding is zero, we can rewrite eq 4 as

$$\Delta G = \Delta H_s + \Delta C_p(T - T_s) - T\Delta C_p \ln(T/T_s) \quad (5)$$

where our reference temperature is taken to be T_s.

In addition, the linear free energy model (*11*) predicts that ΔG varies with respect to denaturant concentration, [D], as

$$\Delta G_d = \Delta G + m_g[D] \quad (6)$$

where m_g is defined as the denaturant-dependent slope of ΔG. In the case in which m_g is linearly dependent on temperature, ΔG_d can be expressed as a function of temperature and denaturant as

$$\Delta G_d = \Delta G + [D][m_g + a(T - T_s)] \quad (7)$$

Finally

$$\Delta G_d = \Delta H_s + \Delta C_p(T - T_s) - T\Delta C_p \ln(T/T_s) + [D][m_g + a(T - T_s)] \quad (8)$$

STRATEGIES FOR INCREASING THERMOSTABILITY

From the explicit relationship for the free energy of protein folding as a function of temperature where the reference temperature is taken as T_s (eq 5), it is apparent that there are three primary parameters which influence the free energy of unfolding as a function of temperature: T_s, ΔH_s, and ΔC_p. Each of these affect the character of the free energy−temperature curve, and each parameter has a different, separable effect on the curve. Figure 3A shows an example of a free energy versus temperature profile for a theoretical protein. By examination of this representation of eq 5, we can discover the following: (1) Because the temperature at which the protein is most stable in its folded form, T_s, is determined by the temperature at which the net entropy change between the folded and unfolded forms of the protein is zero, on a plot of free energy versus temperature, entropy dictates the position of the maximal stability of the protein with respect to the temperature axis. (2) The magnitude of the stability at T_s is dictated entirely by the enthalpy term, ΔH_s. (3) The degree to which the free energy of unfolding diminishes as a function of temperature is approximately inversely related to $\sqrt{\Delta C_p}$. Thus, the lower the difference in heat capacity between the folded and unfolded states of the protein is, the broader the free energy−temperature curve will be.

That these three primary parameters affect the free energy as a function of temperature in distinct ways suggests several thermodynamic mechanisms for the stabilization of a protein at high temperatures (*12*, *13*). Figure 3B shows several potential strategies for increasing the thermostability of a protein. The solid line shows the free energy profile as a function of temperature for a typical mesophilic protein. For a protein from a thermophilic organism, stability at moderate temperatures and the temperature at which cold unfolding occur are of little significance. Thus, the important feature of this curve with regard to creating a thermostable protein

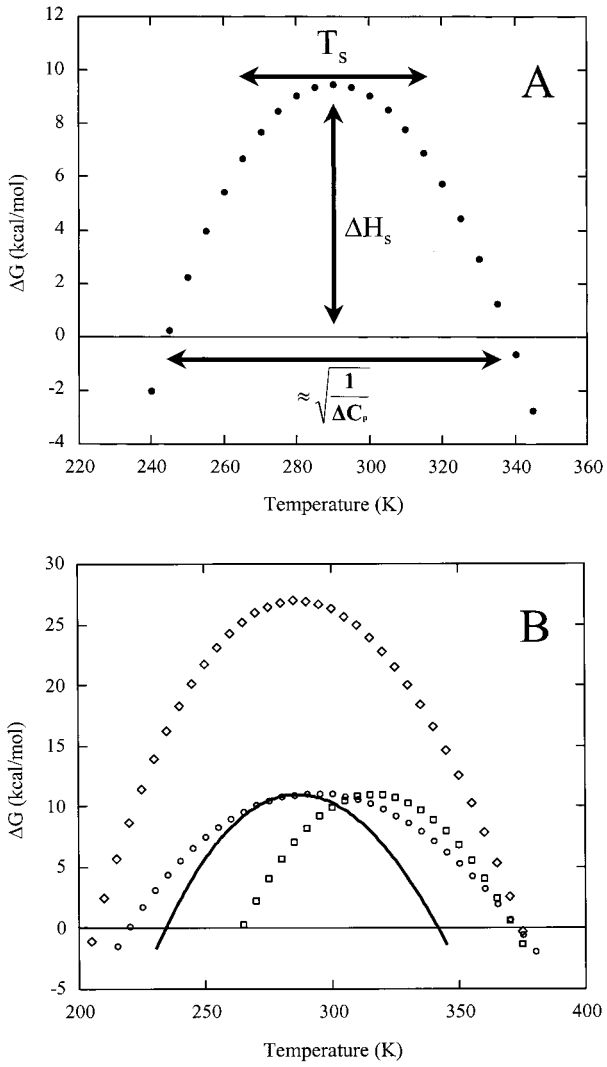


FIGURE 3: (A) Free energy of unfolding plotted as a function of temperature for a typical mesophilic protein. Depicted are the effects of ΔS, ΔH_s, and ΔC_p on the shape and position of the curve. (B) Free energy of unfolding plotted as a function of temperature for a typical mesophilic protein (solid line) and thermophilic proteins stabilized through entropy effects (□), enthalpy effects (◇), and ΔC_p effects (○).

is the temperature at which the free energy of unfolding reaches zero at elevated temperatures. There are three primary strategies by which to increase this transition temperature. Because the temperature of maximal stability for a protein, T_s, is dependent solely on entropic considerations, decreasing the magnitude of ΔS for the folded form of the protein relative to the unfolded form will shift T_s to higher temperatures. To maintain a constant ΔH_s value, this reduction in ΔS must be accompanied by a corresponding and compensating decrease in $\Delta H'$. This is a form of entropy−enthalpy compensation often seen in order−disorder transitions. If the value of $\Delta H'$ does not fully compensate for the change in $\Delta S'$, the height of the maximum will rise or fall. These compensatory changes in enthalpy and entropy will increase the T_m for both thermally induced denaturation and cold unfolding but leave ΔH_s and ΔC_p unaffected (Figure 3B, squares). Conversely, increasing the magnitude of ΔH_s, without changing $\Delta S'$, will result in a similar free energy versus temperature profile but with higher ΔG_d values at all temperatures (Figure 3B, diamonds), thus resulting in an

exceptionally stable protein at moderate temperatures as well as an increase in temperature of thermal unfolding. (e.g., refs *14* and *15*). Finally, by decreasing ΔC_p, the rate at which free energy of unfolding decreases as a function of temperature will be lowered, resulting in a shallower dependence of ΔG_d on temperature (Figure 3B, circles). This will increase the temperature of thermal denaturation without affecting either ΔH_s or T_s.

We find that although TmY exhibits aspects of all of these mechanisms, it is the decrease in ΔC_p that is most significant in the increased thermostability of TmY. Interestingly, a similar decrease in ΔC_p was observed for a mesophilic and thermophilic pair of RNase H proteins, suggesting that changes in ΔC_p may be a relatively common way to generate stability in proteins from thermophiles (*16*).

MATERIALS AND METHODS

Protein Sample Preparation for TmY. TmY was overexpressed in cell line XL1-blue (Stratagene) in LB−H medium. Purification was done as follows: Cells were spun down at 5000*g* for 20 min, resuspended in 50 mM TRIS/5 mM $MgCl_2$ pH 8.0, and lysed with a French press. The lysate was spun down at 14000*g* for 30 min. The supernatant was removed and incubated at 80 °C for 15 min. The precipitate was then spun down at 14000*g* for 15 min. The supernatant was removed and passed over a DE52 anion exchange column equilibrated in lysis buffer. The flow-through was collected and dialyzed overnight against a 25 mM sodium acetate, pH 3.5, buffer and passed over an SP Sepharose cation exchange column. TmY was eluted with a 0−2 M NaCl gradient. Fractions were collected and dialyzed against 50 mM sodium phosphate, pH 6.0.

Protein Sample Preparation for BsY. BsY was overexpressed in BL-21 λDE3 cells in LB−H medium. Purification was done as follows: Cells were resuspended in 8 M urea, 100 mM sodium phosphate, and 10 mM TRIS, pH 8.0, and stirred at room temperature for 1 h. The solution was then spun down at 14000*g* for 15 min. The supernatant was removed and passed over a nickel agarose column (Qiagen). The column was washed with lysis buffer followed with 8 M urea, 100 mM sodium phosphate, and 10 mM TRIS, pH 5.6. BsY was then eluted in 8 M urea/200 mM glacial acetic acid. Fractions were collected, pooled, and refolded overnight into 50 mM TRIS, pH 8.8/10 mM 2-mercaptoethanol. The refolded protein was then dialyzed into 50 mM sodium phosphate, pH 8.0.

Protein Sample Preparation for Lysozyme. WT* T4 lysozyme was prepared according to previously described methods (*17*).

Denaturant-Induced Unfolding Curves. Samples for chemical denaturation of TmY were prepared using serial dilutions of a protein solution composed of 20 μg/mL protein/50 mM sodium phosphate, pH 6.0, with an identical solution of protein containing a high concentration of guanidinium chloride. Samples for BsY chemical unfolding experiments were prepared similarly and were composed of 20 μg/mL protein, 50 mM sodium phosphate, pH 8.0, 100 mM KCl, 0.5 mM DTT, and appropriate concentrations of guanidine. Samples for T4 lysozyme were composed of 20 μg/mL protein and 50 mM sodium phosphate, pH 6.0. Samples were equilibrated at 18 °C for 48 h prior to data acquisition. For each point in the chemical unfolding experiment, the CD signal was measured at 12 temperatures varying from 18 to 50 °C. At each temperature, the sample was allowed to equilibrate until no changes in the CD signal were observed. CD measurements were taken on a Jasco J-720 spectrophotometer with a Peltier temperature-controlled cuvette holder. Each CD measurement was taken at 223 nm and was signal averaged over 4 min in a 1 cm path length cell.

Thermal Unfolding Curves. Samples for thermal unfolding experiments contained 20 μg/mL protein in the buffer systems detailed above, with desired amounts of guanidine added. Samples were heated at 1 °C/min, and the CD signal was averaged over 16 s. Thermal reversibility for TmY and T4L was excellent in all cases, and for BsY thermal reversibility was good provided samples were rapidly cooled after the thermal transition. Extended periods of time at higher temperatures led to apparent degradation of the BsY samples and reversibility suffered.

Data Analysis. Denaturant-induced unfolding curves measured at 12 temperatures between 18 and 50 °C were fit globally to a two-state F→U unfolding model as a function of temperature using the equation

$$\epsilon = \epsilon_f + s_{f,d}[D] + s_{f,t}T + (\epsilon_u + s_{u,d}[D] + s_{u,t}T)\exp[-(\Delta G_d/RT)])/1 + \exp(-\Delta G_d/RT) \quad (9)$$

where ϵ represents the observed CD signal, ϵ_f and ϵ_u represent the CD signal of the fully folded and unfolded forms of the protein, $s_{f,t}$ and $s_{u,t}$ represent the temperature-dependent slopes of the folded and unfolded baselines of the chemical unfolding experiments, $s_{f,d}$ and $s_{u,d}$ represent the denaturant-dependent slopes of the folded and unfolded baselines, and ΔG_d is given by eq 8 above.

The program Mathematica and the Levenberg Marquardt algorithm were used for fitting the required parameters. Six parameters were related to defining the slope and intercept of the base planes of the data for the folded and unfolded forms, whereas the remaining five ($\Delta H'$, $\Delta S'$, ΔC_p, m_g, and a) described the temperature and denaturant dependence of ΔG_d. All parameters were fit with the exception of m_g, which was determined on the basis of the denaturant dependence of ΔG calculated from a series of thermal unfolding experiments taken in the presence of low denaturant concentrations. Trial data fits were done allowing m_g to be fit as well, and agreement was good with the experimentally determined value, although errors in the fit were slightly lower using the calculated value for m_g.

RESULTS

Global Stability Parameters for TmY, BsY, and T4 Lysozyme. The magnitude of ΔG_d for TmY, BsY, and T4L was determined on the basis of denaturant-induced unfolding curves at 12 temperatures between 18 and 50 °C using circular dichroism as a probe of secondary structure. Samples of individual denaturant-induced unfolding experiments for TmY (●), BsY (◆), and T4L (■) are shown in Figure 4. Each protein showed two-state reversible unfolding behavior under the solvent conditions used. Using the modified Gibbs−Helmholtz relationship (eq 8), we have determined ΔC_p and the contribution of ΔH and ΔS for all three proteins as a function of temperature. The free energy of unfolding

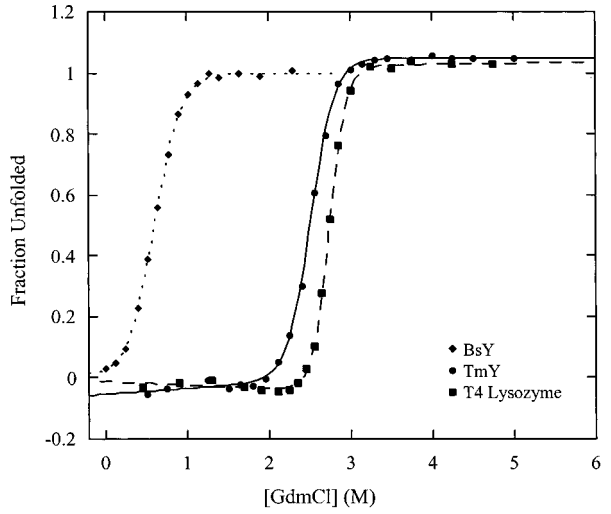


FIGURE 4: Guanidinium chloride induced unfolding of BsY (◆), TmY (●), and T4 lysozyme (■) at 30 °C. ΔG_{unf} energies are 3.1 kcal/mol for BsY, 9.5 kcal/mol for TmY, and 10.0 kcal/mol for T4L.

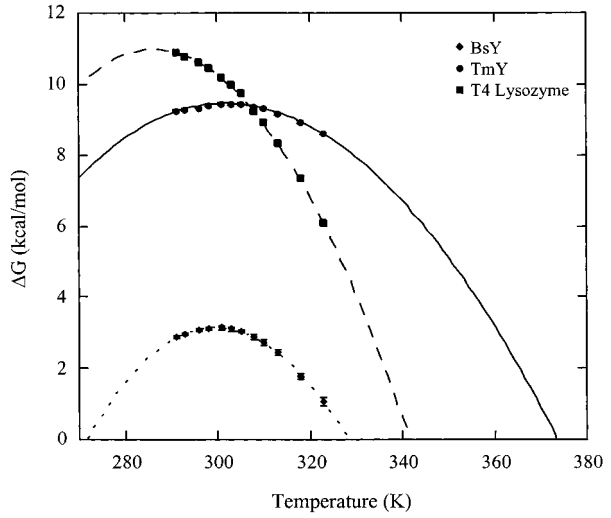


FIGURE 5: Measured values of ΔG_{unf} for BsY (◆), TmY (●), and T4 lysozyme (■) plotted as a function of temperature. The curves depicted are plots of the Gibbs−Helmholtz equation using the values for $\Delta H'$, $\Delta S'$, and ΔC_p derived from the global fits of the data. Estimated errors for measurement of each ΔG data point range from 0.03 to 0.04 kcal/mol for TmY and T4L and from 0.03−0.08 kcal/mol for BsY. Error bars are indicated on the plot. For TmY and T4L, the size of the error bars is less than the print size of each data point.

determined from these solvent-induced unfolding experiments for TmY (●), BsY (◆), and T4 lysozyme (■) is plotted as a function of temperature in Figure 5. The thermodynamic values determined from these fits are reported in Table 1. We have included data from T4 lysozyme both as an example of the behavior of a relatively stable mesophilic protein and as a control to test our data analysis procedure. Although the range of temperatures over which data were acquired was somewhat limited, the values predicted by our fits agree well with the unfolding temperature T_m and $\Delta H°$ (molar free enthalpy at T_m), as determined by previous thermal unfolding experiments. Our fit was done without inclusion of either T_m or $\Delta H°$ data from thermal unfolding experiments, so these independent measurements

Table 1

	T. maritima CheY	*B. subtilis* CheY	T4 lysozyme
T_m (K)	374	328	341
ΔG_s (kcal/mol)	9.54	3.14	10.98
T_s (K)	302	300	287
$\Delta H°$ (kcal/mol·K)	94.47 ± 2.2	70.47 ± 5.4	131.7 ± 2.3
$\Delta S°$ (kcal/mol·K)	.252±.006	.215±.016	.386±.007
ΔC_p (kcal/mol·K)	1.17±.05	2.34±.20	2.18±.09

serve as an indication of the quality of the model used and the fit of the model to the data. The T_m and $\Delta H°$ for BsY predicted here are 55 °C and 65.6 kcal/mol, respectively; from thermal unfolding experiments they have been determined to be 55.5 °C and 65.9 kcal/mol. Previously published values for T_m and $\Delta H°$ for T4L are 65.3 °C and 130 kcal/mol; the data presented here predict 68.6 °C and 132 kcal/mol. For TmY we are unable to directly compare these values, as the thermal unfolding transition is not observable in the absence of denaturant; however, the predicted T_m of 101 °C and $\Delta H°$ of 96.9 kcal/mol are in good agreement with values extrapolated from thermal unfolding experiments done in incrementally decreasing denaturant concentrations, which predict a T_m of 98 °C and a $\Delta H°$ of 97.0 kcal/mol.

The value we have measured for the ΔC_p of T4 lysozyme agrees closely with that previously published by Chen and Schellman (*19*), using a completely different method to establish the value of ΔC_p. The value we measure for BsY agrees well with previously published estimates for CheY from the mesophile *E. coli* (*20*). The ΔC_p value for the thermophilic TmY is substantially less than that of either mesophilic protein and was determined to be 1.14 ± 0.04 and 1.19 ± 0.05 kcal/mol·K in two separate trials.

An additional concern that must be addressed is the two-state nature of the unfolding transition for each of these proteins and the effect of non-two-state behavior on the measurement of the thermodynamic parameters presented here. T4 lysozyme has been extensively characterized, and it has been shown that the unfolding of lysozyme closely follows a two-state process under these conditions. We have examined TmY and BsY using both chemical and thermal unfolding experiments, and TmY has been studied using NMR structural techniques. In no instance has any evidence of non-two-state behavior been observed. Multistate unfolding behavior in chemical unfolding experiments is known to decrease the apparent free energy of unfolding through a decreased apparent *m* value, whereas the midpoint of the unfolding transition remains unaffected. Thus, unaccounted for multistate unfolding behavior in our analysis will lead to an underestimation of ΔH_s, which in turn will cause underestimation of the T_m of the protein for both thermal and cold unfolding. The data presented above from thermal unfolding experiments show that this is not the case for either TmY or BsY with respect to thermal unfolding. In addition to the data for the thermal unfolding of TmY, the midpoint of the cold unfolding transition for TmY has been measured to be ~3 °C in 2 M GdmCl (data not shown). Our analysis here predicts this temperature to be 0 °C. The close agreement of the predicted cold unfolding temperature using the two-state model strongly argues that the system follows this model. Although minor details of the curve distant from the two melting points may vary for TmY, the breadth of

the free energy−temperature curve, and therefore the ΔC_p for TmY, has been well determined.

Temperature of Maximal Stability for TmY, BsY, and T4L. The temperatures of maximal stability for both TmY and BsY are approximately equal, 29 °C for TmY and 27 °C for BsY, whereas the T_s for lysozyme was lower at 14 °C. These temperatures are similar to the T_s values that have been reported for other proteins from both thermophilic and mesophilic hosts, with values ranging from 7 to 44 °C (*14*, *18*). The nearly identical T_s values for TmY and BsY indicate that the strategy of shifting the stability curve and T_s of the protein to a higher temperature, without perturbing ΔH_s, is not significant in the thermal stabilization of TmY, as compared to BsY. This is not a surprising result, as T_s is determined by the temperature at which the net entropy of protein folding is zero. The primary contributions to the entropy of protein folding are related to differences in the configurational entropy and protein−solvent interactions of the unfolded form of the protein as compared to the folded form. The determinants of these entropic contributions are based on the physical and chemical properties of the protein and the solvent it is in contact with. Thus, it appears that in this case neither of these components is easily altered by the magnitude necessary for creating a thermophilic protein within the context of a protein unfolding in an aqueous solvent.

Enthalpic Contributions to Protein Stability. Improving protein thermostability through increased enthalpic contributions is a strategy that has been cited previously in the stabilization of thermophilic proteins (*14*, *15*). This mechanism is responsible for a portion of the thermostabilization of TmY, as can be seen in Figure 5 by the fact that the free energy of unfolding for TmY is significantly greater than that of BsY at all temperatures (∼7 kcal/mol at 27 °C, the T_s for BsY). This, however, is not the entire basis of thermostability in TmY. This can be seen by comparison of the free energy of unfolding curves for TmY and T4L in Figure 5. T4L (■) and TmY (●) have similar stabilities near the T_s of TmY, yet lysozyme unfolds at a comparatively modest 65 °C.

Generating an increased enthalpic stabilization may be a viable strategy in certain cases; however, the magnitude of the ΔH increase necessary for creation of a very thermostable protein is much larger than might be expected. For example, in T4 lysozyme, assuming an unaltered $\Delta S'$ and ΔC_p contribution, increasing the thermostability of the protein to match TmY through entirely enthalpic contributions would require a ΔH_s of nearly 30 kcal/mol (Figure 3B, diamonds). Even in the case of BsY with its higher T_s, the ΔH_s contribution would still need to be nearly 20 kcal/mol. Thus, it appears that generating an enthalpic stabilization of the necessary magnitude in the proteins of the CheY family is not a viable strategy.

ΔC_p Values of TmY, BsY, and T4L. The differences in the thermostability of the thermophile TmY and the mesophilic proteins studied here are a direct result of differences in ΔC_p. Both mesophilic proteins presented here have similar ΔC_p values at 2.18 kcal/mol·K for T4L and 2.34 kcal/mol·K for BsY. The ΔC_p value for the thermophilic TmY is substantially less than that of either mesophilic protein and was determined to be 1.17 kcal/mol·K.

This apparently small decrease in the ΔC_p of TmY results in significant increases in the T_m of the protein. This is demonstrated by the fact that if the ΔH_s contribution of BsY is increased to match that of TmY while maintaining their respective ΔC_p values, we find that TmY is still 24 °C more thermostable than BsY. Conversely, if we decrease the ΔH_s contribution of TmY to match that of BsY, due to its smaller ΔC_p TmY is still 15 °C more stable than BsY and 5 °C more stable than T4L, even though the reduced ΔH contribution of TmY is almost 40 kcal/mol less than that of T4L at the T_s of TmY.

From inspection of eq 5, we can see that the entropic contribution to ΔG is dependent primarily on ΔC_p. Thus, a smaller value of ΔC_p results in a smaller entropic destabilization at higher temperatures. Therefore, the protein will then require a smaller ΔH stabilization to remain folded. The delicate balance of ΔS and ΔH that TmY exhibits through a wide range of temperatures is a direct consequence of its low ΔC_p value.

DISCUSSION

The ΔC_p value reported here for BsY is consistent with that of CheY from *E. coli* (EcY) as reported by DeKoster et al. (*20*), who estimated ΔC_p for EcY to be at least 2.3 kcal/mol·K. This suggests that the principal factor separating the thermophilic TmY from its mesophilic CheY homologues is, in fact, its reduced ΔC_p value. This agrees with previous results seen in RNase H* (*16*). Both CheY and RNase H* from thermophilic hosts have significantly reduced ΔC_p values in comparison to their mesophilic counterparts. Although stabilization of thermophilic proteins through reduced ΔC_p contributions has been reported before, it is by no means a universal feature. Although comparisons of ΔC_p in thermophilic and mesophilic versions of RNase H* showed results similar to those presented here (*16*), studies of archeal histones have shown little or no difference in the ΔC_p values of the thermophilic and mesophilic proteins (*18*). The histones are smaller and highly related in sequence, particularly with respect to their hydrophobic core; therefore, because ΔC_p is thought to result largely from changes in hydration of buried residues (*21*−*23*), this is not an unexpected result. With their highly related core sequences and small molecular mass, the histones are unlikely to exhibit significant differences in ΔC_p. In addition, the primary activity of histones, DNA binding, is likely less dependent on dynamic conformational fluctuations than are the enzymatic activities of RNase or conformational dependent signaling of CheY. Thus, simply making a significantly more stable molecule without concern for its dynamic properties is a viable strategy in the histone case. Conversely, increasing stability through a decreased ΔC_p may be a mechanism available to a wider range of thermophilic proteins than is stabilization through primarily entropic or enthalpic contributions.

Past studies have noted a correlation between protein size and ΔC_p (*21*). Whereas T4L fits this correlation well, BsY and TmY show significant deviations from this correlation. This, in addition to the fact that TmY and BsY are virtually identical in size but have ΔC_p values that differ significantly, indicates that although the heat capacity of a polypeptide chain has an important contribution due to the chemical character of the polypeptide chain and its interactions with

solvent in the unfolded state, it also depends significantly upon detailed structural or packing interactions in the folded state. Noncovalent, nonsolvent-related contributions to ΔC_p have been estimated to contribute <3% to the total ΔC_p for a protein (*22*). Clearly, in the case of TmY, significantly larger contributions from these interactions have been realized.

Lowered ΔC_p can be a powerful tool for stabilizing proteins from thermophilic hosts. The alternate strategy of generating a protein with an extremely large enthalpy of folding in order to offset large entropic costs will likely result in a very stable, yet likely dynamically limited, protein. In the CheY family of proteins, this is an unused compromise, as dynamic fluctuations in CheY are necessary for biological activity. If we compare the ΔG values for TmY and BsY at the optimal growth temperatures for their hosts, we find that TmY is only ~2 kcal/mol more stable, suggesting that a heavily stabilized CheY molecule is unlikely to be biologically active. A similar correlation for RNase H* has been observed (*16*). Furthermore, if we examine the temperature dependence of free energy for a protein stabilized primarily through ΔH contributions as compared to a protein stabilized through ΔC_p, we see a significant difference at temperatures near the unfolding transition (Figure 3B). Because the temperature dependence of the ΔH-stabilized protein (◇) is significantly greater than that of the ΔC_p-stabilized protein (○), minor fluctuations in temperature produce large variations in stability. These dramatic stability differences will affect properties of the protein such as enzymatic activity, dynamic behavior, and the population of the protein accessing partially unfolded forms. Significant variations in these characteristics are likely to negatively impact the ability of the protein to effectively complete its role in the cell and interfere with the ability of a thermophilic organism to appropriately regulate cellular functions in response to environmental fluctuations in temperature.

REFERENCES

1. Yip, K. S., Britton, K. L., Stillman, T. J., Lebbink, J., de Vos, W. M., Robb, F. T., Vetriani, C., Maeder, D., and Rice, D. W. (1998) *Eur. J. Biochem. 255*, 336−346.
2. Elcock, A. H. (1998) *J. Mol. Biol. 284*, 489−502.
3. Tanner, J. J., Hecht, R. M., and Krause, K. L. (1996) *Biochemistry 35*, 2597−2609.
4. Wallon, G., Kryger, G., Lovett, S. T., Oshima, T., Ringe, D., and Petsko, G. A. (1997) *J. Mol. Biol. 266*, 1016−1031.
5. Gromiha, M. M., Oobatake, M., and Sarai, A. (1999) *Biophys. Chem. 82*, 51−67.
6. Bogin, O., Peretz, M., Hacham, Y., Korkhin, Y., Frolow, F., Kalb, A. J., and Burstein, Y. (1998) *Protein Sci. 7*, 1156−1163.
7. Perl, D., Mueller, U., Heinemann, U., and Schmid, F. X. (2000) *Nat. Struct. Biol. 7*, 380−383.
8. Xiao, L., and Honig, B. (1999) *J. Mol. Biol. 289*, 1435−1444.
9. Volz, K. (1993) *Biochemistry 32*, 11741−11753.
10. Usher, K. C., de la Cruz, A. F., Dahlquist, F. W., Swanson, R. V., Simon, M. I., and Remington, S. J. (1998) *Protein Sci. 7*, 403−412.
11. Schellman, J. A. (1987) *Annu. Rev. Biophys. Biophys. Chem. 16*, 115−137.
12. Rees, D. C., and Adams, M. W. (1995) *Structure 3*, 251−254.
13. Jaenicke, R., and Bohm, G. (1998) *Curr. Opin. Struct. Biol. 8*, 738−748.
14. Knapp, S., Karshikoff, A., Berndt, K. D., Christova, P., Atanasov, B., and Ladenstein, R. (1996) *J. Mol. Biol. 264*, 1132−1144.
15. Beadle, B. M., Baase, W. A., Wilson, D. B., Gilkes, N. R., and Shoichet, B. K. (1999) *Biochemistry 38*, 2570−2576.
16. Hollien, J., and Marqusee, S. (1999) *Biochemistry 38*, 3831−3836.
17. Muchmore, D. C., McIntosh, L. P., Russell, C. B., Anderson, D. E., and Dahlquist, F. W. (1989) *Methods Enzymol. 177*, 44−73.
18. Li, W. T., Grayling, R. A., Sandman, K., Edmondson, S., Shriver, J. W., and Reeve, J. N. (1998) *Biochemistry 37*, 10563−10572.
19. Chen, B. L., and Schellman, J. A. (1989) *Biochemistry 28*, 685−691.
20. DeKoster, G. T., and Robertson, A. D. (1995) *J. Mol. Biol. 249*, 529−534.
21. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) *Protein Sci. 4*, 2138−2148.
22. Gomez, J., Hilser, V. J., Xie, D., and Freire, E. (1995) *Proteins 22*, 404−412.
23. Cooper, A. (2000) *Biophys. Chem. 85*, 25−39.

BI010665T