ELSEVIER

Biophysical Chemistry 111 (2004) 235–246

Biophysical Chemistry

www.elsevier.com/locate/bpc

Experiment-guided thermodynamic simulations on reversible two-state proteins: implications for protein thermostability

Sandeep Kumar[a], Ruth Nussinov[b,c,*]

[a] *Department of Biological Sciences and Bioengineering, Indian Institute of Technology Kanpur, Kanpur, U.P. 208016, India*
[b] *Basic Research Program, SAIC-Frederick, Inc., Laboratory of Experimental and Computational Biology, NCI-Frederick Building 469, Room 151, Frederick, MD 21702, USA*
[c] *Department of Human Genetics and Molecular Medicine, Sackler Institute of Molecular Medicine, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel*

Received 1 March 2004; received in revised form 27 May 2004; accepted 1 June 2004
Available online 8 July 2004

Abstract

Here, we perform protein thermodynamic simulations within a set of boundary conditions, effectively blanketing the experimental data. The thermodynamic parameters, melting temperature (T_G), enthalpy change at the melting temperature (ΔH_G) and heat capacity change (ΔC_p) were systematically varied over the experimentally observed ranges for small single domain reversible two-state proteins. Parameter sets that satisfy the Gibbs–Helmholtz equation and yield a temperature of maximal stability (T_S) around room temperature were selected. The results were divided into three categories by arbitrarily chosen T_G ranges. The T_G ranges in these categories correspond to typical values of the melting temperatures observed for the majority of the proteins from mesophilic, thermophilic and hyperthermophilic organisms. As expected, ΔC_p values tend to be high in mesophiles and low in hyperthermophiles. An increase in T_G is accompanied by an up-shift and broadening of the protein stability curves, however, with a large scatter. Furthermore, the simulations reveal that the average ΔH_G increases with T_G up to $\sim$ 360 K and becomes constant thereafter. ΔC_p decreases with T_G with different rates before and after $\sim$ 360 K. This provides further justification for the separate grouping of proteins into thermophiles and hyperthermophiles to assess their thermodynamic differences. This analysis of the Gibbs–Helmholtz equation has allowed us to study the interdependence of the thermodynamic parameters T_G, ΔH_G and ΔC_p and their derivatives in a more rigorous way than possible by the limited experimental protein thermodynamics data available in the literature. The results provide new insights into protein thermostability and suggest potential strategies for its manipulation.

Keywords: Heat capacity; Melting temperature; Thermophiles; Hyperthermophiles; Thermodynamic simulations; Stability; Thermodynamics

Abbreviations: T_G, heat denaturation (melting) temperature; T'_G, cold denaturation temperature; T_S, temperature of maximal protein stability; ΔH_G, molar enthalpy change between native and denatured states of protein at T_G; ΔS_G, molar entropy change between native and denatured states of protein at T_G; ΔC_p, molar heat capacity change between native and denatured states of protein; $\Delta G(T)$, molar Gibbs free energy change between native and denatured states of protein at temperature T; $\Delta G(T_S)$, molar Gibbs free energy change between native and denatured states of protein at T_S; TRange, temperature range of thermodynamic stability; $-\Delta S_G$, slope of the protein stability curve at T_G; $-\Delta C_p/T_G$, curvature of the protein stability curve at T_G; °C, degrees Celsius; kcal/mol, kilocalories per mole; kcal mol^{-1} K^{-1}, kilocalories per mole per Kelvin; DSC, differential scanning calorimetry; CD, circular dichroism; V_T, total volume available for sampling by a simulation in a cuboid defined by vertices T_G, ΔH_G and ΔC_p; V_S, the volume of $T_G\Delta H_G\Delta C_p$ cuboid sampled by a simulation; N_S, the number of parameter sets sampled by a simulation that yield T_S to be around room temperature; G, granularity of a simulation; V_F, fraction of the available volume sampled by a simulation.

* Corresponding author. Basic Research Program, SAIC-Frederick, Inc., Laboratory of Experimental and Computational Biology, NCI-Frederick Building 469, Room 151, Frederick, MD 21702, USA. Tel.: +1-301-846-5579; fax: +1-301-846-5598.
E-mail address: ruthn@ncifcrf.gov (R. Nussinov).

0301-4622/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2004.06.005

1. Introduction

Thermophilic and hyperthermophilic proteins have immense potential for use in several industrial applications [1–4]. The increased thermal stability is beginning to be understood, mainly based on comparisons of sequences and structures from homologous thermophilic and mesophilic proteins [5–7]. Among the several routes to higher protein thermostability, the most frequent is via modification of protein electrostatics, relieving electrostatic repulsions [8], increased formation of salt bridges and their networks [7,9–13] and cation–pi interactions [14,15]. A second approach to study protein thermostability is analysis of experimental thermodynamic data and comparisons among families of homologous thermophilic and mesophilic proteins [16–23]. Available experimental data were compiled in a book [24] and in an electronic database [25].

Single domain proteins which undergo reversible two-state folding↔unfolding transition show the simplest temperature-dependent behavior. For such proteins with positive ΔC_p, the temperature-dependent variation in their thermodynamic stability can be described by the Gibbs–Helmholtz equation [26,27]:

$$\Delta G(T) = \Delta H_G(1 - T/T_G) - \Delta C_p[(T_G - T) + T\ln(T/T_G)] \quad (1)$$

where $\Delta G(T)$ is the Gibbs free energy change between the denatured (D) and the native (N) states of a protein at a given temperature T, ΔH_G is the enthalpy change between the two states at the melting temperature (T_G) and ΔC_p is the heat capacity change between the two states. ΔH_G, ΔC_p and T_G are usually determined using differential scanning calorimetry and spectroscopy. Plotting $\Delta G(T)$ versus T using Eq. (1) yields a skewed parabola with a maximum in protein stability [$\Delta G(T_S)$] at a temperature T_S [26,27].

The thermodynamic properties of homologous thermophilic versus mesophilic proteins were interpreted in terms of the microscopics of protein sequence and structure [21]. The stability curves of thermophilic proteins were up-shifted and broadened as compared to their mesophilic homologues due to reduced ΔC_p and increased per residue enthalpy change (Δh_G), showing greater maximal stabilities. We have further collected thermodynamic data from the literature on proteins that show reversible two-state folding↔unfolding transition at or near neutral pH. The proteins are derived from hyperthermophilic, mesophilic and psychrophilic organisms. Most have large hydrophobic cores and are maximally stable around room temperature ($T_S = 293.8$ K), irrespective of size, fold, number of domains, oligomeric state, melting temperature and the living temperature of the source organisms. Analyses of the single domain proteins revealed interesting correlations, for example, the average free energy contribution by individual residues towards maximal protein stability decreases with the protein size, indicating an upper limit for maximal stability for the single domains [22]. The temperature range over which a single domain protein is stable decreases with ΔC_p but increases with $\Delta G(T_S)$ [23].

Here, we further explore protein thermostability. We simulate protein thermodynamics using the Gibbs–Helmholtz equation to explore the dependencies of the thermodynamic parameters, T_G, ΔH_G and ΔC_p, and their derivatives on each other. The simulations are guided by experimental data on single domain proteins. We perform systematic variations in T_G, ΔH_G and ΔC_p over the experimentally observed ranges, similar to a grid search strategy in molecular simulations. The simulations provide a more rigorous basis to observations from analyses of the limited experimental data on protein thermodynamics and yield further insight that could not be obtained from experimental data alone. They provide a rational basis for segregating hyperthermophilic from thermophilic proteins. The results are divided into three arbitrary categories with T_G ranges typical for proteins from mesophiles, thermophiles and hyperthermophiles. The average ΔH_G increases with T_G up to ~ 360 K and becomes constant thereafter. The average ΔC_p decreases with T_G, but the rate of decrease changes around this temperature. Reduction in ΔC_p is the key to protein thermostability and leads to the broadening and up-shift of the stability curves for the thermostable proteins. Most proteins are maximally stable around room temperature [22,28], irrespective of the living environments of their source organisms. The simulations suggest that this results in a negative relationship between ΔC_p and T_G. Our results have important implications for rational manipulation of protein stability.

2. Methods

2.1. Twelve single domain proteins in our study

We have already described the procedure of the data collection and quality control [22,23]. Briefly, our dataset consists of 12 single domain proteins with two-state folding↔unfolding transitions in the pH range 6–8, with $\geq 90\%$ reversibility. The cooperativity ratio, R, for these proteins lies in the range 0.9–1.10. All are maximally stable around room temperature and contain a single hydrophobic core. These proteins are Sso7d [29], TmCsp [30], λrepressor$_{6-85}$ [31], BsHPr [24,32], Barstar [33], C_L of Igλ [34], FeCyt b_{562} [35], Thioredoxin [36], Snase [37], GDH domain II [38], Apoflavodoxin [39] and K-STI [40]. The average thermodynamic parameters for these 12 proteins are $T_G = 343 \pm 14$ K (range: 325–371 K), $\Delta H_G = 77 \pm 18$ kcal/mol (range: 58–107 kcal/mol) and $\Delta C_p = 1.6 \pm 0.6$ kcal mol^{-1} K^{-1} (range: 0.6–2.6 kcal mol^{-1} K^{-1}). The average maximal stability is 5.5 ± 1.6 kcal/mol (range: 3.5–9.0 kcal/mol), and the average size is 113 ± 40 residues (range: 62–181 residues). The ranges of variations in our dataset have inspired the boundary conditions for the different simulations presented here.

2.2. Calculation of the derived thermodynamic parameters

Using Eq. (1) and the experimentally determined T_G, ΔH_G and ΔC_p for a reversible two-state protein, we can plot the protein stability curves and derive the slope at T_G ($-\Delta S_G$), curvature ($-\Delta C_p/T_G$) at T_G, the temperature of maximal protein stability (T_S), maximal protein stability[($\Delta G(T_S)$] and the temperature range (TRange) of thermodynamic stability of the native state [26,27].

The slope of the protein stability curve at T_G is given by:

$$(\mathrm{d}\Delta G/\mathrm{d}T)_{T=T_G} = -\Delta H_G/T_G = -\Delta S_G \quad (2)$$

Hence, at T_G, the slope of the protein stability curve is determined by ΔH_G. The slope also represents the entropy difference between the folded and unfolded states of the protein at T_G. The curvature of the protein stability curve at T_G is given by:

$$(\mathrm{d}^2\Delta G/\mathrm{d}T^2)_{T=T_G} = -\Delta C_p/T_G \quad (3)$$

Hence, ΔC_p determines the curvature of the protein stability curve at T_G. The temperature at which the protein is maximally stable (T_S) is given by:

$$T_S = T_G \exp[-\Delta H_G/(T_G \Delta C_p)] \quad (4)$$

Note that T_S should not be confused with T_S* which is the temperatures at which the entropy of unfolding (ΔS) is zero. The maximal protein stability [$\Delta G(T_S)$] is given by:

$$\Delta G(T_S) = \Delta H_G - (T_G - T_S)\Delta C_p \quad (5)$$

The temperature range (TRange) over which the native state of the protein is thermodynamically stable is given by:

$$\mathrm{TRange} = T_G - T'_G \quad (6)$$

where T'_G is the cold denaturation temperature at which $\Delta G(T) = 0$. This temperature was estimated using the Gibbs–Helmholtz equation and numerical interpolation.

2.3. Protein thermodynamics simulations

We systematically vary ΔH_G, ΔC_p and T_G over the experimental value ranges for the 12 single domain reversible two-state proteins:

$$\Delta H_G = 50 - 150 \text{ kcal/mol}^{-1}$$

$$\Delta C_p = 0.5 - 3.0 \text{ kcal mol}^{-1} \text{ K}^{-1}$$

$$T_G = 300 - 400 \text{ K}$$

and use Eq. (4) to select the sets of ΔH_G, ΔC_p and T_G which yield T_S in the range of 273–310 K. This is the observed range of T_S in the experimental data [22]. The protein stability curve for a selected parameter set was plotted using Eq. (1) and the TRange of thermodynamic stability was computed using Eq. (6). Other parameters were computed using the equations above.

The thermodynamic simulations described here may be thought of as sampling over the volume of a three-dimensional cuboid with vertices defined by T_G, ΔH_G and ΔC_p, although one should remember that the three parameters are not independent of one another. The total volume (V_T) of the cuboid available for sampling depends upon the ranges used for the three parameters. In the present study, this volume is 2.5×10^4 (kcal/mol)^2 (100 kcal/mol $\times$ 2.5 kcal mol^{-1} K^{-1} $\times$ 100 K). The volume actually sampled by the simulations (V_S) is limited by the requirement for T_S to be around room temperature. The procedure followed is similar to the grid search (conformational sampling) used in molecular simulations. However, unlike molecular simulations, there is no force-field parameterization and the simulations are not atomistic.

In total, three simulations have been performed at different levels of granularity (G) determined by the steps by which the three parameters were varied. The characteristic properties of these simulations are described in Table 1. The steps for variation of the thermodynamic parameters and thus granularity were chosen arbitrarily.

2.3.1. Simulation 1 (S1)

In the first simulation, ΔH_G was varied in steps of 0.5 kcal/mol, ΔC_p in steps of 0.05 kcal mol^{-1} K^{-1} and T_G in steps of 0.5 K. The granularity in this case is 0.0125 (kcal/mol)^2 (0.5 kcal/mol $\times$ 0.05 kcal mol^{-1} K^{-1} $\times$ 0.5 K). A total of 663,295 sets which yield T_S in the range of 273–310 K were selected. These sets cover 33.2% of the total space available to S1 (Table 1).

2.3.2. Simulation 2 (S2)

In the second simulation, ΔH_G, ΔC_p and T_G were varied in steps of 1 kcal/mol, 0.1 kcal mol^{-1} K^{-1} and 1 K, respectively, with $G = 0.1$ (kcal/mol)^2. A total of 84,461 parameter sets that yield T_S in the range of 273–310 K

Table 1
Characteristics of various simulations in the present study[a]

Simulation	G (kcal/mol)^2	N_S	V_S (kcal/mol)^2	V_T (kcal/mol)^2	V_F (%)
S1	0.0125	663,295	8291.19	25,000	33.16
S2	0.1	84,461	8446.10	25,000	33.78
S3	0.005	1,657,777	8288.89	25,000	33.16

[a] For each simulation, G denotes the granularity of the simulation, N_S stands for the number of parameter sets (T_G, ΔH_G, ΔC_p) with temperature of maximal stability being around the room temperature ($273 \le T_S \le 310$K). V_S denotes the volume sampled by a simulation in the $\Delta H_G \Delta C_p T_G$ space and V_T shows the total volume available for sampling to the simulation. V_F measures the fraction of the available volume covered by simulation. Note that all simulations sample approximately one-third of the total volume (V_T).

were selected. S2 samples 33.8% of the available space (Table 1).

2.3.3. Simulation 3 (S3)

For the third simulation (S3), we varied ΔH_G in steps of 0.1 kcal/mol, ΔC_p in steps of 0.01 kcal mol^{-1} K^{-1} and T_G in steps of 5 K. The granularity was 0.005 $(kcal/mol)^2$. A total of 1,657,777 ($\sim 1.7 \times 10^6$) parameter sets with T_S within 273–310 K range were selected. S3 covers 33.2% of the total $\Delta H_G \Delta C_p T_G$ volume available (Table 1).

For each simulation, we have computed the fractional volume (V_F) of the $T_G \Delta H_G \Delta C_p$ cuboid sampled by the simulation, using the following formulas:

$$V_F = V_S \; 100 / V_T \tag{7}$$

and

$$V_S = N_S G \tag{8}$$

V_S, V_T and G are defined above. N_S is the number of the (ΔH_G, ΔC_p, T_G) sets which yield T_S in the range of 273–310 K.

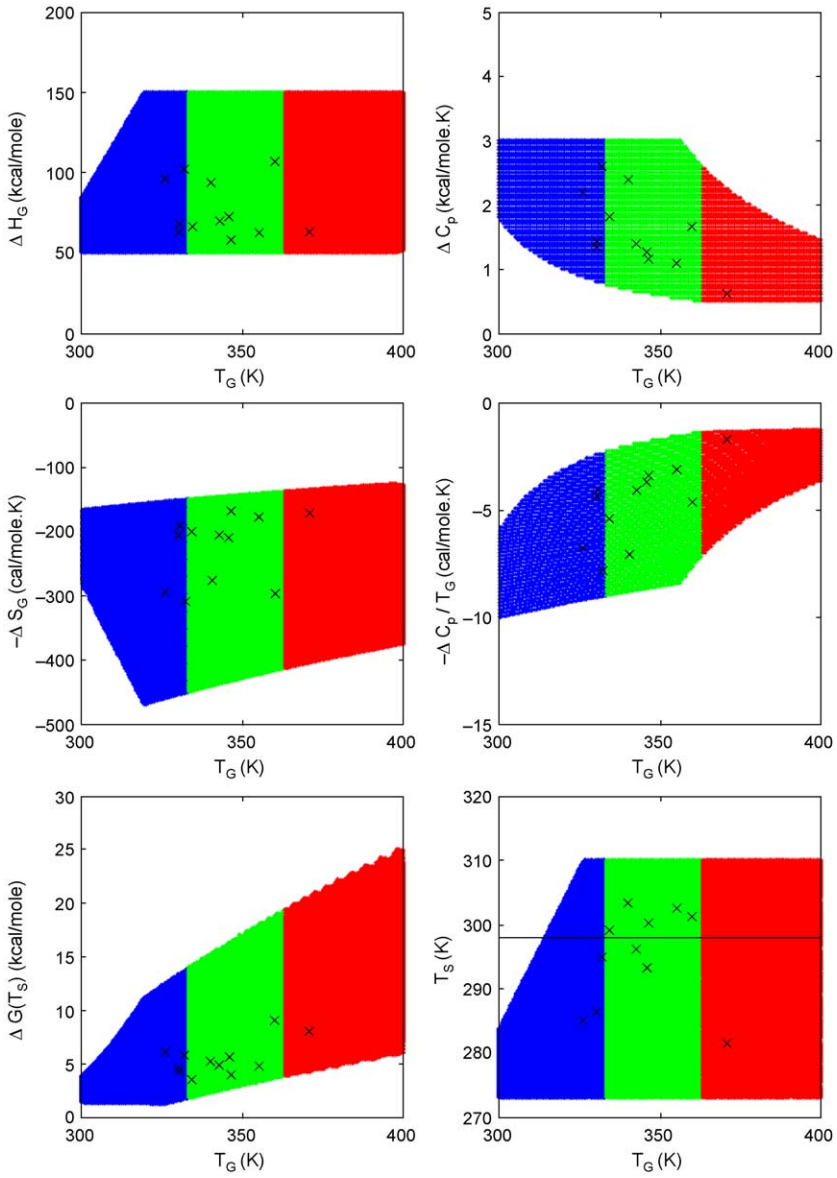

Fig. 1. The results of simulation 1 (S1). Various thermodynamic parameters [ΔH_G, ΔC_p, $-\Delta S_G$ (slope), $-\Delta C_p/T_G$ (curvature), $\Delta G(T_S)$ and T_S] are plotted with respect to T_G. In each plot, the X-axis is represented by T_G. The data shown here correspond to 663,295 combinations (sets of T_G, ΔH_G and ΔC_p) with T_S in the range 273–310 K obtained from S1. The data are divided into three classes with respect to T_G and shown in different color in each plot. Blue refers to class I. The T_G in this class falls in the range 300–333 K (27–60 °C). Green refers to class II with T_G 333.5–363 K (60.5–90 °C) and red refers to class III with T_G 363.5–400 K (90.5 °C–127 °C). The three classes roughly correspond to proteins from mesophiles (blue), thermophiles (green) and hyperthermophiles (red). The data from real proteins are also represented in the same plots by symbols 'x'. Our simulation is able to reproduce all the real data points.

2.3.4. Control simulations

In addition to these three simulations, we have conducted two further simulations. In the first, parameter sets which yield $T_S < 273$ K were selected. In the second, we selected sets which yield T_S above 310 K.

3. Results

3.1. Simulation 1 (S1)

The results of this simulation are shown in Fig. 1. We have categorized the thermodynamic parameter sets obtained by S1 into three classes according to the melting temperature (T_G), namely, class I [237,113 sets (36%) with $300 \le T_G \le 333$ K; shown in blue color], class II [279,156 sets (42%) with $333 < T_G \le 363$ K; green] and class III [147,026 (22%) sets with $363 < T_G \le 400$ K; red]. The temperature ranges for these classes were chosen arbitrarily and correspond to the typical melting temperature ranges of proteins from mesophilic, thermophilic and hyperthermophilic organisms [3–7,21]. Data from the real proteins (12 single domains) are also plotted in this figure. The simulation not only covers the experimental data points, but also explores larger space. Table 2 provides the average values of thermodynamic parameters in these three classes along with the data for the real proteins.

For the mesophilic (class I) and hyperthermophilic (class III) proteins, our simulations do not cover the full range of ΔC_p (0.5–3.0 kcal mol^{-1} K^{-1}). In class I, low values of ΔC_p are not sampled and the simulation visits the higher values more often. The lowest value of ΔC_p is 0.8 kcal mol^{-1} K^{-1} for class I. Similarly, high values of ΔC_p are not sampled in hyperthermophiles and the simulation covers ΔC_p in the range of 0.5–2.55 kcal mol^{-1} K^{-1} for this class. The average ΔC_p decreases from mesophiles → thermophiles → hyperthermophiles (Table 2). The difference between the average ΔC_p values for classes I (mesophiles) and II (thermophiles) is small. However, ΔC_p values decrease to a greater extent between classes II (thermophiles) and III (hyperthermophiles). The changes in the average ΔC_p values in the three classes are not statistically significant (Table 2). However, this trend is consistent with the third simulation (S3), as described below. Analysis of experimental data also showed a negative correlation between T_G and ΔC_p [23]. To investigate whether this negative relationship is due to T_S being restricted to room temperature, we have conducted two control simulations which select parameter sets outside this range for T_S (i.e. with $T_S < 273$ K and $T_S > 310$ K). When $T_S < 273$ K, the selected parameter sets often contained low values for both T_G and ΔC_p. Similarly, when $T_S > 310$ K, the selected parameter sets contained high values for both T_G and ΔC_p. Hence, the negative relationship between T_G and ΔC_p is due to T_S being around room temperature, an outcome of the hydrophobic effect [41,42].

Mesophilic proteins tend to explore smaller values of ΔH_G when T_G is small, but quickly cover the whole range as T_G increases. Fig. 1 shows that there is no apparent correlation between T_G and ΔH_G. Earlier experimental data analysis also did not reveal a significant correlation [23]. However, a correlation between T_G and Δh_G ($= \Delta H_G / N_{res}$), the value of enthalpy change normalized by protein size, was observed on a dataset of homologous thermophilic and mesophilic proteins [21]. In contrast to classes I and III, class II explores the full ranges of ΔH_G and ΔC_p available to the simulation. The slope and curvature of the stability curves show similar behavior to ΔH_G and ΔC_p, respectively (see plots in the middle panel of Fig. 1).

Table 2
Average, standard deviation and range of various thermodynamic parameters in three classes of S1[a]

Thermodynamic parameter	Class I (mesophiles)	Class II (thermophiles)	Class III (hyperthermophiles)	Twelve single domains
T_G (K)	321 ± 9	347 ± 8	379 ± 11	343 ± 14
	(300–333)	(333.5–363)	(363.5–400)	(325–371)
ΔH_G (kcal/mol)	83.6 ± 25	103.5 ± 27.6	109 ± 27.3	77 ± 18
	(50–150)	(50–150)	(50–150)	(58–107)
ΔC_p (kcal mol^{-1} K^{-1})	2.33 ± 0.49	1.89 ± 0.58	1.18 ± 0.38	1.6 ± 0.6
	(0.8–3.0)	(0.5–3.0)	(0.5–2.55)	(0.6–2.6)
$\Delta G(T_S)$ (kcal/mol)	4.83 ± 2.48	8.21 ± 3.18	12.69 ± 3.89	5.3 ± 1.6
	(1.26–13.79)	(1.79–19.36)	(3.83–25.03)	(3.5–9.0)
T_S (K)	286 ± 9	294 ± 10	294 ± 11	294 ± 8
	(273–310)	(273–310)	(273–310)	(281–304)
$-\Delta S_G$ (cal mol^{-1} K^{-1})	−260.1 ± 75.3	−298.4 ± 78.8	−287.9 ± 72.4	−225 ± 53
	(−469 to −150)	(−449.8 to −137.7)	(−412.7 to −125.6)	(−309 to −167)
$-\Delta C_p/T_G$ (cal mol^{-1} K^{-1})	−7.27 ± 1.56	−5.46 ± 1.72	−3.12 ± 1.06	−4.7 ± 1.8
	(−10 to −2.4)	(−9 to −1.38)	(−7.02 to −1.25)	(−7.8 to −1.7)

[a] The average, standard deviation and range for various thermodynamic parameters obtained from S1. The data were classified into three classes according to T_G. Class I contains the thermodynamic parameter sets with T_G in the range 300–333 K. For classes II and III, the T_G ranges are 333.5–363 and 363.5–400 K, respectively. For each parameter, the range of variation is given in parenthesis. The last column presents the data on 12 single domain proteins [23]. The thermodynamic data on the 12 single domain proteins were taken from the experiments reported in literature.

The maximal protein stability ranges visited by the simulation increase with T_G in all the three classes. In class I, the simulation covers narrower ranges of $\Delta G(T_S)$ as compared to classes II and III, with a greater range for class III (Fig. 1). On average, hyperthermophilic and thermophilic classes have greater $\Delta G(T_S)$ than mesophiles (Table 2). The increase in average $\Delta G(T_S)$ between classes II and III is greater than that of $\Delta G(T_S)$ between classes I and II. These trends are consistent with simulations S2 and S3 and previous experimental observations [18,21–23]. The simulation suggests large variabilities in $\Delta G(T_S)$ for a given T_G, although the average trends still hold. This is reasonable. As the difference between T_G and T_S ($T_G - T_S$) increases, enhanced exploration of maximal protein stability may be required to still keep T_S around room temperature.

3.2. Simulation 2 (S2)

The results of S2 were used to plot the stability curves in the three T_G classes (Fig. 2) and to explore the TRange of thermodynamic stability versus thermodynamic parameters that measure the protein energetics (Fig. 3). Fig. 2 plots hypothetical protein stability curves in classes I, II and III for the 84,461 combinations of T_G, ΔH_G and ΔC_p. The curves tend to be up-shifted and broader as the melting temperature increases. Experimental thermodynamic data on five families containing homologous thermophilic and mesophilic proteins yielded similar results [21]. Fig. 3 plots the temperature ranges for the parameter sets obtained from S2 with respect to ΔC_p, ΔH_G, curvature, slope and $\Delta G(T_S)$. Consistent with previous results, TRange depends more on ΔC_p (hence curvature) than on ΔH_G (slope). Proteins with greater TRange values have greater maximal protein stabilities. However, the spread in $\Delta G(T_S)$ increases with the increase in TRange. Experimental data analysis also showed that TRange and maximal protein stability are positively correlated [23]. This observation is consistent with the positive correlation between T_G and $\Delta G(T_S)$ and the requirement for T_S to be around room temperature. Taken together, these observations are related to lower ΔC_ps at higher T_G's and broader and up-shifted stability curves for the thermostable proteins.

3.3. Simulation 3 (S3)

In this simulation, we have computed the averages and standard deviations in the thermodynamic parameters for every 5 K step in T_G in the range of 300–400 K (Table 3). These data have been used to generate Fig. 4. S3 shows that the average stability curves are consistently up-shifted and broadened as T_G increases [Fig. 4], consistent with previous experimental data analysis on homologous thermophilic and mesophilic proteins [21]. For homologous proteins, greater T_G's can be achieved in three ways [17,20,43,44]. The maximal stability can be larger for the thermophile as compared to its homologous mesophile, resulting in an up-shift and broadening of the curve, and hence a higher T_G. This would be an outcome of a larger ΔH_G and/or a smaller ΔC_p. This mechanism is supported by studies on experimental protein data [17,21,23] and by the current simulations. Second, the $\Delta G(T_S)$ of both proteins may be similar, but the curvatures of their stability curves may differ. Third is a left/right shift of the curves. At the same

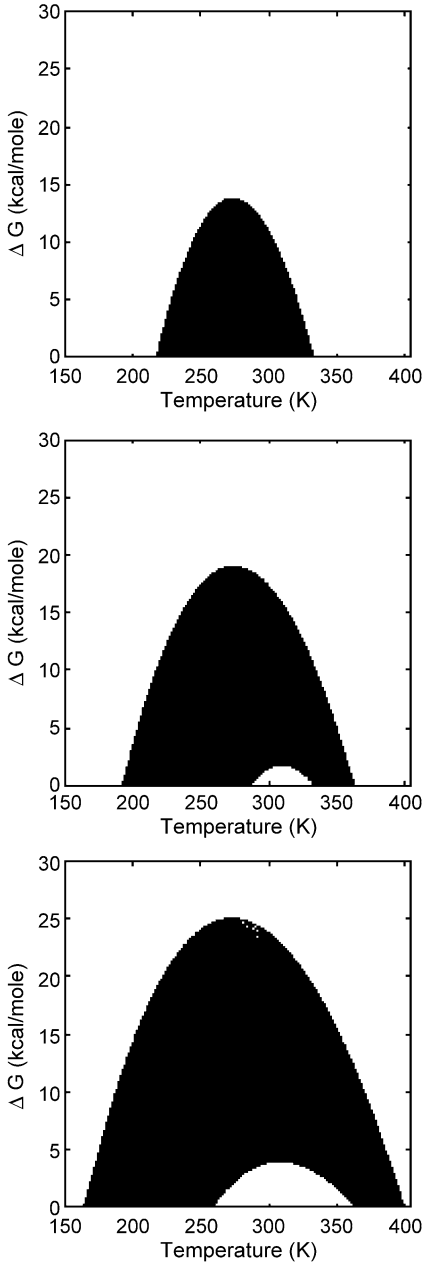

Fig. 2. Protein stability curves plotted using the Gibbs–Helmholtz equation and the thermodynamic parameter sets obtained from simulation 2 (S2) in classes I (mesophile), II (thermophile) and III (hyperthermophile). The *X*-axis denotes the temperature and the *Y*-axis denotes $\Delta G(T)$. Stability curves shown in the top panel correspond to class I ($T_G = 300-333$ K), in the middle panel to class II ($T_G = 333.5-363$ K) and in the bottom panel to class III ($T_G = 363.5-400$ K).

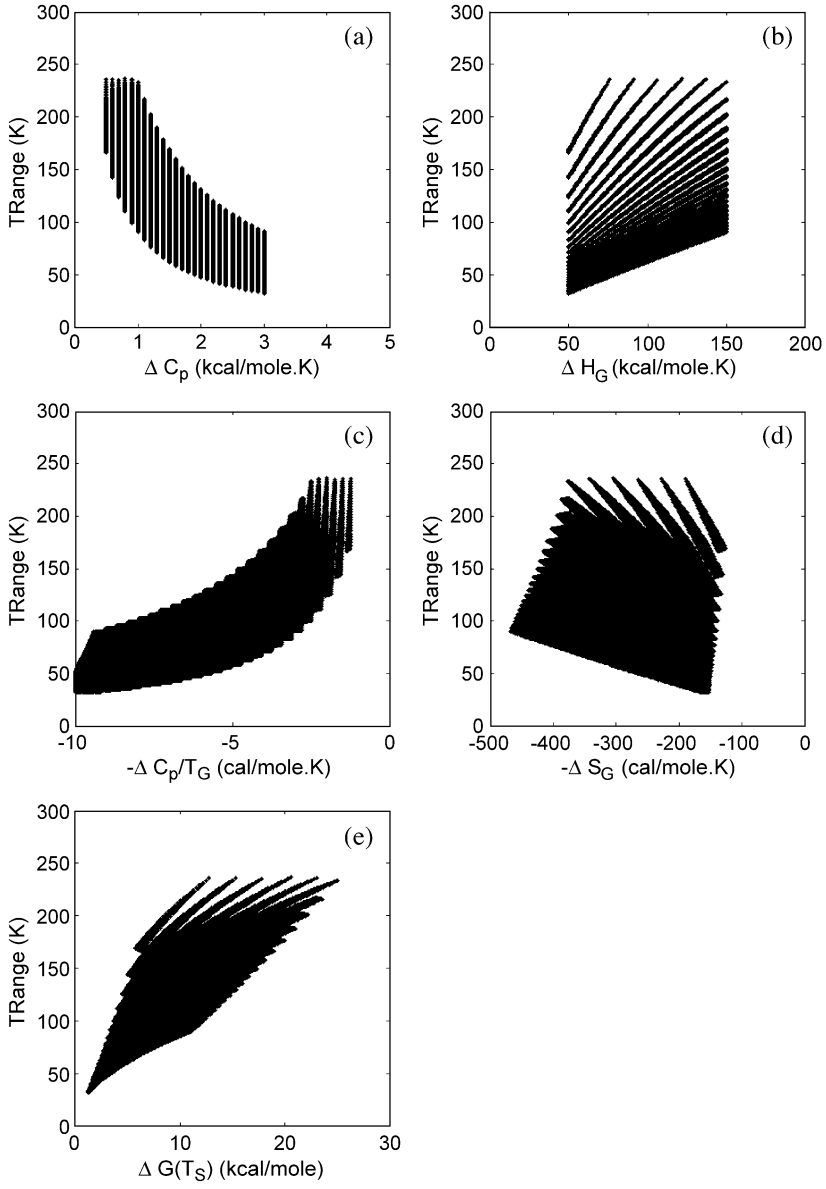

Fig. 3. Temperature range of thermodynamic stability (TRange) for the thermodynamic parameter sets obtained from simulation 2 (S2) plotted with respect to (a) ΔC_p, (b) ΔH_G, (c) curvature = $-\Delta C_p/T_G$, (d) slope = $-\Delta S_G$ and (e) maximal protein stability [$\Delta G(T_S)$].

time, the spread in the stability curves shown in Fig. 2 and the plot of TRange versus $\Delta G(T_S)$ (Fig. 3) suggest that alternative mechanisms may still be observed when larger experimental data become available, although the first mechanism would dominate.

Interestingly, the average enthalpy change (ΔH_G^{av}) shows an initial increase with T_G. However, around 360 K (~ 90 °C), the increase in ΔH_G^{av} becomes insignificant and ΔH_G^{av} converges to approximately 110 kcal/mol thereafter [Fig. 4]. The average slope of the protein stability curves also shows a similar behavior. The slight difference is due to the increase in T_G. The average heat capacity change decreases with T_G, but the decrease appears more rapid from 350 to 390 K (~ 80–120 °C) [Fig. 4]. The average curvature also shows a similar behavior. T_G, ΔH_G and ΔC_p are related to $\Delta G(T_S)$ by Eq. (5). However, the changes in average characteristics of ΔH_G and ΔC_p with respect to T_G do not seem to affect the average $\Delta G(T_S)$ which increases steadily with T_G (Fig. 4 and Table 3). The T_G region where the average characteristics of ΔH_G^{av} and ΔC_p^{av} change coincides with the region which separates thermophiles from hyperthermophiles. These results indicate that the thermodynamic properties may differ among thermophilic and hyperthermophilic proteins.

4. Discussion

Here, we have performed a formal analysis of Gibbs–Helmholtz equation within the experimentally observed

Table 3
Average and standard deviation of various thermodynamic parameters obtained from S3 every 5 K T_G step[a]

T_G (K)	ΔH_G (kcal/mol)	ΔC_p (kcal mol^{-1} K^{-1})	$-\Delta S_G$ (cal mol^{-1} K^{-1})	$-\Delta C_p/T_G$ (cal mol^{-1} K^{-1})	$\Delta G(T_S)$ (kcal/mol)	T_S (K)
300	61.64 ± 8.27	2.59 ± 0.29	− 205.47 ± 27.55	− 8.64 ± 0.97	2.41 ± 0.54	277.02 ± 2.67
305	67.16 ± 12.17	2.50 ± 0.36	− 220.21 ± 39.91	− 8.18 ± 1.18	2.95 ± 0.88	279.05 ± 3.95
310	72.77 ± 16.14	2.43 ± 0.41	− 234.73 ± 52.05	− 7.83 ± 1.32	3.53 ± 1.28	281.14 ± 5.25
315	78.45 ± 20.16	2.37 ± 0.45	− 249.06 ± 63.98	− 7.53 ± 1.42	4.16 ± 1.74	283.27 ± 6.56
320	84.17 ± 24.18	2.33 ± 0.48	− 263.03 ± 75.59	− 7.28 ± 1.49	4.84 ± 2.24	285.44 ± 7.88
325	88.12 ± 26.34	2.28 ± 0.50	− 271.13 ± 81.05	− 7.02 ± 1.53	5.37 ± 2.60	287.99 ± 9.09
330	91.45 ± 27.00	2.21 ± 0.51	− 277.13 ± 81.81	− 6.70 ± 1.56	5.86 ± 2.79	290.36 ± 9.84
335	95.70 ± 26.88	2.12 ± 0.53	− 285.68 ± 80.23	− 6.34 ± 1.59	6.51 ± 2.85	291.92 ± 9.95
340	99.91 ± 26.74	2.03 ± 0.55	− 293.84 ± 78.64	− 5.98 ± 1.62	7.21 ± 2.87	293.14 ± 9.99
345	103.65 ± 26.83	1.94 ± 0.56	− 300.43 ± 77.78	− 5.62 ± 1.63	7.94 ± 2.87	294.13 ± 10.08
350	106.58 ± 27.20	1.83 ± 0.56	− 304.53 ± 77.71	− 5.24 ± 1.61	8.69 ± 2.90	294.83 ± 10.24
355	108.23 ± 27.62	1.70 ± 0.54	− 304.86 ± 77.80	− 4.80 ± 1.52	9.43 ± 2.98	295.10 ± 10.42
360	108.35 ± 27.66	1.55 ± 0.48	− 300.97 ± 76.84	− 4.31 ± 1.34	10.14 ± 3.12	294.87 ± 10.47
365	108.37 ± 27.65	1.42 ± 0.43	− 296.91 ± 75.74	− 3.89 ± 1.19	10.83 ± 3.26	294.64 ± 10.49
370	108.48 ± 27.56	1.31 ± 0.39	− 293.19 ± 74.49	− 3.54 ± 1.06	11.51 ± 3.40	294.47 ± 10.49
375	108.69 ± 27.41	1.22 ± 0.36	− 289.83 ± 73.09	− 3.25 ± 0.95	12.18 ± 3.53	294.35 ± 10.49
380	109.01 ± 27.17	1.14 ± 0.32	− 286.87 ± 71.50	− 3.00 ± 0.85	12.85 ± 3.65	294.26 ± 10.48
385	109.47 ± 26.84	1.07 ± 0.30	− 284.33 ± 69.71	− 2.78 ± 0.77	13.52 ± 3.77	294.21 ± 10.48
390	110.09 ± 26.39	1.01 ± 0.27	− 282.27 ± 67.68	− 2.60 ± 0.70	14.20 ± 3.86	294.18 ± 10.49
395	110.88 ± 25.82	0.96 ± 0.25	− 280.70 ± 65.37	− 2.44 ± 0.63	14.89 ± 3.93	294.15 ± 10.51
400	111.87 ± 25.08	0.92 ± 0.23	− 279.68 ± 62.69	− 2.29 ± 0.57	15.61 ± 3.97	294.10 ± 10.51

[a] The average and standard deviations of the thermodynamic parameters obtained from S3. In this simulation, ΔH_G was varied in 0.1 kcal/mol steps and ΔC_p was varied in 0.01 kcal mol^{-1} K^{-1} steps. The averages and standard deviations were computed every 5 K T_G step. Only those parameter sets were selected which yield T_S around the room temperature. A total of 1,657,777 parameter sets were used in these calculations.

ranges of the thermodynamic parameters seen for single domain reversible two-state proteins. This analysis has allowed us to systematically explore the $T_G\Delta H_G\Delta C_p$ space. Our simulations do not treat the thermodynamic parameters T_G, ΔH_G and ΔC_p as independent variables. Instead, these simulations explore the interdependencies of these and other derived thermodynamic parameters on each other. The simulations are phenomenological in nature and cannot be expected to yield insights into the mechanism of protein thermostability. On the other hand, they have enabled us to trace the variation in the values of various thermodynamic parameters with T_G at the level of specified granularity. This provides much more data to study the interdependencies of these parameters than possible by the available experimental thermodynamic data. This is where the simulations are useful. Hence, observations made from the simulation results have a more rigorous basis than previous analyses of limited experimental data [18,21–23,45–48]. Our simulations are, however, also limited by the available experimental data on single domain proteins.

The Gibbs–Helmholtz equation [Eq. (1)] assumes that ΔC_p is a constant (>0). Hence, our simulations are restricted to ΔC_p being constant. However, there is experimental evidence that ΔC_p may vary with the T for single domain proteins, particularly for those from hyperthermophilic organisms (e.g. see Ref. [44] and subsequent papers from the same group). To qualitatively assess this effect, we have rederived the slope [Eq. (2)] and curvature [Eq. (3)] taking into account the variation of ΔC_p with T and assuming that the Gibbs–Helmholtz equation still holds (i.e. the variation of ΔC_p with T is relatively small and folding↔unfolding transition is still largely two-state and reversible). According to the new equations, the slope of a protein stability curve at T_G is given by:

$$(\mathrm{d}\Delta G/\mathrm{d}T)_{T=T_G} = -H_G/T_G + T_G(\mathrm{d}\Delta C_p/\mathrm{d}T)_{T=T_G} \quad (9)$$

and the curvature of the stability curve at T_G is given by

$$(\mathrm{d}^2\Delta G/\mathrm{d}T^2)_{T=T_G} = -C_p/T_G + [T_G - \ln T_G](\mathrm{d}\Delta C_p/\mathrm{d}T)_{T=T_G} + T_G(\mathrm{d}^2\Delta C_p/\mathrm{d}T^2)_{T=T_G} \quad (10)$$

Hence, in such cases, both the slope and curvature of the protein stability curve need to be corrected by additional factors involving the first- and second-order derivatives of ΔC_p with respect to T. If the values of these derivatives at T_G are small, the corrections of the slope and curvature of the protein stability curve may be quite negligible. Another point that deserves mentioning here is the different values for ΔC_p determined by different experimental methods and the differences in intrinsic versus apparent values of ΔC_p due to the linkage of protonation and anion binding to protein folding at different pH and temperature (e.g. see Ref. [44]). However, this is beyond the scope of present discussion.

The simulations select thermodynamic parameter sets that yield the temperature of maximal stability to be around room temperature. This assumption is based on our previous observation that most proteins with large enough hydrophobic cores are maximally stable around room temperature, irrespective of the protein size, fold and living temperature of the source organism [22]. The major reason is the

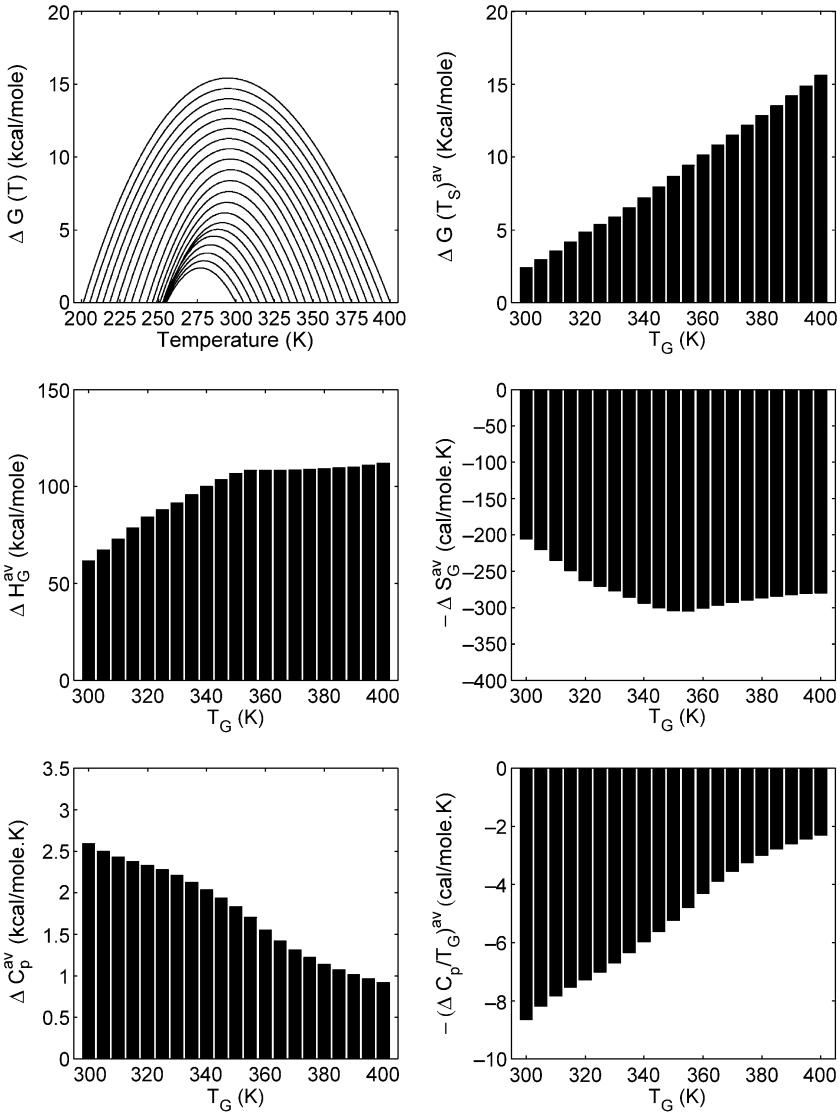

Fig. 4. Results from simulation 3 (S3). In this simulation, ΔH_G and ΔC_p were varied in steps of 0.1 kcal/mol and 0.01 kcal mol^{-1} K^{-1}, respectively. The data were collected and averages computed for every 5 K steps in T_G. (a) Average protein stability curves. *X*-axis represents temperature (T) and *Y*-axis represents $\Delta G(T)$. The average values of various thermodynamic parameters are plotted with respect to T_G. (b) $\Delta G(T_S)^{av}$ versus T_G, (c) ΔH_G^{av} versus T_G, (d) $-\Delta S_G^{av}$ versus T_G, (e) ΔC_p^{av} versus T_G and (f) $-\Delta C_p/T_G^{av}$ versus T_G. The average maximal protein stability increases with T_G. Above 360 K (~ 90 °C), this increase is primarily due to the decrease in ΔC_p because ΔH_G appears to become constant around this temperature region.

dominance of the hydrophobic effect [41] in protein folding [42]. The nonpolar amino acids have low solubilities in water around room temperature [22,28,41]. However, the exact value for the temperature of maximal stability (T_S) of a protein also depends on the charged/polar residues in the protein and the electrostatic interactions among them. Thermophilic proteins often contain increased proportions of charged and polar residues. These are thought to contribute toward their increased stability [3,4]. Hence, it is unclear whether thermophilic proteins will also be maximally stable around room temperature. Only a few proteins from the extremophilic organisms have been thermodynamically fully characterized so far. The initial trends support our assumption. For example, psychrophilic α-amylase, hyperthermophilic DNA binding protein Sso7d and cold shock protein from *Thermotoga maritima* are maximally stable at 290, 282 and 303 K, respectively [22]. Moreover, the thermophilic proteins in four out of the five families of the homologous thermophilic–mesophilic proteins in the dataset of Kumar et al. [21] are maximally stable around room temperature like their mesophilic homologues.

The thermodynamic simulations are similar in spirit to the conformational grid search commonly used in molecular simulations. Here, the effect of systematic variation in the thermodynamic parameters T_G, ΔH_G and ΔC_p is explored using the Gibbs–Helmholtz equation. The problem then reduces to finding parameter sets that yield maximal protein stability around room temperature. The simulations clearly indicate that a significant shift of T_S (temperature of maximal stability) away from room tem-

perature is not required for adaptation to extreme temperatures (Table 3). Broadening and up-shifting the stability curves can achieve similar results. Usually, this requires only minor modulations in structural features, preserving the native fold [4,7].

The thermodynamic parameters varied in the simulations and those obtained from simulation results are interrelated via protein structure, especially the polar and nonpolar surface areas (e.g. see Robertson and Murphy [46] for a review). For example, we have previously observed that both TRange and T_G are anticorrelated with ΔC_p [23]. The current simulations also support this observation. ΔC_p is related to the hydrophobic effect [27,49] and is proportional to protein size and the nonpolar surface area buried in the protein core [23,45–48,50].

Our analyses of the limited experimental thermodynamic data revealed a number of statistically significant linear correlations among various thermodynamic parameter pairs [22,23]. The simulations have suggested that these parameter pairs are indeed related. However, their relationship may be nonlinear in some cases. There is a large scatter in the simulated data due to the complex relationships. Yet, a few observations stand out. The average ΔH_G increases with T_G until $\sim$ 360 K and becoming constant thereafter. Around the same temperature region, the average ΔC_p also decreases more rapidly. This is a new observation revealed by the simulations. The temperature $\sim$ 360 K ($\sim$ 90 °C) separates thermophiles from hyperthermophiles. Hence, this suggests separate groupings of proteins from thermophiles and hyperthermophiles as having different thermodynamic properties. Earlier, Szilagyi and Zavodszky [5] had done this in their study of microscopic parameters responsible for protein stability, but several investigators, including ourselves, tended to club the thermophilic and hyperthermophilic proteins together. The simulations provide the rationale for treating these classes of proteins separately.

Our simulations suggest strategies for manipulating protein stability. Temperature resistance relates to maximal protein stability measured in terms of the free energy difference between the folded and unfolded states at T_S. Hence, increasing the melting temperature raises the maximal stability. A small increment in melting temperature (mesophile → thermophile) might be achieved by increasing ΔH_G or decreasing ΔC_p. For a larger increment (mesophile → hyperthermophile), increasing ΔH_G alone may be insufficient. ΔC_p may need a simultaneous decrease. For thermophile to hyperthermophile, we may decrease ΔC_p keeping ΔH_G constant. Decrease in ΔC_p is a primary consideration when improving protein stability. Low ΔC_p is often cited as a reason for increased protein thermostability (e.g. Motono et al. [51]).

Proteins can achieve lower ΔC_p and higher T_G by reducing their size [45,47,48,50]. Deletion/shortening of loops in the thermophilic proteins are commonly observed [3,15,52]. Lower ΔC_p may also be achieved by increasing the charged and polar residues in thermophilic proteins leading to ion pairs and their networks. This is the most consistent trend seen in thermophile–mesophile comparisons [3,4,7,15,53]. Polar → charged amino acid substitution is the most consistent genomic correlate for hyperthermostability [54]. Loladze et al. [55] have reported that substitutions of aliphatic side chains by polar in the protein interior significantly reduce ubiquitin's ΔC_p while the structural changes are insignificant. Consistently, Zhou [56] has shown that increased electrostatic interactions in thermostable proteins relate to their reduced ΔC_p. Cation–pi interactions [14,15] and aromatic clusters [6] may also lower ΔC_p. Since the thermodynamic parameters are interdependent, lowering ΔC_p will also affect other parameters especially ΔH_G, ΔS_G, TRange and $\Delta G(T_s)$.

5. Conclusions

Protein thermodynamic simulations appear a useful tool in probing protein thermostability. By varying three thermodynamic values (ΔH_G, T_G and ΔC_p) in the Gibbs–Helmholtz equation within a set of experimentally influenced boundary conditions, we are able to quantitatively assess differences between meso-, thermo-and hyperthermophiles over broad numerical ranges. Thus, the simulations complement experiments and are able to validate and provide further insight into the trends observed from the analyses of limited experimental data. The simulations suggest that hyperthermophilic proteins may have different thermodynamic properties than the thermophilic ones. They further suggest that the hydrophobic effect, the origin of maximal stability at room temperature, may lead to a lower ΔC_p at higher T_G, a well-known primary route to greater protein thermostability.

Acknowledgements

We thank Dr. Neeti Sinha, Dr. Chung-Jung Tsai, Dr. K. Gunasekaran, Dr. David Zanuy and, in particular, Dr. Jacob V. Maizel for numerous helpful discussions. The research of R. Nussinov in Israel has been supported in part by the "Center of Excellence in Geometric Computing and its Applications" funded by the Israel Science Foundation, by the Ministry of Science grant and by the Tel Aviv University Basic Research grants. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract number NO1-CO-12400. The content of this publication does not necessarily reflect the view or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products or organization imply endorsement by the U.S. government.

References

[1] R.M. Daniel, R.A. Cowan, H.W. Morgan, M.P. Curran, A correlation between protein thermostability and resistance to proteolysis, Biochem. J. 207 (1982) 641–644.

[2] M.W. Adams, R.M. Kelly, Finding and using hyperthermophilic enzymes, Trends Biotech. 13 (1998) 329–332.

[3] R. Sterner, W. Liebl, Thermophilic adaptation of proteins, Crit. Rev. Biochem. Mol. Biol. 36 (2001) 39–106.

[4] S. Kumar, R. Nussinov, How do thermophilic proteins deal with heat? Cell. Mol. Life Sci. 58 (2001) 1216–1233.

[5] A. Szilagyi, P. Zavodszky, Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey, Struct. Fold. Des. 8 (2000) 493–504.

[6] N. Kannan, S. Vishveshwara, Aromatic clusters: a determinant of thermal stability of thermophilic proteins, Protein Eng. 13 (2000) 753–761.

[7] S. Kumar, C.J. Tsai, R. Nussinov, Factors enhancing protein thermostability, Protein Eng. 13 (2000) 179–191.

[8] D. Perl, F.X. Schmid, Some like it hot: the molecular determinants of protein thermostability, ChemBiochem 3 (2002) 39–44.

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

[10] L. Xiao, B. Honig, Electrostatic contributions to the stability of hyperthermophilic proteins, J. Mol. Biol. 289 (1999) 1435–1444.

[11] J.M. Sanchez-Ruiz, G.I. Makhatadze, To charge or not to charge? Trends Biotechnol. 19 (2001) 132–135.

[12] S. Kumar, B. Ma, C.J. Tsai, R. Nussinov, Electrostatic strengths of salt bridges in thermophilic and mesophilic glutamate dehydrogenase monomers, Proteins: Struct., Funct., Genet. 38 (2000) 368–383.

[13] S. Kumar, R. Nussinov, Close range electrostatic interactions in proteins, ChemBiochem 3 (2002) 604–617.

[14] M.M. Gromiha, S. Thomas, C. Santosh, Role of cation–pi interactions to the stability of thermophilic proteins, Prep. Biochem. Biotechnol. 32 (2002) 355–362.

[15] S. Chakravarty, R. Varadarajan, Elucidation of factors responsible for enhanced thermal stability of proteins: a structural genomics based study, Biochemistry 41 (2002) 8152–8161.

[16] J. Hollien, S. Marqusee, A thermodynamic comparison of mesophilic and thermophilic ribonuclease H, Biochemistry 38 (1999) 3831–3836.

[17] B.M. Beadle, W.A. Baase, D.R. Wilson, N.R. Gilkes, B.K. Shoichet, Comparing the thermodynamic stabilities of a related thermophilic and mesophilic enzyme, Biochemistry 38 (1999) 2570–2576.

[18] D.C. Rees, A.D. Robertson, Some thermodynamic implications for the thermostability of proteins, Protein Sci. 10 (2001) 1187–1194.

[19] W. Li, R.A. Grayling, K. Sandman, S. Edmondson, J.W. Shriver, J.N. Reeve, Thermodynamic stability of archaeal histones, Biochemistry 37 (1998) 10563–10572.

[20] D.C. Rees, M.W. Adams, Hyperthermophiles: taking the heat and loving it, Structure 3 (1995) 251–254.

[21] S. Kumar, C.J. Tsai, R. Nussinov, Thermodynamic differences among homologous thermophilic and mesophilic proteins, Biochemistry 40 (2001) 14152–14165.

[22] S. Kumar, C.J. Tsai, R. Nussinov, Maximal stabilities of reversible two state proteins, Biochemistry 41 (2002) 5359–5374.

[23] S. Kumar, C.J. Tsai, R. Nussinov, Temperature range of thermodynamic stability for the native state of reversible two-state proteins, Biochemistry 42 (2003) 4864–4873.

[24] W. Pfeil, Protein Stability and Folding. A Collection of Thermodynamic Data, Springer-Verlag, Heidelberg, 1988.

[25] M.M. Gromiha, J. An, H. Kono, M. Oobatake, H. Uedaira, P. Prabakaran, A. Sarai, ProTherm, version 2.0: thermodynamic database for proteins and mutants, Nucleic Acids Res. 28 (2000) 283–285.

[26] W. Becktel, J.A. Schellman, Protein stability curves, Biopolymers 26 (1987) 1859–1877.

[27] P.L. Privalov, Cold denaturation of proteins, Crit. Rev. Biochem. Mol. Biol. 25 (1990) 281–305.

[28] J.A. Schellman, Temperature, stability and hydrophobic interaction, Biophys. J. 73 (1997) 2960–2964.

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

[30] D. Wassenberg, C. Welker, R. Jaenicke, Thermodynamics of the unfolding of the cold shock protein from *Thermotoga maritima*, J. Mol. Biol. 289 (1999) 187–193.

[31] G.S. Huang, T.G. Oas, Heat and cold denatured states of monomeric λ repressor are thermodynamically and chemically equivalent, Biochemistry 35 (1996) 6173–6180.

[32] J.M. Scholtz, Conformational stability of HPr: The histidine-containing phosphocarrier protein from *Bacillus subtilis*, Protein Sci. 4 (1995) 35–43.

[33] A. Schoppe, H.J. Hinz, V.R. Agashe, S. Ramachandran, J.B. Udgaonkar, DSC studies of the conformational stability of barstar wild-type, Protein Sci. 6 (1997) 2196–2202.

[34] T. Okajima, Y. Kawata, K. Hamaguchi, Chemical modification of tryptophan residues and stability changes in proteins, Biochemistry 29 (1990) 9168–9175.

[35] Y. Feng, S.G. Sligar, Effect of heme binding on the structure and stability of *Escherichia coli* apocytochrome b_{562}, Biochemistry 30 (1991) 10150–10155.

[36] M.M. Santoro, D.W. Bolen, A test of the linear extrapolation of unfolding free energy changes over an extended denaturant concentration range, Biochemistry 31 (1992) 4901–4907.

[37] D. Shortle, A.K. Meeker, E. Freire, Stability mutants of Staphylococcal nuclease: large compensating enthalpy–entropy changes for the reversible denaturation reaction, Biochemistry 27 (1988) 4761–4768.

[38] V. Consalvi, R. Chiaraluce, L. Giangiacomo, R. Sandurra, P. Christova, A. Karshikoff, S. Knapp, R. Ladenstein, Thermal unfolding and conformational stability of recombinant domain II of glutamate dehydrogenase from the hyperthermophile *Thermotoga maritima*, Protein Eng. 13 (2000) 501–507.

[39] C.G. Genzor, A. Beldarrain, C. Gomez-Moreno, J.L. Lopez-Lacomba, J. Cortijo, M. Sancho, Conformational stability of apoflavodoxin, Protein Sci. 5 (1996) 1376–1388.

[40] H. Fukada, S. Kitamura, K. Takahashi, Calorimetric study of the thermal unfolding of Kunitz-type soybean trypsin inhibitor at pH 7.0, Thermochim. Acta 266 (1995) 365–372.

[41] W. Kauzmann, Some factors in interpretation of protein denaturation, Adv. Protein Chem. 14 (1959) 1–63.

[42] K.A. Dill, Dominant forces in protein folding, Biochemistry 29 (1990) 7133–7155.

[43] P. Alexander, S. Fahnestock, T. Lee, J. Orban, P. Bryan, Thermodynamic analysis of the folding of the streptococcal protein G IgG-binding domains B1 and B2. Why small proteins tend to have high denaturation temperatures, Biochemistry 31 (1992) 3597–3603.

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

[45] G.I. Makhatadze, P. Privalov, Energetics of protein structure, Adv. Protein Chem. 47 (1995) 307–425.

[46] A.D. Robertson, K.P. Murphy, Protein structure and the energetics of protein stability, Chem. Rev. 97 (1997) 1251–1267.

[47] S.P. Edgcomb, K.P. Murphy, Structural energetics of protein folding and binding, Curr. Opin. Biotech. 11 (2000) 62–66.

[48] P.L. Privalov, S.J. Gill, Stability of protein structure and hydrophobic interaction, Adv. Prot. Chem. 39 (1988) 191–234.

[49] C.J. Tsai , J.V. Maizel Jr., R. Nussinov, The hydrophobic effect: a new insight from cold denaturation and a two state water structure, Crit. Rev. Biochem. Mol. Biol. 37 (2002) 55–69.

[50] J.K. Myers, C.N. Pace, J.M. Scholtz, Denaturant *m* values and heat

capacity changes: relation to changes in accessible surface areas of protein unfolding, Protein Sci. 4 (1995) 2138–2148.
[51] C. Motono, T. Oshima, A. Yamagishi, High thermal stability of 3-isopropylmalate dehydrogenase from *Thermus thermophilus* resulted from low ΔC_p of unfolding, Protein Eng. 14 (2001) 961–966.
[52] M.J. Thomson, D. Eisenberg, Transproteomic evidence of a loop-deletion mechanism for enhancing protein thermostability, J. Mol. Biol. 290 (1999) 595–604.
[53] K.S. Yip, T.J. Stillman, K.L. Britton, P.J. Artymiuk, P.J. Baker, S.E. Engel, P.C. Engel, A. Pasquo, R. Chiaraluce, V. Consalvi, The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures, Structure 3 (1995) 1147–1158.
[54] K. Suhre, J.M. Claverie, Genomic correlates of hyper-thermostability, an update, J. Biol. Chem. 278 (2003) 17198–17202.
[55] V.V. Loladze, D.N. Ermolenko, G.I. Makhatadze, Heat capacity changes upon burial of polar and non-polar groups in proteins, Protein Sci. 10 (2001) 1343–1352.
[56] H.X. Zhou, Towards physical basis of thermophilic proteins: linking of enriched polar interactions and reduced heat capacity of unfolding, Biophys. J. 83 (2002) 3126–3133.