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 (ΔC_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 ΔC_p. We thus suggest that the reduced ΔC_p of thermophilic proteins could partly be attributed to enriched polar interactions. A reduced ΔC_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 (ΔC_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; Filimonov et al., 1999). The results of ΔC_p for the thermophiles are all lower than those for the mesophilic proteins. In addition, values of $\Delta C_p = 0.75$ kcal/mol/K for *A. ambivalens* ferredoxin (Moczygemba et al., 2001) and $\Delta C_p = 2.86$ kcal/mol-trimer/K for *S. acdidocaldarius* adenylate kinase (Backmann et al., 1998) were considered low based on estimates of ΔC_p from the buried surface areas upon folding. Table 1 also shows that both mesophilic and thermophilic proteins have maximal stability around room temperature. The thermophiles typically show higher maximal stability than their mesophilic counterparts.

A large positive ΔC_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 ΔC_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 ΔC_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 ΔC_p (Loladze et al., 2001). It should be noted that, in this case, the reduced ΔC_p values were accompanied by decreased melting temperatures (and thus decreased folding stability).

If ΔC_p is assumed to be temperature independent, the unfolding free energy ΔG at any temperature T is given by

$$\Delta G = \Delta G_s + \Delta C_p(T - T_s) - \Delta C_p T \ln (T/T_s), \quad (1)$$

in which T_s is the temperature at which ΔG takes its maximal value ΔG_s. A plot of ΔG 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 ΔC_p controls the broadness of the curve. A reduced ΔC_p will broaden the curve such that the melting temperature T_m (at which $\Delta G = 0$) will increase. That reduced ΔC_p values are indeed observed in thermophilic proteins is intriguing. What is the physical origin for the reduced ΔC_p?

Here we suggest that the reduced ΔC_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 ΔG) and, at the same time, decreases ΔC_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.

0006-3495/02/12/3126/08 $2.00

TABLE 1 Thermodynamic properties of the unfolding of thermophilic and mesophilic proteins

Protein*	T_m (°C)	T_s (°C)	ΔG_s (kcal/mol)	ΔC_p (kcal/mol/K)
T. thermophilus RNase H	86	20	12.7	1.8
E. coli RNase H	66	24	7.5	2.7
T. maritama CheY	101	29	9.54	1.17
B. subtilis CheY	55	27	3.14	2.34
T. thermophilus IPMDH	109	31	15.8	1.73
E. coli IPMDH	90	34	17.8	3.69
T. kodakaraensis MGMT	98.6	29.5	10.2	1.2
E. coli AdaC	43.8	7.4	4.0	1.8
T. thermophilus PGK	~80	~30	~12	~0.06
S. cerevisiae PGK	60	25	5.3	1.6
S. solfataricus Sso7d	99	10	8.4	0.65
Average of 6 SH3 domains	71 ± 4	16	3.8	0.77 ± 0.04

*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); Sso7d, Knapp et al. (1996); and 6 SH3 domains, Knapp et al. (1998) and Filimonov et al. (1999).

THEORY

Electrostatic model

Fig. 2 *A* 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) that contains 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.

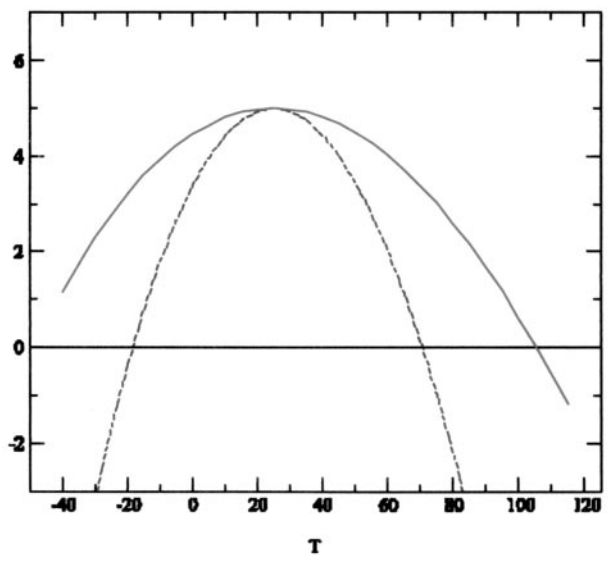

FIGURE 1 Temperature dependence of the unfolding free energy. The maximal stability is set to 5 kcal/mol, and the temperature at which this occurs is set to 25°C. By just reducing the heat capacity of unfolding from 1.5 kcal/mol/K (*dashed curve*) to 0.5 kcal/mol/K (*solid curve*), the melting temperature is increased from 70°C to 105°C.

Electrostatic contribution to ΔG

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)

$$U_0(q, r) = -166\left(\frac{1}{\varepsilon_p} - \frac{1}{\varepsilon_s}\right)\frac{q^2}{r}, \quad (2)$$

in which ϵ_p and ϵ_s are the dielectric constants of the protein medium and water, respectively. When s is not zero, the solvation energy is

$$U_{solv}(q, s, r) = -166\left(\frac{1}{\varepsilon_p} - \frac{1}{\varepsilon_s}\right)\frac{q^2}{r} \times \sum_{l=0}^{\infty}\frac{l+1}{l+1+(\varepsilon_p/\varepsilon_s)l}(s/r)^{2l}. \quad (3)$$

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

$$U_{int}(q, q', s, s', d, r) = \frac{332qq'}{\varepsilon_p r} - 332\left(\frac{1}{\varepsilon_p} - \frac{1}{\varepsilon_s}\right)\frac{qq'}{r} \times \sum_{l=0}^{\infty}\frac{l+1}{l+1+(\varepsilon_p/\varepsilon_s)l} \times (ss'/r^2)^l P_l(\cos\gamma), \quad (4)$$

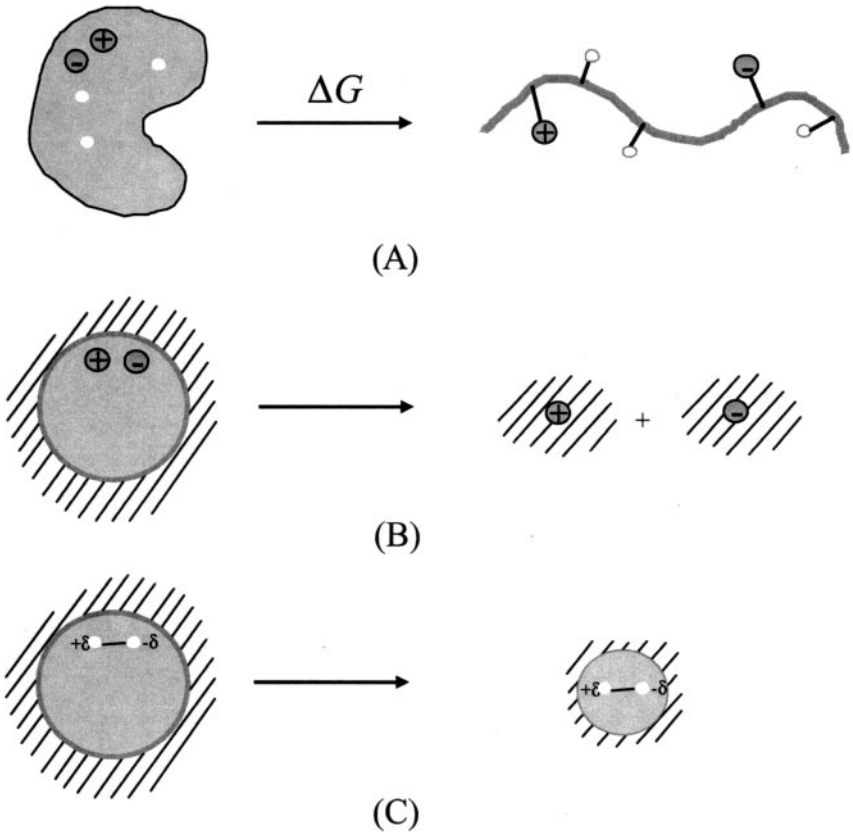

FIGURE 2 (*A*) Model of protein unfolding. In *B* and *C*, hashes represent the infinite solvent dielectric. A small circle with + or − inside represents an ionized group, whereas two small white circles connected by a line and with $+\delta$ and $-\delta$ attached represent a polar group.

$\cos\gamma = (s^2 + s'^2 - d^2)/2ss'$ and $P_l(x)$ are the Legendre polynomials.

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

$$\Delta G_{el} = [2U_0(e, a) - U_{solv}(e, s, R) - U_{solv}(e, s', R)] - U_{int}(e, -e, s, s', d, R), \quad (5)$$

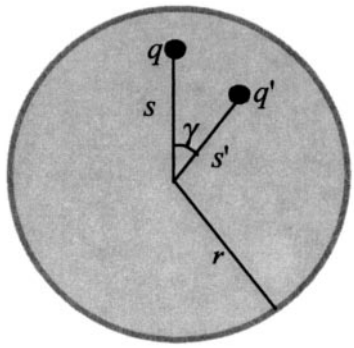

FIGURE 3 Spherical electrostatic model. It applies both to the folded protein (for which the radius $r = R$) and to ionized and polar groups in the unfolded state (in which $r = a$). When more than two charges are present, the electrostatic free energy can be calculated by considering one pair of charges at a time.

in which s and s' are the radial distances of the two charges in the folded proteins. Thus, ΔG_{el} consists of a solvation term ΔG_{solv} and an interaction term ΔG_{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

$$\Delta G_{solv} = 2U_{solv}(\delta, d/2, a) - U_{solv}(\delta, s, R) - U_{solv}(\delta, s', R) + U_{int}(\delta, -\delta, d/2, d/2, d, a) - U_{int}(\delta, -\delta, s, s', d, R). \quad (6)$$

Other charge distributions can be similarly accounted for.

Electrostatic contribution to ΔC_p

A standard thermodynamic relation leads to

$$\Delta C_p^{el} = -T\frac{\partial^2 \Delta G_{el}}{\partial T^2} \quad (7a)$$

$$= \Delta C_p^{solv} + \Delta C_p^{int}. \quad (7b)$$

The two terms in Eq. 7b arise from the solvation and interaction components of ΔG_{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 ϵ_s are (Archer and Wang, 1990):

$$\frac{\partial \ln \varepsilon_s}{\partial \ln T} = -1.37, \tag{8a}$$

$$\frac{\partial^2 \ln \varepsilon_s}{\partial (\ln T)^2} = -1.43. \tag{8b}$$

In particular, we have

$$-T\frac{\partial^2(1/\varepsilon_s)}{\partial T^2} = \frac{1}{T\varepsilon_s}\left[\frac{\partial^2 \ln \varepsilon_{sl}}{\partial (\ln T)^2} - \left(\frac{\partial \ln \varepsilon_{sl}}{\partial \ln T}\right)^2 - \frac{\partial \ln \varepsilon_{sl}}{\partial \ln T}\right]$$
$$= -\frac{1.94}{T\varepsilon_s}. \tag{9}$$

The negative sign of the value in Eq. 9 is the source of the main result (i.e., reduced ΔC_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 (NH_4^-, HCO_2^-, and $H_2PO_4^-$).

Choice of parameters

The protein dielectric constant ϵ_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.

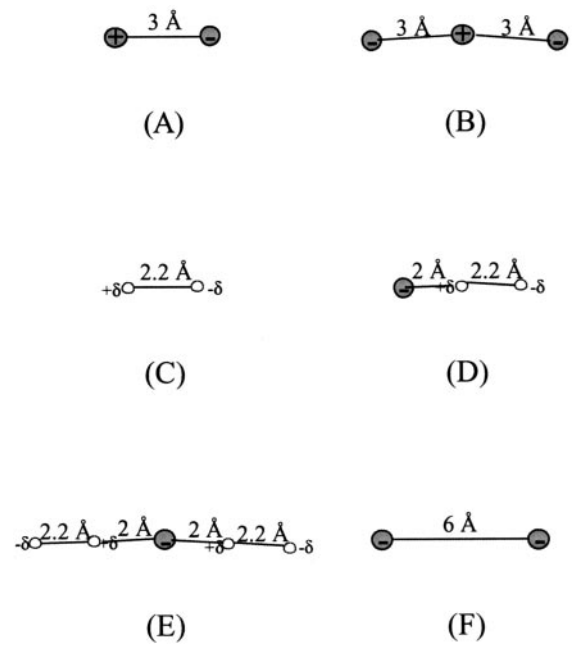

FIGURE 4 Different charge distributions considered in the present work: (*A*) a salt bridge, (*B*) a positive ion forming two salt bridges, (*C*) a polar group, (*D*) a positive ion forming a hydrogen bond with a polar group, (*E*) a positive ion forming hydrogen bonds with two polar groups, and (*F*) a pair of negative charges. All charges have the same radial distances of 14 Å in the folded state, except in *F*, where the radial distances are 14.7 Å. The 6-Å separation between the two negative ions in *F* is roughly the distance between residues E3 and E66 in *B. subtilis* CspB (PDB entry 1csp; Schindelin et al., 1993). In this case, the two charges are moved closer to the protein surface to reduce the destabilizing effect (desolvation cost plus charge-charge repulsion).

RESULTS AND DISCUSSION

Contributions of a salt-bridge network to ΔG and ΔC_p

The various charge distributions considered in the present study are shown in Fig. 4. The calculated results of their contributions to ΔG and ΔC_p are listed in Table 2. For an ion pair (i.e., distribution A), the desolvation cost ($-\Delta G_{solv}$) 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 ΔG_{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 ΔC_p by ~10 cal/mol/K.

Contributions of a hydrogen-bonding network to ΔG and ΔC_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 ΔC_p by ~5 cal/mol/K.

Reduction of ΔC_p by polar interactions

The spherical model yields a potentially important result: Polar interactions around an ionized group in the folded state significantly reduce ΔC_p. Although the contribution of the solvent exposure of polar groups to ΔC_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 implicated 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 ΔC_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 ΔC_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 ΔC_p.

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

$$U_{int} = -\frac{332e^2}{\varepsilon_s d}. \qquad (10)$$

The contribution of the interaction energy to ΔC_p is (see Eqs. 5 and 7a)

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

The second factor is given by Eq. 9 and is negative, thus the interaction reduces ΔC_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 ε_p) sharing a planar boundary with the high dielectric (having dielectric constant ε_s). The image charge of charge $+e$ is $-(\varepsilon_s - \varepsilon_p)/(\varepsilon_s + \varepsilon_p)e$. The interaction energy is thus

$$U_{int} = -\frac{332e^2}{\varepsilon_p d} + \frac{332e^2}{\varepsilon_p d'} - \frac{664e^2}{(\varepsilon_s + \varepsilon_p)d'}, \qquad (12)$$

in which d' is the distance between the image charge and charge $-e$. The only term contributing to ΔC_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 ε_p to ε_s (note $\varepsilon_p \ll \varepsilon_s$). Again, a negative contribution to ΔC_p is obtained.

If polar interactions around ionized groups in the folded state reduce ΔC_p, to what extent do these interactions contribute to the lower ΔC_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 ΔC_p, then the 10 charged residues will reduce ΔC_p by 0.2 kcal/mol/K. This is a significant fraction of the average of 1 kcal/mol/K for the difference in ΔC_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 ΔC_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 ΔC_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 ΔC_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 ΔC_p is perhaps an underestimate by the spherical model, but there might be an additional source for the gap between the resulting value of

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

Residue*	Polar interactions†
R2L	E64 (NH2-OE1: 3.3 Å); K3 (NH2-N: 3.0 Å); R4 (NH2-N: 2.7 Å)
R4Q	D66 (NH2-OD1: 3.3 Å; NH2-OD2: 3.3 Å); E64 (NE-OE2: 3.7 Å)
A6E	*R27* (*OE2-NH1: 3.3 Å*)
E39Y	R46 (OE1-NH2: 3.4 Å); K50 (OE2-NZ: 3.7 Å)
S41R	*None*
K50M	E39 (NZ-OE2: 3.7 Å)
E54V	K57 (OE1-NZ: 3.6 Å; OE2-NZ: 3.7 Å)
H62P	*Q113* (*NE2-NE2: 3.4 Å*)
D66I	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)
H72Q	E48 (ND1-OE2: 3.8 Å); D70 (ND1-O: 3.6 Å)
K76Q	W81 indole ring: 3.5 Å
E80Q	T77 (OE2-OG1: 3.8 Å)
G95K	*None*
R101V	V98 (NE-O: 3.0 Å); P97 (NH1-O: 3.9 Å)
E105Q	R101 (OE2-O: 3.5 Å)
A106R	*E57* (*NE-OE2: 2.9 Å; NH1-OE1: 2.9 Å*)
L108D	*K86* (*OD1-NZ: 2.9 Å; OD2-NZ: 3.1 Å*)
R115Q	E64 (NH1-OE1: 3.1 Å)
R135Δ	None
R138A	D134 (NH1-OD1: 3.5 Å)
K146A	Q144 (NZ-O: 2.8 Å)
R152L	No coordinates
A153E	*None*
P154D	*No coordinates*
H158Q	No coordinates
E159V	No coordinates

*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/Å^2 for ΔC_p 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 ΔC_p.

All of our calculation results are for room temperature. Both thermophilic and mesophilic show maximal stability around this temperature, and the maximal stability of thermophilic proteins is typically higher (Table 1). We illustrated that a salt-bridge or hydrogen-bonding network around an ionized group can increase ΔG and decrease ΔC_p at the same time. The reduced ΔC_p is due in part to the decrease of ε_s with temperature (see Eqs. 9 and 8a). The decrease of ε_s at high temperatures will decrease the desolvation cost and increase the strength of charge-charge interactions, resulting in more favorable contributions to folding stability. This fact was noted by Elcock (1998). However, our calculations indicate that, even at room temperature, a salt-bridge or hydrogen-bonding network around a charged residue can contribute to the typically observed higher stability of thermophilic proteins.

Enriched polar interactions in *Thermus thermophilus* RNase H

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 *Escherichia 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 ΔC_p (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

TABLE 3 Contributions of salt-bridge and hydrogen bonding networks to ΔG (in kcal/mol) and ΔC_p (in cal/mol/K) at room temperature

Charge distribution	ΔG_{solv}	ΔG_{int}	ΔG_{el}	ΔC_p^{solv}	ΔC_p^{int}	ΔC_p
A	−12.4	11.6	−0.8	−1.4	−7.6	−9.0
B	−18.6	20.4	1.8	−2.1	−10.4	−12.5
C	−2.8		−2.8	−1.3		−1.3
D	−9.0	8.7	−0.3	−2.0	−3.1	−5.1
E	−11.8	16.7	4.9	−3.3	−6.1	−9.4
F	−2.3	−1.9	−4.2	4.4	5.2	9.6

have more polar interactions around charged residues than ECTO.

Exceptions to reduced ΔC_p of thermophilic proteins

Although we have presented a trend of reduced ΔC_p in thermophilic proteins, there are exceptions. At 0.2 M KCl, archaeal histones HMfA, HMfB, and HPyA1 from thermophilic *M. fervidus* and *Pyrococcus* strain GB-3a have average ΔC_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 ΔC_p of 2.8 kcal/mol/K. However, at a salt concentration of 1 M, the difference in ΔC_p disappears: HMfA has an average ΔC_p of 2.0 kcal/mol/K, whereas HFoB has an average ΔC_p of 2.1 kcal/mol/K. The difference in ΔC_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 ΔC_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 ΔC_p. The technical difficulty in the accurate measurement of ΔC_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 ΔC_p of thermophilic 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 ΔC_p.

Linking of enriched polar interactions and reduced ΔC_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 ΔC_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 ΔC_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 ΔC_p.

In the past, residual structure in the unfolded state has been suggested as a possible explanation of the reduced ΔC_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 ΔC_p. Possibly, a reduced ΔC_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 ΔC_p, it is of interest to see whether charge mutations will restore ΔC_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 ΔC_p of *T. maritima* CheY. This work was supported in part by the National Institutes of Health Grant GM58187.

REFERENCES

Abraham, M. H., and Y. Marcus. 1986. The thermodynamics of solvation of ions: Part 1. The heat capacity of hydration at 298.15 K. *J. Chem. Soc. Faraday Trans.* 82:3255–3274.

Archer, D. A., and P. Wang. 1990. The dielectric constant of water and Debye-Huckel limiting law slopes. *J. Phys. Chem. Ref. Data.* 19: 371–411.

Backmann, J., G. Schafer, L. Wyns, and H. Bonisch. 1998. Thermodynamics and kinetics of unfolding of the thermostable trimeric adenylate kinase from the archaeon *Sulfolobus acidocaldarius*. *J. Mol. Biol.* 284:817–833.

Baldwin, R. L. 1986. Temperature dependence of the hydrophobic interaction in protein folding. *Proc. Natl. Acad. Sci. U. S. A.* 83:8069–8072.

Born, M. 1920. Volumen u. hydratationswarme der ionen. *Z. Phys.* 1:45–48.

Creighton, T. E. 1993. Proteins: Structures and Molecular Properties, 2nd ed. Freeman, New York.

Delbruck, H., U. Mueller, D. Perl, F. X. Schmid, and U. Heinemann. 2001. Crystal structures of mutant forms of the *Bacillus caldolyticus* cold shock protein differing in thermal stability. *J. Mol. Biol.* 313:359–369.

Deutschman, W. A., and F. W. Dahlquist. 2001. Thermodynamic basis for the increased thermostability of CheY from the hyperthermophile *Thermotoga maritima*. *Biochemistry.* 40:13107–13113.

Dominy, B. N., D. Perl, F. X. Schmid, C. L. Brooks III. 2002. The effects of ionic strength on protein stability: the cold shock protein family. *J. Mol. Biol.* 319:541–554.

Elcock, A. H. 1998. The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. *J. Mol. Biol.* 284:489–502.

Filimonov, V. V., A. I. Azuaga, A. R. Viguera, L. Serrano, and P. L. Mateo. 1999. A thermodynamic analysis of a family of small globular proteins: SH3 domains. *Biophys. Chem.* 77:195–208.

Gallagher, K., and K. Sharp. 1998. Electrostatic contributions to heat capacity changes of DNA-ligand binding. *Biophys. J.* 75:768–776.

Goedken, E. R., J. L. Keck, J. M. Berger, and S. Marqusee. 2000. Divalent metal cofactor binding in the kinetic folding trajectory of *Escherichia coli* ribonuclease HI. *Protein Sci.* 9:1914–1921.

Hollien, J., and S. Marqusee. 1999. A thermodynamic comparison of mesophilic and thermophilic ribonucleases H. *Biochmeistry*. 38: 3831–3836.

Ishikawa, K., M. Okumura, K. Katayanagi, and S. Kimura. 1993. Crystal structure of ribonuclease H from *Thermus thermophilus* HB8 refined at 2.8 A resolution. *J. Mol. Biol.* 230:529–542.

Jaenicke, R., and G. Bohm. 1998. The stability of proteins in extreme environments. *Curr. Opin. Struct. Biol.* 8:738–748.

Knapp, S., A. Karshikoff, K. D. Berndt, P. Christova, B. Atanasov, and R. Ladenstein. 1996. Thermal unfolding of the DNA-binding protein Sso7d from the hyperthermophile *Sulfolobus solfataricus*. *J. Mol. Biol.* 264: 1132–1144.

Knapp, S., P. T. Mattson, P. Christova, K. D. Berndt, A. Karshikoff, M. Vihinen, C. I. E. Smith, and R. Ladenstein. 1998. Thermal unfolding of small proteins with SH3 domain folding pattern. *Proteins.* 31:309–319.

Li, W.-t., R. A. Grayling, K. Sandman, S. Edmondson, J. W. Shriver, and J. N. Reeve. 1998. Thermodynamic stability of archaeal histones. *Biochmeistry.* 37:10563–10572.

Livingstone, J. R., R. S. Spolar, and M. T. Record, Jr. 1991. Contribution to the thermodynamics of protein folding from the reduction in water-accessible nonpolar surface area. *Biochemistry.* 30:4237–4244.

Loladze, V. V., D. N. Ermolenko, and G. I. Makhatadze. 2001. Heat capacity changes upon burial of polar and nonpolar groups in proteins. *Protein Sci.* 10:1343–1352.

Madan, B., and K. Sharp. 1996. Heat capacity changes accompanying hydrophobic and ionic solvation: a monte carlo and random network model study. *J. Phys. Chem.* 100:7713–7721.

Madan, B., and K. Sharp. 2001. Hydration heat capacity of nucleic acid constituents determined from the random network model. *Biophys. J.* 81:1881–1887.

Madura, J. D., J. M. Briggs, R. C. Wade, M. E. Davis, B. A. Luty, A. Ilin, J. Antosiewicz, M. K. Gilson, B. Bagheri, L. R. Scott, and J. A. McCammon. 1995. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian Dynamics program. *Comput. Phys. Commun.* 91:57–95.

Makhatadze, G. I., and P. L. Privalov. 1995. Energetics of protein structure. *Adv. Protein Chem.* 47:307–425.

McCrary, B. S., S. P. Edmondson, and J. W. Shriver. 1996. Hyperthermophile protein folding thermodynamics: differential scanning calorimetry and chemical denaturation of Sac7d. *J. Mol. Biol.* 264:784–805.

Moczygemba, C., J. Guidry, K. L. Jones, C. M. Gomes, M. Teixeira, and P. Wittung-Stafshede. 2001. High stability of a ferredoxin from the hyperthermophilic archaeon *A. ambivalens*: involvement of electrostatic interactions and cofactors. *Protein Sci.* 10:1539–1548.

Motono, C., T. Oshima, and A. Yamagishi. 2001. High thermal stability of 3-isopropylmalate dehydrogenase from *Thermus thermophilus* resulting from low $\Delta\Delta C_p$ of unfolding. *Protein Eng.* 14:961–966.

Murphy, K. P., and E. Freire. 1992. Thermodynamics of structural stability and cooperative folding behavior in proteins. *Adv. Protein Chem.* 43: 313–361.

Myers, J. K., C. N. Pace, and J. M. Scholtz. 1995. Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* 4:2138–2148.

Nojima, H., A. Ikai, T. Oshima, and H. Noda. 1977. Reversible thermal unfolding of thermostable phosphoglycerate kinase: thermostability associated with mean zero enthalpy change. *J. Mol. Biol.* 116:429–442.

Perl, D., U. Mueller, U. Heinemann, and F. X. Schmid. 2000. Two exposed amino acid residues confer thermostability on a cold shock protein. *Nature Struct. Biol.* 7:380–383.

Perutz, M. F. 1978. Electrostatic effects in proteins. *Science.* 201: 1187–1191.

Perutz, M. F., and H. Raidt. 1975. Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2. *Nature.* 255:256–259.

Petrosian, S. A., and G. I. Makhatadze. 2000. Contribution of proton linkage to the thermodynamic stability of the major cold-shock protein of Escherichia coli CspA. *Protein Sci.* 9:387–394.

Petsko, G. A. 2001. Structural basis of thermostability in hyperthermophilic proteins, or "there's more than one way to skin a cat." *Methods Enzymol.* 334:468–478.

Privalov, P. L., and G. I. Makhatadze. 1990. Heat capacity of proteins: II. Partial molar heat capacity of the unfolded polypeptide chain of proteins: protein unfolding effects. *J. Mol. Biol.* 213:385–391.

Robertson, A. D., and K. P. Murphy. 1997. Protein structure and the energetics of protein stability. *Chem. Rev.* 97:1251–1267.

Robic, S., J. M. Berger, and S. Marqusee. 2002. Contributions of folding cores to the thermostabilities of two ribonucleases H. *Protein Sci.* 11:381–389.

Sanchez-Ruiz, J. M., and G. I. Makhatadze. 2001. To charge or not to charge? *Trends Biotechnol.* 19:132–135.

Schindelin, H., M. A. Marahiel, and U. Heinemann. 1993. Universal nucleic acid-binding domain revealed by crystal structure of the *B. subtilis* major cold-shock protein. *Nature.* 364:164–168.

Shiraki, K., S. Nishikori, S. Fujiwara, H. Mashimoto, Y. Kai, M. Takagi, and T. Imanaka. 2001. Comparative analyses of the conformational stability of a hyperthermophilic protein and its mesophilic counterpart. *Eur. J. Biochem.* 268:4144–4150.

Spolar, R. S., J. R. Livingstone, and M. T. Record, Jr. 1992. Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. *Biochemistry.* 31:3947–3955.

Szilagyi, A., and P. Zavodszky. 2000. Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure.* 8:493–504.

Vijayakumar, M., and H.-X. Zhou. 2001. Salt bridges stabilize the folded structure of barnase. *J. Phys. Chem. B.* 105:7334–7340.

Vogt, G., and P. Argos. 1997. Protein thermal stability, hydrogen bonds, and ion pairs. *Fold Des.* 2:S40–S46.

Wassenberg, D., C. Welker, and R. Jaenicke. 1999. Thermodynamics of the unfolding of the cold-shock protein from *Thermotoga maritima*. *J. Mol. Biol.* 289:187–193.

Xiao, L., and B. Honig. 1999. Electrostatic contributions to the stability of hyperthermophilic proteins. *J. Mol. Biol.* 289:1435–1444.

Zhou, H.-X. 2002. A Gaussian-chain model for treating residual charge-charge interactions in the unfolded state of proteins. *Proc. Natl. Acad. Sci. U. S. A.* 99:3569–3574.