The stability of proteins in extreme environments

Rainer Jaenicke* and Gerald Böhm†

Three complete genome sequences of thermophilic bacteria provide a wealth of information challenging current ideas concerning phylogeny and evolution, as well as the determinants of protein stability. Considering known protein structures from extremophiles, it becomes clear that no general conclusions can be drawn regarding adaptive mechanisms to extremes of physical conditions. Proteins are individuals that accumulate increments of stabilization; in thermophiles these come from charge clusters, networks of hydrogen bonds, optimization of packing and hydrophobic interactions, each in its own way. Recent examples indicate ways for the rational design of ultrastable proteins.

Addresses
*Institute of Biophysics and Physical Biochemistry, University of Regensburg, D-93040 Regensburg, Germany;
e-mail: rainer.jaenicke@biologie.uni-regensburg.de
†Institute for Biotechnology, Molecular Bioinformatics, University of Halle-Wittenberg, D-06099 Halle (Saale), Germany;
e-mail: gerald.boehm@biochemtech.uni-halle.de

Current Opinion in Structural Biology 1998, **8**:738–748

http://biomednet.com/elecref/0959440X00800738

Abbreviations

CS	citrate synthase
DHFR	dihydrofolate reductase
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GluDH	glutamate dehydrogenase
Hs	*Halobacterium salinarium*
IPMDH	3-isopropylmalate dehydrogenase
LDH	lactate dehydrogenase
PGK	phosphoglycerate kinase
PRAI	phosphoribosyl anthranilate isomerase
TIM	triosephosphate isomerase
Tm	*Thermotoga maritima*

Introduction

Life on earth exhibits an enormous adaptive capacity. Except for centers of volcanic activity, the surface of our planet is 'biosphere'. In quantitative terms, the limits of the biologically relevant physical variables are –40 to +115°C (in the stratosphere and hydrothermal vents, respectively), ≤120 MPa (for hydrostatic pressures in the deep sea), $a_w \approx 0.6$ (for the activity of water in salt lakes) and ≈1<pH<11 (for acidic or alkaline biotopes). During evolution, organisms achieved viability under extreme conditions either by 'escaping' or 'compensating' the stress or by enhancing the stability of their cellular inventory. In the case of temperature and pressure, there is no alternative to mutative adaptation for survival [1]. Here, we shall review the recent progress in research on protein stabilization, focusing on thermophiles with optimum temperatures of growth of more than 60°C (for hyperthermophiles, more than 80°C) and halophiles with optimum water activities around 0.6. Studies on proteins from acidophiles and alkalophiles have been scarce. Strict barophiles have recently been isolated — thousands of microbes were isolated from the first samples collected from the Challenger Deep at ~110 MPa [2], but very few of them were truly barophilic [3•]. Their proteins are still *terra incognita*.

Limits of stability and growth

Proteins, independent of their mesophilic or extremophilic origin, consist exclusively of the 20 canonical natural amino acids. In the multicomponent system of the cytosol, these are known to undergo covalent modifications at extremes of temperature, pH and pressure (deamidation, β elimination, disulfide interchange, oxidation, Maillard reactions, hydrolysis, etc. [4]). Extremophiles must compensate for amino acid degradation either by using compatible protectants or by enhanced synthesis and repair. Little is known about the chemistry involved, for example, in the hydrothermal decomposition of proteins, and even less is known about protection and repair. Applying temperatures beyond 100°C, the thermal stabilities of the common amino acids are (Val,Leu)>Ile>Tyr>Lys>His>Met>Thr>Ser>Trp>(Asp,Glu, Arg,Cys). In many cases, the half-lives of the degradation reactions are significantly shorter than the generation time of hyperthermophilic microorganisms [5]; to this limit, biomolecules could still be resynthesized at biologically feasible rates. The temperature at which ATP hydrolysis becomes the limiting factor for viability lies between 110 and 140°C [6]. This temperature limit coincides with the temperature range at which the hydrophobic hydration of proteins vanishes and water becomes an 'ordinary solvent' [1]. Apparently, both the integrity of the natural amino acids and the formation of the hydrophobic core upon protein folding are essential for viability. Extrinsic factors and compatible solutes may enhance the stability and shift the limits of growth of prokaryotes as well as eukaryotes [7].

Fundamentals of protein stability

Proteins exhibit marginal stabilities that are equivalent to only a small number of weak intermolecular interactions [1,8]. In this respect, proteins from extremophiles do not differ strongly from their mesophilic counterparts. Their adaptation, either intrinsic or through interaction with extrinsic factors, is accompanied by only marginal increases in the free energy of stabilization. No general strategy of stabilization has yet been established. In recent years, however, well-defined increments of stability have been elucidated by analyzing ultrastable proteins and verifying their specific anomalies by rational design. As indicated by these studies, stabilization may involve all levels of the hierarchy of protein structure: local packing of the polypeptide chain, secondary and supersecondary structural elements, domains and subunits [4]. Taking thermal stability as an example, several experimental approaches have been used to assign specific structural alterations to changes in stability: selection of temperature-sensitive

mutants; systematic variations of amino acid residues in the core or in the periphery of model proteins; fragmentation of domain proteins or modifications of connecting peptides between domains; and alteration of subunit interactions by mutagenesis or solvent perturbation [1,9].

Stability refers to the maintenance of a defined functional state under extreme conditions. High-resolution structures in the crystalline state and in solution have shown that the atomic coordinates of proteins can be determined down to a resolution better than 1 Å. Even this precision, however, does not allow the calculation of the free energy of stabilization from coordinates, nor does it consider the dynamics as an essential prerequisite of protein function. The polypeptide chain may fluctuate between preferred conformations with amplitudes and angles up to 50 Å and 20°, respectively [10].

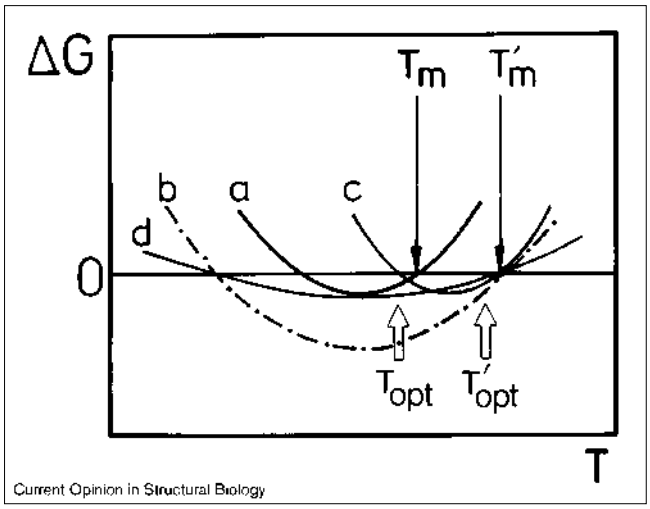
Considering extremophiles in comparison with their mesophilic counterparts, evolutionary adaptation is nothing more than the conservation of functionally important motions in such a way that, under altered physical conditions, the protein inventories of extremophiles and mesophiles are in 'corresponding states' [1]. In this context, the stability of an individual protein refers to the native state, as well as the intermediates on its pathway from the nascent or unfolded ensemble of states (U) to the functional entity (N). Evidently, in order 'to be extremophilic', a protein has to cope with the extreme conditions at all stages along its folding pathway.

Stability and folding

The driving forces that are responsible for protein folding reflect the hierarchy of contributions involved in protein stabilization, that is, on the one hand, the nearest neighbor and through-space short-range interactions that optimize packing and minimize cavity volume and, on the other hand, the entropy effects due to water release from hydrophobic surfaces [10,11]. Both the enthalpic and entropic contributions to the free energy of stabilization are affected by the extreme conditions we are dealing with. The difference in the stabilities of mesophilic and (hyper-)thermophilic proteins, $\Delta\Delta G_{N\rightarrow U}$, does not exceed ~100 kJ/mol, that is, the equivalent of a few noncovalent interactions [12•,13•]; often it is even lower, rendering the definition of general 'strategies' of thermal adaptation extremely difficult. This is especially true because the enthalpic contributions, in terms of additional stabilizing interactions, necessarily lead to the previously mentioned decrease in flexibility, which naturally corresponds to a decrease in conformational entropy. Thus, evolution had to find a balance between rigidity as a prerequisite of stability and specificity, on the one hand, and flexibility, for example, in connection with ligand interactions and degradation, on the other.

As a consequence of the parabolic temperature dependence of the free energy of stabilization, proteins exhibit heat and cold denaturation (Figure 1). Commonly, the latter becomes detectable only under moderately destabilizing conditions [14,15•]. In the case of proteins from thermophiles, the ΔG versus temperature profile is either flattened or increased to larger $\Delta G_{N\rightarrow U}$ levels, rather than being shifted to higher temperatures. The ΔG maximum is always far below the optimal growth temperature; this holds also true for the (hyper-)thermophilic proteins [12•,16].

Figure 1

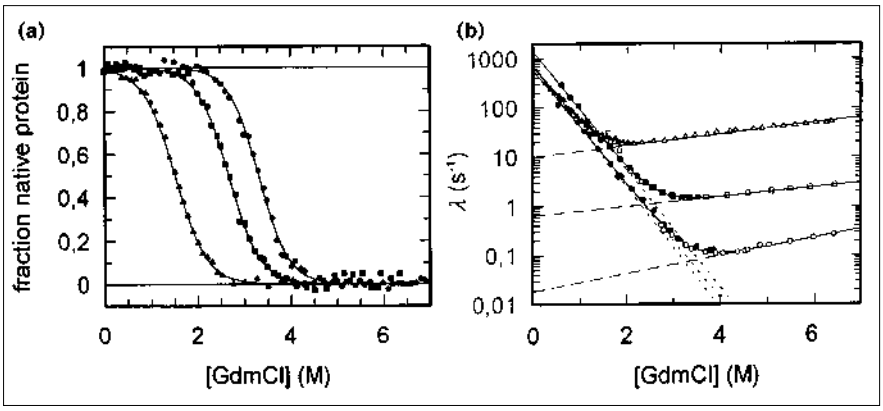


Hypothetical temperature profile of the free energy of **(a)** mesophilic and **(b–d)** thermophilic proteins. ΔG is defined as the difference in the free energies between the native and denatured proteins. T_m and T'_m are the melting temperatures of the mesophilic and thermophilic variants, respectively. The minimum of the ΔG parabola for a given protein (i.e. maximum stability) is observed at a temperature that is much below the optimal growth temperature (T_{opt} and T'_{opt}) of the respective mesophilic or thermophilic organism.

In order to simulate the effect of temperature on folding, the *in vitro* denaturation/renaturation of hyperthermophilic glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was studied at 0–100°C. Refolding over a wide temperature range was found to yield the native state, even beyond the physiological temperature range, indicating that thermal stability refers not only to the native state, but also to intermediates on the folding pathway, independent of their states of association. At 0°C, the hyperthermophilic enzyme is trapped as a tetrameric intermediate with molten globule-like properties; upon shifting the temperature beyond ~10°C, the native state is reached instantaneously [16,17].

Regarding the folding kinetics, available data allow the conclusion that increasing intrinsic stability is reflected by a decrease in the rate of unfolding. In this context, mutant studies have shown that enhanced stability may be determined kinetically rather than thermodynamically [18•]. Comparing the unfolding and folding kinetics of the all-β, single-domain cold-shock proteins from *Bacillus subtilis*,

Figure 2



The conservation of the unfolding/folding mechanism of cold-shock proteins. **(a)** Equilibrium unfolding transitions of cold-shock proteins from *B. subtilis* (▲), *B. caldolyticus* (■) and *T. maritima* (●) induced by guanidinium chloride (GdmCl) at 25°C and monitored by intrinsic fluorescence. Least-squares fit analyses based on the two-state model U $\rightleftharpoons$ N (full lines) yield stabilization energies ΔG_{stab} = 11.3, 20.1 and 26.2 kJ/mol for *B. subtilis*, *B. caldolyticus* and *T. maritima* Csp, respectively. **(b)** Kinetics of unfolding (open symbols) and refolding (closed symbols) of *B. subtilis* (△,▲), *B. caldolyticus* (□,■) and *T. maritima* Csp (○,●), respectively. The apparent rate constant, λ, is plotted against the concentration of GdmCl. The fits are on the basis of the linear two-state model. Reproduced with permission from [20•].

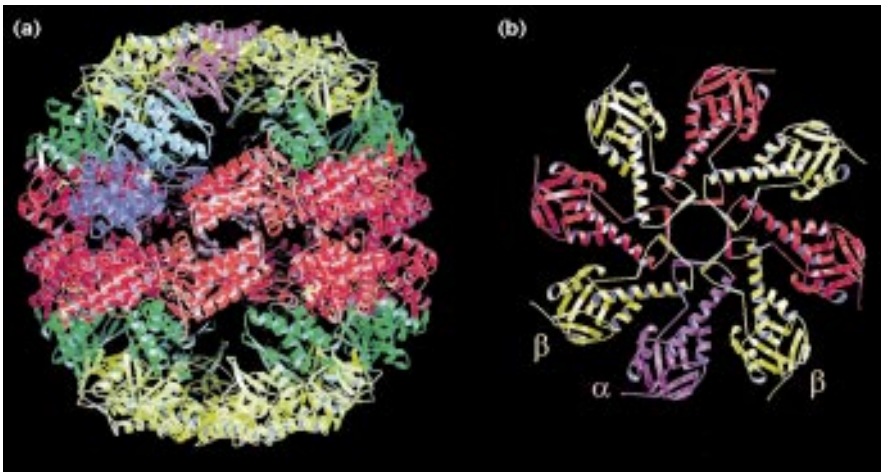
B. caldolyticus and *Thermotoga maritima* (Tm) (denaturation temperature T_m = 52, 72 and ~90°C, respectively), unfolding was shown to exhibit significantly slower kinetics with strongly decreasing rates, in accordance with the increase in stability. On the other hand, folding occurred extremely fast (τ ~1 ms), in a simple two-state reaction with closely similar kinetics, despite numerous sequence variations among the three proteins (Figure 2) [19,20•].

In the case of large proteins, generally, kinetic partitioning, that is, aggregation as a side reaction, competes with proper folding [11]. As a consequence, due to (partial) irreversibility, neither chemical denaturants nor temperature allow thermodynamic stability data to be determined. This is especially true for multidomain proteins and, to an even greater extent, for oligomers and multifunctional fusion proteins [16,18•,21,22•]. In certain cases, full reversibility was accomplished by choosing appropriate solvent conditions [23]. Evidently, kinetic partitioning *in vivo* is taken care of by molecular chaperones, if it is allowed to interpret the heat-stress response of, for example, *Pyrodictium occultum*, *Thermus thermophilus*, *Sulfolobus shibatae* and *Methanopyrus kandleri* along this line. All four species have been shown to express ATP-dependent 'thermosomes', sometimes to extreme levels [24,25]. The three-dimensional structure of these thermosomes exhibits a relationship with bacterial and eukaryal chaperones, with the only distinction being that the hyperthermophilic archaeal thermosomes obviously do not require co-chaperones; in the case of the thermosome from *T. acidophilum*, the apical domain seems to adopt this function [26•,27••] (Figure 3).

Forces and mechanisms involved in protein stabilization

The spatial structure of proteins is determined by electrostatic forces between polar and ionized groups and by hydrophobic effects involving nonpolar residues. The

Figure 3



Structure of the hexadecameric *T. acidophilum* thermosome (class II) seen in **(a)** side view and **(b)** top view, illustrating the 'cage-like' assembly with its eightfold symmetry. The apical domains close off both end cavities. Protrusions extend centrally from each thermosome apical domain, forming side-by-side contacts and a central β-sheet ring, comprising a built-in equivalent of GroES in the GroE system. Domains are colored in red (equatorial), green (intermediate) and yellow (apical). Within each complex, domains of aligned subunits are highlighted in blue (equatorial), light blue (intermediate) and violet (apical). (b) Thermosome α (red/violet) and β (yellow) apical domains. Reproduced with permission from [38••].

physical nature of the latter was recently interpreted as being entropic and enthalpic due to significant contributions from van der Waals' forces [1,8,28]. Attempts to obtain quantitative estimates of the different types of weak interactions have been based both on a comparison of known X-ray structures of mesophilic and extremophilic homologs, and on thermodynamic studies of point mutants. The $\Delta\Delta G_{N\rightarrow U}$ for extremophilic proteins is of the same order of magnitude as the overall free energy of stabilization $\Delta G_{N\rightarrow U}$ observed for mesophilic proteins. The following conclusions may be drawn regarding the significance of the various types of interactions:

1. $\Delta G_{N\rightarrow U}$ is equivalent to the energy required to break, at most, five hydrogen bonds, corresponding to about 1% of their total number in the folded structure of an average single-domain protein [11,29].

2. In the unfolded state, a 10 kDa protein exposes about 440 polar sites, half of them involved in internal hydrogen bonds in the native state. As a consequence, even a marginal difference in hydrogen-bond strength between water–water and water–protein hydrogen bonds will be magnified to an energy change that may well exceed $\Delta G_{N\rightarrow U}$ [29].

3. Water release from polar and nonpolar sites leads to an increase in entropy, which is supposed to be the driving force in many folding and assembly processes [30].

4. α helices and extended β structures contribute significantly to protein stability; in this context, helix-dipole interactions with charged groups in their vicinity are highly significant [9,31].

5. Contributions may also come from multiple hydrogen bonds between arginine groups and backbone carbonyl oxygens [32].

6. Only about 70% of the theoretically available hydrophobic contributions are realized as a consequence of the balance of favorable contributions to $\Delta G_{N\rightarrow U}$ on protein folding, leaving ample space for additional optimization [29].

7. As charged groups are commonly exposed to the aqueous solvent, intramolecular coulombic interactions cannot be of major importance in protein stabilization, unless they form clusters. Most of the polar sites in the cores of proteins are internally hydrogen bonded. The lack of this kind of internal saturation is found to be strongly destabilizing unless there are structured water molecules available to compensate for geometrical constraints [9,33]. Such 'ordered clathrate hydrates' have also been reported for patches of nonpolar residues exposed to the aqueous medium. Whether and how they contribute to stability is still unresolved [34].

The effects of temperature, pressure, charge and water activity on weak intermolecular interactions are highly complex. In the case of temperature, direct measurements of intermolecular forces have shown that the energy per polar group can exceed the thermal energy [35]. Hydrogen bonds are favored at low temperature and become weaker as the temperature is increased. Due to the compensatory effects in the total energy balance, predictions with respect to the significance of any specific type of interaction cannot be made.

Regarding the structural levels contributing to protein stability, increments may originate from local nearest neighbor interactions, secondary and supersecondary structural elements, subdomains, domains and subunits. The cumulative effect in terms of, at least, qualitative additivity has been illustrated by fragment and mutant studies [9,10,21,36,37]. At the quaternary level, a remarkable feature of hyperthermophiles is the occurrence of anomalous states of association and fused multifunctional proteins. Obviously, the reduction in the accessible surface area is associated with extreme thermophilicity [38••,39–47,48••,49]. Phosphoribosyl anthranilate isomerase (TmPRAI) and the bifunctional phosphoglycerate kinase–triosephosphate isomerase (TmPGK–TIM) fusion protein from *T. maritima* may serve as examples [39,47]. TmPRAI is a dimer with a complete $(\alpha\beta)_8$-barrel fold. In the monomeric *Escherichia coli* enzyme, the α_5 helix is replaced by a loop. The increase in stability comes from two long protrusions, which fit into cavities in the other subunit, thus favoring dimerization. Moreover, the sidechains of the N-terminal methione and the C-terminal leucine are immobilized in a hydrophobic cluster and, finally, the number of ion pairs is increased [48••]. In the case of the TmPGK–TIM fusion protein, a (–1) frameshift leads to the expression of both monomeric 43 kDa TmPGK and tetrameric 286 kDa TmPGK–TIM [39]. The dissection of the gene and the subsequent cloning, expression and characterization of the separate TmTIM showed that, in holding the complex together through three-dimensional domain swapping [50], the dimer of dimers gains significant stability within the fusion protein [12•,21,22•,51•]. The isolated entities are stabilized by the above increments. From a structural point of view, TmPGK in the fusion protein is practically indistinguishable from its thermophilic and mesophilic counterparts [39,51•,52]. In contrast, TmTIM (as the core of the fusion protein) differs significantly as a result of the additional intersubunit contacts (D Maes, RK Wierenga, personal communication). As one would expect for a complex multidomain oligomer, the folding of the fusion protein is hampered by kinetic partitioning. *In vivo*, folding might be assisted by thermosomal chaperone activity. Lactate dehydrogenase (LDH) has been used as a paradigm to summarize the various strategies of intrinsic stabilization. In this case, octamer formation is an 'artifact' of overexpression in *E. coli*, rather than a means of stabilization [53••,54]. Apart from intrinsic protein stabilization and

Table 1

The relative amino acid compositions of mesophiles and thermophiles [38••].

Amino acid*	Mesophiles	Thermophiles
Charged residues (DEKRH)	24.11%	29.84%
Polar/uncharged residues (GSTNQYC)	31.15%	26.79%
Hydrophobic residues (LMIVWPAF)	44.74%	43.36%

*One-letter abbreviations of the amino acid residues are given in brackets.

chaperoning, extrinsic factors, such as ions, carbohydrates or cofactors, may have additional stabilizing effects [7,55].

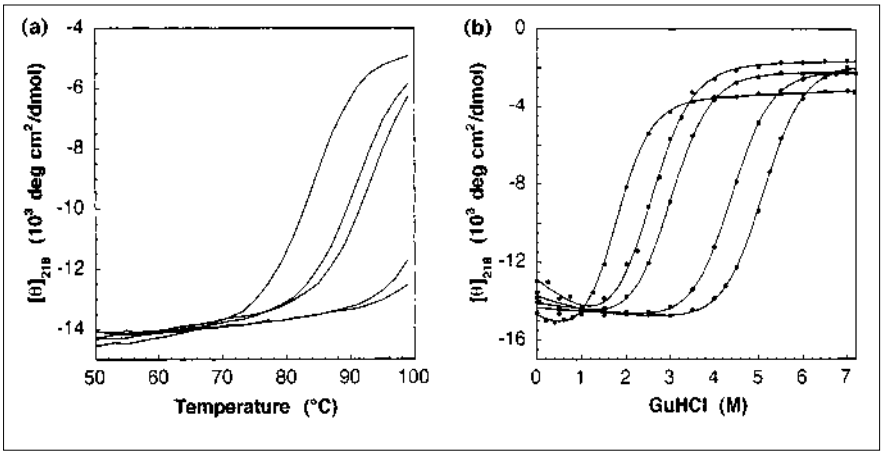
Genomes, phylogeny and general 'rules' of extremophilic adaptation

Inspecting available genome sequences and protein structures, there seems to be no way to unambiguously correlate extremes of physical conditions with either the amino acid composition or the three-dimensional structure of a given set of homologous proteins. What has been gained with increasing detail is an understanding of the phylogeny of microorganisms and their genes [56].

16S ribosomal RNA has been used for more than a decade to analyze evolutionary relationships between organisms [57••]. As a result, the three domains of life — bacteria, archaea and eukarya — were defined, with the general conclusion being that bacterial and archaeal hyperthermophiles are close to the root of the phylogenetic tree, preceding their mesophilic counterparts. At this point, it is important to note that such phylogenetic relationships say nothing about the temperature at which life started [58]. The advent of complete genome sequences made it clear that the phylogenetic tree has more complex roots than expected so far. Considering one single genome, single genes may or may not agree with the rRNA tree. Even more perplexing is the fact that genomes contain a mix of DNAs, some are close to archaea, while others are close to bacteria. The recently determined complete genome sequence of the hyperthermophilic bacterium *Aquifex aeolicus* [38••] shows the phenomenon in a unique way. A comparison of several of its genes with their counterparts in a range of species from all three domains of life reveals that *Aquifex* gets different pylogenetic placements, depending on which gene is being considered. It looks as if each gene has its own history, possibly due to lateral gene transfer or the 'swapping of genes' among organisms. Although the mechanism of gene swapping is still unknown, there seems to be no better explanation for the observation that 17 out of 34 families of eukaryotic proteins that date back to early cell evolution look as if they come from bacteria, while only eight families show a greater similarity to archaea, the supposed ancestor of eukarya. In spite of these inconsistencies, presently available genome sequences still fit the three-domain hypothesis [59•].

Among the 14 bacterial and archaeal genomes sequenced so far, five belong to hyperthermophiles: *Aquifex aeolicus*, *Archaeoglobus fulgidus*, *Methanobacterium thermoautotrophicum*, *Methanococcus jannaschii* and *Pyrococcus horikoshii* (http://www.ncbi.nlm.nih.gov/Entrez/Genome/org.html). A comparison of these genomes with respect to specific genes from mesophiles provides a data set that is sufficiently large to extract certain trends in amino acid usage (Table 1). Some correlations seem to hold — compared to mesophiles, genomes of thermophiles encode higher levels of charged amino acids, primarily at the expense of uncharged polar residues. Glutamine, in particular, seems to be significantly discriminated against in hyperthermophiles (compare [60]). This observation might be rationalized by an increased rate of deamidation of this residue at higher temperatures. One might expect the same difference in the number of asparagines; however, this residue does not appear to be subject to similar discrimination [38••].

Figure 4



Circular dichroism measurements of the *Streptococcal* protein Gβ1 domain and its mutants with optimized core packing, a decreased hydrophobic surface area and improved secondary structure propensity. For the nomenclature of the mutants, 'c' refers to residues in the core, whereas 'b' stands for bordering residues. Profiles from left to right refer to: Gβ1 (wildtype); Gβ1-c3 (Y3F, L7I, V39I); Gβ1-c3b1 (Y3F, L7I, V39I, T25E); Gβ1-c3b2 (Y3F, L7I, V39I, T61I, T18I); Gβ1-c3b4 (Y3F, L7I, V39I, T61I, T18I, T25E, V29I). **(a)** Thermal unfolding. **(b)** Chemical denaturation by guanidinium chloride at 50°C [66••]. Reproduced with permission from [66••].

Earlier attempts to find 'traffic rules' of protein stabilization were based on sequence comparisons in relatively small sample sizes. A critical analysis proved these early predictions to be statistically insignificant [61]. Examining the increasing number of high-resolution structures, a variety of strategies of thermal adaptation have emerged [62–64]. Recently, Vogt and Argos [65], in an attempt to rank the strategies proposed so far, examined 16 protein families, each with at least one known thermophilic and one known mesophilic structure, focusing on hydrogen bonds, ion pairs, polar surface composition, internal cavities, packing densities and secondary structure composition. The results show that enhanced thermostability is correlated with a consistent increase in the number of hydrogen bonds and ion pairs, apart from an increase in polar surface area.

The picture that is now emerging of the stabilization effects of hyperthermophilic proteins, showing a mosaic of different strategies of thermal adaptation, has been tested and supported by the prediction of mutagenesis effects, as, for example, in the case of the *Streptococcal* protein Gβ1 [66••]. A sevenfold mutant of the protein was designed by an objective computer algorithm that modeled in parallel several complex contributions to stability into the protein, incorporating optimized core packing, increased burial of hydrophobic surface area, more favorable helix-dipole interactions and the improvement of secondary-structure propensities. The stabilization effects were found to be additive, with a ΔΔG of 18 kJ/mol and a shift in the thermal transition from 83°C for the wildtype protein to more than 99°C for the sevenfold mutant (Figure 4).

In a limited number of cases, enzymes from psychrophiles have been used in order to expand the temperature scale and confirm present ideas on protein stabilization. Summarizing the state-of-the-art, it appears that the adaptive mechanisms involve weakening of intramolecular interactions and/or increasing interactions with the solvent; both these tendencies lead to enhanced flexibility, that is, catalysis at a lower energy cost [67•]. TIM [68•], 3-isopropylmalate dehydrogenase (IPMDH) [69] and citrate synthase (CS) [70] have been studied in attempts to recognize specific adaptive strategies. In the latter two cases, a comparison of the amino acid compositions, as well as sequence alignments and homology modeling, have shown that there are no traffic rules, except for an increase in extended surface loops and a decrease in the number of proline residues in loops, equivalent to an increase in the configurational entropy of the denatured state. The observed changes in charge distribution are difficult to classify.

A pairwise comparison of homologous proteins with different thermal stabilities has been widely applied in order to discover strategies of thermal adaptation. Based on known three-dimensional structures, stabilizing features were deduced. In order to test the proposals, site-directed

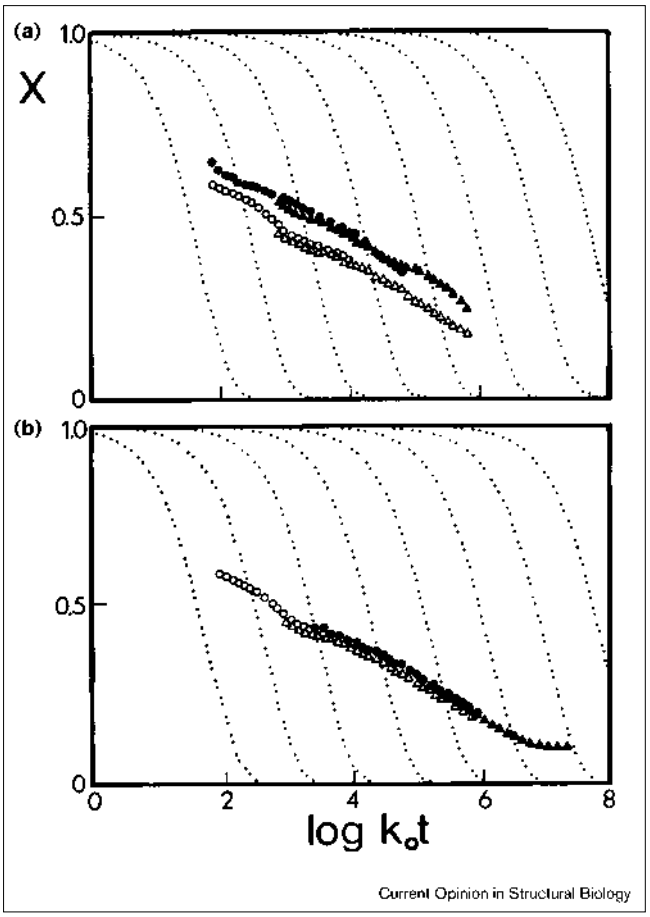
Table 2

Strategies of thermal stabilization of selected thermophilic enzymes.

Protein	Major cause(s) of thermostability*
CS (*P. furiosus*)	Increased compactness, enhanced subunit interactions, increased number of intersubunit ion pairs, shortening of loops [88••].
Ferredoxin (*T. maritima*)	Structurally: stabilization of α helices; replacement of conformationally strained residues by glycines; strong docking of the N-terminal methionine; increase in the number of hydrogen bonds [89]. Thermodynamically: flat ΔG versus temperature profile caused by a low ΔCp of unfolding [90•].
Ferredoxin (*Synechococcus elongatus*)	Extension of the hydrophobic core, a unique hydrophobic patch on the surface of a β sheet, two unique ion-pair networks [91].
GluDH (*P. furiosus*)	A series of extended ion-pair networks on protein subunit surfaces and ion-pair networks buried at interdomain and intersubunit interfaces, enhanced packing within the inner core, amino acid replacements increasing the hydrophobicity and sidechain branching (Val→Ile) [92,93•].
GAPDH (*T. maritima*)	Large number of additional salt bridges [94].
Indole 3-glycerol phosphate synthase (*Sulfolobus solfataricus*)	Large number of additional salt bridges (partly cross-linking adjacent helices), increased helix capping, dipole stabilization, increased hydrophobic interactions, strengthening of chain termini and solvent-exposed loops [95].
IPMDH (*T. thermophilus*)	Increased number of ion pairs and hydrogen bonds, extended hydrophobic subunit interactions and improved packing of the hydrophobic core, shortened chain termini [96••,97•,98•].
LDH (*T. maritima*)	Increased number of ion pairs, decreased hydrophobic surface area, increased helicity, less cavity volume [53••].
PGK (*T. maritima*)	Increased rigidity by additional ion pairs, stabilization of α helix and loop regions, stabilization by fusion with TIM [23,51•,52].

*Listed according to the authors' priorities (if described). ΔCp, heat capacity change.

Figure 5



Hydrogen–deuterium exchange of GAPDH from *T. maritima* (closed symbols) and rabbit (open symbols), measured at pH 6.0 (●,○) and 7.0 (▲,△) plotted as relaxation spectra. For experimental details, see [72]. **(a)** Measurements at constant temperature (25°C); increasing X values reflect increased rigidity. **(b)** Measurements at 68°C for TmGAPDH and 25°C for rabbit GAPDH. Coincidence of the curves indicates similar flexibility (P Závodszky, C Böthe, J Kardos, A Svingor, R Jaenicke, unpublished data).

mutagenesis or similar control experiments were performed. Examples of strategies of thermal adaptation are given in Table 2.

On going from lower to higher growth temperatures, numerous differences are most frequently reported: the clustering of (intrasubunit and/or intersubunit) ion pairs; improved packing of the hydrophobic core (increased van der Waals' interactions); additional networks of hydrogen bonds and enhanced secondary structure propensity; increased helix-dipole stabilization; an increased polar surface area; a decreased number and total volume of cavities; and burying hydrophobic surface area by either tightening interdomain and intersubunit contacts or by increasing the state of association. For a collection of 'recipes' to improve thermal stability, see [71]. The common denominator of all the adaptive changes is the conservation and optimization of the functional state of the given protein. Experimentally, this can be defined either by biological activity or by physical parameters connected with the flexibility or rigidity of the polypeptide chain. A suitable approach is hydrogen–deuterium exchange kinetics at various temperatures. They confirm the hypothesis that under optimum physiological conditions, proteins are in corresponding states [1,72,73] (Figure 5).

Halophiles

Halophilic proteins require multimolar salt concentrations for activity and denature at low salt. Malate dehydrogenase from *Haloarcula marismortui* was the first extremely halophilic enzyme structure to be solved by X-ray crystallography, ending a long physicochemical odyssey from a 'dimeric core with loops' to the real tetrameric 'LDH homolog' [74]. Anomalous features of the protein include an excess of acidic residues over basic residues and an increase in the number of intramolecular salt bridges compared to nonhalophilic homologs [75]. An explanation for its halophilicity is its tendency to increase water binding to the protein surface, in order to compete with the highly concentrated salt solution for water of hydration. This idea was confirmed by the significant increase in peripheral hydrogen bonds in the case of the 1.9 Å X-ray structure of 2Fe–2S ferredoxin from *H. marismortui* [76]. On the other hand, the observed significant increase in halophilicity after mutating a single glutamic acid position (unique in halophilic malate dehydrogenases) to arginine sheds some doubt on the hypothesis [77]. An additional feature of halophilic adaptation, which emerged from homology modeling of glutamate dehydrogenase (GluDH) from *Halobacterium salinarium* (Hs), is a significant reduction in 'exposed hydrophobic character', due to a decrease in the number of surface-exposed lysine residues [78•]; whether this finding will stand the experimental test of X-ray analysis remains to be seen.

Considering the X-ray structure of a moderately halophilic protein, monomeric dihydrofolate reductase (DHFR) from *Haloferax volcanii*, no striking differences from its mesophilic counterpart have been detected, except for some charge clusters of negatively charged residues [79•]. Similar observations were reported for the dimeric dihydrolipoamine dehydrogenase from the same source. In this case, site-directed mutagenesis of four coordinated glutamate residues involved in a charge cluster at the subunit interface was shown to have a profound effect on the salt dependence of the activity of the enzyme [80•].

In summarizing available data, halophilic enzymes do not exhibit specific structural properties. The most notable feature is the presence of clusters of negatively charged residues; their repulsion at low salt may account for the instability of halobacterial proteins at low salt. A similar structural motif, with the associated salt effect on stability, is also present in nonhalophilic proteins, such as ribonuclease T1 [81].

Conclusions

The significant enhancement in our understanding of extremophilic adaptation has lately come from four directions: the discovery of more extreme extremophiles; complete genome sequences; new genetic methods; and rapidly increasing numbers of high-resolution, three-dimensional structures of proteins. Regarding thermophiles, a maximum growth temperature of 113°C has been the latest news for the Guinness Book of Records [82]; apart from this minor sensation, successful expeditions to the abyssal depth of Challenger Deep brought thousands of microorganisms to the surface, although very few of them were barophilic [2,3•]. Time will show whether any common properties exist within the broad barotolerant microbial population and its nucleic acids and proteins. The second direction has opened new horizons into phylogenetics, both unveiling incongruities of the phylogenetic tree, from its root to its major branchings, and generating new scenarios that will soon lead to a boom of experiments both in genetics and in biocomputing. The third point, that is, new techniques, such as directed evolution, will become essential, because simulating evolution needs fast selection methods, which are becoming increasingly available. Finally, X-ray analysis has become a common tool in many laboratories, so the number of high-resolution structures is growing at an unexpected pace. It may sooner or later reach the level at which the sample size becomes sufficiently large for significant statistical analyses. With the increase in information from the regime of thermophiles, a limited number of strategies of stabilization has emerged, which may be used for careful predictions, as well as in industrial applications. A recent illustration is the cumulative mutation approach underlying the engineering of a thermolysine-like protease that combines low temperature activity with thermal stability beyond the boiling point of water [83]. In this context, the potential of chemistry both to expand the protein alphabet by including noncanonical amino acids or to increase chelating affinity for extrinsically stabilizing ligands promises to be a gold mine for future protein chemistry and bioengineering studies [84–87].

Restricting ourselves to the canonical amino acids, thermostability has been shown to be the cumulative effect of packing efficiency (mainly through van der Waals' interactions), networks of ion pairs and/or hydrogen bonds (including α-helix stabilization), the reduction of conformational strain (loop stabilization) and resistance to chemical modification. Comparing mesophilic and thermophilic enzymes, the increase in conformational rigidity suggests that evolutionary adaptation tends to maintain corresponding states with respect to conformational flexibility, that way optimizing biological function under specific physiological conditions. In halophiles, the correspondence is assumed to refer to the state of hydration of the protein. The complexity of multicomponent thermodynamics does not, however, allow clear-cut adaptive strategies to be defined. The same holds true for barophiles, acidophiles and alkalophiles.

Obviously, extremophilic adaptation refers to proteins in their native state, as well as in their nascent state. Since misfolding is expected to be an important side reaction under extreme conditions, the discovery of molecular chaperones in hyperthermophiles did not come as a surprise. Given their complex assembly structure, however, the determination of their spatial structure at high resolution has been a most amazing success. The elucidation of their function as a cage or an excluded volume component might be the subject of a future article in *Current Opinion in Structural Biology* dealing, for example, with hyperthermophilic heat-shock proteins.

Acknowledgment

This review is dedicated to Professor Alfred Schellenberger on the occasion of his 70th birthday.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Jaenicke R: **Protein stability and molecular adaptation to extreme conditions.** *Eur J Biochem* 1991, **202**:715-728.

2. Takami H, Akira I, Fuji F, Horikoshi K: **Microbial flora in the deepest sea mud of the Mariana trench.** *FEMS Microbiol Lett* 1997, **152**:279-285.

3. • Kato C, Li L, Nogi Y, Nakamura Y, Tamaoka J, Horikoshi K: **Extremely barophilic bacteria isolated from the Mariana trench, Challenger Deep, at a depth of 11,000 meters.** *Appl Environ Microbiol* 1998, **64**:1510-1513.

Obligatory barophilic proteobacteria (genus *Moritella*) with optimal growth at more than 70 MPa and no growth at less than 50 MPa were isolated from a sample taken from the Mariana Trench at a depth of 10,898 m.

4. Jaenicke R: **What ultrastable globular proteins teach us about protein stability.** *Biochemistry* (*Moscow)* 1998, **63**:312-321.

5. Bernhardt G, Lüdemann H-D, Jaenicke R, König H, Stetter KO: **Biomolecules are unstable under 'Black smoker' conditions.** *Naturwissenschaften* 1984, **71**:583-586.

6. Leibrock E, Bayer P, Lüdemann H-D: **Non-enzymatic hydrolysis of ATP at high temperatures and high pressures.** *Biophys Chem* 1995, **54**:175-180.

7. Carpenter JF, Clegg JS, Crowe JH, Somero GN (Eds): **Compatible solutes and macromolecular stability.** *Cryobiology* 1993, **30**:201-241.

8. Dill KA: **Dominant forces in protein folding.** *Biochemistry* 1990, **29**:7133-7155.

9. Matthews BW: **Structural and genetic analysis of the folding and function of T4 lysozyme.** *FASEB J* 1996, **10**:35-41.

10. Jaenicke R: **Stability and folding of domain proteins.** *Progr Biophys Mol Biol* 1999, in press.

11. Jaenicke R: **Protein folding and association: *in vitro* studies for self-organization and targeting in the cell.** *Curr Top Cell Reg* 1997, **34**:213-334.

12. • Grättinger M, Dankesreiter A, Schurig H, Jaenicke R: **Recombinant PGK from the hyperthermophilic bacterium *Thermotoga maritima*: catalytic, spectral and thermodynamic properties.** *J Mol Biol* 1998, **280**:525-533.

As derived from the temperature dependence of $\Delta G_{N\rightarrow U}$, optimum stability is observed at 30°C. The extrapolated temperatures of cold and heat denaturation are −10 and +85°C, respectively. Comparing the free energy of stabilization of the *T. maritima* enzyme with its counterpart from yeast, ΔΔG amounts to 85 kJ/mol.

13. Pfeil W: *Protein Stability and Folding: a Collection of Thermodynamic*
• *Data.* Heidelberg: Springer Berlin; 1998.
This book makes the vast amount of data in the literature from 1964 to 1996 accessible to energy calculations of natural and modified proteins, tabulating Gibbs energy changes, enthalpy changes and heat capacity changes.

14. Franks F: **Protein destabilization at low temperatures.** *Adv Protein Chem* 1995, **46**:105-139.

15. Graziano G, Catanzano F, Riccio A, Barone G: **A reassessment of**
• **the molecular origin of cold denaturation.** *J Biochem* 1997, **122**:395-401.
Based on Ikegamis' statistical thermodynamic model of the structural fluctuations of proteins, Gibbs energy changes associated with the transfer of *N*-alkyl-amides to water were used to devise a model of protein stability. The results indicate that the amphiphilic nature of the polypeptide chain and the presence of water are fundamental to cold denaturation.

16. Jaenicke R: **GAPDH from *Thermotoga maritima*: strategies of protein stabilization.** *FEMS Microbiol Rev* 1996, **18**:215-224.

17. Jaenicke R, Schurig H, Beaucamp N, Ostendorp R: **Structure and stability of hyperstable proteins: glycolytic enzymes from hyperthermophilic bacterium *Thermotoga maritima.*** *Adv Protein Chem* 1996, **48**:181-269.

18. Pappenberger G, Schurig H, Jaenicke R: **Disruption of an ionic**
• **network leads to accelerated thermal denaturation of GAPDH from the hyperthermophilic bacterium *Thermotoga maritima.*** *J Mol Biol* 1997, **274**:676-683.
Replacing the central arginine residue of an ionic network between the N and C-terminal helices (αB and α3) by alanine or asparagine causes a shift of 7°C in the temperature of half-denaturation. The decrease in the free energy of activation $\Delta G^{\neq}$ for thermal unfolding amounts to 4 kJ/mol (at 100°C); at physiological temperature, this yields a strong increase of the half-life of the enzyme.

19. Welker C, Böhm G, Schurig H, Jaenicke R: **Cloning, overexpression, purification and physicochemical characterization of a cold shock protein homolog from *Thermotoga maritima.*** *Protein Sci* 1998, in press.

20. Perl D, Welker C, Schindler T, Schröder K, Marahiel MA, Jaenicke R,
• Schmid FX: **Conservation of rapid two-state folding in mesophilic, thermophilic, and hyperthermophilic cold shock proteins.** *Nat Struct Biol* 1998, **5**:229-235.
Cold-shock proteins from *B. subtilis*, *B. caldolyticus* and *T. maritima* differ significantly in both their sequences and thermal stabilities ($\Delta G_{N\rightarrow U} = 11$, 20 and 26 kJ/mol, respectively), but show similar, extremely fast folding kinetics ($k_{U\rightarrow N} = 690$, 1370, 565 s^{-1}). The lack of correlation between the thermodynamic stability and folding rate indicates that proteins with a more stable folded state do not necessarily fold faster, as suggested from lattice models.

21. Beaucamp N, Schurig H, Jaenicke R: **The PKG-TIM fusion protein from *Thermotoga maritima* and its constituent parts are intrinsically stable and fold independently.** *Biol Chem* 1997, **378**:679-685.

22. Beaucamp N, Hofmann A, Kellerer B, Jaenicke R: **Dissection of the**
• **gene of the bifunctional PGK-TIM fusion protein from the hyperthermophilic bacterium *Thermotoga maritima*: design and characterization of the separate TIM.** *Protein Sci* 1997, **6**:2159-2165.
The structural properties of recombinant separate *T. maritima* triosephosphate isomerase (TmTIM) resemble those of its mesophilic homologs. Its catalytic activity at 80°C comes close to that of higher vertebrates at their physiological temperature, in accordance with the idea that, under *in vivo* conditions, enzymes occupy 'corresponding states'. Within the tetrameric phosphoglycerate kinase (PGK)–TIM fusion protein, the intrinsic thermal stability of TmTIM is enhanced ($\Delta T_m = 8$°C); the same is true for the catalytic efficiency.

23. Zaiss K, Schurig H, Jaenicke R: **Differential scanning calorimetry of PGK from the hyperthermophilic bacterium *Thermotoga maritima.*** In *Biocalorimetry: Application of Calorimetry in Biological Sciences.* Edited by Ladbury J, Chowdhry BZ. Sussex: John Wiley; 1998:283-293.

24. Trent JD, Kagawa HK, Yaoi T, Olle E, Zaluzec NJ: **Chaperonin filaments: the archaeal cytoskeleton?** *Proc Natl Acad Sci USA* 1997, **94**:5383-5388.

25. Minuth T, Frey G, Lindner P, Jaenicke R, Stetter KO: **Recombinant homo- and hetero-oligomers of an ultrastable chaperonin from *Pyrodictium occultum* show chaperone activity *in vitro.*** *Eur J Biochem* 1998, in press.

26. Klumpp M, Baumeister W, Essen L-O: **Structure of the substrate**
• **binding domain of the thermosome, an archaeal group II chaperonin.** *Cell* 1997, **91**:263-270.
Based on the 2.3 Å crystal structure of the (Ser214–Asn365) substrate-binding domain at the α subunit from the *T. acidophilum* thermosome and molecular modeling of the hexadecameric complex from the tetradecameric GroEL structure, a structural model of the intact thermosome was presented; the model is consistent with cryoelectron microscopy data.

27. Ditzel L, Löwe J, Stock D, Stetter KO, Huber H, Huber R, Steinbacher
•• S: **Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT.** *Cell* 1998, **93**:125-138.
The determination of the structure of the complete hexadecameric thermosome from *T. acidophilum* at 2.6 Å resolution confirms its topological homology with the GroEL system, with a novel type of inter-ring contact. Parts of the apical domain form a lid, substituting for the absent GroES-like chaperonin 10. The central cavity has a polar surface that is involved in folding assistance. The active site geometry allows a mechanism for ATP hydrolysis to be proposed.

28. Makhatadse GI, Privalov PL: **Energetics of protein structure.** *Adv Protein Chem* 1996, **47**:307-425.

29. Pace CN, Shirley BA, McNutt M, Gajiwala K: **Forces contributing to the conformational stability of proteins.** *FASEB J* 1996, **10**:75-83.

30. Jaenicke R: **Folding and association of proteins.** *Progr Biophys Mol Biol* 1987, **49**:117-237.

31. Colacino F, Crichton RR: **Enzyme thermostabilization: the state of the art.** *Biotechnol Genet Eng Rev* 1997, **14**:211-277.

32. Borders CL Jr, Broadwater JA, Bekeny PA, Salmon JE, Lee AS, Elridge AM, Pett VB: **A structural role for arginine in proteins: multiple H-bonds to backbone carbonyl oxygens.** *Protein Sci* 1994, **3**:541-548.

33. Stigter D, Dill KA: **Charge effects on folded and unfolded proteins.** *Biochemistry* 1990, **29**:1262-1271.

34. Teeter MM: **The water structure surrounding proteins.** In *Protein Folding.* Edited by Gierasch L, King J. Washington DC: AAA; 1990:44-54.

35. Leikin S, Parsegian VA: **Temperature-induced complementarity as a mechanism for biomolecular assembly.** *Proteins* 1994, **19**:73-76.

36. Matthews BW: **Studies on protein stability with T4 lysozyme.** *Adv Protein Chem* 1995, **46**:249-278.

37. Vita C, Fontana A, Jaenicke R: **Folding of thermolysin fragments: hydrodynamic properties of isolated domains and subdomains.** *Eur J Biochem* 1989, **183**:513-518.

38. Deckert G, Warren PV, Gaasterland T, Young WG, Lenox AL, Graham
•• DE, Overbeek R, Snead MA, Keller M, Aujay M *et al.*: **The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus.*** *Nature* 1998, **392**:353-358.
According to 16S rRNA analysis, *A. aeolius* is one of the earliest diverging bacteria known. With 1,551,335 base pairs, its genome is only one-third the size of the *E. coli* genome. It contains 93% protein coding regions, so that most of the genes are expected to be related to known open reading frames characterized elsewhere. In spite of its hyperthermophilic nature, only a few specific indications of thermophilicity are apparent from the genome.

39. Schurig H, Beaucamp N, Ostendorp R, Jaenicke R, Adler E, Knowles J: **PGK and TIM from the hyperthermophilic bacterium *Thermotoga maritima* form a covalent bifunctional enzyme complex.** *EMBO J* 1995, **14**:442-451.

40. Gabelsberger J, Liebl W, Schleifer KH: **Purification and properties of recombinant β-glucosidase of the hyperthermophilic bacterium *Thermotoga maritima.*** *Appl Microbiol Biotechnol* 1993, **40**:46-52.

41. Kengen SW, Luesink EJ, Stams AJ, Zehnder AJ: **Purification and characterization of an extremely thermostable β-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus.*** *Eur J Biochem* 1993, **213**:305-312.

42. Schurig H, Rutkat K, Rachel R, Jaenicke R: **Octameric enolase from the hyperthermophilic bacterium *Thermotoga maritima*: purification, characterization and image processing.** *Protein Sci* 1995, **4**:228-236.

43. Hess D, Krüger K, Knappik A, Halm P, Hensel R: **Dimeric PGKs from hyperthermophilic archaea. Cloning, sequencing and expression of the PGK gene of *Pyrococcus woesei* in *E. coli* and characterization of the protein. Structural and functional comparison with the PGK of *Methanothermus fervidus.*** *Eur J Biochem* 1996, **233**:227-237.

44. Kohlhoff M, Dahm A, Hensel R: **Tetrameric TIM from hyperthermophilic archaea.** *FEBS Lett* 1996, **383**:245-250.

45. Dams T, Böhm G, Auerbach G, Bader G, Schurig H, Jaenicke R: **Homo-dimeric dihydrofolate reductase from *Thermotoga maritima* shows extreme intrinsic stability.** *Biol Chem* 1998, **379**:367-371.

46. Aguilar CF, Sanderson I, Moracci M, Ciaramella M, Nucci R, Rossi M, Pearl LH: **Crystal structure of the β-galactosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*: resilience as a key factor in thermostability.** *J Mol Biol* 1997, **271**:789-802.

47. Sterner R, Kleemann GR, Szadkowski H, Lustig A, Hennig M, Kirschner K: **PRAI from *Thermotoga maritima* is an extremely stable and active homo-dimer.** *Protein Sci* 1996, **5**:2000-2008.

48. •• Hennig M, Sterner R, Kirschner K, Jansonius JN: **Crystal structure at 2.0 Å resolution of PRAI from the hyperthermophile *Thermotoga maritima*: possible determinants of protein stability.** *Biochemistry* 1997, **36**:6009-6016.
In contrast to most mesophilic PRAIs, hyperthermophilic TmPRAI is a homodimer (as a *caveat*, in *E. coli*, the standard mesophile, PRAI is part of a bifunctional enzyme). In spite of the increased overall rigidity of TmPRAI, its activity at 80°C is about 35-fold higher compared to its mesophilic counterpart at 37°C, this way successfully competing with the spontaneous hydrolysis of the substrate at physiological temperature. Stability increments come from multiple hydrophobic interactions and additional ion pairs. Their elimination by protein engineering leads to the fully active monomer exhibiting drastically reduced kinetic stability (R Thoma, K Kirschner, personal communication).

49. Villeret V, Clantin B, Tricot C, Legrain C, Roovers M, Stalon V, Glansdorff N, van Beeumen J: **The crystal structure of *Pyrococcus furiosus* ornithine carbamoyltransferase reveals a key role for oligomerization in enzyme stability at extremely high temperatures.** *Proc Natl Acad Sci USA* 1998, **95**:2801-2806.

50. Schlunegger MP, Bennett MJ, Eisenberg D: **Three-dimensional domain swapping.** *Adv Protein Chem* 1997, **50**:61-122.

51. • Auerbach G, Huber R, Grättinger M, Zaiss K, Schurig H, Jaenicke R, Jacob U: **Closed structure of PGK from *Thermotoga maritima* reveals the catalytic mechanism and determinants of thermal stability.** *Structure* 1997, **5**:1475-1483.
This 2.0 Å crystal structure determination adds to a series of four known phosphoglycerate kinase structures from mesophiles and moderate thermophiles, allowing the structural determinants of thermostability to be analyzed. In addition, the abortive ternary complex (with the substrate and AMP–adenyl-imino diphosphate) gives a detailed picture of the hinge bending 'phosphoryl gripper' mechanism. Thermophilic characteristics include increased rigidity as a consequence of enhanced intramolecular interactions involving ion pairs, α helices and loop regions.

52. Auerbach G, Jacob U, Grättinger M, Zaiss K, Schurig H, Jaenicke R: **PGK from *Thermotoga maritima*.** *Biol Chem* 1997, **378**:327-329.

53. •• Auerbach G, Ostendorp R, Prade L, Korndörfer I, Dams T, Huber R, Jaenicke R: **LDH from the hyperthermophilic bacterium *Thermotoga maritima*: the crystal structure at 2.1 Å resolution reveals strategies for intrinsic stabilization.** *Structure* 1998, **6**:769-781.
Five lactate dehydrogenases (LDH) from moderate thermophiles and mesophiles (prokaryotes and eukaryotes) were compared with the structure of the quaternary complex of TmLDH (with NADH, the allosteric activator fructose 1,6-bisphosphate and the substrate analog oxamate as ligands). The three-dimensional folds of all five enzymes are very similar, allowing a correlation between structural parameters and thermal stability to be established: 7 out of 18 arginine residues are engaged in ion pairs and four form hydrogen bonds to mainchain carbonyl groups; the number of ion pairs correlates with the growth temperature; hydrogen bonds do not make a significant contribution; there is a decrease in hydrophobic surface area and an increase in helicity (including an additional 'thermo helix'), but no change in the amount of surface area buried in subunit contacts; TmLDH has fewer and smaller cavities than its less thermostable counterparts. The high cysteine content is not involved in metal binding; the relevance for stability of a four cysteine cluster is still to be explained.

54. Dams T, Ostendorp R, Ott M, Rutkat K, Jaenicke R: **Tetrameric and octameric LDH from the hyperthermophilic bacterium *Thermotoga maritima*: structure and stability of the two active forms.** *Eur J Biochem* 1996, **240**:274-279.

55. Jaenicke R: **Oligomeric proteins.** In *Molecular Chaperones in the Life Cycle of Proteins*. Edited by Fink AL, Goto Y. New York: Marcel Dekker; 1998:35-70.

56. Koonin EV, Tatusov RL, Galperin ML: **Beyond complete genomes: from sequence to structure and function.** *Curr Opin Struct Biol* 1998, **8**:355-363.

57. •• Woese C: **The universal ancestor.** *Proc Natl Acad Sci USA* 1998, **95**:6854-6859.
In analogy to physical annealing, a 'genetic annealing model' of evolution is presented. In the beginning, simple cellular entities evolved with high rates of mutation and lateral gene transfer. Later, both rates dropped and finally approached the evolutionary dynamics existing today; organismic lineages still did not exist. The universal phylogenetic tree emerges as its peripheral branches develop. Thus, the universal ancestor is a diverse community of cells evolving as biological unit; over time, it was refined into a decreasing number of cell types, finally leading to the present 'kingdoms of life'.

58. Miller SL, Lazcano A: **The origin of life — did it occur at high temperatures?** *J Mol Evol* 1995, **41**:689-692.

59. • Doolittle RF: **Microbial genomes opened up.** *Nature* 1998, **392**:339-342.
Evidence from completely sequenced genomes seems to indicate that each gene has its own history due to the combined effects of mutations and lateral gene transfers. Available genome sequences (twelve bacteria and one yeast) are discussed in the context of phylogenetic trees, minimal gene content, the common ancestor and evolutionary mechanisms.

60. Catanzano F, Graziano G, Capasso S, Barone G: **Thermodynamic analysis of the effect of selective monodeamidation at Asp67 in ribonuclease A.** *Protein Sci* 1997, **6**:1682-1693.

61. Böhm G, Jaenicke R: **On the relevance of sequence statistics for the properties of extremophilic proteins.** *Int J Peptide Prot Res* 1994, **43**:97-106.

62. Russell RJM, Taylor GL: **Engineering thermostability: lessons from thermophilic proteins.** *Curr Opin Biotechnol* 1995, **6**:370-374.

63. Querol E, Perez-Pons JA, Mozo-Villarias A: **Analysis of protein conformational characteristics related to thermostability.** *Protein Eng* 1996, **9**:265-271.

64. Vieille C, Zeikus J: **Thermoenzymes: identifying molecular determinants of protein structural and functional stability.** *Trends Biotechnol* 1996, **14**:183-191.

65. Vogt G, Argos P: **Protein thermal stability: hydrogen bonds or internal packing?** *Fold Des* 1997, **2**:S40-S46.

66. •• Malakauskas SM, Mayo SL: **Design, structure and stability of a hyperthermophilic protein variant.** *Nat Struct Biol* 1998, **5**:470-475.
The design algorithm models several proposed stability increments simultaneously, using empirical potential functions and a combinatorial optimization procedure based on the dead-end elimination theorem. The designed sevenfold mutant shows an increased stability (18 kJ/mol) and enhanced thermostability (> 99°C). This method may potentially be applied to the stabilization of other unrelated protein folds.

67. • Gerday C, Aittaleb M, Arpigny JL, Baise E, Chessa J-P, Garsoux G, Petrecu I, Feller G: **Psychrophilic enzymes: a thermodynamic challenge.** *Biochim Biophys Acta* 1997, **1342**:119-131.
Although cold is the most widespread challenge in the biosphere, very little is known about cold adapted and psychrophilic proteins. The authors give an overview of 'cold enzymes' from Antarctic bacteria, focusing on α-amylase, β-lactamase, lipase and subtilisin.

68. • Alvarez M, Zeelen JP, Mainfroid V, Rentier-Delrue F, Martial JA, Wyns L, Wierenga RK, Maes D: **TIM of the psychrophilic bacterium *Vibrio marinus*.** *J Biol Chem* 1998, **273**:2199-2206.
Comparing the crystal structures of *V. marinus* TIM (T_{opt} = 15°C) and *E. coli* TIM, the root mean square differences were found to be less than 0.6 Å, in accordance with the high sequence identity (66%). No conclusions can be drawn from the structural differences, except for residue Ala238 in *Vibrio* TIM; altering this to serine (present in all other TIMs) shifts the stability toward the mesophilic enzyme.

69. Wallon G, Lovett ST, Magyar C, Svingor A, Szilagyi A, Závodszky P, Ringe D, Petsko GA: **Sequence and homology model of IPMDH from the psychrophilic bacterium *Vibrio sp.15* suggest reasons for thermal instability.** *Protein Eng* 1997, **10**:665-672.

70. Gerike U, Danson MJ, Russell NJ, Hough DW: **Sequencing and expression of the gene encoding a cold-active citrate synthase from an Antarctic bacterium, strain DS2-3R.** *Eur J Biochem* 1997, **248**:49-57.

71. Querol E, Perez-Pons JA, Mozo-Villarias A: **Analysis of protein conformational characteristics related to thermostability.** *Protein Eng* 1996, **9**:265-271.

72. Wrba A, Schwaiger A, Schultes V, Jaenicke R, Závodszky P: **GAPDH from the extreme thermophilic eubacterium *Thermotoga maritima*.** *Biochemistry* 1990, **29**:7585-7592.

73. Závodszky P, Kardos J, Svingor A, Petsko GA: **Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins.** *Proc Natl Acad Sci USA* 1998, **95**:7406-7411.

74. Dym O, Mevarech M, Sussman JL: **Structural features that stabilize halophilic MDH from an archaebacterium.** *Science* 1995, **267**:1344-1346.

75. Eisenberg H: **Life in unusual environments: progress in understanding the structure and function of enzymes from extreme halophilic bacteria.** *Arch Biochem Biophys* 1995, **318**:1-5.

76. Frolow F, Harel M, Sussman JL, Mevarech M, Shoham M: **Insights into protein adaptation to a saturated salt environment from the crystal structure of a halophilic 2Fe–2S ferredoxin.** *Nat Struct Biol* 1996, **3**:452-458.

77. Madern D, Pfister C, Zacchai G: **Mutation at a single acidic amino acid enhances the halophilic behaviour of malate dehydrogenase from *Haloarcula marismortui* in physiological salts.** *Eur J Biochem* 1995, **230**:1088-1095.

78. • Britton KL, Stillman TJ, Yip KSP, Forterre P, Engel PC, Rice DW: **Insights into the molecular basis of salt tolerance from the study of glutamate dehydrogenase from *Halobacterium salinarium*.** *J Biol Chem* 1998, **273**:9023-9030.
Comparative modeling of HsGluDH based on GluDH from *Pyrococcus furiosus* (47% sequence identity) with other related GluDH structures as a guide for small insertions or deletions, was used to provide insight into the molecular basis of halophilic adaptation. Features that may contribute to the stability of HsGluDH in high salt include decoration of the surface with acidic residues and a reduction in the number of surface-exposed lysine residues.

79. • Pieper U, Kapadia G, Mevarech M, Herzberg O: **Structural features of halophilicity derived from the crystal structure of DHFR from the Dead Sea halophilic archaeon, *Haloferax volcanii*.** *Structure* 1998, **6**:75-88.
The crystal structure of dihydrofolate reductase (DHFR) at 2.6 Å resolution reveals the well-known overall fold of other DHFRs. The moderately halophilic nature of the enzyme, which retains its native state at monovalent salt concentrations as low as 0.5 M, is characterized by a lack of the structural features that are known to occur in extremely halophilic proteins. Differences in the percentage of carboxylate groups are statistically insignificant; similarly, no significant increase in ion-pair interactions is detectable.

80. • Jolley KA, Russell JM, Hough DW, Danson MJ: **Site-directed mutagenesis and halophilicity of dihydrolipoamide dehydrogenase from the halophilic archaeon, *Haloferax volcanii*.** *Eur J Biochem* 1997, **248**:362-368.
A homology-modeled structure, based on the crystal structure of dihydrolipoamide dehydrogenase from *P. fluorescence*, allowed the prediction of a putative K^+-binding site at the dimer interface. This binding site seems to be involved in the stabilization of the enzyme at high salt and high temperature. Similar local effects have been reported for malate dehydrogenase from *H. marismortui* [74], supporting the view that the structural basis of halophilicity can only be understood at the level of the precise three-dimensional structure.

81. Kiefhaber T, Schmid FX, Renner M, Hinz H-J, Hahn U: **Stability of recombinant Lys25-RNase T1.** *Biochemistry* 1990, **29**:8250-8257.

82. Blöchl E, Rachel R, Burggraf S, Hafenbradl D, Jannasch HW, Stetter KO: ***Pyrolobus fumarii* represents a novel group of archaea, extending the upper temperature limit for life to 113°C.** *Extremophiles* 1997, **1**:14-21.

83. van den Burg B, Vriend G, Veltman OR, Venema G, Eijsink VGH: **Engineering an enzyme to resist boiling.** *Proc Natl Acad Sci USA* 1998, **95**:2056-2060.

84. de Filippis V, de Antoni F, Frigo M, Polverino de Loreto P, Fontana A: **Enhanced protein thermostability by Ala→Aib replacement.** *Biochemistry* 1998, **37**:1686-1696.

85. Mendel D, Cornish VW, Schultz PG: **Site-directed mutagenesis with an expanded genetic code.** *Annu Rev Biophys Biomol Struct* 1995, **24**:435-462.

86. Veltman OR, Vriend G, Berendsen HJC, van den Burg B, Venema G, Eijsink VGH: **A single Ca^{2+} binding site is crucial for the Ca^{2+}-dependent thermal stability of thermolysin-like proteases.** *Biochemistry* 1998, **37**:5312-5319.

87. Noll KM, Vargas M: **Recent advances in genetic analyses of hyperthermophilic archaea and bacteria.** *Arch Microbiol* 1997, **168**:73-80.

88. •• Russell RJM, Ferguson JMC, Hough DW, Danson MJ, Taylor GL: **The crystal structure of citrate synthetase from the hyperthermophilic archaeon *Pyrococcus furiosus* at 1.9 Å resolution.** *Biochemistry* 1997, **36**:9983-9994.
Comparing the *Pyrococcus* citrate synthetase (CS) enzyme with CS from pig and *Thermoplasma acidophilum*, all three share a similar fold and the same acid-base catalysis. Given these similarities, trends in going from mesophilic CS to hyperthermophilic CS are detectable. These include increased compactness, more intimate subunit interactions, an increase in clustered intersubunit ion pairs, the shortening of loops and a reduction in the number of thermolabile residues.

89. Macedo-Ribeiro S, Darimont B, Sterner R, Huber R: **Small structural changes account for the high thermostability of 1[4Fe–4S] ferredoxin from *Thermotoga maritima*.** *Structure* 1996, **4**:1291-1301.

90. • Pfeil W, Gesierich U, Kleemann GR, Sterner R: **Ferredoxin from the hyper-thermophile *Thermotoga maritima* is stable beyond the boiling point of water.** *J Mol Biol* 1997, **272**:591-596.
The 60 residue *T. maritima* ferredoxin contains a single [4Fe–4S] cluster that is irreversibly released upon unfolding. Since there are no further heat-absorbing processes, the reaction can be interpreted thermodynamically. Polar interactions contribute significantly to the stability. This shows its maximum at 45°C, when ΔG amounts to 39 kJ/mol. Thus, the protein is thermodynamically no more stable than average mesophilic proteins. Due to its small size and the low heat capacity change ΔC_p of unfolding (3.6 kJ/molK), the ΔG versus temperature profile is extremely shallow, thus explaining the wide range of stability between −20 and +125°C.

91. Hatanaka H, Tanimura R, Katoh S, Inagaki F: **Solution structure of ferredoxin from the thermophilic cyanobacterium *Synechococcus elongatus* and its thermostability.** *J Mol Biol* 1997, **268**:922-933.

92. Yip KSP, Stillman TJ, Britton KL, Artymiuk PJ, Baker PJ, Sedelnikova SE, Engel PC, Pasquo A, Chiaraluce R, Consalvi V *et al.*: **The structure of *Pyrococcus furiosus* GluDH reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures.** *Structure* 1995, **3**:1147-1158.

93. • Yip KSP, Britton KL, Stillman TJ, Lebbink J, de Vos WM, Robb FT, Vetriani C, Maeder D, Rice DW: **Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the GluDH family.** *Eur J Biol* 1998, **255**:336-346.
A comparison of the crystal structures of glutamate dehydrogenases (GluDHs) from *P. furiosus*, *Clostridium symbiosum* and *E. coli* gives clear evidence for the formation of extended networks of ion pairs being a significant contribution to thermal stability. Homology modeling indicates that a decrease in temperature stability correlates with increased 'fragmentation' of ion-pair networks.

94. Korndörfer I, Steipe B, Huber R, Tomschy A, Jaenicke R: **The crystal structure of holo-GAPDH from the hyperthermophilic bacterium *Thermotoga maritima* at 2.5 Å resolution.** *J Mol Biol* 1995, **246**:512-521.

95. Hennig M, Darimont B, Sterner R, Kirschner K, Jansonius JN: **2.0 Å structure of indole-3-glycerol phosphate dehydrogenase from the hyperthermophile *Sulfolobus solfataricus*: possible determinants of protein stability.** *Structure* 1995, **3**:1295-1306.

96. •• Wallon G, Kryger G, Lovett ST, Oshima T, Ringe D, Petsko GA: **Crystal structures of *E. coli* and *S. typhimurium* 3-isopropyl MDH and comparison with their thermophilic counterpart from *Thermus thermophilus*.** *J Mol Biol* 1997, **266**:1016-1031.
All three 3-isopropyl malate dehydrogenase (IPMDH) structures show a resolution < 2.1 Å. The main stabilizing features in thermophilic IPMDH are an increased number of ion pairs, additional hydrogen bonds, a proportionately larger and more hydrophobic subunit interface, shortened N and C termini, and a larger number of proline residues. Aromatic–aromatic interactions, helical capping and cavities seem to be of lesser importance. The overall flexibility (monitored by the B factors of equivalent residues) is reduced in the case of *S. typhimurium* IPMDH.

97. • Akanuma S, Yamagishi A, Tanaka N, Oshima T: **Serial increase in the thermal stability of IPMDH from *Bacillus subtilis* by experimental evolution.** *Protein Sci* 1998, **7**:678-705.
The mesophilic IPMDH from *B. subtilis* was expressed in an IPMDH-deficient mutant of *T. thermophilus*. Variants growing at temperatures higher than the stability limit of wildtype *B. subtilis* IPMDH (56°C) were selected at 61–70°C. Three point mutants, I95L, M292I and T308I, and double and triple mutants thereof were isolated and characterized. The stability of the variants was improved, showing additivity in their stabilization effects, and the triple mutant had also higher specific activity than the wildtype protein.

98. • Aoshima M, Oshima T: **Stabilization of *E. coli* IPMDH by single amino acid substitutions.** *Protein Eng* 1997, **10**:249-254.
In attempting to define the key positions responsible for the high stability of IPMDH from *Thermus thermophilus* (Tt) and *Thermus aquaticus* (Ta), sequence comparisons of all 20 known IPMDHs were carried out. As a result, a four-residue motif, which is present in the Tt and Ta sequences only, was found located between two highly conserved stretches of the sequence. Introducing this motif into *E. coli* IPMDH, one by one, enhanced thermostability, whereas introducing the whole motif caused destabilization. The three-dimensional structure allows a tentative explanation of the mechanism.