Sequence-Based Analysis of Protein Energy Landscapes Reveals Nonuniform Thermal Adaptation within the Proteome

Jenny Gu[1] *and Vincent J. Hilser*

Department of Biochemistry and Molecular Biology, and Sealy Center for Structural Biology and Molecular Biophysics, University of Texas Medical Branch

Thermal adaptation of individual proteins is often achieved through modulating protein stability, with proteins that are adapted to extreme cold environments having increased conformational flexibility when brought to mesophilic conditions. Conversely, proteins adapted to higher temperatures appear less dynamic and are found to be much more stable against thermal denaturation than their mesophilic counterparts. According to the current paradigm, the adaptation of an organism for survival at higher or lower temperatures is facilitated by the adaptation of the component proteins. We note, however, that these observations have been carried out on relatively few proteins. The extent to which the conformational stabilities of all members of the proteome have been modulated for thermal adaptation remains unclear, with no direct experimental strategies to address this issue. Adapted extremophilies are likely to use a multitude of molecular and biophysical strategies for survival and, therefore, evolution of specific biophysical properties of proteins for optimal function may not be necessary for all proteins in the proteome. Using a sequence-based predictor of protein stability, eScape, an in silico examination of several extremophilic proteomes shows a correlation between the collective stability of the proteins and the thermal range of survival for the organism as expected. Unexpectedly, however, the analysis shows that protein thermostability is modified to different extents across the proteome and depends on the functional role for which the protein is involved. Identification of these differences provides unique opportunities to study interdependence within the proteome as well as the role that the proteome plays in the process of evolutionary thermal adaptation.

Introduction

Modification of protein conformational flexibility (CF) has long been posited as a mechanism by which organisms can thermally adapt to optimize function at a given ambient temperature (Fields and Somero 1998; Zavodszky et al. 1998; Beadle et al. 1999; Shlyk-Kerner et al. 2006; Feller 2007; Tadokoro et al. 2007). However, investigations of thermal adaptation mechanisms, often studied through structure-based comparative analyses of psychro-, meso-, and thermophilic orthologs, are often limited to small data sets (>30 proteins) (Gianese et al. 2002; Berezovsky and Shakhnovich 2005; Jahandideh et al. 2007; Spiwok et al. 2007), and these studies are heavily dominated by proteins with known enzymatic function. Although important insights into possible strategies for conferring thermoadaptability have been gleaned from such data sets, whether such trends can be generally applied across the entire proteome remains unclear. Proteomic and genomic analysis of sequence compositions has identified significant correlation with the corresponding optimal growth temperature of organisms (Zeldovich et al. 2007). Nonetheless, these findings, do not address the question as to whether the stability or CF of proteins is uniformly or selectively modified.

Proteostasis (Balch et al. 2008), or protein homeostasis, is the process of maintaining the balance of native and denatured states in an organism such that active conformers are available for cellular function. The extent to which the proteome is thermally adapted is an important issue that will contribute to understanding the most effective evolutionary mechanism for an organism to maintain proteostasis in different environments. In principle, adjustment of proteostasis can be achieved either by changing protein thermostability through mutation, adjusting stabilizing osmolyte concentrations, or regulating pathways influencing protein synthesis, folding, trafficking, disaggregation, or degradation. With such a variety of potential strategies for modulating proteostasis, the assumption that all proteins within the proteome undergo the same degree of thermal adaptation requires investigation. Do all proteins thermally adapt, and is the mechanism of adaptation uniform across all proteins?

In this paper, we use a recently developed sequence-based prediction algorithm, eScape (Gu and Hilser 2008), to explore these questions. Because this approach is rooted in an experimentally tested ensemble-based model of proteins, it provides a means to access the position-specific stability within proteins without the need for experimental verification for all sequences examined. The ability of eScape to provide a position-specific stability profile of a protein from sequence information alone allows us to understand the extent to which protein conformational stabilities are modified in proteomes of extremophiles. Previous proteomic comparisons with eScape have revealed that organisms use different portions of the available sequence space (Gu and Hilser 2008). Our goal here is to explore whether proteomes are uniformly modulated for thermal adaptation to extreme thermal environments or whether the adaptation of the proteome as a whole relies on the selective adaptation of a subset of the proteins. In addition we wish to know whether adaptation is facilitated by modulating local fluctuations (i.e., flexibility) or through changes in the global stability.

Materials and Methods

The eScape Algorithm

eScape (Gu and Hilser 2008) is a machine learning trained model that predicts position-specific stability

[1] Present address: Institute of Evolution and Biodiversity, University of Münster, Münster, Germany.

Key words: thermal adaptation, proteome stability, protein stability, conformational flexibility, metabolism evolution.

E-mail: vjhilser@utmb.edu; j.gu@uni-muenster.de.

Mol. Biol. Evol. 26(10):2217–2227. 2009
doi:10.1093/molbev/msp140
Advance Access publication July 10, 2009

profiles for proteins using only sequence information with an adjusted R-squared value of 0.70 and an average Pearson correlation coefficient of 83.63%. eScape was trained with 10-fold cross validation using a nonredundant set of 122 human proteins (Larson and Hilser 2004). The Gibbs free energy of protein stability, as well as the separate enthalpic and entropic contributions, can be predicted in cal/mol for each position of a natively folded protein using the following model:

$$\Delta G = ((0.8195 * \min_{t,\Delta G}) + (0.7492 * \max_{t,\Delta G})) + 4,696;$$

$$\Delta H_{\mathrm{ap}} = ((0.7665 * \min_{t,\Delta H_{\mathrm{ap}}}) + (0.7632 * \max_{t,\Delta H_{\mathrm{ap}}})) - 5,068;$$

$$\Delta H_{\mathrm{p}} = ((0.7791 * \min_{t,\Delta H_{\mathrm{p}}}) + (0.7524 * \max_{t,\Delta H_{\mathrm{p}}})) + 6,195;$$

$$T\Delta S = ((0.7047 * \min_{t,T\Delta S}) + (0.7507 * \max_{t,T\Delta S})) + 1,998;$$

where $\min_i$ and $\max_i$ (t = the representing tripeptide) are values corresponding to the range of energetic values observed for the respective thermodynamic descriptor in a library of tripeptides with associated observed energetic values that are needed for the predictions. The library of energetic values reflects the contributions of both global and local effects as calculated using COREX, a structure-based statistical thermodynamic model (D'Aquino et al. 1996; Hilser and Freire 1996; Hilser et al. 2006). The implicitly represented global contributions to stability allow eScape to make position-specific predictions.

Briefly, COREX uses an ensemble of conformational states that can be adopted by a protein to calculate the energetics of the system. The ensemble of conformational states is generated through the systematic unfolding of the native protein to yield a range from the fully folded to denatured conformation. Under equilibrium conditions, the probability of any given conformational microstate, i, in the ensemble is given by

$$P_i = \frac{K_i}{\sum_{i=1}^{N_{\mathrm{states}}} K_i} = \frac{K_i}{Q},$$

where $K_i = \mathrm{e}^{(-\Delta G_i/RT)}$ is the statistical weight of each microstate, where R is the gas constant, for a given absolute temperature T. The summation in the denominator is the partition function, Q, for the system. The ratio of the summed probability of all states in the ensemble in which a particular residue j is in a folded conformation ($\Sigma P_{f,j}$) to the summed probability of all states in which j is in an unfolded conformation ($\Sigma P_{nf,j}$) yields a statistical descriptor defined as the residue stability constant $\kappa_{f,j}$ (D'Aquino et al. 1996):

$$\kappa_{f,j} = \frac{\sum P_{f,j}}{\sum P_{nf,j}},$$

which is then used to obtain the position-specific free energy expressed in units of cal/mol written as

$$[\Delta G]_j = -RT \cdot \ln \frac{\sum P_{f,j}}{\sum P_{nf,j}} = \langle \Delta G_{f,j} \rangle - \langle \Delta G_{nf,j} \rangle.$$

Position-specific thermodynamic descriptors are also calculated by taking the difference in folded and unfolded subensemble quantities:

$$[\Delta H]_{\mathrm{pol},\ j} = \langle \Delta H_{\mathrm{pol},f,j} \rangle - \langle \Delta H_{\mathrm{pol},nf,j} \rangle,$$

$$[\Delta H]_{\mathrm{apol},\ j} = \langle \Delta H_{\mathrm{apol},f,j} \rangle - \langle \Delta H_{\mathrm{apol},nf,j} \rangle,$$

$$[\Delta S]_{\mathrm{conf},\ j} = \langle \Delta S_{\mathrm{conf},f,j} \rangle - \langle \Delta S_{\mathrm{conf},nf,j} \rangle.$$

We ask the readers to refer to the respective reference for more details about eScape (Gu and Hilser 2008) and COREX (D'Aquino et al. 1996; Hilser and Freire 1996; Hilser et al. 2006).

Proteomes of Extremophiles

Thermo- and psychrophiles were used to identify significant changes in thermostability within the proteome that may be important mechanisms for evolutionary adaptation. The proteomes downloaded from Intergr8 (Kersey et al. 2005) and used in this study are *Cenarchaeum symbiosum* (Preston et al. 1996), *Chlorobium tepidum*, *Methanococcoides burtonii* (Franzmann et al. 1997), *Methanocaldococcus jannaschii*, *Methanobacterium thermoautotrophicum*, *Methanosaeta thermophila*, *Methanococcus vannielii*, *Pyrococcus abyssi*, *Psychrobacter arcticus*, *Psychrobacter cryohalolentis*, *Pyrococcus furiosus* (Fiala and Stetter 1986), *Pyrococcus horikoshii*, *Psychromonas ingrahamii*, *Pyrobaculum islandicum*, *Pyrococcus kodakaraensis*, *Polaromonas naphthalenivorans*, *Polaromonas* sp., *Sulfolobus acidocaldarius*, *Sulfolobus solfataricus*, *Sulfolobus tokodaii*, *Thermoplasma acidophilum*, *Thermobifida fusca*, *Thermotoga maritima* (Huber et al. 1986), *Thermoproteus neutrophilus*, *Thermoanaerobacter pseudethanolicus*, *Thermoanaerobacter tengcongensis*, *Thermus thermophilus*, and *Thermoplasma volcanium*.

Comparison of Thermostability between Proteomes

eScape was applied to all sequences in the proteomes to obtain predicted position-specific stability values. The mean average stabilities of all proteins within proteomes were compared using the standard t-test and the differences between them are significant ($\alpha = 0.01$). Four randomly selected extremophiles (*M. burtonii*, *P. furiosus*,

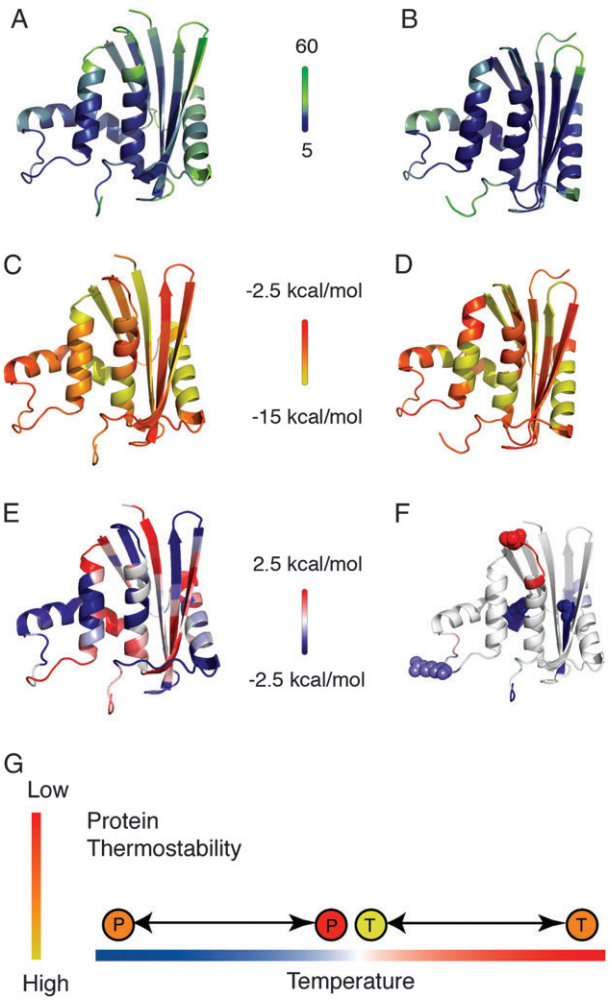


FIG. 1.—Detection of thermal adaptation through changes in CF with eScape. eScape is useful for the qualitative comparison of changes in protein stability as demonstrated with various thermovariants of the RNAse H1. Structure temperature factors (B factors) are displayed for both the (*A*) psychrophilic and (*B*) mesophilic RNAse H1. Different stability predictions by eScape are reported for the (*C*) psychrophilic and (*D*) mesophilic variants. The (*E*) difference in stability between the psychrophilic and mesophilic variants are mapped as well as the (*F*) difference between wildtype mesophilic and a variant containing five mutations for a hyperstabilized variant. (*G*) Stabilty predictions with eScape for psychrophiles (P) and thermophiles (T) are expected to reflect stability values under mesophilic conditions.

C. symbiosum, and *T. maritima*) were used for a more detailed analysis of thermostability changes within the proteome for thermal adaptation. Although the analysis can be performed on all proteomes with available data, the analysis is limited to four proteomes such that a sufficient number of shared homologs can be obtained for a subsequent analysis to make direct conclusive observations regarding proteomic evolution to thermal stress. For the comparison of extreme values, the single tail *t*-test was used instead to compare the collective global stability and local flexibility of proteins in the proteome. The analysis was conducted by averaging the 20 highest and lowest stability values that will serve as the statistic representing local flexibility and global stability, respectively. Domains have been found to peak at 100 residues with a number of domains containing 200 residues; therefore, the average of 20 residues at the extreme spectrum of stability will serve as the statistic for ~10–20% of the stability values for a single domain.

Comparison of Thermostability Clusters Partitioned Based on Gene Ontology Functional Annotation

Four randomly selected extremophiles (*M. burtonii*, *P. furiosus*, *C. symbiosum*, and *T. maritima*) were partitioned into the respective Gene Ontology (Ashburner et al. 2000) functional classification and the average stability within this partition was calculated. For comparison of stability differences between proteomes for each GO function partition, the average of absolute differences between proteomes partitioned similarly was calculated.

Comparison of Orthologs to Identify Differences in Thermoadaptability within the Proteome

A total of 183 ortholog clusters, identified using Blast (Altschul et al. 1997), is shared between *M. burtonii*, *P. furiosus*, *C. symbiosum*, and *T. maritima* and used for a direct comparative analysis. Multiple sequence alignments (MSA) for each cluster of orthologs were generated and position-specific magnitudes of conformational change were calculated and normalized. Sequences for each cluster were first aligned using ProbCons (Do et al. 2005). A score to compare the magnitude of change in CF for each ortholog clusters normalized by alignment size was calculated as:

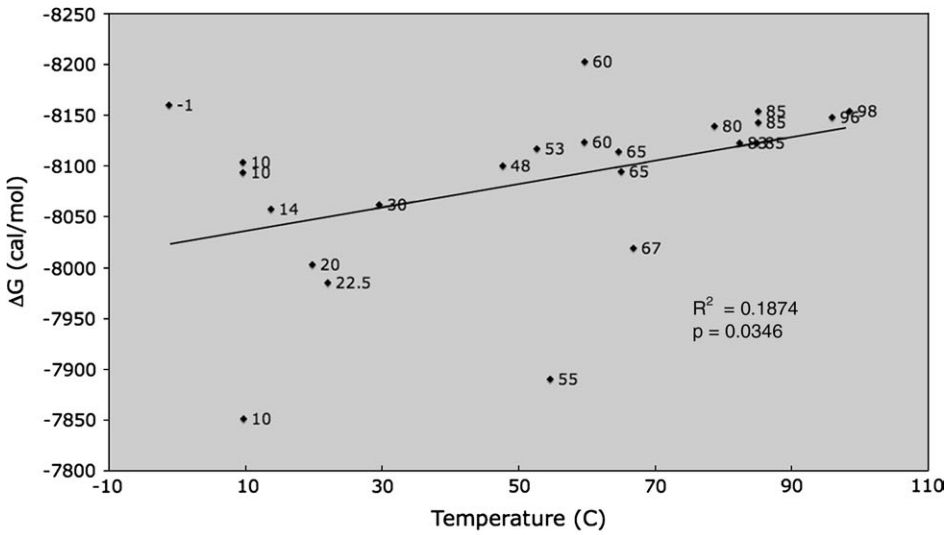
$$\Delta\mathrm{CF} = \frac{\sum_{i=1}^{N_{\mathrm{pos}}} \left(\sum_{\substack{m=1,n=1,n>m, \\ n \text{ or } m \neq \text{gap}}}^{N_{\mathrm{MSA}}} (\Delta G_{i,m} - \Delta G_{i,n})^2 \right)}{N_{\mathrm{pos}}},$$

where ΔG is the predicted eScape values, i = MSA alignment position, N_{pos} = length of MSA, N_{MSA} = number of sequences in the alignment, with the condition that a gap is not present in either alignment position n or m. This score was used as a measure for comparative analysis to identify nonuniform thermostability modulation within the proteomes for thermal adaptation. The measure was applied for each identified ortholog and to investigate changes in local flexibility and global stability by investigating the top and bottom 20 ranked predicted position-specific stability for each protein. The final ranking did not change significantly when using other methods of scores like Euclidean distance.

Results and Discussion

eScape Predicts Correct Changes in Conformational Modulation for Thermal Adaptation

To determine whether eScape can capture mutational effects, the impact of mutations on the stability of RNAse H1 was examined (Akasako et al. 1995; Haruki et al. 2007; Tadokoro et al. 2007). Shown in figure 1 is the position-specific conformational stability of psychrophilic and mesophilic isoforms of RNAse H1, as well as engineered hyperstablized variants. As is evident, the temperature



Species	Average ΔG (cal/mol)	Temperature Range (C)	Median Temperature (C)
C. symbiosum	-7852.62	10	10
T. fusca	-7891.23	55	55
Polarmomonas sp.	-7987.57	20 - 25	22.5
P. naphthalenivorans	-8003.83	20	20
T. thermophilus	-8020.06	50 - 82	65
M. burtonii	-8058.04	1.7 - 25.9	14
M. vannielii	-8061.44	20 - 40	30
P. arcticus	-8094.31	-10 - 30	10
M. thermoautotrophicum	-8095.35	40 - 70	65
T. pseudethanolicus	-8096.11	65	65
C. tepidum	-8100.01	48	48
P. cryohalolentis	-8102.72	-10 - 40	10
T. tengcongensis	-8114.96	50 - 80	65
T. neutrophilus	-8118.06	53	53
M. thermophila	-8121.54	Diverse Habitat	60
M. jannaschii	-8122.54	85	85
P. kodakaraensis	-8123.05	65 - 100	83
T. maritima	-8134.80	80	80
P. furiosus	-8137.31	70 - 103	85
P. abyssi	-8145.98	70 - 103	85
P. islandicum	-8146.86	95	96
P. horikoshii	-8151.85	98	98
P. ingrahamii	-8159.76	-12 - 10	-1
T. volcanium	-8202.59	60	60

FIG. 2.—Average mean stability of proteomes correspond to ambient temperature of extremophiles. eScape was used to make themostability predictions for proteins in the proteomes of 24 species. The average mean stability for each proteome is plotted against the median of the thermal range tolerated by the organism. The general trend indicates that the proteomes of thermophiles are relatively more stable than those of psychrophiles.

factors from X-ray structures (B factors) do not reflect the differences in conformational stability between the mesophilic and psychrophilic isoforms (fig. 1*A* and *B*). eScape predictions, however, show the psychrophilic isoform to be less thermally stable than the mesophilic one, which is expected because psychrophilic proteins are generally less stable under mesophilic conditions. Of note is that the regions, where conformational stability is predicted to be significantly different, correspond to mutational sites used to construct a hyperstabilized variant of RNAse H1. As such, eScape

Table 1
Additional Physical Selective Pressure in Addition to Thermal Stress Impacts the Collective Stability of Proteomes

Species	ΔG (cal/mol)	Temperature Range (°C)	pH Range	Median Temperature (°C)
Thermobifida fusca	−7,891.23	55	4–10	55
Methanobacterium thermoautotrophicum	−8,095.35	40–70	7.2–7.6	65
Thermoanaerobacter tengcongensis	−8,114.96	50–80	7.0–7.5	65
Pyrococcus furiosus	−8,137.31	70–103	5–9	85
Pyrococcus abyssi	−8,145.98	70–103	7	85
Pyrococcus horikoshii	−8,151.85	98	5–8	98
Sulfolobus acidocaldarius	−8,174.31	75–80	2–3	80
Thermoplasma volcanium	−8,202.59	60	2	60
Sulfolobus tokodaii	−8,212.08	70–85	2.5–3	80
Sulfolobus solfataricus	−8,224.91	78–87	3	82
Thermoplasma acidophilum	−8,234.48	55–59	2	59

Thermoacidophiles (gray) are significantly more stable than thermophiles not found in highly acidic conditions. The average mean ΔG free-energy stability values (cal/mol) of proteins in the proteome are reported for each species.

successfully reflects the qualitative stability changes associated with thermal adaptation, and identifies regions wherein conformational stability can be affected by mutation.

The stability values predicted by eScape reflect those that are expected when measured under mesophilic conditions due to the method in which this algorithm was trained. Consequently, we do expect to observe thermophiles to have higher predicted thermostability compared with mesophiles and conversely for psychrophiles. The importance of the RNAse H1 analysis is that it suggests that stability information can be inferred from sequence alone, and that large-scale analysis of eScape predictions can be conducted on entire proteomes, providing a venue for protein class-specific comparisons between organisms.

Collective Stabilities for Proteomes Rank with Optimal Temperature of Survival

Twenty-four proteomes of extremophiles have been selected for comparative analysis in order to identify if a trend in the collective stability of proteomes can be observed for thermal adaptation to extreme environments. As evident in figure 2, the average mean stability for each protein in the proteome correlates with the median temperature of the thermal range in which the organisms are found (with the exception of several outliers as described below). In short, the trend shows that the collective stability of proteins from thermophilic proteomes is more stable than those of psychrophiles. Although the adjusted R-squared value ($R^2 = 0.1874$) is low, when four of the extreme outliers were removed this score improved significantly to $R^2 = 0.4497$. The improvement in score with the removal of outliers suggests that other factors, apart from thermal adaptation, may be impacting the stability of the proteomes. Indeed, when thermophiles are further subdivided to account for the acidity of their natural environment, differences in stability are once again observed (table 1). Specifically, organisms that have evolved to survive in highly acidic environments are found to have higher predicted protein stability compared with those found in environments closer to a neutral pH. It is unclear whether this correlation arises due to the fact that protonation effects are not accounted for explicitly in eScape or whether there is a true difference in the stability of proteins from organisms that habituate acidic environments. Other abiotic factors that may impact the stability of proteins include pressure and viscosity, parameters that play an increasing role as the temperature decreases. Finally, the different tolerance ranges to abiotic stress should also be considered and will also contribute to proteome stability. Nonetheless, the results reveal

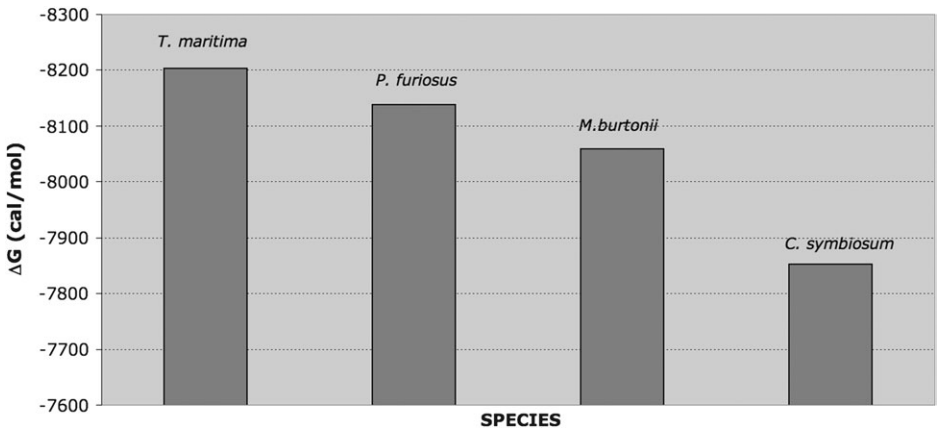


FIG. 3.—Average mean stability of proteomes for four extremophiles used for more a more detailed analysis of protein stability changes within the proteome. Two thermo- and psychrophilic proteomes were randomly chosen for a more detailed examination of CF modulation within the proteome for thermophilic adaptation. The four species are *Methanococcoides burtonii*, *Pyrococcus furiosus*, *Cenarchaeum symbiosum*, and *Thermatoga maritima*.

Table 2
One-Tailed *t*-Test Comparison of eScape Stability Values between Proteomes for (A) Global Stability and (B) Local Stability

	Thermotoga maritima	*Pyrococcus furiosus*	*Methanococcoides burtonii*	*Cenarchaeum symbiosum*
A. Comparison of global stability				
Thermotoga maritima		6.17E−05	2.10E−08	1.18E−07
Pyrococcus furiosus	1		5.82E−02	0.07263
Methanococcoides burtonii	1	0.9418		0.511
Cenarchaeum symbiosum	1	0.9274	0.489	
B. Comparison of Local Flexibility				
T. maritima		0.3481	0.4283	1
P. furiosus	0.6519		0.5939	1
M. burtonii	5.72E−01	0.4061		1
C. symbiosum	7.71E−16	2.20E−16	2.20E−16	

P values reported, statistical significance defined as $\alpha = 0.01$.

that the proteomes of thermophilic organisms are significantly more stable than those of psychrophilic origins, a result that is not unexpected. Whether these differences are manifested homogenously across the proteome remain unknown.

Comparisons of Energetics Associated with Global and Local Stability Show Consistent Trends

The proteomes of two psychrophiles and two thermophiles were randomly selected to conduct a more detailed analysis of conformational stability differences within proteomes that result from changes in global and local stability. The species used for this analysis were *M. burtonii* (2,242 sequences), *P. furiosus* (2,045 sequences), *C. symbiosum* (2,014 sequences), and *T. maritima* (1,852 sequences). As expected, the calculated average of the mean stability for each protein in the proteome follows the rank order of the optimal temperature of survival, as previously observed (fig. 3). *Thermatoga maritima* and *P. furiosus* are both thermophiles, whereas *M. burtonii* is eurypsychrophilic (psychrotolerant) and *C. symbiosum* is a psychrophile. The correlation to the respective ambient temperature confirms that correct qualitative predictions can be facilitated with eScape.

In order to ascertain the origins of the stability differences between the proteomes, the predicted stabilities for each protein were divided into the extreme high (top 20 least stable residues for each protein) and extreme low (bottom 20 for each protein), which correspond to those regions that are affected by changes in local and global stability, respectively (Hilser and Freire 1996; Hilser et al. 2006). By averaging the 20 highest and lowest values for each protein, the statistics will report, on average, the stability value of 10–20% of the residues for a single domain (note: the distribution of domain sizes have been found to peak around 100 residues with a sizeable number of domains containing 200 residues; Sowdhamini et al. 1996; Jones et al. 1998; Wheelan et al. 2000), and will provide a means of investigating two possible strategies for thermal adaptation. The importance of studying these effects separately is exemplified in a recent observation of a psychrophilic phenylalanine hydroxylase from *Colwellia psychrerythraea* 34H (Leiros et al. 2007). This cold-adapted enzyme has been observed to have increased flexibility and mobility around the active site without markedly affecting the thermostability of the protein. In other words, the local, and not the global stability, was modified. This is in contrast to previous general observations wherein the catalytic activity of cold-adapted proteins is often associated with decreases in local stability (i.e., increases in flexibility) as well as decreases in global stability.

To evaluate differences between the proteomes, we first conducted single tail *t*-tests between proteomes to identify significant differences ($\alpha = 0.01$) in the global and local stability of proteins in the proteomes (table 2). For global stability, *T. maritima* was found to be significantly more stable than all the other proteomes, whereas *P. furiosus* was only marginally more stable than *M. burtonii*. The average local stability of proteins in *C. symbiosum* was significantly less than that of the other proteomes.

In general, the results show that thermal adaptation through modulation of global and local stabilization does not necessarily follow the same correlating trend between stability and the optimal temperature for survival as was observed for the mean stability of all residues in the proteomes (table 2). For example, the global stability of *P. furiosus* is not significantly more stable than *C. symbiosum* as might be expected for this thermophile compared with the psychrophile. The lack of significant differences does not necessary imply that the conformational stability differences between the proteomes are not significantly different. Instead, an alternative hypothesis would be that the magnitude of change is not uniform across the entire proteome, thus obscuring the signal. To address this issue, a direct comparison of conformational stability between orthologs shared between each proteome is needed to identify subsets that may be subjected to more conformational modulation than others.

Comparison of Orthologs Shows Function-Specific Variations in Conformational Stability Modulation within the Proteome

Differences in thermoadaptability between proteins have been experimentally observed for psychrophilic and thermophilic elongation factor 2 (EF2) proteins (Thomas et al. 2001; Thomas and Cavicchioli 2002). A comparison of the in vitro activity and stability profiles of the two EF2 isoforms does not match at their

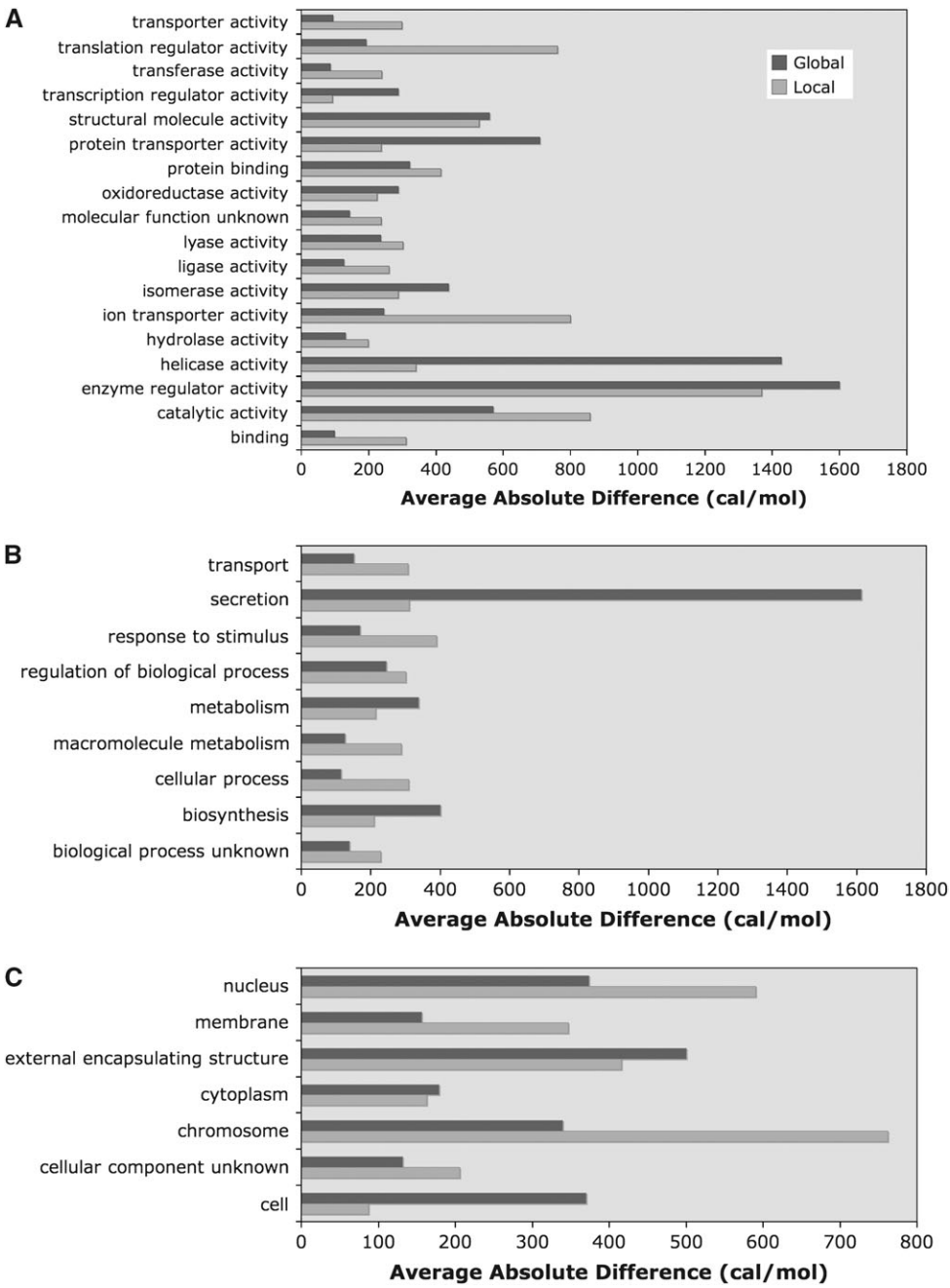


FIG. 4.—CF modulation varies significantly depending on the function of the protein. The average of absolute differences in local flexibility and global stability of proteins between extremophiles on opposite ends of the thermal spectrum vary significantly based on the (*A*) function of the proteins, (*B*) biological processes, (*C*) and cellular localization. Proteomes of four randomly selected extremophiles (*Methanococcoides burtonii*, *Pyrococcus furiosus*, *Cenarchaeum symbiosum*, and *Thermatoga maritima*) were partitioned based on GO functional annotation.

corresponding optimal growth temperatures. Instead, intracellular components affecting thermostability are involved in thermal adaptation of these proteins, thus illustrating that modulation of CF through sequence changes is not the only resource for adaptation. Furthermore, previous proteomic studies using sequence-based algorithms to survey the stability of proteins reveal variability within the proteome (Linding et al. 2004; Tartaglia et al. 2005; Monsellier et al. 2008). To investigate whether proteomes show class-specific degrees of thermal stabilization, we examined the entire proteomes and segregated the effects based on the GO Gene Ontology for protein annotation (Ashburner et al. 2000).

The results of the analysis show different magnitudes of change between the four proteomes (fig. 4). Several noteworthy observations can be made. First and most obvious, proteins annotated for enzyme regulation and catalytic activity (fig. 4*A*) are found to have the largest conformational modulation, impacting both global stability and local flexibility, a result that agrees with the large body of experimental comparisons, which is heavily biased toward proteins with enzymatic function. Second, proteins in other functional

Table 3
Stability Changes between Orthologs

Protein	ΔCF
Proteins with the largest[a] amount of observed conformational changes between the selected extremophiles	
ABC transporter	5.65E+06
Glycosyltransferase	5.18E+06
Dehydrogenase	5.12E+06
Fe–S oxidoreductase	5.01E+06
EndoIII-related endonuclease	4.98E+06
Dehydrogenase	4.81E+06
Nucleoside–diphosphate–sugar epimerase	4.80E+06
2-Methylthioadenine synthetase	4.80E+06
Methionyl-tRNA synthetase	4.79E+06
Cysteine desulfurase/selenocysteine lyase	4.78E+06
Adenosylmethionine-8-amino-7-oxononanoate aminotransferase	4.67E+06
Elongation factor 1-alpha	4.65E+06
Pyruvate–formate lyase-activating enzyme	4.62E+06
Nucleoside–diphosphate–sugar pyrophosphorylase	4.59E+06
Geranylgeranyl pyrophosphate synthase/ geranyltranstransferase	4.51E+06
2-Oxoglutarate ferredoxin oxidoreductase	4.51E+06
Threonine phosphate decarboxylase	4.48E+06
ABC-type phosphate transport system	4.48E+06
Type IV secretory pathway component	4.44E+06
Glutamyl-tRNA synthetase	4.32E+06
Proteins with the least[b] amount of observed conformational changes between the selected extremophiles	
30S ribosomal protein S5P	2.46E+06
Xanthosine triphosphate pyrophosphatase	2.46E+06
Chromosome segregation ATPase	2.45E+06
Translation factor	2.43E+06
Indole-3-glycerol phosphate synthase	2.43E+06
30S ribosomal protein S12P	2.42E+06
Putative uncharacterized protein	2.42E+06
Mn-dependent transcriptional regulator	2.39E+06
2′–5′ RNA ligase	2.32E+06
30S ribosomal protein S13P	2.31E+06
Chorismate synthase	2.31E+06
30S Ribosomal protein S19	2.30E+06
CTP synthase	2.26E+06
Phosphoribosyl-AMP cyclohydrolase/ phosphoribosyl-ATP pyrophosphohydrolase	2.26E+06
Putative uncharacterized protein	2.23E+06
Ketol-acid reductoisomerase	2.08E+06
Phosphoribosylcarboxyaminoimidazole	2.03E+06
Pyridoxine/pyridoxal 5-phosphate biosynthesis enzyme	2.01E+06
Enolase	1.97E+06
50 Ribosomal protein L11	1.87E+06

[a] The top 10% (20 clusters) of orthologs with the most observed changes in CF are listed.
[b] The top 10% (20 clusters) of orthologs with the least observed changes in CF are listed.

categories have varying degrees of modulation in their thermostability, with some groups showing more pronounced effects on the global stability, whereas other groups show more pronounced effects in local flexibility. For example, proteins with helicase activity ranked second highest in the extent to which global stability differs between proteomes, although changes in local flexibility were less obvious. Inspection of stability changes to proteins based on the GO annotation of biological processes shows those involved in secretion have the largest change in global stability, followed by metabolism and biosynthesis, when compared between the extremophiles (fig. 4*B*). Largest changes in local flexibility were observed for proteins found to localize at the chromosome and in the nucleus (fig. 4*C*). This observation is consistent with previous studies showing an increase in disordered regions associated with transcription factors (Ward et al. 2004; Liu et al. 2006; McEwan et al. 2007).

To examine these results in more detail, direct comparisons were made between orthologs shared by the four proteomes identified with Blast (Altschul et al. 1997) and aligned with ProbCons (Do et al. 2005). Limiting the data set to an examination of four proteomes allowed us to retain a sizable number of shared orthologs for direct comparison of thermostability changes that would otherwise be difficult in an exhaustive analysis. The ranking of 183 identified ortholog clusters, using a normalized measure for the magnitude of change, shows that metabolic enzymes are subject to more variance in modulation of CF than proteins involved in core translational and transcriptional processes (table 3), confirming our previous result. Focusing on changes specific to modulation of global stability (20 most stable residues) and local flexibility (20 least stable residues) shows the same observations compared with an examination of all positions (table 4). The majority of the top 40 ortholog clusters having large observed changes in local and global stability across the four proteomes corresponded to enzymes and metabolism associated proteins. The bottom 40 ortholog clusters with the least amount of conformational modulation impacting local and global stability largely included ribosomal proteins.

Metabolic proteins have direct impact on the intracellular milieu of metabolites and this may be the reason why proteins of this class may be more subjected to thermal adaptation via changes in protein stability. By targeting this select group with the proteome, impacts from abiotic stress on core translational and transcriptional processes can therefore be reduced by altering the protective properties of metabolomes. The findings do not imply that proteins associated with translation and transcription are exempted from stability modulation, but such observances occur to a lesser extent and with smaller changes in thermostability, comparatively. Instead, their stability can be further modulated by the products of metabolic enzymes, which may be subjected to stronger selection pressures for thermal adaptation, as suggested in this analysis.

Conclusions

The biochemical basis for thermal adaptation to extreme environments has been the subject of intensive study. Although comparative analyses of proteins have identified structural and sequence changes associated with adaptation, it has yet to be established that all proteins are stabilized in the same way in response to changes in environmental conditions. The analysis presented here shows that there is nonuniform modulation of CF and stability across components of the proteome. Indeed, not only are different proteins affected to different extents, the apparent mechanism of thermal adaptation (i.e., stabilizing local vs. global stability) differs from protein to protein. In general, our

Table 4
Top 40 Rankings of Ortholog Clusters with the Most and Least Changes in Local Flexibility and Global Stability

Protein	ΔCF
A) Top 40 proteins with the most modulation to local flexibility	
ABC transporter, ATPase subunit	1.21E+06
Glycosyltransferase	2.89E+05
Nucleotidyl transferase	2.49E+05
Nucleoside–diphosphate–sugar epimerase	2.14E+05
AAA ATPase	2.00E+05
Translation elongation factor	1.97E+05
Acetylornithine aminotransferase	1.97E+05
Geranylgeranyl reductase	1.88E+05
Threonine–phosphate decarboxylase	1.81E+05
Dehydrogenase	1.76E+05
Fe–S oxidoreductase	1.70E+05
Cysteine desulfurase/selenocysteine lyase	1.68E+05
Isopropylmalate/homocitrate/citramalate synthase	1.64E+05
Type IV secretory pathway component	1.64E+05
3-Isopropylmalate dehydratase	1.61E+05
Methionyl-tRNA synthetase	1.51E+05
2-methylthioadenine synthetase	1.44E+05
ATP synthase subunit beta	1.43E+05
UDP–glucose/GDP-mannose dehydrogenase	1.41E+05
2-Oxoglutarate ferredoxin oxidoreductase	1.37E+05
Anthranilate phosphoribosyltransferase	1.35E+05
Glucosamine 6-phosphate synthetase	1.34E+05
NADH–ubiquinone oxidoreductase	1.28E+05
Homoserine dehydrogenase	1.28E+05
Inosine-5′-monophosphate dehydrogenase	1.26E+05
Signal recognition particle GTPase	1.25E+05
Threonine synthase	1.25E+05
Phosphomannomutase	1.23E+05
Phosphoglycerate dehydrogenase	1.23E+05
Pyridoxal phosphate-dependent enzyme	1.22E+05
Alanyl-tRNA synthetase	1.17E+05
2-Oxoglutarate ferredoxin oxidoreductase	1.17E+05
NADH–ubiquinone oxidoreductase	1.16E+05
Geranylgeranyl pyrophosphate synthase/ geranyltranstransferase	1.14E+05
Pyruvate–formate lyase-activating enzyme	1.12E+05
ABC-type phosphate transport system	1.11E+05
Ornithine carbamoyltransferase	1.11E+05
EndoIII-related endonuclease	1.11E+05
Aspartyl/asparaginyl-tRNA synthetase	1.10E+05
Superfamily II DNA/RNA helicase	1.07E+05
B) Top 40 proteins with the most modulation to global stability	
ABC transporter, ATPase subunit	1.14E+06
Glycosyltransferase	2.68E+05
Nucleotidyl transferase	2.23E+05
Nucleoside–diphosphate-sugar epimerase	2.02E+05
Translation elongation factor	1.92E+05
Dehydrogenase	1.70E+05
AAA ATPase	1.64E+05
Geranylgeranyl reductase	1.63E+05
Acetylornithine aminotransferase	1.63E+05
Isopropylmalate/homocitrate/citramalate synthase	1.61E+05
Fe–S oxidoreductase	1.59E+05
Threonine–phosphate decarboxylase	1.56E+05
3-Isopropylmalate dehydratase	1.53E+05
Cysteine desulfurase/selenocysteine lyase	1.46E+05
Methionyl-tRNA synthetase	1.44E+05
Inosine-5′-monophosphate dehydrogenase	1.40E+05
Type IV secretory pathway component	1.36E+05
ATP synthase subunit beta	1.34E+05
UDP–glucose/GDP–mannose dehydrogenase	1.32E+05
Pyruvate–formate lyase-activating enzyme	1.25E+05
Homoserine dehydrogenase	1.24E+05
Threonine synthase	1.21E+05

Table 4
Continued

Protein	ΔCF
2-Methylthioadenine synthetase	1.20E+05
Anthranilate phosphoribosyltransferase	1.20E+05
2-Oxoglutarate ferredoxin oxidoreductase	1.19E+05
Glucosamine 6-phosphate synthetase	1.18E+05
Phosphomannomutase	1.17E+05
2-Oxoglutarate ferredoxin oxidoreductase	1.17E+05
Pyridoxal phosphate-dependent enzyme	1.14E+05
EndoIII-related endonuclease	1.13E+05
Signal recognition particle GTPase	1.10E+05
NADH–ubiquinone oxidoreductase	1.08E+05
NADH–ubiquinone oxidoreductase	1.08E+05
Phosphoglycerate dehydrogenase	1.06E+05
Geranylgeranyl pyrophosphate synthase/ geranyltranstransferase	1.01E+05
Aspartyl/asparaginyl-tRNA synthetase	9.94E+04
Ornithine carbamoyltransferase	9.89E+04
ABC-type phosphate transport system	9.58E+04
Chaperonin GroEL	9.38E+04
Thioredoxin reductase	9.24E+04
C) Top 40 proteins with the least modulation to local flexibility	
Ribosomal protein L11	3.38E+04
30S ribosomal protein S19	3.53E+04
Putative uncharacterized protein	3.59E+04
Divalent cation tolerance protein	3.68E+04
Phosphoribosyl-AMP cyclohydrolase/ phosphoribosyl-ATP pyrophosphohydrolase	3.73E+04
Phosphoribosylaminoimidazole carboxylase	3.75E+04
2′–5′ RNA ligase	3.91E+04
Ribosomal protein L5	3.94E+04
Ferritin-like protein	3.98E+04
30S ribosomal protein S13P	3.99E+04
Mn-dependent transcriptional regulator	4.03E+04
30S ribosomal protein S11P	4.08E+04
Xanthosine triphosphate pyrophosphatase	4.09E+04
Deoxycytidylate deaminase	4.17E+04
Enolase	4.18E+04
Tryptophanyl-tRNA synthetase	4.23E+04
6,7-Dimethyl-8-ribityllumazine synthase	4.24E+04
Glycerol dehydrogenase	4.25E+04
Ketol-acid reductoisomerase	4.26E+04
Ribosomal protein S12P	4.32E+04
30S ribosomal protein S5P	4.32E+04
Acetylglutamate kinase/acetylaminoadipate kinase	4.33E+04
Ribosomal protein L6P/L9E	4.35E+04
Pyridoxine/pyridoxal 5-phosphate biosynthesis enzyme	4.38E+04
Undecaprenyl pyrophosphate synthase	4.44E+04
Orotate phosphoribosyltransferase	4.49E+04
30S ribosomal protein S9P	4.53E+04
GTPase	4.54E+04
Aspartate–semialdehyde dehydrogenase	4.55E+04
Chorismate synthase	4.57E+04
Imidazole glycerol-phosphate dehydratase	4.58E+04
Phosphoribosylpyrophosphate synthetase	4.62E+04
Aspartate aminotransferase	4.64E+04
30S ribosomal protein S10	4.65E+04
NADH–ubiquinone oxidoreductase	4.66E+04
Imidazole glycerol phosphate synthase subunit hisH	4.68E+04
Dihydroorotate dehydrogenase	4.72E+04
Indole-3-glycerol phosphate synthase	4.74E+04
Glycine/serine hydroxymethyltransferase	4.80E+04
Ribosomal protein S7	4.84E+04
D) Top 40 proteins with the least modulation to global stability	
Ribosomal protein L11	2.54E+04
Ribosomal protein S12P	2.97E+04

Table 4
Continued

Protein	ΔCF
Iron-dependent transcriptional repressor	3.21E+04
Phosphoribosylaminoimidazole carboxylase	3.23E+04
30S ribosomal protein S11P	3.26E+04
30S ribosomal protein S19	3.30E+04
Xanthosine triphosphate pyrophosphatase	3.36E+04
Divalent cation tolerance protein	3.39E+04
30S ribosomal protein S13P	3.40E+04
Imidazole glycerol–phosphate dehydratase	3.43E+04
30S ribosomal protein S10	3.44E+04
50S ribosomal protein L2P	3.45E+04
Chorismate synthase	3.59E+04
NADH–ubiquinone oxidoreductase	3.65E+04
Enolase	3.69E+04
30S ribosomal protein S3P	3.76E+04
Phosphoribosyl-AMP cyclohydrolase/ phosphoribosyl-ATP pyrophosphohydrolase	3.81E+04
Ribosomal protein L6P/L9E	3.89E+04
2′–5′ RNA ligase	3.93E+04
Ferritin-like protein	3.98E+04
Ribosomal protein L5	4.04E+04
6,7-Dimethyl-8-ribityllumazine synthase	4.04E+04
Dihydroorotate dehydrogenase	4.07E+04
Imidazole glycerol phosphate synthase subunit hisH	4.07E+04
Indole-3-glycerol phosphate synthase	4.08E+04
Shikimate 5-dehydrogenase	4.12E+04
3-Phosphoglycerate kinase	4.12E+04
O-Sialoglycoprotein endopeptidase	4.12E+04
Ketol-acid reductoisomerase	4.13E+04
Phosphopantothenoylcysteine synthetase/ decarboxylase	4.17E+04
Orotate phosphoribosyltransferase	4.19E+04
Putative uncharacterized protein	4.25E+04
30S ribosomal protein S5P	4.26E+04
ATPases involved in chromosome partitioning	4.29E+04
Deoxycytidylate deaminase	4.32E+04
Dimethyladenosine transferase	4.33E+04
Pyridoxine/pyridoxal 5-phosphate biosynthesis enzyme	4.34E+04
Phosphoribosylpyrophosphate synthetase	4.34E+04
CTP synthase	4.34E+04
Ribosomal protein S7	4.37E+04

UDP, uridine diphosphate; GDP, guanidine diphosphate; NADH, Nicotinamide adenine dinucleotide

results indicate that the subsets of proteins that undergo thermal adaptation are not necessarily those involved in core translational and transcriptional processes. Instead, adaptation through increases in both local and global stability appear to be concentrated in proteins that have catalytic activity as well as those that regulate proteins with catalytic activity.

Although we currently have no explanation as to why proteins involved in metabolism are targeted for thermal adaptation through mutation, whereas other proteins are not, there are at least two possibilities, which may occur to different extents. The first is that the presence of chaperones in the cell provides a level of tolerance to thermal stress, and those proteins that do not interact with or are not regulated by chaperones will require adaptation to occur at the level of the individual protein. Those proteins that do interact with chaperones, however, could presumably rely on a more systemic strategy through increasing either the levels or the stability properties of the proteostasis machinery itself, for example. This is an especially attractive mechanism for proteins whose functions are highly regulated and intertwined with the activities of other proteins. Second, proteins whose stability appears most affected by thermal adaptation are enzymes that may be involved in biological processes that impact the intracellular milieu of organic and inorganic solutes. In particular are proteins involved in the synthesis of osmolytes leading to the general stabilization of proteins (Bolen and Baskakov 2001; Yancey 2005). This chain of events reduces the need to modify the stability of all components of the proteome.

Regardless of the reasons, the results presented here reveal significant class-specific differences in how the proteome responds to thermal adaptation, indicating that although adaptive constraints are often manifested at the level of the individual proteins, molecular adaptation also appears in the context of biological processes within the organism. The implication of these findings is that the mutational search space for thermal adaptation of the proteome is therefore significantly reduced with selection acting on subsets of the proteome. The findings suggests that rather than uniformly changing the thermostability of all proteins in the proteome, a more efficient strategy would be to stabilize a subset, and that subset may be most efficient in providing alternative solutions to stabilize the proteome.

Acknowledgments

This study was supported by NIH grant GM63747 and the Robert A. Welch Foundation (H-1461). J.G. would like to thank the Jeanne Kempner Scholar Foundation and the NSF Antarctic Biology Training Program.

Literature Cited

Akasako A, Haruki M, Oobatake M, Kanaya S. 1995. High resistance of *Escherichia coli* ribonuclease HI variant with quintuple thermostabilizing mutations to thermal denaturation, acid denaturation, and proteolytic degradation. Biochemistry. 34:8115–8122.

Altschul S, Madden T, Schaffer A, Zhang JH, Zhang Z, Miller W, Lipman D. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25(17):3389–3402.

Ashburner M, Ball CA, Blake JA, et al. (19 co-authors). 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 25:25–29.

Balch WE, Morimoto RI, Dillin A, Kelly JW. 2008. Adapting proteostasis for disease intervention. Science. 319:916–919.

Beadle BM, Baase WA, Wilson DB, Gilkes NR, Shoichet BK. 1999. Comparing the thermodynamic stabilities of a related thermophilic and mesophilic enzyme. Biochemistry. 38:2570–2576.

Berezovsky IN, Shakhnovich EI. 2005. Physics and evolution of thermophilic adaptation. Proc Natl Acad Sci USA. 102:12742–12747.

Bolen DW, Baskakov IV. 2001. The osmophobic effect: natural selection of a thermodynamic force in protein folding. J Mol Biol. 310:955–963.

D'Aquino JA, Gomez J, Hilser VJ, Lee KH, Amzel LM, Freire E. 1996. The magnitude of the backbone conformational entropy change in protein folding. Proteins. 25:143–156.

Do CB, Mahabhashyam MS, Brudno M, Batzoglou S. 2005. ProbCons: probabilistic consistency-based multiple sequence alignment. Genome Res. 15:330–340.
Feller G. 2007. Life at low temperatures: is disorder the driving force? Extremophiles. 11:211–216.
Fiala G, Stetter KO. 1986. *Pyrococcus–Furiosus* Sp-Nov represents a novel genus of marine heterotrophic Archaebacteria growing optimally at 100-degrees C. Arch Microbiol. 145:56–61.
Fields PA, Somero GN. 1998. Hot spots in cold adaptation: localized increases in conformational flexibility in lactate dehydrogenase A(4) orthologs of Antarctic notothenioid fishes. Proc Natl Acad Sci USA. 95:11476–11481.
Franzmann PD, Liu Y, Balkwill DL, Aldrich HC, Conway de Macario E, Boone DR. 1997. *Methanogenium frigidum* sp. nov., a psychrophilic, H2-using methanogen from Ace Lake, Antarctica. Int J Syst Bacteriol. 47:1068–1072.
Gianese G, Bossa F, Pascarella S. 2002. Comparative structural analysis of psychrophilic and meso- and thermophilic enzymes. Proteins. 47:236–249.
Gu J, Hilser VJ. 2008. Predicting the energetics of conformational fluctuations in proteins from sequence: a strategy for profiling the proteome. Structure. 16(11):1627–1637.
Haruki M, Tanaka M, Motegi T, Tadokoro T, Koga Y, Takano K, Kanaya S. 2007. Structural and thermodynamic analyses of Escherichia coli RNase HI variant with quintuple thermostabilizing mutations. Febs J. 274:5815–5825.
Hilser VJ, Freire E. 1996. Structure-based calculation of the equilibrium folding pathway of proteins. Correlation with hydrogen exchange protection factors. J Mol Biol. 262:756–772.
Hilser VJ, Garcia-Moreno EB, Oas TG, Kapp G, Whitten ST. 2006. A statistical thermodynamic model of the protein ensemble. Chem Rev. 106:1545–1558.
Huber R, Langworthy TA, Konig H, Thomm M, Woese CR, Sleytr UB, Stetter KO. 1986. *Thermotoga–Maritima* Sp-Nov represents a new genus of unique extremely thermophilic Eubacteria growing up to 90-degrees-C. Arch Microbiol. 144:324–333.
Jahandideh S, Abdolmaleki P, Jahandideh M, Barzegari Asadabadi E. 2007. Sequence and structural parameters enhancing adaptation of proteins to low temperatures. J Theor Biol. 246:159–166.
Jones S, Stewart M, Michie A, Swindells MB, Orengo C, Thornton JM. 1998. Domain assignment for protein structures using a consensus approach: characterization and analysis. Protein Sci. 7:233–242.
Kersey P, Bower L, Morris L, et al. (19 co-authors). 2005. Integr8 and genome reviews: integrated views of complete genomes and proteomes. Nucleic Acids Res. 33:D297–D302.
Larson SA, Hilser VJ. 2004. Analysis of the "thermodynamic information content" of a *Homo sapiens* structural database reveals hierarchical thermodynamic organization. Protein Sci. 13:1787–1801.
Leiros HK, Pey AL, Innselset M, Moe E, Leiros I, Steen IH, Martinez A. 2007. Structure of phenylalanine hydroxylase from *Colwellia psychrerythraea* 34H, a monomeric cold active enzyme with local flexibility around the active site and high overall stability. J Biol Chem. 282:21973–21986.
Linding R, Schymkowitz J, Rousseau F, Diella F, Serrano L. 2004. A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins. J Mol Biol. 342:345–353.
Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, Dunker AK. 2006. Intrinsic disorder in transcription factors. Biochemistry. 45:6873–6888.
McEwan IJ, Lavery D, Fischer K, Watt K. 2007. Natural disordered sequences in the amino terminal domain of nuclear receptors: lessons from the androgen and glucocorticoid receptors. Nucl Recept Signal. 5:e001.
Monsellier E, Ramazzotti M, Taddei N, Chiti F. 2008. Aggregation propensity of the human proteome. PLoS Comput Biol. 4:e1000199.
Preston CM, Wu KY, Molinski TF, DeLong EF. 1996. A psychrophilic crenarchaeon inhabits a marine sponge: *cenarchaeum symbiosum* gen. nov., sp. nov. Proc Natl Acad Sci USA. 93:6241–6246.
Shlyk-Kerner O, Samish I, Kaftan D, Holland N, Sai PS, Kless H, Scherz A. 2006. Protein flexibility acclimatizes photosynthetic energy conversion to the ambient temperature. Nature. 442:827–830.
Sowdhamini R, Rufino SD, Blundell TL. 1996. A database of globular protein structural domains: clustering of representative family members into similar folds. Fold Des. 1:209–220.
Spiwok V, Lipovova P, Skalova T, Duskova J, Dohnalek J, Hasek J, Russell NJ, Kralova B. 2007. Cold-active enzymes studied by comparative molecular dynamics simulation. J Mol Model. 13:485–497.
Tadokoro T, You DJ, Abe Y, Chon H, Matsumura H, Koga Y, Takano K, Kanaya S. 2007. Structural, thermodynamic, and mutational analyses of a psychrotrophic RNase HI. Biochemistry. 46:7460–7468.
Tartaglia GG, Cavalli A, Pellarin R, Caflisch A. 2005. Prediction of aggregation rate and aggregation-prone segments in polypeptide sequences. Protein Sci. 14:2723–2734.
Thomas T, Cavicchioli R. 2002. Cold adaptation of archaeal elongation factor 2 (EF-2) proteins. Curr Prot Pept Sci. 3:223–230.
Thomas T, Kumar N, Cavicchioli R. 2001. Effects of ribosomes and intracellular solutes on activities and stabilities of elongation factor 2 proteins from psychrotolerant and thermophilic methanogens. J Bacteriol. 183:1974–1982.
Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. 2004. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J Mol Biol. 337:635–645.
Wheelan SJ, Marchler-Bauer A, Bryant SH. 2000. Domain size distributions can predict domain boundaries. Bioinformatics. 16:613–618.
Yancey PH. 2005. Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. J Exp Biol. 208:2819–2830.
Zavodszky P, Kardos J, Svingor , Petsko GA. 1998. Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. Proc Natl Acad Sci USA. 95:7406–7411.
Zeldovich KB, Berezovsky IN, Shakhnovich EI. 2007. Protein and DNA sequence determinants of thermophilic adaptation. PLoS Comput Biol. 3:62–72.

Douglas Crawford, Associate Editor

Accepted June 23, 2009