Toward the Physical Basis of Thermophilic Proteins: Linking of Enriched Polar Interactions and Reduced Heat Capacity of Unfolding

Huan-Xiang Zhou
Institute of Molecular Biophysics and Department of Physics, Florida State University, Tallahassee, Florida 32306 USA

ABSTRACT The enrichment of salt bridges and hydrogen bonding in thermophilic proteins has long been recognized. Another tendency, featuring lower heat capacity of unfolding ($\Delta C_p$) than found in mesophilic proteins, is emerging from the recent literature. Here we present a simple electrostatic model to illustrate that formation of a salt-bridge or hydrogen-bonding network around an ionized group in the folded state leads to increased folding stability and decreased $\Delta C_p$. We thus suggest that the reduced $\Delta C_p$ of thermophilic proteins could partly be attributed to enriched polar interactions. A reduced $\Delta C_p$ might serve as an indicator for the contribution of polar interactions to folding stability.

INTRODUCTION

Thermophilic proteins offer a new opportunity to examine our understanding of the physical basis of protein stability. So far a number of mechanisms have been proposed to explain the enhanced thermostability of these proteins relative to their mesophilic counterparts. These include enriched salt bridges and other types of polar interactions, better packing, differing amino acid distributions, and smaller loop sizes (Perutz and Raidt, 1975; Perutz, 1978; Vogt and Argos, 1997; Jaenicke and Bohm, 1998; Szilagyi and Zavodszky, 2000; Petsko, 2001). Whereas thermostability likely results from optimizations of all these mechanisms, the presence of enriched polar interactions has been a common theme among thermophilic proteins.

The focus of the present paper is a potential new tendency, characterized by lower heat capacity of unfolding ($\Delta C_p$) than found in mesophilic proteins that appears to be emerging from the recent literature on thermophilic proteins. Table 1 lists thermodynamic properties of the unfolding of six thermophilic proteins and their mesophilic counterparts (Hollien and Marqusee, 1999; Deutschman and Dahlquist, 2001; Motono et al., 2001; Shiraki et al., 2001; Nojima et al., 1977; Knapp et al., 1996, 1998; Szilagyi and Zavodszky, 2000; Petsko, 2001). Whereas thermostability likely results from optimizations of all these mechanisms, the presence of enriched polar interactions has been a common theme among thermophilic proteins.

A large positive $\Delta C_p$ has long been recognized as an important character of protein unfolding. It is taken to indicate the dominance of hydrophobic interactions in driving protein folding, because of the well known fact that exposure of nonpolar compounds to water also gives rise to a large positive $\Delta C_p$ (Baldwin, 1986; Privalov and Makhatadze, 1990; Livingstone et al., 1991; Spolar et al., 1992; Murphy and Freire, 1992; Creighton, 1993; Myers et al., 1995; Makhatadze and Privalov, 1995; Robertson and Murphy, 1997). Based on heat capacity data for transferring model compounds to water, it was also contended that the exposure of polar groups to water gives rise to a negative $\Delta C_p$ (Spolar et al., 1992; Murphy and Freire, 1992; Myers et al., 1995; Makhatadze and Privalov, 1995). A recent experiment has shown that replacing buried nonpolar sidechains by a polar one reduces $\Delta C_p$ (Loladze et al., 2001). It should be noted that, in this case, the reduced $\Delta C_p$ values were accompanied by decreased melting temperatures (and thus decreased folding stability).

If $\Delta C_p$ is assumed to be temperature independent, the unfolding free energy $\Delta G$ at any temperature $T$ is given by

$$
\Delta G = \Delta G_s + \Delta C_p(T - T_s) - \Delta C_p T \ln (T/T_s),
$$

in which $T_s$ is the temperature at which $\Delta G$ takes its maximal value $\Delta G_s$. A plot of $\Delta G$ as a function of temperature, as given by Eq. 1, shows a nearly parabolic curve that, for $\Delta C_p > 0$, decreases at high (and low) temperatures (Fig. 1). From this plot, one can immediately recognize that $\Delta C_p$ controls the broadness of the curve. A reduced $\Delta C_p$ will broaden the curve such that the melting temperature $T_m$ (at which $\Delta G = 0$) will increase. That reduced $\Delta C_p$ values are indeed observed in thermophilic proteins is intriguing. What is the physical origin for the reduced $\Delta C_p$?

Here we suggest that the reduced $\Delta C_p$ is related to the enriched polar interactions found in thermophilic proteins. Using a simple electrostatic model, we illustrate that a salt-bridge or hydrogen-bonding network around an ionized group stabilizes the folded state (increasing $\Delta G$) and, at the same time, decreases $\Delta C_p$. 

Submitted June 26, 2002, and accepted for publication August 2, 2002.

Address reprint requests to Huan-Xiang Zhou, Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306. Tel.: 850-644-4764; Fax: 850-644-7244; E-mail: zhou@sb.fsu.edu.

© 2002 by the Biophysical Society
0006-3495/02/12/3126/08 $2.00
TABLE 1 Thermodynamic properties of the unfolding of thermophilic and mesophilic proteins

<table>
<thead>
<tr>
<th>Protein*</th>
<th>( T_m ) (°C)</th>
<th>( T_c ) (°C)</th>
<th>( \Delta G_0 ) (kcal/mol)</th>
<th>( \Delta C_p ) (kcal/mol/K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. thermophilus RNase H</td>
<td>86</td>
<td>20</td>
<td>12.7</td>
<td>1.8</td>
</tr>
<tr>
<td>E. coli RNase H</td>
<td>66</td>
<td>24</td>
<td>7.5</td>
<td>2.7</td>
</tr>
<tr>
<td>T. maritima CheY</td>
<td>101</td>
<td>29</td>
<td>9.54</td>
<td>1.17</td>
</tr>
<tr>
<td>B. subtilis CheY</td>
<td>55</td>
<td>27</td>
<td>3.14</td>
<td>2.34</td>
</tr>
<tr>
<td>T. thermophilus IPMDH</td>
<td>109</td>
<td>31</td>
<td>15.8</td>
<td>1.73</td>
</tr>
<tr>
<td>E. coli IPMDH</td>
<td>90</td>
<td>34</td>
<td>17.8</td>
<td>3.69</td>
</tr>
<tr>
<td>T. kodakaraensis MGMT</td>
<td>98.6</td>
<td>29.5</td>
<td>10.2</td>
<td>1.2</td>
</tr>
<tr>
<td>E. coli AdaC</td>
<td>43.8</td>
<td>7.4</td>
<td>4.0</td>
<td>1.8</td>
</tr>
<tr>
<td>T. thermophilus PGK</td>
<td>~80</td>
<td>~30</td>
<td>~12</td>
<td>~0.06</td>
</tr>
<tr>
<td>S. cerevisiae PGK</td>
<td>60</td>
<td>25</td>
<td>5.3</td>
<td>1.6</td>
</tr>
<tr>
<td>T. thermophilus RNase H</td>
<td>86</td>
<td>20</td>
<td>12.7</td>
<td>1.8</td>
</tr>
<tr>
<td>S. cerevisiae Smo7d</td>
<td>99</td>
<td>10</td>
<td>8.4</td>
<td>0.65</td>
</tr>
<tr>
<td>Average of 6 SH3 domains</td>
<td>71 ± 4</td>
<td>16</td>
<td>3.8</td>
<td>0.77 ± 0.04</td>
</tr>
</tbody>
</table>

*References: RNase H, Hollien and Marqusee (1999); CheY, Deutschman and Dahlquist (2001); IPMDH, Motono et al. (2001); MGMT and AdaC, Shiraki et al. (2001); PGK, Nojima et al. (1977); Smo7d, Knapp et al. (1996); and 6 SH3 domains, Knapp et al. (1998) and Filimonov et al. (1999).

**THEORY**

**Electrostatic model**

Fig. 2A illustrates the contrasts between the folded state of a protein and the unfolded state. The folded state is compact with groups enjoying specific interactions and solvated to a lesser extent. In the unfolded state, the protein molecule samples different conformations and has all its groups highly exposed to the solvent. In this article, we treat only the electrostatic aspect of the folding process. Specifically, the folded state will be modeled as a sphere (with radius \( R \)) containing whole or partial charges (from ionized and polar groups, respectively) and is solvated in water (Fig. 2, B and C). In the unfolded state, an ionized group will be represented by a small sphere (with radius \( a \)) containing a whole charge \((\pm e)\) at the center, whereas a polar group will be treated as a small sphere containing partial charges \(\pm \delta\) (Fig. 2, B and C). Interactions among ionized and polar groups in the unfolded state, which have been treated elsewhere (Zhou, 2002), will be ignored here for simplicity.

**Electrostatic contribution to \( \Delta G \)**

The various contributions to the unfolding free energy from the interactions between the charges and with the solvent can be obtained from the electrostatic potential of a charge \( q \) embedded at a radial distance \( s \) in a sphere with radius \( r \) (Fig. 3). When \( s = 0 \), the interaction with the solvent results in a free energy (Born, 1920)

\[
U_0(q, r) = -166 \left( \frac{1}{e_p} - \frac{1}{e_s} \right) \frac{q^2}{r}, \tag{2}
\]

in which \( e_p \) and \( e_s \) are the dielectric constants of the protein medium and water, respectively. When \( s \) is not zero, the solvation energy is

\[
U_{\text{solv}}(q, s, r) = -166 \left( \frac{1}{e_p} - \frac{1}{e_s} \right) \frac{q^2}{r} \times \sum_{l=1}^{\infty} \frac{l + 1}{l + 1 + (e_p/e_s)l} (s/r)^{2l}. \tag{3}
\]

If a second charge \( q' \) is also present inside the sphere at a radial distance \( s' \) and a distance \( d \) from charge \( q \) (Fig. 3), the free energy of interaction is

\[
U_{\text{int}}(q, q', s, s', d, r) = \frac{332qq'}{e_p r} - 332 \left( \frac{1}{e_p} - \frac{1}{e_s} \right) \frac{qq'}{r} \times \sum_{l=1}^{\infty} \frac{l + 1}{l + 1 + (e_p/e_s)l} \times (ss'/r^2)^l P_l(\cos \gamma), \tag{4}
\]

![FIGURE 1](image-url)  Temperature dependence of the unfolding free energy. The maximal stability is set to 5 kcal/mol, and the temperature at which this occurs is set to 25°C. By just reducing the heat capacity of unfolding from 1.5 kcal/mol/K (dashed curve) to 0.5 kcal/mol/K (solid curve), the melting temperature is increased from 70°C to 105°C.
\[
\cos \phi = (s^2 + s'^2 - d^2)2ss' \quad \text{and} \quad P_l(s) \text{ are the Legendre polynomials.}
\]

The electrostatic component of the unfolding free energy, \( \Delta G_{el} \), can now be calculated. For example, if the protein has two ionized groups (with charges \( +e \) and \( -e \)), we have

\[
\Delta G_{el} = \left[ 2U_0(e, a) - U_{solv}(e, s, R) - U_{solv}(e, s', R) \right] - U_{int}(e, -e, s, s', d, R),
\]

in which \( s \) and \( s' \) are the radial distances of the two charges in the folded proteins. Thus, \( \Delta G_{el} \) consists of a solvation term \( \Delta G_{solv} \) and an interaction term \( \Delta G_{int} \). The solvation term for a polar group represented by partial charges \( \pm \delta \) at a distance \( d \) inside a sphere with radius \( a \) can be calculated as

\[
\Delta G_{solv} = 2U_{solv}(\delta, d/2, a) - U_{solv}(\delta, s, R) - U_{int}(\delta, s, R) + U_{int}(\delta, -\delta, d/2, d/2, d, a) - U_{int}(\delta, -\delta, s, s', d, R).
\]

Other charge distributions can be similarly accounted for.

**Electrostatic contribution to \( \Delta C_p \)**

A standard thermodynamic relation leads to

\[
\Delta C_p^{el} = -T \frac{\partial^2 \Delta G_{el}}{\partial T^2} = \Delta C_p^{solv} + \Delta C_p^{int}.
\]

The two terms in Eq. 7b arise from the solvation and interaction components of \( \Delta G_{el} \). In evaluating Eq. 7a, we assume that the only temperature-dependent parameter is
the dielectric constant of water. The derivative can be evaluated analytically. At room temperature $T = 298$ K, $\epsilon_s = 78.4$, and the derivatives of $\epsilon_s$ are (Archer and Wang, 1990):

$$\frac{\partial \ln \epsilon_s}{\partial \ln T} = -1.37,$$

$$\frac{\partial^2 \ln \epsilon_s}{\partial (\ln T)^2} = -1.43.$$  

(8a)  

(8b)

In particular, we have

$$-T \frac{\partial^2 (1/\epsilon_s)}{\partial T^2} = \frac{1}{T \epsilon_s} \left[ \frac{\partial^2 \ln \epsilon_s}{\partial (\ln T)^2} - \frac{\partial \ln \epsilon_s}{\partial \ln T} \right]$$

$$= - \frac{1.94}{T \epsilon_s}.$$  

(9)

The negative sign of the value in Eq. 9 is the source of the main result (i.e., reduced $\Delta C_p$) of the present study. For an ion with a charge $+e$ or $-e$ and a radius of 2 Å solvated in water, Eqs. 2, 7a, and 9 predict a heat capacity of hydration of $-7$ cal/mol/K at room temperature. This value nearly falls within the range of experimental results for univalent ions, $-10$ to $-20$ cal/mol/K (Abraham and Marcus, 1986). Thus, the simple model actually yields results that are not unreasonable. Gallagher and Sharp (1998) have shown that the continuum model can yield reasonable results for the heat capacity of hydration of more complicated ions ($\text{NH}_4^+$, $\text{HCO}_2^-$, and $\text{H}_2\text{PO}_4^-$).

**Choice of parameters**

The protein dielectric constant $\epsilon_p$ is set to 4 and assumed to be temperature independent. The radius of an ionized group is set to $a = 2.4$ Å. The solvation energy of such an ion at room temperature, calculated according to Eq. 2, is $-16.4$ kcal/mol, which is close to what one obtains by applying the UHBD program (Madura et al., 1995) to a charged residue. A polar group is modeled as two partial charges $0.5e$ and $-0.5e$ at a distance of 2.2 Å inside a sphere with a radius of 2.4 Å. This set of parameters yields a solvation energy of $-3.5$ kcal/mol, which is nearly what one obtains by applying the UHBD program to an Asn or Gln residue.

The radius of the protein is set to $R = 16$ Å. Inside the protein, the distance between the whole charges of two ionized groups is set to 3 Å (a typical value in a salt-bridge situation), whereas the distance between a whole charge and a partial charge of a polar group is set to 2 Å (a typical value in a hydrogen-bonding situation). The radial distances of all charges inside the protein are set to 14 Å unless otherwise indicated.

**RESULTS AND DISCUSSION**

**Contributions of a salt-bridge network to $\Delta G$ and $\Delta C_p$**

The various charge distributions considered in the present study are shown in Fig. 4. The calculated results of their contributions to $\Delta G$ and $\Delta C_p$ are listed in Table 2. For an ion pair (i.e., distribution A), the desolvation cost ($-\Delta_{\text{sol}}$) calculated with the spherical model is slightly larger than the free energy of electrostatic interaction. Thus, the ion pair alone destabilizes the folded structure by 0.8 kcal/mol. However, when a second salt-bridge partner is added (distribution B), the free energy of electrostatic interactions now outweighs the desolvation cost, and the salt-bridge network as a whole stabilizes the folded structure by 1.8 kcal/mol. The influence of the electrostatic environment, in the form of a salt-bridge network or other favorable polar interactions, on the contribution of a charged residue to protein stability has been noted previously (Vijayakumar and Zhou, 2001; Xiao and Honig, 1999).

Both the solvation and the interaction terms of $\Delta_{\text{el}}$ reduce the heat capacity of unfolding with the interaction term playing a dominant role. According to the spherical
model, each salt-bridge interaction decreases $\Delta C_p$ by $-10$ cal/mol/K.

**Contributions of a hydrogen-bonding network to $\Delta G$ and $\Delta C_p$**

Burial of a single polar group alone in the folded state (distribution C) is found to be destabilizing. However, when the polar group forms a hydrogen bond with an ionized group in the folded state (distribution D), the favorable interaction almost offsets the desolvation cost. When the ionized group forms hydrogen bonds simultaneously with two polar groups (distribution E), a significant stabilizing effect (4.9 kcal/mol) is found.

The polar interactions between an ionized group and polar groups are also found to have a major role in reducing the heat capacity of unfolding, with each such interaction reducing $\Delta C_p$ by $-5$ cal/mol/K.

**Reduction of $\Delta C_p$ by polar interactions**

The spherical model yields a potentially important result: Polar interactions around an ionized group in the folded state significantly reduce $\Delta C_p$. Although the contribution of the solvent exposure of polar groups to $\Delta C_p$ is widely accepted, the contribution of polar interactions in the folded state does not appear to have received much attention. Of course the result must be viewed with the caveat that the spherical model is undoubtedly oversimplified. From a molecular point of view, the heat capacity of unfolding arises from the differences in solvent reorganization and in solute-solvent, solvent-solvent, and as implied by the spherical model, intra-solute interactions between the folded and unfolded state. However, quantitative modeling of such effects based on a more fundamental approach remains a challenge (Abraham and Marcus, 1986; Madan and Sharp, 1996, 2001). In a continuum model, all solvent effects are attributed to the dielectric constant of water. The calculation of $\Delta C_p$ entails evaluating second derivatives with respect to temperature. The spherical shape of the model used allows these derivatives to be evaluated analytically. Gallagher and Sharp (1998) have developed a numerical algorithm to evaluate heat capacity for DNA-ligand binding based on the Poisson-Boltzmann equation. This algorithm potentially can be applied to calculate $\Delta C_p$ using more realistic models for the folded and unfolded states. Our main interest here is the qualitative aspects of the contributions of charge-solvent and charge-charge interactions to $\Delta C_p$.

To see why a favorable charge-charge interaction in the folded state reduces $\Delta C_p$, consider two opposite charges interacting in water:

$$ U_{\text{int}} = -\frac{332e^2}{e_p d}. \quad (10) $$

The contribution of the interaction energy to $\Delta C_p$ is (see Eqs. 5 and 7a)

$$ -T \frac{\partial^2(-U_{\text{int}})}{\partial T^2} = \left( \frac{e^2}{d} \right) \left[ -T \frac{\partial^2(1/e_p)}{\partial T^2} \right]. \quad (11) $$

The second factor is given by Eq. 9 and is negative, thus the interaction reduces $\Delta C_p$. A better model for two opposite charges interacting in the folded protein is obtained by embedding the charges in the low dielectric (having dielectric constant $e_p$) sharing a planar boundary with the high dielectric (having dielectric constant $e_s$). The image charge of charge $+e$ is $- (e_p - e_s)/(e_s + e_p)e$. The interaction energy is thus

$$ U_{\text{int}} = -\frac{332e^2}{e_p d} + \frac{332e^2}{e_s d} - \frac{664e^2}{(e_s + e_p)d^2}, \quad (12) $$

in which $d'$ is the distance between the image charge and charge $-e$. The only term contributing to $\Delta C_p$ is the last one, which, aside from a factor of 2, differs from Eq. 10 only by the replacement of $d$ by $d'$ and the addition of $e_s$ to $e_p$ (note $e_p \ll e_s$). Again, a negative contribution to $\Delta C_p$ is obtained.

If polar interactions around ionized groups in the folded state reduce $\Delta C_p$, to what extent do these interactions contribute to the lower $\Delta C_p$ values observed on thermophilic proteins? Consider a thermophilic protein with 10 additional charged residues relative to its mesophilic counterpart. If each of the charged residues makes two polar interactions, and each interaction contributes $-10$ cal/mol/K to $\Delta C_p$, then the 10 charged residues will reduce $\Delta C_p$ by 0.2 kcal/mol/K. This is a significant fraction of the average of 1 kcal/mol/K for the difference in $\Delta C_p$ among the six pairs of thermophilic and mesophilic proteins listed in Table 1. The spherical model may underestimate the magnitude of the contributions of polar/charged group burial and polar interactions (see also the result for an ion given after Eq. 9 and discussion in the following paragraph). In addition, if all the 10 charged residues are substituted by nonpolar residues in the mesophilic protein, the nonpolar residues will be expected to increase $\Delta C_p$ of the mesophilic protein by $-0.2$ kcal/mol/K on account of burying nonpolar surfaces (Spolar et al., 1992; Murphy and Freire, 1992; Myers et al., 1995; Makhatadze and Privalov, 1995). However, we note that charged residues typically substitute for polar residues.

According to the spherical model, burial of a single polar group reduces $\Delta C_p$ by just $1.3$ cal/mol/K. If the group is assumed to have a surface area of 50 Å$^2$, then the contribution per unit area is $-0.03$ cal/mol/K/Å$^2$. The contribution of the burial of polar groups to $\Delta C_p$ has been estimated to range from $-0.09$ to $-0.26$ cal/mol/K/Å$^2$ (Spolar et al., 1992; Murphy and Freire, 1992; Myers et al., 1995; Makhatadze and Privalov, 1995). The $1.3$ cal/mol/K reduction in $\Delta C_p$ is perhaps an underestimate by the spherical model, but there might be an additional source for the gap between the resulting value of
The residues before and after each position number are for T. thermophilus RNase H are in italic. The two dashed lines enclose residues in the core.

The enrichment of charged residues and the resulting extra polar interactions in thermophilic proteins have been well documented (Perutz and Raidt, 1975; Perutz, 1978; Vogt and Argos, 1997; Jaenicke and Bohm, 1998; Szilagyi and Zavodszky, 2000; Petsko, 2001). In particular, surveys by Szilagyi and Zavodszky (2000) found that: 1) the percentage of charged residues is higher in thermophilic proteins than in their mesophilic counterparts; 2) buried surfaces are more polar; and 3) a 300-residue thermophile is expected to have ~4 strong and 14 weaker extra ion pairs. To further illustrate the enrichment of polar interactions around charged residues in thermophilic proteins, in Table 2 we list all the charged-to-neutral and neutral-to-charged substitutions between T. thermophilus and E. coli RNases H, respectively. Changes to charged residues in E. coli RNase H are in italic. The two dashed lines enclose residues in the core.

### TABLE 2 Differences between T. thermophilus and E. coli RNases H involving charged residues

<table>
<thead>
<tr>
<th>Residue*</th>
<th>Polar interactions†</th>
</tr>
</thead>
<tbody>
<tr>
<td>R2L</td>
<td>E64 (NH2-OE1: 3.3 Å); K3 (NH2-N: 3.0 Å); R4 (NH2-N: 2.7 Å)</td>
</tr>
<tr>
<td>R4Q</td>
<td>D66 (NH2-OD1: 3.3 Å; NH2-OD2: 3.3 Å); E64 (NE-OE2: 3.7 Å)</td>
</tr>
<tr>
<td>A6E</td>
<td>R27 (OE2-NH1: 3.3 Å)</td>
</tr>
<tr>
<td>E39Y</td>
<td>R46 (OE1-NH2: 3.4 Å); K50 (OE2-NZ: 3.7 Å)</td>
</tr>
<tr>
<td>S41R</td>
<td>None</td>
</tr>
<tr>
<td>K50M</td>
<td>E39 (NZ-OE2: 3.7 Å)</td>
</tr>
<tr>
<td>E54V</td>
<td>K57 (OE1-NZ: 3.6 Å;OE2-NZ: 3.7 Å)</td>
</tr>
<tr>
<td>H62P</td>
<td>Q113 (NE2-NE2: 3.4 Å)</td>
</tr>
<tr>
<td>D66I</td>
<td>R4 (OD1-NH2: 3.3 Å; OD2-NH2: 3.3 Å); R117 (OD1-NH2: 3.0 Å; OD2-NH2: 3.2 Å); H119 (OD1-NE2: 2.7 Å)</td>
</tr>
<tr>
<td>H72Q</td>
<td>E48 (ND1-OE2: 3.8 Å); D70 (ND1-O: 3.6 Å)</td>
</tr>
<tr>
<td>K76Q</td>
<td>W81 indole ring: 3.5 Å</td>
</tr>
<tr>
<td>E80Q</td>
<td>T77 (OE2-OG1: 3.8 Å)</td>
</tr>
<tr>
<td>G95K</td>
<td>None</td>
</tr>
<tr>
<td>R101V</td>
<td>V98 (NE-O: 3.0 Å); P97 (NH1-O: 3.9 Å)</td>
</tr>
<tr>
<td>E105Q</td>
<td>R101 (OE2-O: 3.5 Å)</td>
</tr>
<tr>
<td>A106R</td>
<td>E57 (NE-OE2: 2.9 Å; NH1-OE1: 2.9 Å)</td>
</tr>
<tr>
<td>L108D</td>
<td>K86 (OD1-NZ: 2.9 Å; OD2-NZ: 3.1 Å)</td>
</tr>
<tr>
<td>R115Q</td>
<td>E64 (NH1-OE1: 3.1 Å)</td>
</tr>
<tr>
<td>R135A</td>
<td>None</td>
</tr>
<tr>
<td>R138A</td>
<td>D134 (NH1-OD1: 3.5 Å)</td>
</tr>
<tr>
<td>K146A</td>
<td>Q144 (NZ-O: 2.8 Å)</td>
</tr>
<tr>
<td>R152L</td>
<td>No coordinates</td>
</tr>
<tr>
<td>A153E</td>
<td>None</td>
</tr>
<tr>
<td>P154D</td>
<td>No coordinates</td>
</tr>
<tr>
<td>H158Q</td>
<td>No coordinates</td>
</tr>
<tr>
<td>E159V</td>
<td>None</td>
</tr>
</tbody>
</table>

*†The residues before and after each position number are for T. thermophilus and E. coli RNases H, respectively. Changes to charged residues in E. coli RNase H are in italic. The two dashed lines enclose residues in the core.

Distances are from x-ray structures of the proteins (PDB entries 1ril and 1f21; Ishikawa et al., 1993; Goedken et al., 2000).

−0.03 cal/mol/K/Å² for ΔCp per unit area of polar surface and previous estimates. A buried polar group typically forms hydrogen bonds with other polar groups. Such hydrogen-bonding interactions may further reduce ΔCp.

All of our calculation results are for room temperature. Both thermophilic and mesophilic proteins have a higher stability around a charged residue, and the maximal stability of thermophilic proteins occurs at temperatures higher than those of mesophilic proteins (Table 1). We illustrated that a buried polar group typically forms hydrogen bonds or salt bridges. The enrichment of charged residues and the resulting extra polar interactions around charged residues in thermophilic proteins have been well documented (Perutz and Raidt, 1975; Perutz, 1978; Vogt and Argos, 1997; Jaenicke and Bohm, 1998; Szilagyi and Zavodszky, 2000; Petsko, 2001). In particular, surveys by Szilagyi and Zavodszky (2000) found that: 1) the percentage of charged residues is higher in thermophilic proteins than in their mesophilic counterparts; 2) buried surfaces are more polar; and 3) a 300-residue thermophile is expected to have ~4 strong and 14 weaker extra ion pairs. To further illustrate the enrichment of polar interactions around charged residues in thermophilic proteins, in Table 2 we list all the charged-to-neutral and neutral-to-charged substitutions between T. thermophilus and E. coli RNases H. In all, T. thermophilus RNase H has 10 more charged residues. Except for the insertion R135, all the charged residues replacing neutral ones in E. coli RNase H, when coordinates are reported (Ishikawa et al., 1993; Goedken et al., 2000), form salt bridges or hydrogen bonds.

Marqusee and co-workers (Robic et al., 2002) recently conducted an interesting experiment. They swapped residues 43 to 120 (the core) of T. thermophilus and E. coli RNases H, resulting in two new proteins: TCEO and ECTO. The protein with the thermophilic core, TCEO, is found to have a lower ΔCp (1.6 kcal/mol/K) than the protein with the mesophilic core (2.4 kcal/mol/K). It can be seen from Table 2 that most of the additional polar interactions around charged residues in T. thermophilus RNase H occur in the core. That is, TCEO still
have more polar interactions around charged residues than ECTO.

Exceptions to reduced $\Delta C_p$ of thermophilic proteins

Although we have presented a trend of reduced $\Delta C_p$ in thermophilic proteins, there are exceptions. At 0.2 M KCl, archaean histones HMfA, HMfB, and HPyA1 from thermophilic *M. fervidus* and *Pyrococcus* strain GB-3a have average $\Delta C_p$ of 2.2, 1.9, and 2.2 kcal/mol/K (over pH 2.5 to 7.5) (Li et al., 1998). Under the same conditions, the histone HFoB from mesophilic *M. formicicum* does have a higher average $\Delta C_p$ of 2.8 kcal/mol/K. However, at a salt concentration of 1 M, the difference in $\Delta C_p$ disappears: HMfA has an average $\Delta C_p$ of 2.0 kcal/mol/K, whereas HFoB has an average $\Delta C_p$ of 2.1 kcal/mol/K. The difference in $\Delta C_p$ between HMfA and HFoB at high salt concentrations could be suppressed by salt screening of electrostatic interactions and by specific ion binding.

Both thermophilic and mesophilic cold-shock proteins (Csps) have heat capacities of unfolding around 1 kcal/mol/K (Wassenberg et al., 1999; Petrosian and Makhatadze, 2000; Perl et al., 2000). The difference in stability between *B. caldolyticus* and *B. subtilis* Csps has been attributed in part to the relief of an electrostatic repulsion between residues E3 and E66 in *B. subtilis* Csp (Perl et al., 2000; Delbruck et al., 2001). The role of electrostatic interactions in the increased stability of the thermophilic protein has been investigated in a number of recent theoretical studies (Sanchez-Ruiz and Makhatadze, 2001; Dominy et al., 2002; D. Feng and H.-X. Zhou, submitted manuscript). The pairing of two like charges should raise $\Delta C_p$ (Fig. 4 F; the last row in Table 3) according to the spherical model. However, *B. subtilis* Csp also has two other neutral-to-charged mutations (S24D and Q53E). These two charges might lower $\Delta C_p$. The technical difficulty in the accurate measurement of $\Delta C_p$ should also be noted (Wassenberg et al., 1999; Petrosian and Makhatadze, 2000; McCrary et al., 1996). This difficulty might raise doubt about the reduced $\Delta C_p$ of thermostable proteins, the focus of the present study. However, the repeated observations (Table 1) make us feel confident that there is a real trend of reduced $\Delta C_p$.

### Linking of enriched polar interactions and reduced $\Delta C_p$

Both the enrichment of polar interactions in thermophilic proteins (Perutz and Raidt, 1975; Perutz, 1978; Vogt and Argos, 1997; Jaenicke and Bohm, 1998; Szilagyi and Zavodszky, 2000; Petsko, 2001) and the reduction in $\Delta C_p$ by exposing buried polar groups to water upon unfolding (Spolar et al., 1992; Murphy and Freire, 1992; Myers et al., 1995; Makhatadze and Privalov, 1995; Loladze et al., 2001) have been noted. However, it appears that the reduced $\Delta C_p$ of thermophilic proteins has not previously been linked to the enriched polar interactions around charged residues. Calculations based on the simple electrostatic model illustrate the plausibility of such a link. They suggest that a salt-bridge or hydrogen-bonding network around an ionized group stabilizes the folded state and, at the same time, decreases $\Delta C_p$.

In the past, residual structure in the unfolded state has been suggested as a possible explanation of the reduced $\Delta C_p$ of thermophilic proteins (Motono et al., 2001; Shiraki et al., 2001; Nojima et al., 1977; Robic et al., 2002). This explanation was mainly based on the consideration that a residual structure will keep some nonpolar surfaces buried (thus lowering the heat capacity of the unfolded state), rather than based on concrete experimental evidence. It is open to question in two respects. First, why would thermophilic proteins tend to have more residual structures in the unfolded state (with some nonpolar groups buried)? It should be kept in mind that thermophilic proteins typically have more polar surfaces buried in the folded state than mesophilic ones. Second, a protein with an unfolded state that retains residual structures would be expected to have a smaller unfolding free energy, because not all the structural elements have to be totally destroyed. This scenario is contradictory to the increased stability of thermophilic proteins.

The present study suggests additional investigations into the physical basis of thermophilic proteins. It is of interest to see whether thermophilic proteins that use enriched or optimized polar interactions around charged residues as a mechanism for increased stability will consistently have reduced $\Delta C_p$. Possibly, a reduced $\Delta C_p$ will serve as an indicator for the contribution of polar interactions to folding stability. In cases where thermophilic proteins have been observed to have reduced $\Delta C_p$, it is of interest to see whether charge mutations will restore $\Delta C_p$ to the levels of the mesophilic counterparts.

I thank Robert L. Baldwin for careful reading of the manuscript and encouragement and Frederick Dahlquist for bringing my attention to the reduced $\Delta C_p$ of *T. maritima* CheY. This work was supported in part by the National Institutes of Health Grant GM58187.

### REFERENCES
