

Conformational States and Protein Stability from a Proteomic Perspective

Vesna Prošinecki, Patrícia F.N. Faísca and Cláudio M. Gomes*

Instituto Tecnologia Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal

Abstract: Protein structure and stability rely on the interplay of a large number of weak molecular interactions working in concert to assure a stable and unique native fold. Throughout evolution, different strategies have been devised to modulate protein conformational stability and enhance function and survival of proteins even under adverse conditions. The increasing number of characterized genomes and proteomes, especially those from thermophiles, provides a unique resource to study protein conformations at a wider scale. An integrated proteome-level perspective of protein conformational states in different cellular contexts is likely to contribute to a better understanding of functioning and control of biological systems. This review will address recent proteomic approaches, which allow screening and profiling proteins according to particular conformational features. We will discuss emerging methodologies that allow screening proteomes for unstructured or conformationally altered proteins, and novel approaches that profile and identify proteins within complete proteomes on the basis of their differential resistances to temperature, chemicals, or proteolysis. In particular, the profiling of proteins from thermophiles according to their thermostability will be highlighted as these studies may contribute to elicit general strategies accounting for protein stability and thermostable cellular processes.

Key Words: Protein folding, protein stability, thermostability, thermophiles, proteolysis, disordered proteins.

INTRODUCTION

Proteins fold to a particular three-dimensional structure using information which is encoded in their primary structure. The conformational stability of the protein structure results from the combined contributions of a large number of molecular interactions, which altogether account for the maintenance of the native state. Within the cell, several factors may drive the native conformation towards misfolded, partly unfolded or unstructured states. Therefore, the native state itself has a dynamic nature. In spite of the intervention of the chaperon rescue mechanisms, some of these altered conformations have an enhanced aggregation propensity and may lead to the formation of fibrils or β -amyloid structures. This suggests that the landscape of conformational states accessible to a protein in a living cell is rather vast, and that the possibility of interconversion between these states is intrinsic to the protein's ability to function (Dobson, 2000). An integrated, proteome-wide perspective of protein conformational states in different cellular contexts is expected to contribute to a better understanding of functioning and control of biological systems. This review will start by briefly addressing the molecular factors underlying the conformation and stability of a protein, and then move to the discussion of recent proteomic approaches, which allow screening and profiling proteins according to particular conformational features, such as enhanced thermal stability or intrinsic structural disorder.

PROTEIN FOLDING, STRUCTURE AND STABILITY

Protein folding is the process by which a polypeptide chain acquires a specific three-dimensional structure. The

fact that some chemically denatured proteins are able to spontaneously refold to their respective native conformations suggests that the native state is a minimum of the Gibbs free energy, and that the information required for the protein to fold is comprised within the sequence itself (Anfinsen, 1973). The highly crowded environment in the cell cytosol, with macromolecule (DNA, RNA and proteins) concentrations ranging from 300 to 400 mg/ml, makes it a very hostile place for the folding reaction to occur efficiently. In fact, the crowding increases the propensity of incompletely folded chains to aggregate (Ellis, 2003; Minton, 2005). For this reason, protein folding *in vivo* is often assisted by chaperons (Hartl, 2002). These proteins do not determine the native structure, but rather have several roles: they catalyse the folding reaction, recognize and unfold proteins which are conformationally altered or misfolded, and control the sub-cellular localization of the functional proteins.

Diversity of Protein Conformational States

Protein stability is defined by imprecise cancellation of two large effects, namely, the hydrophobic effect, favouring folding, and chain entropy, disfavouring it. Because these two effects are similar in magnitude the net stability of proteins is marginal. Indeed, the average Gibbs free energies of denaturation (ΔG_D) ranges from 20 to 60 kJ.mol⁻¹ (at 25°C), a value which is comparable to the magnitude of the weak forces that stabilise the native conformation (discussed below). The fact that protein stability relies on a minor free energy difference between the native and unfolded states implies a certain conformational flexibility of the structure. Indeed, due to sufficiently low kinetic barriers, other alternative conformations, with considerably low energy, are accessible to the protein (Uversky, 2002). One such example is the molten globule state, which is almost as compact as the native form and have a loosely packed core, while retaining

*Address correspondence to this author at the Instituto Tecnologia Química e Biológica. Universidade Nova de Lisboa., Av. República 127. 2780-756 Oeiras. Portugal; Tel: +351 214469332; Fax: +351 214411277; E-mail: gomes@itqb.unl.pt

some of its native secondary structure (Pitsyn, 1995). Consequently, each protein has a particular energetic landscape of conformations that it can adopt under physiological conditions. Ultimately this may correspond to a protein in which the most populated ensemble of structural conformations is disordered or that contains highly disordered regions (Dunker *et al.*, 2005; Uversky, 2002). This possibility impacts on the structure-function paradigm, as disordered proteins are biologically active in functions related to regulation of transcription and translation, protein phosphorylation, storage of small molecules, and regulation of the self assembly of multiprotein complexes (Dyson and Wright, 2005). Disordered proteins often fold into an ordered structure upon binding to a protein partner. That is, for example, the case of thyroid hormone and retinoid receptors, which occur as unstructured ensembles, and fold upon interacting with a nuclear receptor binding domain. The same occurs upon interaction with a ligand, such as a metal ion or a nucleic acid, as in the zinc-finger containing transcription factor TFIIIA or the translation initiation factor eIF4E (reviewed e.g., in (Dyson and Wright, 2005; Tompa, 2002). Interestingly, disordered domains may adopt distinct ordered conformations depending on the interacting partner, thus reflecting a significant functional flexibility, in agreement with the fact that disordered proteins are mainly involved in signalling and regulatory pathways. This has led to the proposal of generically coining disordered proteins as 'pliable' (Tompa, 2002). Overall, some proteins comprising intrinsically disordered domains or segments have a plethora of accessible ordered conformations that they can adopt, upon protein-ligand interaction(s). In thermodynamic terms, coupling folding to binding results in a strategy to use the binding enthalpy to pay the entropic cost to fold a disordered protein.

Molecular Determinants of Protein Stability

The dominant forces that fold proteins and stabilize the native states have long been identified (Kauzmann, 1959). These comprise residue-residue interactions and water-residue interactions, and are typically classified into the following classes: van der Waals interactions (present between any group), hydrogen bonds, salt bridges (bonds between oppositely charged residues that are sufficiently close to each other to experience electrostatic attraction), disulfide and hydrophobic forces (Chandler, 2005). The magnitudes of these forces in proteins, as determined from mutagenesis and unfolding experiments, are very low. For example, a stabilising contribution of 4-5 kJ.mol⁻¹ has been estimated per hydrogen bond (Shirley *et al.*, 1992); a single ion pair may be responsible for a 12.5-20 kJ.mol⁻¹ stabilisation (Anderson, Becktel, and Dahlquist, 1990). Short distance (<7Å) aromatic pairs interactions, such as Tyr-Tyr and Phe-Phe, contribute approximately 5.5 kJ.mol⁻¹ towards thermodynamic stabilisation (Serrano *et al.*, 1991). However, the fact that the free energy of folding is always very low (20 to 60 kJ.mol⁻¹), makes these low-magnitude forces relevant contributors to the overall stability, as the presence or absence of one, or a few, of these interactions may be enough to shift the equilibrium towards the native or unfolded state. Among these, hydrophobic interactions are widely believed to be the main driving forces behind the folding of globular proteins (Dill, 1990) and result from the incapacity of non-polar residues to

make hydrogen bonds with water molecules. The resulting water-residue interactions are thermodynamically unfavourable and drive the non-polar residues into the interior of proteins (Perutz, 1965; Sandelin, 2004) where they arrange into densely packed clusters. The contribution of the hydrophobic effect to the stability of globular proteins has been estimated from cavity-creating mutations in which a small aliphatic chain in the interior of the protein replaced a larger one with identical characteristics: this has shown that an energetic gain of ~5.5 kJ.mol⁻¹ is obtained per buried methyl group (Pace, 1992). Many proteins are also stabilised by covalent interactions which crosslink segments of secondary structure in the native state: these comprise disulfide bridges, coordination of a metal ion or attachment of an organic or organometallic cofactor (Petsko and Ringe, 2004). Secreted and periplasmic proteins, or those that are present at external surfaces such as cell walls, have an increased content of disulfide bridges. These confer an increased conformational rigidity to the protein in the oxidising conditions found outside the cell. Metal ion cross linked domains are found in zinc fingers (Auld, 2001), which occur in transcription factors and constitute the most abundant domain encoded in the human genome, or in iron-sulfur proteins, namely ferredoxins. In both examples, removal of these ions frequently results in protein unfolding or in the formation of non-native states (Leal and Gomes, 2007; Lee *et al.*, 1991). While the major contributors to protein stability are non-covalent interactions, metal ion and disulfide cross-linking interactions assume a particularly relevant role in small proteins comprising irregular domains. Although proteins frequently lack a large hydrophobic core and have minimal secondary structure; covalent interactions allow attaching and stabilising different parts of the protein. This is what happens, e.g., with scorpion toxin, allergy factor Ra5, and several protease inhibitors (Petsko and Ringe, 2004).

Proteins with Enhanced Conformational Stability

Some proteins need to be particularly stable as a result of molecular adaptation to a particular physiological condition or to a harsh environmental factor, such as high salinity, extreme pH and high temperatures. Among thermophiles, enhanced protein stability encompasses both thermodynamic and kinetic stability. While the kinetic stability depends on the energy barrier to unfolding, i.e., on the activation energy of unfolding, the thermodynamic stability is reflected in the conformational stabilities (i.e., ΔG_D), which may be up to 100 kJ.mol⁻¹ larger than those from mesophilic proteins (Jaenicke, 2000), and in the midpoint transition temperatures for unfolding (i.e., T_m), which are typically between 20-30°C above those of mesophiles (Razvi, 2006; Rees, 2001). A recent study suggested that the thermodynamic strategy leading to the higher denaturation temperatures exhibited by thermophiles relies on the elevation of the stability curve (i.e., ΔG vs T) (Razvi, 2006), rather than on broadening, or shifting it toward higher temperatures. A relevant problem in protein chemistry is that of identifying and understanding the intrinsic factors that are responsible for the functional stability exhibited by thermophilic proteins at very high temperatures. This fundamental problem of paramount importance was initially investigated by Perutz and others in the 1970s (Perutz, 1978, 1975), and since then, numerous studies have

tackled this issue. The following structure-based approaches can be loosely identified: (i) comparison of structures from thermophilic proteins with those of their mesophilic homologues and (ii) systematic structure-based sequence comparisons for a group of proteins. (iii) More recently, due to progress in genome sequence projects, it has become possible to perform large-scale comparison between genome sequences from thermophiles and mesophiles. Based on experimental and theoretical analyses, diverse stabilizing strategies have been suggested, as putative intrinsic drivers for the enhanced stability exhibited by thermophilic proteins (see e.g. (Vieille and Zeikus, 2001) for a comprehensive review). Structural properties like better core (and secondary structure) packing (Gromiha, 1999; Pack, 2005; Querol, 2006), deletion or shortening of loops and increased helical content have been typically ascribed to thermophilic proteins (Chakravarty, 2002). On the physicochemical side, it is generally claimed that thermophilic proteins have more hydrophobic residues (Britton, 1995; Chen, 1994; Sadeghi, 2006; Vanhove, 1995), a larger amount of main-chain hydrogen bonds (Sadeghi, 2006) and a higher number of proline residues. However, a recent study, which analysed 18 non-redundant families of thermophilic and mesophilic proteins, reported that these factors do not show consistent, substantial variations between mesophiles and thermophiles (Kumar, 2000b). The higher number of salt bridges among thermophilic proteins suggests that these interactions play a role in stability enhancement (Chakravarty, 2002; Frankenberg, 1999; Grimsley, 1999; Karshikoff, 2001; Kumar, 2000b; Sadeghi, 2006; Szilagy, 2000). Nevertheless, there is not a general physical mechanism able to rationalize the stability enhancement upon increasing the number of salt bridges, because the net electrostatic free energy of salt bridges can be either stabilizing or destabilizing (Bosshard, 2004). This results from the fact that energetically favorable Coulombic charge-charge interaction forming in the protein core, is opposed by the unfavorable desolvation of interacting charges – the transfer of a salt bridge from water to nonpolar environment costs $\sim 42\text{-}67 \text{ kJ.mol}^{-1}$ (Honig, 1984). Therefore, it is mostly the surface, solvent-exposed, salt bridges that effectively lead to an increase in protein stabilization (Loladze, 1999; Makhatazde, 2003; Sarakatsannis, 2005; Wang, 2000). Moreover, there is also some evidence that extended networks of salt-bridges (formed by residues that participate in more than one salt-bridge) between protein subunits are critical for achieving the superior thermostability of hyperthermostable proteins (Kumar, 2000a; Marqusee, 1994; Yip, 1998).

Intrinsically Disordered Proteins

Whereas some proteins require an organised fold and ultimately a certain structural rigidity to attain an increased stability, others are disordered under physiological conditions. These are called intrinsically disordered or unstructured proteins, and their discovery has shaken the grounds of a long-standing paradigm in structural biology, namely that the function of a protein is strictly determined by an ordered three-dimensional fold (Dyson and Wright, 2005; Tompa, 2002; Uversky, 2002). As noted previously, while some of these proteins remain permanently unstructured under physiological conditions, others become partially ordered.

This suggests that the coupled folding and binding operating mechanism is a strategy employed by unfolded proteins to achieve *in vivo* stabilization and functionality (Uversky, 2002). Indeed, these unstructured molecules are biologically functional (Plaxco, 1997), typically being involved in critical steps of the cell cycle and regulatory processes (Dyson and Wright, 2005), and mutant variants are often associated to cancer and neurodegeneration (Csizmok *et al.*, 2007). For example, that is the case of the tumor suppressor protein p53, a DNA binding transcription factor that is regulated by a complex network of proteins, and is inactivated due to missense mutations in half of all human cancers. The protein has a modular organization with disordered regions, and the lack of rigid structure combined with a low conformational stability have been hypothesized to be determinant for the regulation of its turn over and for the interaction with a multitude of partner proteins (Bell *et al.*, 2002).

The catalogue of proteins expected to comprise intrinsic disorder has been increasing, mostly due to the use of prediction algorithms that rely on the analysis of particular sequence features that lead to low compositional complexity (Vucetic *et al.*, 2005). The reason why some proteins adopt a structurally unfolded conformation seems to be encoded in the primary sequence in a relatively simple and straightforward way. A combination of low overall hydrophobicity coupled with a large net charge represents a pre-requisite for the absence of compact structure in proteins under physiological conditions. In this case, the absence of an apolar set of residues whose packing would promote folding, and the fact that upon folding, a significant electrostatic repulsion would arise from charged residues put in close proximity, make these sequence features unique and typical of disordered proteins (Uversky, 2002). Disordered proteins are frequently heat stable, in the sense that they remain in solution and do not aggregate upon thermal challenge. In contrast with the features of the highly stable proteins discussed in the previous section, this stems from the fact that these proteins lack a significant hydrophobic core. Therefore, thermal perturbation does not lead to exposure of apolar residues that would trigger aggregation.

PROTEOMIC ANALYSIS OF PROTEIN CONFORMATIONAL CHANGES

Within a cell, conformational changes in a protein may result from a functional modification, for example, as a result of an interaction with an activity modulator (such as another protein, a nucleic acid or a metal ion) or they may reflect the onset of more drastic alterations, such as misfolding, degradation or aggregation. In pure protein solutions, these alterations can be monitored using different biophysical methods such as circular dichroism, fluorescence, nuclear magnetic resonance, infrared spectroscopies, dynamic light scattering and many other approaches (Plaxco, 1996; Schmid, 2005). However, when it comes to the analysis of complex mixtures of proteins, the scenario is different. The experimental global analysis of conformational changes and assessment of protein disorder at a whole proteome scale has become recently facilitated with the development of two novel methodologies, which will be overviewed in this section.

In-Gel Detection of Protein Surface Hydrophobicity Changes

Modifications on the conformation of a protein frequently involve a variation on the surface hydrophobicity as a result of the exposure of otherwise solvent-shielded apolar groups. Chaudhuri and co-workers have established a procedure that allows monitoring changes in the exposure of surface hydrophobic domains in proteins as a result of a conformational change (Pierce *et al.*, 2006). In their study they took the advantage of the ability of BisANS (4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic acid) to interact with the protein and covalently bind to the site of interaction after being exposed to ultraviolet (UV) irradiation. This apolar fluorescent probe binds preferentially to hydrophobic patches, which become solvent accessible as a result of a moderate conformational change. Therefore, the rationale behind this approach is that the intensity of BisANS fluorescence will depend on the variation of the protein surface hydrophobicity, which can be correlated to a change in the protein conformation. The fact that this molecule can also be photolabelled to targets makes it a useful conformational probe. This approach has the advantage of being able to discriminate partially altered conformational states, as shown through validation studies performed on rhodanase (Pierce *et al.*, 2006), Fig. (1); However, it does not identify unfolded states and aggregates, as in these cases the hydrophobic regions

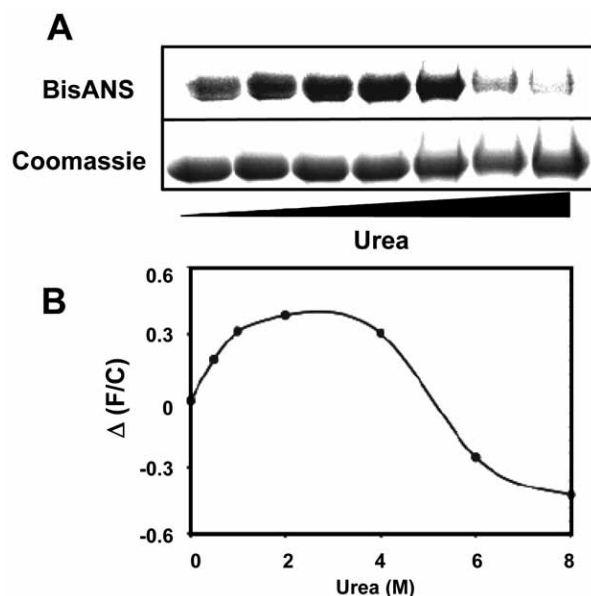


Fig. (1). In gel detection of changes in protein conformation. The known effect of urea on rhodanase surface hydrophobicity was used to validate the BisANS labelling and detection procedure. Rhodanase ($1 \text{ mg}\cdot\text{ml}^{-1}$) was incubated with different urea concentration and labelled with 0.1 nM BisANS, and analysed by SDS-PAGE. Densitometry quantitated the BisANS fluorescence (F), which was normalized for the Coomassie (C) staining (A). The dependence of fluorescence intensity $\Delta F/C$ was plotted as a function of urea concentration (B): the maximum at 2 M indicates the formation of stable intermediate complexes with increased surface hydrophobicity. For concentrations of urea above 4 M the protein is likely to be unfolded, as suggested by the dramatic decrease in BisANS incorporation. Reprinted in part with permission from *Biochemistry* (Pierce *et al.*, 2006), Copyright (2006) American Chemical Society.

necessary for the interactions are either absent or unavailable. The ultimate goal of this methodology is its application at a proteomic scale to screen for conformational changes occurring *in vivo* as result of a pathological process or a stress factor. This assay was used to analyse the effect of oxidative stress on a complex mixture of cytosolic proteins originating from skeletal muscle. The proteome was exposed to increasing levels of oxidative stress using either *in vitro* metal-catalysed oxidation, or *in vivo* denervation, a process that induces oxidative stress. Subsequent photoincorporation of BisANS and resolution of the proteome by 2DE led to the observation that, in both circumstances, a significant variation of the surface hydrophobicity was noted in two major proteins from skeletal muscle proteins, creatine kinase and glyceraldehyde-3-phosphate dehydrogenase. These two proteins are among those that undergo a conformational change upon oxidative stress, serving therefore to illustrate the validity of the approach (Pierce *et al.*, 2006). While fluorophore labelling approaches have the merit of allowing an evaluation of the conformational change of a single protein within a complex proteome, some aspects should be noted. For example, this approach fails to detect conformational states lacking a substantial amount of the necessary superficial hydrophobic reporter regions, such as aggregates or denatured states. Additionally, the fact that the extent of labelling differs among proteins reflects different propensities towards the probe, which may result in misleading results. Considering these aspects, some of which are intrinsic to the nature of the fluorophore-protein interaction, this methodology elegantly allows a direct analysis of a conformational change in a particular set of proteins, in the background of a complex proteome.

Electrophoretic Detection of Intrinsically Disordered Proteins

The initial discovery of structurally disordered proteins resulted from studies on isolated expressed proteins, whose properties resembled those of the denatured states of globular proteins with respect to structural content, conformational flexibility, and hydrodynamic radius (Tompa, 2002). A series of subsequent proteomic approaches based on the use of acid, organic solvents and heat treatment to selectively enrich extracts for disordered proteins have been established. These strategies and respective results, which have been addressed in a recent review (Csizmek *et al.*, 2007), rely on the principle that under these harsh conditions globular proteins will precipitate whereas intrinsically disordered protein will remain in solution. The combined use of 2DE and MS has allowed the identification of a series of putative disordered proteins. These approaches constitute valid starting points, but the fact that some apparent positives may comprise proteins with an intrinsic high stability or proteins that do not necessarily precipitate upon denaturation and have the ability to refold upon cooling, can not be fully discarded.

More recently, a novel technique was established by Tompa and colleagues that enables detection of proteins with intrinsic disorder (Csizmek *et al.*, 2006). This methodology provides direct evidence for protein structural disorder, thus circumventing some of the above-mentioned limitations of the previous strategies. The outlined methodology results from the combination of native and denaturing electrophore-

sis of heat-treated samples. The rationale of the approach is that the heat treatment step will simplify the initial protein mixture by removing globular proteins that precipitate under these conditions. The subsequent first dimension native electrophoresis will resolve proteins according to their charge/mass ratios. In the second dimension in 8 M urea, disordered proteins that are unaffected by the denaturant will migrate the same as in the first dimension, and end along the diagonal of the gel; as a result of an increased friction coefficient heat resistant proteins will migrate less and stay above the diagonal. This methodology was tested and validated using known disordered proteins, and control globular proteins, Fig. (2), and later expanded to the analysis of cellular extracts of *Saccharomyces cerevisiae* and *Escherichia coli*. Again, some limitations must be taken into consideration, namely the fact that the native first dimension has a poor resolving power and that basic proteins are lost under standard conditions (Csizmok *et al.*, 2006). Overall, this methodology is an important step forward towards the identification of the “disorderome” (Csizmok *et al.*, 2006, 2007).

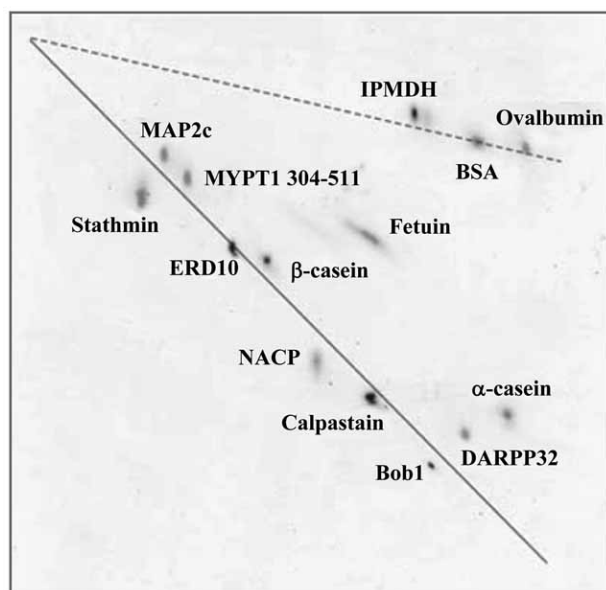


Fig. (2). Electrophoretic detection of intrinsically disordered proteins. A mixture of unstructured and globular proteins (1 μ g each) was resolved on a native 7.5% gel in the first dimension and on a 7.5% denaturing 8 M urea gel in the second dimension. For experimental details see (Csizmok *et al.*, 2007; Csizmok *et al.*, 2006). Proteins migrating along the solid diagonal line those with disordered structure. Globular proteins: fetuin, IPMDH, BSA, ovalbumin; Disordered proteins: stathmin, MAP2c, MYPT1 304-511, ERD10, β -casein, NACP, CSD1, Bob1, DARPP32, α -casein. Reprinted from *Current Protein and Peptide Science* (Csizmok *et al.*, 2007), with permission from Bentham Science Publishers Ltd.

PROFILING HYPERSTABLE PROTEINS AT A PROTEOMIC SCALE

Proteins with a very high conformational stability can be labelled as hyperstable. Profiling these proteins at a proteomic level is of interest within numerous perspectives. Hyperstable proteins can be used, for example, in processes relying on biological catalysis (Cowan, 1992; Li *et al.*, 2005). Also,

the profiling of proteins of a given pathogen according to their stabilities is a potentially valuable approach as this could allow the identification of potential therapeutic targets. Hyperstable proteins also constitute excellent working models in more fundamental studies aiming at the elucidation of the molecular determinants of the stability of a particular fold (Fitter, 2005; Perl and Schmid, 2001), or as targets for protein structure determination.

Proteins with Extreme Thermal and Chemical Resistance

Recently, Gomes and co-workers, have established a novel approach aimed at profiling a proteome for its most intrinsically stable proteins (Prosinecki *et al.*, 2006). The methodology was implemented on the proteome of a hyperthermophile, as the high optimal growth temperatures of these organisms (>80°C) make them valuable sources of proteins with intrinsically high stability. Under such extreme conditions, the stability of the proteome must rely on the combined effect of several factors, in which intrinsic molecular determinants are pivotal. However, the methodology can be applied to any proteome. The outlined procedure consists on thermal and chemical perturbation of the proteome under study as a function of time. The impact of the destabilization during the perturbation time is assessed by periodic sampling, fractionation and electrophoretic analysis. A comparison of the chromatographic and electrophoretic profiles allows the identification of a set of surviving proteins, which are subsequently identified by mass spectrometry methods. Whenever the biological activity of the identified proteins is known, activity assays are performed after the incubation and along the perturbation period to verify if the selected protein retains biological function, a further indication of structural maintenance.

This procedure was employed to profile the cytosolic proteome of the thermoacidophile *Sulfurisphaera* ($T_{opt} = 85^\circ\text{C}$). The identification of hyperstable proteins in such a drastic thermophilic background required extreme perturbation protocols relying in extensive thermal perturbation (up to 96h at 90°C) or combined thermal and chemical perturbation (up to 48h at 90°C and 1M GuHCl). The cytosolic fractions perturbed during 96h at 90°C were resolved by two-dimensional electrophoresis, Fig. (3): it became clear that the number of proteins decreases substantially and that even after such extensive incubation, there is still a large number of proteins which can be detected. MALDI peptide mass fingerprinting analysis led to the identification of proteins comprised in three distinct functional categories: (1) cellular processes and detoxification, (2) DNA binding, translation and protein modification, and (3) energy metabolism. The proteome perturbation method was validated by verifying the extent to which the subset of proteins obtained after the perturbation corresponds to biologically active molecules. For this purpose, two of the identified proteins were selected from the perturbed subset for activity assays, namely peroxiredoxin (Prx) and superoxide dismutase (SOD). The activity of these enzymes was measured in the native and perturbed proteome. Both activities were not only present but a specific activity of the enzymes was also observed, Fig. (4). This observation is compatible with a significant enrichment of these two enzymes in the thermally treated extract, as a result of their enhanced thermostability.

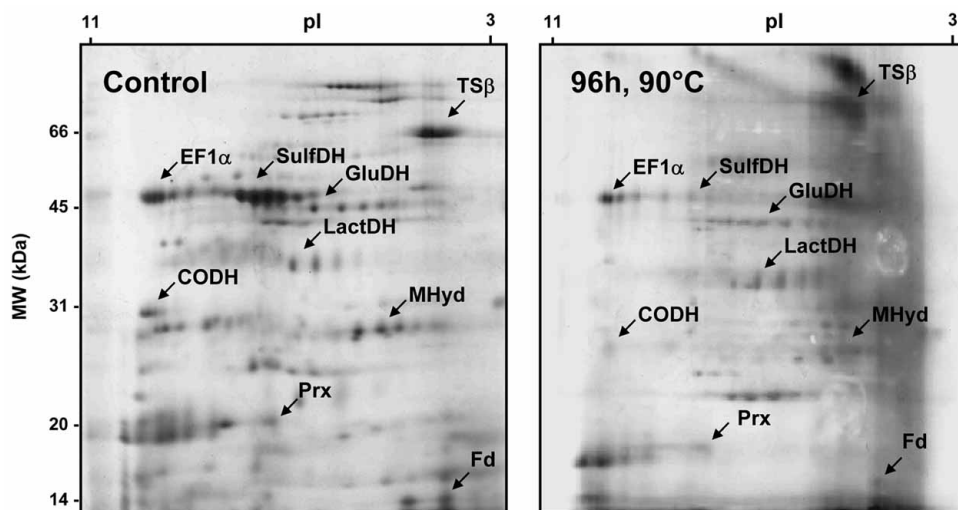


Fig. (3). Profiling proteins according to high thermal stability. A 2-DE Silver stained gel (12.5% SDS-PAGE, 500 μ g protein), corresponding to the cytosolic proteome of *Sulfurisphaera sp.*, before (A) and after (B) 96 h incubation at 90 $^{\circ}$ C. Arrows indicate some of the identified proteins: EF1 α , elongation factor 1-alpha; SulfDH, sulphide dehydrogenase; TS β , thermosome β -subunit; GluDH, Glutamate dehydrogenase; LactDH, Lactate dehydrogenase; MHyd, metal-dependent hydrolase; Prx, peroxiredoxin; Fd, Ferredoxin [3Fe4S][4Fe4S]. See (Prosinecki *et al.*, 2006) for further details.

Proteins Resistant to Proteolysis

Recently, Marqusee and co-workers described an investigation on proteolytic susceptibility/resistance of proteins from a proteomics point of view using the *E. coli* proteome as model system (Park *et al.*, 2007). The rationale of the approach was that proteolytic susceptibility reflects the accessibility of cleavable states determined by the protein conformation, and not by the overall protein stability. However, it

should be noted that *in vivo*, or in the context of a complex mixture of proteins, proteolysis may be controlled by formation of complexes. In order to profile proteins according to their proteolytic resistance, an assay was developed where proteins originating from the cell lysate were submitted to an extensive proteolytic treatment with trypsin or thermolysin during long period of time. It is clear that digestions made up to 4 days lead to a successive enrichment on the so called survivors, Fig. (5), which were further identified *via* the mass spectrometry of spots excised from 2DE gels (Park *et al.*, 2007). The profiled proteins were then compared according to several criteria among which are amino acid composition, protein folds, and biological function. An important control of this approach arose from the analysis of the amino acid composition of the identified survivors, which demonstrated that proteolytic resistance is not arising from the absence of potential cleavage site residues. Further, this analysis showed that the identified proteins do not share any common structural features that could account for their proteolytic resistance.

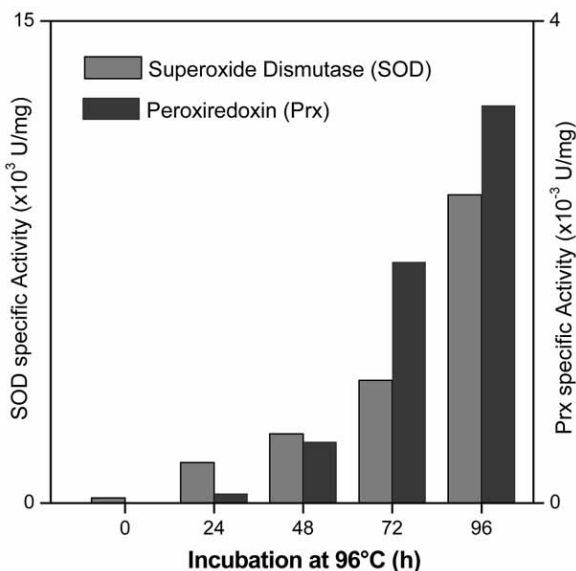


Fig. (4). Activity profiles of oxidative stress proteins during thermal perturbation. The specific activity of superoxide dismutase (SOD) and peroxiredoxin (Prx) were measured as function of incubation at 90 $^{\circ}$ C. The increase in the specific activity shows that upon 96 h perturbation the proteins retain their biological function and presumably their native fold. See (Prosinecki *et al.*, 2006) for further details.

CONCLUSIONS AND OUTLOOK

The methodologies discussed here represent non-exhaustive approaches to address protein conformation, stability and structural disorganization at a proteomic scale. These approaches illustrate efforts to snapshot the conformational state and structure of individual proteins within a complex proteome. On the one hand, the identification of yet unknown disordered proteins and its subsequent biophysical characterisation is likely to lead to important findings in key cellular processes. Within eukaryotes, whose proteomes are predicted to contain from 5 to 15% of disordered proteins, this may have an impact on several human diseases, from cancer to neurodegeneration, as typified by p53 and α -synuclein (Dobson, 2000; Dyson and Wright, 2005; Tompa, 2002). On the other hand, a less discussed aspect relates to

the recent attempts to profile proteins at a proteomic scale according to their enhanced conformational stability, which suggests that some metabolic and cellular processes may

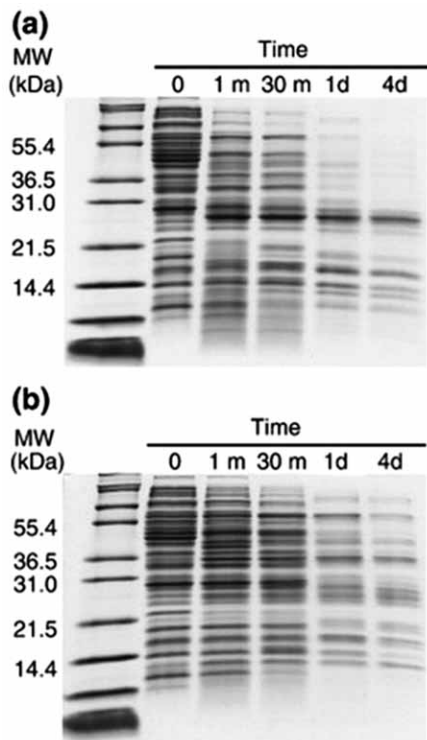


Fig. (5). Profiling proteins according to proteolytic resistance. A cell lysate from *E. coli* was digested with trypsin (a) and thermolysin (b) up to four days, and periodically sampled for SDS-PAGE gel analysis. A significant number of proteins were digested within the first 30 min; some proteins, however, the so-called survivors, persisted up to 4 days. Reprinted with permission from the *Journal of Molecular Biology* (Park *et al.*, 2007), Copyright Elsevier (2007).

require particularly stable proteins. This hypothesis has been originally put forward following the profiling of *Sulfolobus* protein 'stabilome': the identified proteins participate in cellular processes (e.g., defence against reactive oxygen species, nucleic acid protection and energy production) in which some key proteins have enhanced thermal stabilities, which may relate to their importance on the metabolism of the organism. For example, finding proteins involved in cellular detoxification processes to be hyperstable, such as iron superoxide dismutase and peroxiredoxin, may be rationalized in terms of the fact that proteins dealing with reactive oxygen species need enhanced conformational stabilities, namely to withstand and minimize possible oxidative modifications which could lead to protein inactivation (Prosinecki *et al.*, 2006). Interestingly, oxidative stress related proteins such as Fe and Mn superoxide dismutase, catalase and glutathione reductase were also identified as survivors in proteolysis resistance studies in *E. coli* proteome (Park *et al.*, 2007). Another process eventually requiring stable components involves nucleic acids. For example, at temperatures above 70°C, DNA is particularly susceptible to chemical modifications, mainly depurination followed by cleavage of the nearby phosphodiesteric bond. This makes DNA binding proteins particularly important as they protect DNA by wind-

ing and compacting it (Madigan and Oren, 1999). In fact, several DNA-binding and protecting proteins were identified as hyperstable (Park *et al.*, 2007; Prosinecki *et al.*, 2006). Extreme resistance may have also arisen from selective pressure exerted by harsh environmental factors over proteins that are either secreted or are present at the surface of cellular structures. This was elicited from the finding that a significant number of periplasmic proteins were identified as proteolysis resistant in *E. coli* (Park *et al.*, 2007). Another example is the assembly of the bacterial spore coat (Henriques and Moran Jr, 2006), whose components also likely comprise a repertoire of highly stable proteins, as spores have to be extraordinarily durable and stable. In agreement, several spore coat proteins have been reported to be highly resistant (Martins *et al.*, 2002; Wenk and Mayr, 1998). Overall, the reviewed approaches pave way to further proteome-level studies aiming at characterising the structural and conformational properties of proteins in a proteomic context. Future challenges encompass framing this knowledge in the context of the cell physiology, regulation of protein function and metabolic networks.

ACKNOWLEDGMENTS

Kevin Plaxco and Adriano O. Henriques are gratefully acknowledged for critically reading of the manuscript and for insightful comments. The work reported was supported by grants POCTI/QUI/45758 and POCTI/BIO/58465 (to CMG) from the Fundação para a Ciência e a Tecnologia (FCT/MCES, Portugal). VP is a recipient of a PhD fellowship (SFRH/BD/18746/2004), and PFNF is supported by a postdoctoral grant (SFRH/BPD/21492/2005), both from FCT/MCES, Portugal.

REFERENCES

- Anderson, D. E., Becktel, W. J. and Dahlquist, F. W. (1990). pH-induced denaturation of proteins: a single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* **29**: 2403-8.
- Anfinsen, C. (1973). Principles that govern the folding of protein chains. *Science* **181**: 223-30.
- Auld, D. S. (2001). Zinc coordination sphere in biochemical zinc sites. *Biometals* **14**: 271-313.
- Bell, S., Klein, C., Muller, L., Hansen, S. and Buchner, J. (2002). p53 contains large unstructured regions in its native state. *J. Mol. Biol.* **322**: 917-27.
- Bosshard, H. R., Marti, D. N. and Jelesarov, I. (2004). Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings *J. Mol. Recognit.* **17**: 1-16.
- Britton, K. L., Baker, P. J., Borges K. M., Engel, P. C., Pasquo, A., Rice, D. W., Robb, F. T., Scandurra, R., *et al.* (1995). Insights into thermal stability from a comparison of the glutamate-dehydrogenases from *pyrococcus-furiosus* and *thermococcus-litoralis*. *Eur. J. Biochem.* **229**: 688-95.
- Chakravarty, S. and Varadarajan, R. (2002). Elucidation of factors responsible for enhanced thermal stability of proteins: A structural genomics based study. *Biochemistry* **41**: 8152-61.
- Chandler, D. (2005). Interfaces and the driving force of hydrophobic assembly. *Nature* **437**: 640-47.
- Chen, C. H., Roth, L. G., Maccoll, R. and Berns, D. S. (1994). Thermodynamics elucidation of the structural stability of a thermophilic protein. *Biophys. Chem.* **50**: 313-21.
- Cowan, D. A. (1992). Biotechnology of the archaea. *Trends Biotechnol.* **10**: 315-23.
- Csizmok, V., Dosztanyi, Z., Simon, I. and Tompa, P. (2007). Towards proteomic approaches for the identification of structural disorder. *Curr. Protein Pept. Sci.* **8**: 173-9.

- Csizmok, V., Szollosi, E., Friedrich, P. and Tompa, P. (2006). A novel two-dimensional electrophoresis technique for the identification of intrinsically unstructured proteins. *Mol. Cell. Proteomics* **5**: 265-73.
- Dill, K. A. (1990). Dominant forces in protein folding. *Biochemistry* **29**: 7133-55.
- Dobson, C. M. (2000). The nature and significance of protein folding. In mechanisms of protein folding (R. H. Pain, Ed.), Oxford University Press Oxford. pp. 1-28.
- Dunker, A. K., Cortese, M. S., Romero, P., Iakoucheva, L. M. and Uversky, V. N. (2005). Flexible nets. The roles of intrinsic disorder in protein interaction networks. *FEBS J.* **272**: 5129-48.
- Dyson, H. J. and Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **6**:197-08.
- Ellis, R. J. and Minton, A.P. (2003). Join the crowd. *Nature* **425**: 27-28.
- Fitter, J. (2005). Structural and dynamical features contributing to thermostability in alpha-amylases. *Cell Mol. Life Sci.* **62**: 1925-37.
- Frankenberg, N., Welker, C. and Jaenicke, R. (1999). Does the elimination of ion pairs affect the thermal stability of cold shock protein from the hyperthermophilic bacterium *Thermotoga maritima*? *FEBS Lett.* **454**: 299-01.
- Grimsley, G. R., Shaw, K. L., Fee, L. R., Alston, R. W., Huyghues-Despointes, B. M., Thurlkill, R. L., Scholtz, J. M. and Pace, C. N. (1999). Increasing protein stability by altering long-range coulombic interactions. *Prot. Sci.* **8**: 1843-9.
- Gromiha, M. M., Oobatake, M. and Sarai, A. (1999). Important amino acid properties for enhanced thermostability from mesophilic to thermophilic proteins. *Biophys. Chem.* **82**: 51-67.
- Hartl, F. U. and Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* **295**: 1852-8.
- Henriques, A. O. and Moran Jr, C. P. (2007). Structure, Assembly, and function of the spore surface layers. *Annu. Rev. Microbiol.* **61**:555-88.
- Honig, B. and Hubell, W. L. (1984). Stability of salt-bridges in membrane proteins. *Proc. Natl. Acad. Sci. USA* **81**: 5412-16.
- Jaenicke, R. (2000). Stability and stabilization of globular proteins in solution. *J. Biotechnol.* **79**: 193-203.
- Karshikoff, A. and Ladenstein, R. (2001). Ion pairs and the hermotolerance of proteins from hyperthermophiles: a 'traffic rule' for hot roads. *Trends Biochem. Sci.* **26**: 550-6.
- Kauzmann, W. (1959). Chemical specificity in biological systems. *Rev. Mod. Phys.* **31**: 549-6.
- Kumar, S., Ma, B. Y., Tsai, C. J. and Nussinov, R. (2000a). Electrostatic strengths of salt bridges in thermophilic and mesophilic glutamate dehydrogenase monomers. *Proteins* **38**: 368-83.
- Kumar, S., Tsai C. J. and Nussinov, R. (2000b). Factors enhancing protein thermostability. *Protein Engineering* **13**: 179-91.
- Leal, S. S. and Gomes, C. M. (2007). Studies of the molten globule state of ferredoxin: Structural characterization and implications on protein folding and iron-sulfur center assembly. *Proteins* **68**: 606-16.
- Lee, M. S., Gottesfeld, J. M. and Wright, P. E. (1991). Zinc is required for folding and binding of a single zinc finger to DNA. *FEBS Lett.* **279**: 289-94.
- Li, W. F., Zhou, X. X. and Lu, P. (2005). Structural features of thermozymes. *Biotechnol. Adv.* **23**: 271-81.
- Loladze, V. V., Ibarra-Molero, B., Sanchez-Ruiz, J. M. and Makhatadze, G. I. (1999). Engineering a thermostable protein via optimization of charge-charge interactions on the protein surface. *Biochemistry* **38**:16149-423.
- Madigan, M. T. and Oren, A. (1999). Thermophilic and halophilic extremophiles. *Curr. Opin. Microbiol.* **2**: 265-9.
- Makhatadze, G. I., Loladze, V.V., Ermolenko, D.N., Chen, X.F. and Thomas, S.T. (2003). Contribution of surface salt bridges to protein stability: Guidelines for protein engineering. *J. Mol. Biol.* **327**: 1135-48.
- Makhatadze, G. I. and Privalov, P.L. (1995). Energetics of protein structure. *Adv. Protein Chem.* **47**: 307-425.
- Marqusee, S. and Sauer, R.T. (1994). Contributions of a hydrogen-bond salt bridge network to the stability of secondary and tertiary structure in lambda-repressor. *Prot. Science* **3**: 2217-25.
- Martins, L. O., Soares, C. M., Pereira, M. M., Teixeira, M., Costa, T., Jones, G. H. and Henriques, A. O. (2002). Molecular and biochemical characterization of a highly stable bacterial laccase that occurs as a structural component of the bacillus subtilis endospore coat. *J. Biol. Chem.* **277**: 18849-59.
- Minton, A. P. (2005). Influence of macromolecular crowding upon the stability and state of association of proteins: Predictions and observations. *J. Pharmaceutical Sci.* **94**:1668-75.
- Pace, C. N. (1992). Contribution of the hydrophobic effect to globular protein stability. *J. Mol. Biol.* **226**: 29-35.
- Pack, S. P. and Yoo, Y.J. (2005). Packing-based difference of structural features between thermophilic and mesophilic proteins. *Intern. J. Biol. Macromol.* **35**: 169-74.
- Park, C., Zhou, S., Gilmore, J. and Marqusee, S. (2007). Energetics-based protein profiling on a proteomic scale: Identification of proteins resistant to proteolysis. *J. Mol. Biol.* **368**:1426-37
- Perl, D. and Schmid, F. X. (2001). Electrostatic stabilization of a thermophilic cold shock protein. *J. Mol. Biol.* **313**: 343-57.
- Perutz, M. F., Kendrew J. C. and Watson H. C. (1965). Structure and function of haemoglobin II. Some relations between polypeptide chain configuration and amino acid sequence. *J. Mol. Biol.* **13**: 669-78.
- Perutz, M. F. and Raidt, H. (1975). Stereochemical basis of heat stability in bacterial ferredoxins and in haemoglobin A2. *Nature* **255**: 256-159.
- Perutz, M. F. (1978). Electrostatic effects in proteins. *Science* **201**: 1187-91.
- Petsko, G. and Ringe, D. (2004). Protein structure and function. primers in biology new science press.
- Pierce, A., deWaal, E., VanRemmen, H., Richardson, A. and Chaudhuri, A. (2006). A novel approach for screening the proteome for changes in protein conformation. *Biochemistry* **45**: 3077-85.
- Plaxco, K. W. and Dobson, C. M. (1996). Time-resolved biophysical methods in the study of protein folding. *Curr. Opin. Struct. Biol.* **6**: 630-36.
- Plaxco, K. W. and Gross, M. (1997). The importance of being unfolded. *Nature* **386**: 657-59.
- Prosiniecki, V., Botelho, H. M., Francese, S., Mastrobuoni, G., Moneti, G., Urich, T., Kletz, A. and Gomes, C. M. (2006). A proteomic approach toward the selection of proteins with enhanced intrinsic conformational stability. *J. Proteome Res.* **5**: 2720-6.
- Pitsyn, O. B. (1995). Structures of folding intermediates. *Curr. Opin. Struct. Biol.* **5**: 74-8.
- Querol, E., PerezPons J. A. and MozoVillarias A. (2006). Analysis of protein conformational characteristics related to thermostability. *Protein Engineering* **9**: 265-71.
- Razvi, A. and Scholtz, J. M. (2006). Lessons in stability from thermophilic proteins. *Protein Sci.* **15**:1569-78.
- Rees, D. C. and Robertson, A. D. (2001). Some thermodynamic implications for the thermostability of proteins. *Protein Sci.* **10**: 1187-94.
- Sadeghi, M., Naderi-Manesh, H., Zarrabi and M., Ranjbar, B. (2006). Effective factors in thermostability of thermophilic proteins. *Biophysical Chemistry* **119**: 256-70.
- Sandelin, E. (2004). On hydrophobicity and conformational specificity in proteins. *Biophys. J.* **86**: 23-30.
- Sarakatsannis J. N. and Duan, Y. (2005). Statistical characterization of salt bridges in proteins. *Proteins* **60**: 732-39.
- Schmid, F. (2005). Spectroscopic techniques to study protein folding and stability. In protein folding handbook (J. Buchner, and T. Kiefhaber, Eds.), Wiley-VCH, Darmstadt. Vol. 1: pp. 22-43.
- Serrano, L., Bycroft, M. and Fersht, A. R. (1991). Aromatic-aromatic interactions and protein stability. Investigation by double-mutant cycles. *J. Mol. Biol.* **218**: 465-75.
- Shirley, B. A., Stanssens, P., Hahn, U. and Pace, C. N. (1992). Contribution of hydrogen bonding to the conformational stability of ribonuclease T1. *Biochemistry* **31**: 725-32.
- Szilagyí, A. and Zavodszky, P. (2000). Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure* **8**: 493-504.
- Tompa, P. (2002). Intrinsically unstructured proteins. *Trends Biochem. Sci.* **27**: 527-33.
- Uversky, V. N. (2002). Natively unfolded proteins: a point where biology waits for physics. *Protein Sci.* **11**: 739-56.
- Vanhove, M., Houba, S., Lamottebrasseur J. and Frere, J. M. (1995). Probing the determinants of protein stability-comparison of Class A Beta lactamases. *Biochem. J.* **308**: 859-64.
- Vieille, C. and Zeikus, G. J. (2001). Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* **65**: 1-43.

- Vucetic, S., Obradovic, Z., Vacic, V., Radivojac, P., Peng, K., Iakoucheva, L. M., Cortese, M. S., Lawson, J. D., *et al.* (2005). DisProt: a database of protein disorder. *Bioinformatics* **21**: 137-40.
- Wang, M., Carp, S. A., Robblee, J., Hendsch, Z. S., Fairman, R., Tidor, B. and Raleigh, D. P. (2000). Rational modification of protein stability by the mutation of charged surface residues. *Biochemistry* **39**: 872-79.
- Wenk, M. and Mayr, E. M. (1998). Myxococcus xanthus spore coat protein S, a stress-induced member of the betagamma-crystallin superfamily, gains stability from binding of calcium ions. *Eur. J. Biochem.* **255**: 604-10.
- Yip, K. S., Britton, K. L., Stillman, T. J., Lebbink, J., de Vos, V. M., Robb, F. T., Vetriani, C., Maeder, D., *et al.* (1998). Insights into the molecular basis of thermal stability from the analysis of ion-pair networks in the glutamate dehydrogenase family. *Eur. J. Biochem.* **255**: 336-46.

Received: June 5, 2007

Revised: July 9, 2007

Accepted: July 12, 2007