A Thermodynamic Comparison of Mesophilic and Thermophilic Ribonucleases H[†]

Julie Hollien and Susan Marqusee*

Department of Molecular and Cell Biology, University of California, Berkeley, California 94720-3206

Received November 10, 1998; Revised Manuscript Received January 28, 1999

ABSTRACT: The mechanisms by which thermophilic proteins attain their increased thermostability remain unclear, as usually the sequence and structure of these proteins are very similar to those of their mesophilic homologues. To gain insight into the basis of thermostability, we have determined protein stability curves describing the temperature dependence of the free energy of unfolding for two ribonucleases H, one from the mesophile *Escherichia coli* and one from the thermophile *Thermus thermophilus*. The circular dichroism signal was monitored as a function of temperature and guanidinium chloride concentration, and the resulting free energies of unfolding were fit to the Gibbs−Helmholtz equation to obtain a set of thermodynamic parameters for these proteins. Although the maximal stabilities for these proteins occur at similar temperatures, the heat capacity of unfolding for *T. thermophilus* RNase H is lower, resulting in a smaller temperature dependence of the free energy of unfolding and therefore a higher thermal melting temperature. In addition, the stabilities of these proteins are similar at the optimal growth temperatures for their respective organisms, suggesting that a balance of thermodynamic stability and flexibility is important for function.

Proteins from thermophilic organisms offer a unique opportunity to study the determinants of thermostability (for review, see refs *1*−*3*). Although these proteins are often very similar in sequence and structure to their mesophilic homologues, they are much more resistant to thermal denaturation and inactivation. Efforts to determine the origin of this thermostability have led to several hypotheses, such as stabilization by an increased number of ionic interactions, an increased extent of hydrophobic surface burial, an increased number of prolines, and smaller surface loops (*2*). Although evidence for these and other modes of stabilization can be found in specific examples, none apply to all or even most thermostable proteins. If there are general rules for how thermophilic proteins attain their stability, it is clear that they do not lie in individual interactions; they may lie in properties of the whole molecule, such as how the stability is distributed and coupled throughout the structure or how it is divided between enthalpy and entropy.

In this paper, we compare the global stability parameters for two RNase H[1] homologues, one from the mesophilic bacterium *Escherichia coli* and one from the thermophilic bacterium *Thermus thermophilus*. RNase H (*4*), which cleaves RNA from RNA−DNA hybrids, is a small, single-domain protein with no prosthetic groups required for folding, making it a good model for studying protein thermodynamics. The *E. coli* enzyme has been well characterized, and is significantly less stable than the *T. thermophilus* enzyme, despite their 52% sequence identity (see Figure 1). In addition, crystal structures (*5*, *6*), shown in Figure 1, reveal very few differences; the rms deviation between α-carbons in secondary structural regions is 0.95 Å (*6*). Mutagenesis studies (*7*−*9*) have uncovered several specific sites in *E. coli* RNase H that, when replaced with the corresponding residues of *T. thermophilus* RNase H, increase the T_m of the resulting protein. These stabilizing substitutions include a cavity-filling mutation (Val74Leu) (*7*), the introduction of a proline in a turn region (His62Pro) (*8*), and the replacement of a lysine in a left-handed helical conformation with a glycine residue (Lys95Gly) (*9*). The effects of these substitutions are roughly additive (*10*), indicating that they contribute individually to the enhanced thermostability of *T. thermophilus* RNase H. However, they do not appear to account for the entire difference in T_m between the two RNases H.

To fully describe the thermodynamic stability of *T. thermophilus* and *E. coli* RNases H and how this stability depends on temperature, we have determined stability curves for both proteins by measuring the CD signal as a function of temperature and GdmCl concentration. The resulting curves, when fit to the Gibbs−Helmholtz equation, allow the determination of the heat capacity of unfolding (ΔC_p) and describe the enthalpic and entropic contributions to the free energy of unfolding. Stability parameters for *T. thermophilus* and *E. coli* RNases H are compared, and implications for the mechanism of thermostabilization are discussed.

MATERIALS AND METHODS

Design and Construction of a Synthetic Gene for T. thermophilus RNase H. A DNA sequence encoding *T. thermophilus* RNase H was designed to contain useful unique restriction sites and codons preferred for expression

[†] Supported by the NIH (GM50945) and the Helman Faculty Fund.

* To whom correspondence should be addressed. E-mail: marqusee@uclink4.berkeley.edu. Fax: (510) 643-9290.

[1] Abbreviations: RNase H, ribonuclease HI; CD, circular dichroism; GdmCl, guanidinium chloride; NaAc, sodium acetate; EDTA, ethylenediaminetetraacetic acid; DTT, dithiothreitol; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; ΔG_{unf}, free energy of unfolding; T_m, thermal melting temperature; $\Delta H°$, enthaply of unfolding at the T_m; $\Delta S°$, entropy of unfolding at the T_m; ΔC_p, heat capacity of unfolding; ΔASA, difference in accessible surface area between the folded and unfolded states.

10.1021/bi982684h CCC: $18.00
Published on Web 03/04/1999

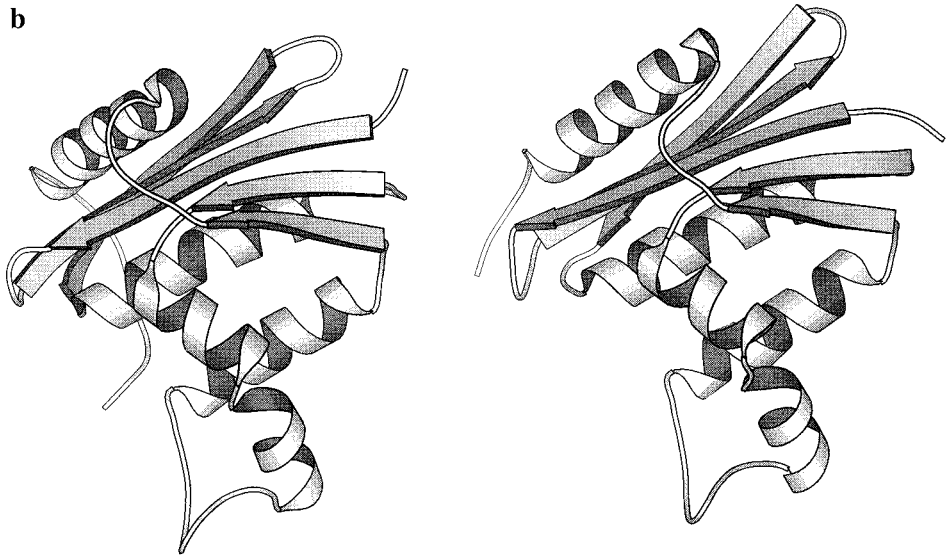
a

```
E. coli RNase HI              MLKQVE IFTDGSCLGN PGPGGYGAIL RYRGREKTFS
T. thermophilus RNase H   MNPSPRKRVA LFTDGACLGN PGPGGWAALL RFHAHEKLLS
                                          *

AGYTRTTNNR MELMAAIVAL EALKEHCEVI LSTDSQYVRQ GITQ-WIHNW KKRGWKTADK
GGEACTTNNR MELKAAIEGL KALKEPCEVD LYTDSHYLKK AFTEGWLEGW RKRGWRTAEG
    *                         *

KPVKNVDLWQ RLDAALGQHQ IKWEWVKGHA GHPENERCDE LARAAAMNPT LEDTGYQVEV
KPVKNRDLWE ALLLAMAPHR VRFHFVKGHT GHPENERVDR EARRQAQSQA KTPCPPRAPT LFHEEA
                                                      *
```

b



E. coli **RNase HI**

T. thermophilus **RNase H**

FIGURE 1: (a) Sequence alignment of wild-type *E. coli* and *T. thermophilus* RNases H. Identical residues are shown in bold, and asterisks indicate cysteine residues in *T. thermophilus* RNase H that were mutated to serine or alanine. (b) Ribbon diagrams of the crystal structures of *E. coli* (*5*) and *T. thermophilus* (*6*) RNases H.

in *E. coli*. We synthesized 13 oligonucleotides covering the entire coding and anticoding strands of the synthetic gene, with 42 base pair overlaps between oligonucleotides of opposite strands. These oligonucleotides were combined in equimolar ratios, annealed, and ligated together. The correct product was amplified by PCR and cloned into a pAED4 vector (*11*), and both strands of the coding region in the resulting plasmid (pJH105) were sequenced. Plasmids encoding three cysteine-free variants of *T. thermophilus* RNase H were created by site-directed mutagenesis of pJH105 using the Kunkel method (*12*). Details of all plasmids are available upon request. The variants are described below using a numbering system in which the amino acid residues of *T. thermophilus* RNase H are numbered sequentially starting at the amino terminus of the protein; this differs from the numbering system used for the sequence in the protein data bank, in which residues are numbered to correspond to residues in *E. coli* RNase H.

Protein Expression and Purification. Plasmids encoding the wild-type or mutant *T. thermophilus* RNases H were transformed into *E. coli* DE3 pLysS (Novagen) cells, and the cells were grown at 37 °C in Luria broth with 100 μg/mL ampicillin. Expression was induced at an OD_{600} of ~0.6 (1 cm cuvette) by the addition of 1 mM IPTG, and cells were grown for an additional 2 h before harvesting by centrifugation. To keep the four free cysteines in the wild-type *T. thermophilus* RNase H reduced, the entire purification for this protein was carried out at pH 5.5 in the presence of fresh 1 mM DTT. Cell pellets were resuspended in 20 mM NaAc, 0.1 mM EDTA, and 1 mM DTT (pH 5.5) and lysed by sonication. The soluble fraction was loaded onto a heparin column and eluted with a linear NaCl gradient (0 to 0.6 M NaCl). Fractions containing RNase H were pooled, loaded onto a Source 15S column [in 20 mM NaAc, 200 mM NaCl, 0.1 mM EDTA, and 1 mM DTT (pH 5.5)], and eluted with a linear NaCl gradient (0.2 to 0.5 M NaCl). Pooled fractions were judged to be pure by SDS−PAGE. Cysteine-free *T. thermophilus* RNase H variants were purified similarly; however, the sonication and heparin column purification steps were carried out at a higher pH [50 mM Tris-HCl, 20 mM NaCl, and 0.1 mM EDTA (pH 8.0)], and DTT was omitted from all buffers. Cysteine-free *E. coli* RNase H (*E. coli* RNase H*) was a gift from J. Kho and was purified as described previously (*13*).

GdmCl-Induced and Thermal Denaturation. Circular dichroism (CD) measurements were carried out on an Aviv 62DS spectrometer with a Peltier temperature-controlled cell holder.

For all samples described below, each CD measurement is an average of the signal at 225 nm for 1 min, in a 1 cm path length cuvette.

For experiments with the wild-type *T. thermophilus* RNase H, samples containing 40 μg/mL protein, 5 mM NaAc, 50 mM KCl, 1 mM DTT (pH 5.5), and the appropriate concentration of GdmCl were allowed to equilibrate overnight before the CD signal was measured. For thermal denaturation, the CD signal was recorded every 3 °C, with a 3 min equilibration time at each temperature, and a 1 min averaging time. Reversibility was determined by returning to the beginning temperature and comparing the CD signal to the premelt signal.

Thermal denaturation experiments with cysteine-free RNase H variants were carried out as described above, and contained 40 μg/mL protein, 5 mM NaAc, and 50 mM KCl (pH 5.5). The cysteine-free *T. thermophilus* RNase H with the highest T_m (C17A/C45S/C67A/C154S) was chosen for further analysis and is referred to as *T. thermophilus* RNase H*.

GdmCl-induced denaturation of *E. coli* and *T. thermophilus* RNases H* was performed by the serial addition of a solution containing a high GdmCl concentration, 40 μg/mL protein, 5 mM NaAc, and 50 mM KCl (pH 5.5) into a starting solution containing the above without GdmCl. For each protein, the same starting and titrant solutions were used to measure the GdmCl-induced equilibrium unfolding at eight or nine different temperatures ranging from 5 to 60 °C. The time necessary to reach equilibrium (3 min to several hours) varied with protein, denaturant concentration, and temperature, and was ensured by monitoring the CD signal over time until no further change was observed.

Denaturation and Stability Curve Data Analysis. We determined free energies of unfolding (ΔG_{unf}) from GdmCl denaturation of *T. thermophilus* and *E. coli* RNases H* at several temperatures, assuming a two-state model and a linear relationship between ΔG_{unf} and [GdmCl] (*14*). Thermal melts were fit to determine the T_m for each protein, using a two-state model and the Gibbs−Helmholtz relationship between ΔG_{unf} and temperature (*15*).

The resulting free energies of unfolding were plotted as a function of temperature. For *T. thermophilus* RNase H*, an additional point at the T_m ($\Delta G_{unf} = 0$) was included; the T_m was not included for *E. coli* RNase H* as thermal denaturation of this protein in the absence of denaturant was not reversible. These data (ΔG_{unf} vs T) were fit to the Gibbs−Helmholtz equation (eq 1)

$$\Delta G_{unf} = \Delta H^\circ - T\Delta S^\circ + \Delta C_p[T - T^\circ - T\ln(T/T^\circ)] \quad (1)$$

where ΔH° and ΔS° refer to the enthalpy and entropy of unfolding at the reference temperature T°, respectively, and ΔC_p is the heat capacity of unfolding, assumed to be constant in this temperature range. When $\Delta G_{unf} = 0$ (i.e., at the T_m), $\Delta S^\circ = \Delta H^\circ/T_m$, so using the T_m as the reference temperature allowed us to fit the data to a modified version of the Gibbs−Helmholtz equation (eq 2) to determine the ΔH°, ΔC_p, and T_m for each protein.

$$\Delta G_{unf} = \Delta H^\circ - T\Delta H^\circ/T_m + \Delta C_p[T - T_m - T\ln(T/T_m)] \quad (2)$$

For *T. thermophilus* RNase H*, the raw data (CD signal as a function of temperature and GdmCl concentration) were also fit globally to a model including all of the above considerations, with the additional assumption that the *m* value does not change with temperature. The program SigmaPlot (Jandel Scientific) was used for this fit, and 10 parameters were included. Six described the intercepts of the unfolded and folded baseplanes (i_u and i_n) and their slopes with respect to temperature ($s_{u,t}$ and $s_{n,t}$) and GdmCl concentration ($s_{u,g}$ and $s_{n,g}$), as shown in eqs 3 and 4. These parameters were used to relate the CD signal of each data point to an equilibrium constant and ΔG_{unf}.

$$\text{unfolded signal} = i_u + s_{u,t}T + s_{u,g}[\text{GdmCl}] \quad (3)$$

$$\text{native signal} = i_n + s_{n,t}T + s_{n,g}[\text{GdmCl}] \quad (4)$$

The other four parameters, ΔH°, ΔC_p, T_m, and *m*, described the dependence of ΔG_{unf} on temperature and GdmCl concentration (eq 5).

$$\Delta G_{unf}(T, [\text{GdmCl}]) = \Delta H^\circ - T\Delta H^\circ/T_m + \Delta C_p[T - T_m - T\ln(T/T_m)] - m[\text{GdmCl}] \quad (5)$$

Accessible Surface Area Calculations. Accessible surface areas for the crystal structures of *E. coli* (*5*, *16*) and *T. thermophilus* (*6*) RNases H were calculated using the program PQMS (*17*) with a probe radius of 1.4 Å. Unfolded surface areas were calculated as the sum of surface areas for individual residues, based on model Ala-X-Ala tripeptide studies by Lee and Richards (*18*).

RESULTS

Stability of Wild-Type T. thermophilus RNase H. We have constructed a synthetic gene encoding the wild-type *T. thermophilus* RNase H (see Materials and Methods). The expressed protein, which contains four cysteines, was purified and characterized at pH 5.5 in the presence of 1 mM DTT, to avoid cysteine oxidation. The stability was evaluated by GdmCl-induced and thermal denaturation, monitored by circular dichroism. The resulting ΔG_{unf} at 25 °C was 13.3 kcal/mol, assuming a two-state transition and a linear dependence of ΔG_{unf} on GdmCl concentration (*14*); and the T_m was 89 °C in the absence of denaturant (77.5 °C in 1 M GdmCl), also assuming a two-state transition, and using the Gibbs−Helmholtz equation to describe the temperature dependence of ΔG_{unf} (*15*). The stability of the protein at 25 °C was also calculated indirectly with the Gibbs−Helmholtz equation, using the T_m of 89 °C and a ΔC_p of 1.8 kcal mol^{-1} K^{-1}, that of the cysteine-free *T. thermophilus* RNase H* (see below). The calculated ΔG_{unf} was identical to that determined by GdmCl denaturation, lending support to the assumptions and analysis.

Cysteine-Free Variants of RNase H. When purified in the absence of a reducing agent, wild-type *T. thermophilus* RNase H tends to unfold in a noncooperative manner by both GdmCl-induced and thermal denaturation, presumably due to unwanted cysteine chemistry. To avoid this problem, as well as to increase expression levels, we created variants in which all four free cysteines were replaced by either serine or alanine. Three variants were made. One contains four serines in place of the cysteines; one contains four alanines, and one contains alanines in place of the two buried cysteines

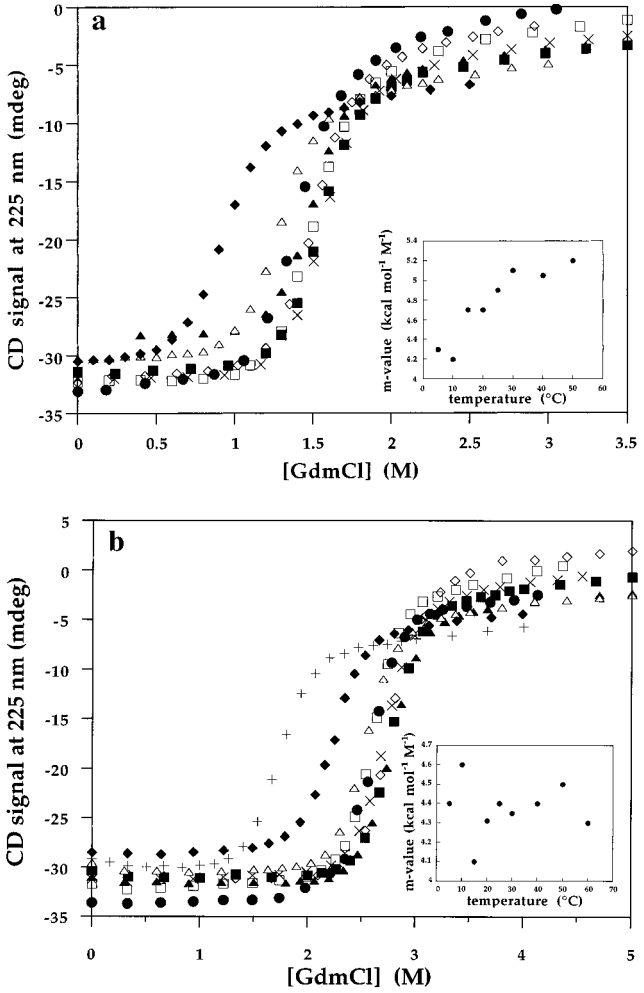


FIGURE 2: GdmCl-induced denaturation of *E. coli* (a) and *T. thermophilus* (b) RNases H* at various temperatures: (●) 5, (□) 10, (◇) 15, (×) 20, (■) 25, (▲) 30, (△) 40, (◆) 50, and (+) 60 °C. Denaturation was performed by the serial addition of a GdmCl-containing solution [40 μg/mL protein, 6.6 M GdmCl for *T. thermophilus* RNase H* or 7.2 M GdmCl for *E. coli* RNase H*, 5 mM NaAc, and 50 mM KCl (pH 5.5)] into a starting solution [40 μg/mL protein, 5 mM NaAc, and 50 mM KCl (pH 5.5)]. After equilibration, the CD signal at 225 nm (1 cm cuvette) was averaged for 1 min for each point. Each curve was fit to a two-state model, assuming a linear dependence of ΔG_{unf} on GdmCl concentration. *m* values shown in the insets were determined from these fits and represent the slope of ΔG_{unf} vs GdmCl concentration.

(Cys17 and Cys67) and serines in place of two more exposed cysteines (Cys45 and Cys145). Although the thermostabilities of these variants are similar (their T_m's are within 2 °C), the mixed Ser/Ala variant is slightly more thermostable than the others, with a T_m of 86 °C. We chose to continue work with this protein, referred to as *T. thermophilus* RNase H*. A similar variant of *E. coli* RNase H, in which the three free cysteines have been replaced with alanine, was used for all thermodynamic analyses here and is referred to as *E. coli* RNase H*.

*Stability Curves of E. coli and T. thermophilus RNases H**. The stabilities of both *E. coli* and *T. thermophilus* RNases H* at several temperatures were determined by GdmCl-induced denaturation, monitored by circular dichroism. The GdmCl denaturation profiles, shown in Figure 2, were fit to a two-state model (*14*), as described in Materials and Methods. For each protein, ΔG_{unf} was plotted as a function of temperature, resulting in the protein stability

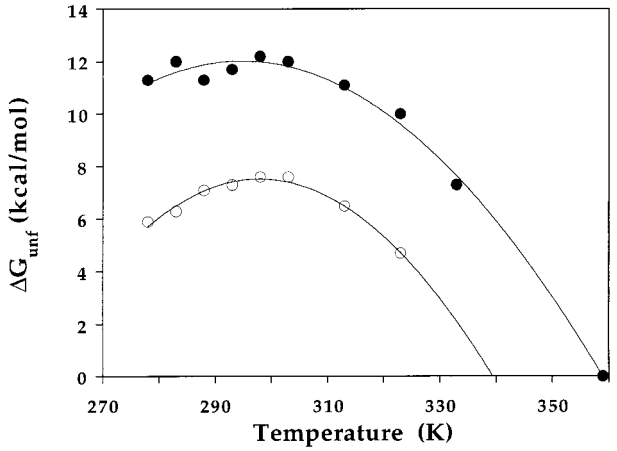


FIGURE 3: Protein stability curves for *E. coli* (○) and *T. thermophilus* (●) RNases H*. Each point represents a ΔG_{unf} determined from an isothermal GdmCl-induced denaturation experiment, with an additional point for the T_m of *T. thermophilus* RNase H*, determined by reversible thermal denaturation. Lines represent fits to the Gibbs−Helmholtz equation.

Table 1: Thermodynamic Parameters for *E. coli* and *T. thermophilus* RNases H*[a]

	E. coli RNase H*	*T. thermophilus* RNase H*
T_m	339 ± 1 K (66 °C)	359 ± 1 K (86 °C)
ΔC_p	2.7 ± 0.2 kcal mol^{-1} K^{-1}	1.8 ± 0.1 kcal mol^{-1} K^{-1}
$\Delta H°$	120 ± 4 kcal/mol	131 ± 5 kcal/mol
T_s	297 K (24 °C)	293 K (20 °C)
ΔG_s	7.5 kcal/mol	12.7 kcal/mol
ΔG_{unf} (at optimum growth temp)	6.8 kcal/mol (at 37 °C)	5.6 kcal/mol (at 68.5 °C)

[a] T_m, $\Delta H°_{unf}$, and ΔC_p were determined from fits of the protein stability curves to the Gibbs−Helmholtz equation; errors shown are for the fits. T_s (temperature of maximal stability), ΔG_s (maximal stability), and ΔG_{unf} at the optimal growth temperature were calculated on the basis of the stability curve fit results.

curves shown in Figure 3. The T_m determined by thermal denaturation was also included as a single point (ΔG_{unf} = 0) for *T. thermophilus* RNase H*. However, as thermal denaturation for *E. coli* RNase H* was not reversible, the T_m for this protein might not reflect equilibrium unfolding and was not included in our analysis. The two stability curves were fit to the Gibbs−Helmholtz equation (eq 2) to obtain the thermodynamic parameters ΔC_p, $\Delta H°$, and T_m (Table 1).

The *m* values describing the GdmCl concentration dependence of ΔG_{unf} for *T. thermophilus* RNase H* do not vary significantly with temperature (Figure 2b). This allowed us to fit the raw data globally, i.e., in a single step, as described in Materials and Methods. This analysis resulted in thermodynamic parameters for *T. thermophilus* RNase H* that were identical within error to those determined by the two-step fitting method. A similar global analysis of the *E. coli* RNase H* data was not performed, as the *m* values for this protein appeared to increase with temperature (Figure 2a).

T. thermophilus RNase H* is more stable than *E. coli* RNase H* at all temperatures, as seen in Figure 3 and Table 1. The T_m for *T. thermophilus* RNase H* (86 °C) is 20 °C higher than that for *E. coli* RNase H* (66 °C), and the ΔC_p is 0.9 kcal mol^{-1} K^{-1} lower (1.8 vs 2.7 kcal mol^{-1} K^{-1}), indicating a shallower dependence of the ΔG_{unf} on temper-

ature. Despite these changes, the temperatures of maximal stability for the two proteins both occur near room temperature.

The parameters determined from fitting these protein stability curves, ΔC_p, $\Delta H°$, and T_m, can be used to calculate the stability of these proteins at any temperature. The ΔG_{unf} values for *T. thermophilus* and *E. coli* RNases H* calculated at the optimal growth temperatures of their respective host organisms are shown in Table 1. These stabilities (5.6 kcal/mol at 68.5 °C for *T. thermophilus* RNase H* and 6.8 kcal/mol at 37 °C for *E. coli* RNase H*) are quite similar, despite the 30 °C difference between the bacterial growth temperatures (*1*).

DISCUSSION

Protein stability curves (ΔG_{unf} vs temperature) describe in detail the thermodynamics of protein unfolding. We have determined stability curves for two RNases H, one from the mesophile *E. coli* and one from the thermophile *T. thermophilus*, by measuring the CD signal as a function of GdmCl concentration and temperature. The results indicate that *T. thermophilus* RNase H achieves its increased thermostability through both a higher maximal ΔG_{unf} and a lower ΔC_p.

Our initial characterization of *T. thermophilus* RNase H was carried out with the wild-type protein, which contains four cysteines. We determined the stability of this protein by GdmCl denaturation at 25 °C and pH 5.5, in the presence of 1 mM DTT. The ΔG_{unf} determined here (13.3 kcal/mol) is much lower than the previously published value (20.9 ± 0.8 kcal/mol) (*19*). Our confidence in the stability we have determined stems from the reproducibility of our GdmCl-induced denaturation, and the agreement between this stability and the ΔG_{unf} calculated on the basis of the T_m of wild-type *T. thermophilus* RNase H (89 °C) and the ΔC_p (1.8 kcal mol^{-1} K^{-1}) of *T. thermophilus* RNase H*, a variant in which the cysteines have been replaced. In addition, when we fit the published CD data to a two-state model for GdmCl-induced denaturation using the same method we used to fit our data, we obtain a ΔG_{unf} of 12.5 ± 1.8 kcal/mol, which is much more similar to our result. This fitting method entails a global analysis of the CD data, in which the entire denaturation profile is used to fit parameters for the folded and unfolded baselines as well as the transition region. Last, the previously published T_m for this protein in 1 M GdmCl under reducing conditions [77.5 °C (*20*)] is identical to our T_m under similar conditions, suggesting that the discrepancies in ΔG_{unf} are not due to the protein itself but to differences in data analysis. Despite being lower than the previously reported data, the ΔG_{unf} determined here is significantly higher than that of both the wild-type *E. coli* RNase H (*21*) and *E. coli* RNase H*, in which the cysteines have been replaced by alanines.

Wild-type *T. thermophilus* RNase H contains four free cysteines, and to avoid cysteine chemistry that could complicate our thermodynamic analyses as well as storage of the protein, we designed a cysteine-free variant, *T. thermophilus* RNase H*. This variant (C17A/C45S/C67A/C154S), although slightly less stable than the wild-type protein, remains a good model for a thermophilic RNase H, with a ΔG_{unf} of 12.2 kcal/mol at 25 °C and a T_m of 86 °C. This protein is more thermostable than variants in which all of the cysteines were replaced with serine or all with alanine, indicating that the more conservative substitution for solvent-exposed cysteine residues is serine, and for buried cysteines is alanine.

The collection of denaturant and thermal melts presented here has allowed a complete description of the thermodynamic stability of *T. thermophilus* and *E. coli* RNases H*. The protein stability curves shown in Figure 3, and the parameters obtained from fitting these curves to the Gibbs−Helmholtz equation shown in Table 1, indicate that *T. thermophilus* RNase H* achieves its thermostabilization through two mechanisms. First, the thermophilic RNase H* has a higher maximal stability, resulting in an upward shift of the entire stability curve to a higher ΔG_{unf} range. Second, the ΔC_p for *T. thermophilus* RNase H* is lower than that of its mesophilic counterpart by 0.9 kcal mol^{-1} K^{-1}, resulting in a smaller dependence of the ΔG_{unf} on temperature, or a flattening of the stability curve.

The lower ΔC_p for *T. thermophilus* RNase H* has interesting implications in the evaluation of the entropic and enthalpic contributions to the stabilization of this protein. Because ΔC_p is related to the temperature dependencies of both ΔS_{unf} and ΔH_{unf}, these two parameters vary less with temperature for *T. thermophilus* RNase H* than for *E. coli* RNase H*. In addition, the maximal stabilities of the two proteins, where $T\Delta S_{unf}$ is zero, occur at similar temperatures, indicating that $T\Delta S_{unf}$ and ΔH_{unf} for *T. thermophilus* RNase H* are similar to $T\Delta S_{unf}$ and ΔH_{unf} for *E. coli* RNase H* near this peak in the stability curves. Specifically, the data from the stability curve fits show that at 32 °C, the $T\Delta S_{unf}$ values for the two proteins are equivalent, and at 37 °C, their ΔH_{unf} values are equivalent, with both terms being lower for *T. thermophilus* RNase H* above these temperatures. Therefore, at higher temperatures (>32 °C), *T. thermophilus* RNase H* achieves its stabilization by entropic means (lower $T\Delta S_{unf}$), whereas at lower temperatures (<37 °C), it achieves its stabilization through enthalpy (higher ΔH_{unf}). This pattern of stabilization indicates a smaller hydrophobic effect, which tends to favor folding at high temperatures by enthalpy and at low temperatures by entropy, and which is responsible for the large ΔC_p of proteins in general.

The thermodynamic parameter ΔC_p is thought to arise chiefly from the hydration of surface area that is exposed upon unfolding of the protein (*22*). However, the changes in accessible surface area (ΔASA) between the unfolded and folded states of the two RNases H are quite similar [12 916 Å^2 for *E. coli* RNase H (*5*) vs 12 687 Å^2 for *T. thermophilus* RNase H (*6*)]; their difference is smaller than the difference in ΔASA obtained merely by using various crystal structures of *E. coli* RNase H for the calculation [12 916 Å^2 using the structure determined by Katayanagi et al. (*5*) vs 12 306 Å^2 using that determined by Yang et al. (*16*)]. In addition, although the ratio of nonpolar to polar ΔASA is consistently lower for *T. thermophilus* RNase H, predicting a lower ΔC_p, the difference is again small. Equations developed by Gomez et al. (*22*) and Myers et al. (*23*) to determine ΔC_p from such calculations predict differences of less than 0.2 kcal mol^{-1} K^{-1} in the ΔC_p between the two proteins. This lack of agreement between the ΔASA calculations and the measured ΔC_p values is difficult to interpret, and may be due to the relatively low resolution of the *T. thermophilus* RNase H crystal structure, which does not include 19 of the 166 residues, due to disorder in the crystal (*6*).

It has been observed previously that thermophilic proteins often have increased thermodynamic stability at low temperatures, but in very few instances has the ΔC_p for a thermophilic protein been directly compared to that of a mesophilic homologue. Protein stability curves have been determined for two small thermophilic DNA binding proteins, Sac7d from *Sulfolobus acidocaldarius* (*24*) and Sso7d from *Sulfolobus solftaricus* (*25*). Although both display shallow temperature dependencies of their stabilities, their ΔC_p values are not significantly lower than those for mesophilic proteins of similar size, and there are no direct mesophilic homologues available for comparison. Stability curves have also been determined for phosphoglycerate kinases from yeast (*26*) and from the hyperthermophilic *Thermotoga maritima* (*27*). Although no ΔC_p was reported for the *Th. maritima* protein, the temperature dependence of its ΔG_{unf} appears to be larger than that for the yeast protein; in this case, the hyperthermophilic protein seems to gain the bulk of its thermostability by an upward shift of the stability curve to a higher ΔG_{unf}. Last, in a recent study of archaeal histones (*28*), stability curves for a mesophilic and three thermophilic homologues were compared, and the proteins were found to have similar ΔC_p values. Thus, the lower ΔC_p of *T. thermophilus* RNase H* is one of the first examples of this mechanism of thermostabilization.

Last, the stabilities of *T. thermophilus* and *E. coli* RNases H* at the optimal growth temperatures of their host organisms are quite similar (5.6 and 6.8 kcal/mol, respectively). This observation suggests that a balance between stability and flexibility is necessary for enzyme function, and is consistent with the lower activity of *T. thermophilus* RNase H* at 25 °C (*19*). This phenomenon has also been observed in thermophilic and mesophilic archaeal histones (*28*).

Overall, these experiments indicate that the relative importance of entropy and enthalpy in stabilizing *T. thermophilus* RNase H* over its mesophilic homologue varies with temperature, but that at all temperatures the thermophilic protein has a higher stability. It will be interesting to see how this stability is distributed structurally and how tightly coupled the structural regions are. Toward this end, we have begun native state hydrogen exchange experiments (*29*, *30*), to determine the stability of different regions of *T. thermophilus* RNase H* for comparison with *E. coli* RNase H*, in hopes of gaining further insight into the mechanisms of thermostabilization.

ACKNOWLEDGMENT

We thank Joan Kho for providing the *E. coli* RNase H* protein, Kael Fischer for helpful discussion and critical reading of the manuscript, Giulietta Spudich for critical reading of the manuscript, and Eric Goedken for help with the figures.

REFERENCES

1. Vielle, C., Burdette, D. S., and Zeikus, G. (1996) *Biotechnol. Annu. Rev. 2*, 1−83.
2. Vogt, G., and Argos, P. (1997) *Folding Des. 2*, S40−S46.
3. Jaenicke, R., and Bohm, G. (1998) *Curr. Opin. Struct. Biol. 8*, 738−748.
4. Hostomsky, Z., Hostomska, Z., and Matthews, D. A. (1993) in *Nucleases* (Linn, S. M., Lloyd, S., and Robert, R. J., Eds.) pp 341−376, Cold Spring Harbor Laboratory Press, Plainview, NY.
5. Katayanagi, K., Miyagawa, M., Matsushima, M., Ishikawa, M., Kanaya, S., Nakamura, H., Ikehara, M., Matsuzaki, T., and Morikawa, K. (1992) *J. Mol. Biol. 223*, 1029−1052.
6. Ishikawa, K., Okumura, M., Katayanagi, K., Kimura, S., Kanaya, S., Makamura, H., and Morikawa, K. (1993) *J. Mol. Biol. 230*, 529−542.
7. Ishikawa, K., Nakamura, H., Morikawa, K., and Kanaya, S. (1993) *Biochemistry 32*, 6171−6178.
8. Kimura, S., Nakamura, H., Hashimoto, T., Oobatake, M., and Kanaya, S. (1992) *J. Biol. Chem. 267*, 21535−21542.
9. Kimura, S., Kanaya, S., and Nakamura, H. (1992) *J. Biol. Chem. 267*, 22014−22017.
10. Akasako, A., Haruki, M., Oobatake, M., and Kanaya, S. (1995) *Biochemistry 34*, 8115−8122.
11. Doering, D. S., and Matsudaira, P. (1996) *Biochemistry 35*, 12677−12685.
12. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488−492.
13. Dabora, J. M., and Marqusee, S. (1994) *Protein Sci 3*, 1401−1408.
14. Santoro, M. M., and Bolen, D. W. (1988) *Biochemistry 27*, 8063−8068.
15. Becktel, W. J., and Schellman, J. A. (1987) *Biopolymers 26*, 1859−1877.
16. Yang, W., Hendrickson, W. A., Crouch, R. J., and Satow, Y. (1990) *Science 249*, 1398−1405.
17. Connolly, M. L. (1993) *J. Mol. Graphics. 11*, 139−141.
18. Lee, B., and Richards, F. M. (1971) *J. Mol. Biol. 55*, 379−400.
19. Kanaya, S., and Itaya, M. (1992) *J. Biol. Chem. 267*, 10184−10192.
20. Hirano, N., Haruki, M., Morikawa, M., and Kanaya, S. (1998) *Biochemistry 37*, 12640−12648.
21. Kanaya, S., Katsuda, C., Kimura, S., Nakai, T., Kitakuni, E., Nakamura, H., Katayanagi, K., Morikawa, K., and Ikehara, M. (1991) *J. Biol. Chem. 266*, 6038−6044.
22. Gomez, J., Hilser, V. J., Xie, D., and Freire, E. (1995) *Proteins: Struct., Funct., Genet. 22*, 404−412.
23. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) *Protein Sci. 4*, 2138−2148.
24. McCrary, B. S., Edmondson, S. P., and Shriver, J. W. (1996) *J. Mol. Biol. 264*, 784−805.
25. Knapp, S., Karishikoff, A., Berndt, K. D., Christova, P., Atanasov, B., and Ladenstein, R. (1996) *J. Mol. Biol. 264*, 1132−1144.
26. Nojima, H., Ikai, A., Oshima, T., and Noda, H. (1977) *J. Mol. Biol. 116*, 429−442.
27. Grattinger, M., Dankesreiter, A., Schurig, H., and Jaenicke, R. (1998) *J. Mol. Biol. 280*, 525−533.
28. Li, W., Grayling, R. A., Sandman, K., Ednondson, S., Shriver, J. W., and Reeve, J. N. (1998) *Biochemistry 37*, 10563−10572.
29. Bai, Y., Sosnick, T. R., Mayne, L., and Englander, S. W. (1995) *Science 269*, 192−197.
30. Chamberlain, A. K., Handel, T. M., and Marqusee, S. (1996) *Nat. Struct. Biol. 3*, 782−787.

BI982684H