Extremophiles (2011) 15:327–336
DOI 10.1007/s00792-011-0363-6

ORIGINAL PAPER

Contribution of inter-subunit interactions to the thermostability of *Pyrococcus furiosus* citrate synthase

Vicky Moore · Ami Kanu · Olwyn Byron · Gordon Campbell · Michael J. Danson · David W. Hough · Susan J. Crennell

Received: 16 July 2010 / Accepted: 28 February 2011 / Published online: 20 March 2011

Abstract Using citrate synthase from the hyperthermophile *Pyrococcus furiosus* (*Pf*CS) as our test molecule, we show through guanidine hydrochloride-induced unfolding that the dimer separates into folded, but inactive, monomers before individual subunit unfolding takes place. Given that forces across the dimer interface are vital for thermostability, a robust computational method was derived that uses the University of Houston Brownian Dynamics (UHBD) program to calculate both the hydrophobic and electrostatic contribution to the dimerisation energy at 100°C. The results from computational and experimental determination of the lowered stability of interface mutants were correlated, being both of the same order of magnitude and placing the mutant proteins in the same order of stability. This computational method, optimised for hyperthermophilic molecules and tested in the laboratory, after further testing on other examples, could be of widespread use in the prediction of thermostabilising mutations in other oligomeric proteins for which dissociation is the first step in unfolding.

Communicated by L. Huang.

V. Moore · A. Kanu · M. J. Danson · D. W. Hough · S. J. Crennell (✉)
Centre for Extremophile Research,
Department of Biology and Biochemistry,
University of Bath, Bath BA2 7AY, UK
e-mail: s.j.crennell@bath.ac.uk

O. Byron · G. Campbell
FBLS Division of Infection and Immunity,
Glasgow Biomedical Research Centre,
University of Glasgow, Glasgow G12 8TA, UK

Keywords Analytical ultracentrifugation · Fluorescence · UHBD · Dimerisation energy · Oligomeric stability · Citrate synthase · Archaeon

Abbreviations

CS	Citrate synthase
DTNB	5,5′-dithiobis(2-nitrobenzoic acid)
EPPS	4-(2-Hydroxyethyl)-1-piperazinepropanesulfonic acid
GdnHCl	Guanidine hydrochloride
*Pf*CS	*Pyrococcus furiosus* citrate synthase
$s^0_{20,w}$	Sedimentation coefficient
SASA	Solvent-accessible surface area
UHBD	University of Houston Brownian Dynamics

Introduction

Understanding the means by which the thermal stabilities of homologous proteins can differ greatly, yet the proteins retain similar activities at their temperature optima, would enable the design of thermostable molecules from mesophilic counterparts, which would be valuable in the expansion of industrial biotechnology. In general, no single feature has been identified as being responsible for protein thermostability (Vielle and Zeikus 2001; Sterner and Liebl 2001). Rather, protein thermostability is thought to arise from a combination of mechanisms, such as reduction in the number of thermolabile amino acids, shortening of flexible loops, improved packing of the hydrophobic core making thermostable enzymes smaller than mesophilic counterparts, an increase in oligomeric state and, in particular, increased ion pairs, both buried (Matsui and Harata

2007) and solvent-accessible (Torrez et al. 2003; Strickler et al. 2006). Not all these features are found in every case, but the importance of ion pairs, particularly when incorporated into more extensive networks of ionic interactions, has become evident from a number of genome-scale studies (Karshikoff and Ladenstein 2001; Chakravarty and Varadarajan 2002; Robinson-Rechavi et al. 2006); moreover, the electrostatic contribution to stability has been shown to become increasingly favourable at higher temperatures (Elcock 1998; Xiao and Honig 1999). Engineering any of these putative stabilising features into proteins and achieving increased stability has not been straightforward, although there have been successes (Eijsink et al. 2004).

A number of computational methods have been devised to probe the contribution of electrostatic interactions to the thermostabilisation of macromolecules, and there have been many reviews on the subject (Baker et al. 2003; Koehl 2005; Baker 2005). One method that has successfully predicted the significant stabilisation of mutated proteins over the wild-type, optimises the energy of charge-charge interactions over the whole surface using a genetic algorithm (Strickler et al. 2006; Gribenko et al. 2009). However, the most popular methods are based on a solution of the Poisson–Boltzmann equation where, to reduce the computational burden and enable these calculations to be applied to macromolecules, the solvent is treated implicitly (Baker 2004). Early studies were purely theoretical, but experimental verification followed. Spector et al. (2000) predicted improved stability of a 41-amino acid protein by replacing a single solvent-exposed amino acid with a non-natural hydrophobic side chain of similar size and shape. In calculating the contribution of surface amino acids to stability using three different proteins, Torrez et al. (2003) were able to achieve an average correlation between calculation and experiment of 0.86. The contribution of ion pairs to thermal stability is generally calculated by finding the difference between the solutions of the Poisson–Boltzmann equation for the folded and unfolded states of a protein, to take into account that mutations may also stabilise the denatured state and thus have little effect on the equilibrium between the two states (Elcock 1998). However, in almost all cases, these approaches were applied to small, monomeric proteins.

In larger proteins, thermal unfolding may not be reversible since unfolding may be followed by irreversible aggregation (Eijsink et al. 2004, and references therein), and in these cases it is the stability of the native state relative to the intermediate state that is required, rather than to a completely unfolded form of the protein. This is the case for our test enzyme, citrate synthase (*Pf*CS), from the hyperthermophilic archaeon, *Pyrococcus furiosus*, which grows optimally at 100°C. The crystal structure of the dimeric enzyme has been determined, showing that the active site contains residues from both monomers; for example, R353 from one monomer extends into the active site of the other subunit (Fig. 1) and serves to bind the α–β phosphates of acetyl-CoA (Russell et al. 1997). Electrostatic interactions across the dimer interface have been implicated in the enzyme's thermostability and thermoactivity (Russell et al. 1997), and this conclusion has been tested experimentally by site-directed mutagenesis (Arnott et al. 2000). A previous study of the contribution of electrostatics to the stability of citrate synthase (CS) (Kumar and Nussinov 2004) compared the hyperthermophilic *Pf*CS with a psychrophilic CS and a mesophilic CS. This work was purely computational and compared amino acids with their hydrophobic isosteres rather than experimentally accessible amino acids. Moreover, although the study focussed on psychrophilicity, the calculation parameters (dielectric constants, atomic radii) were fixed at the values corresponding to 25°C, not the optimal growth temperature for any of the organisms considered, and most distant from *P. furiosus*.

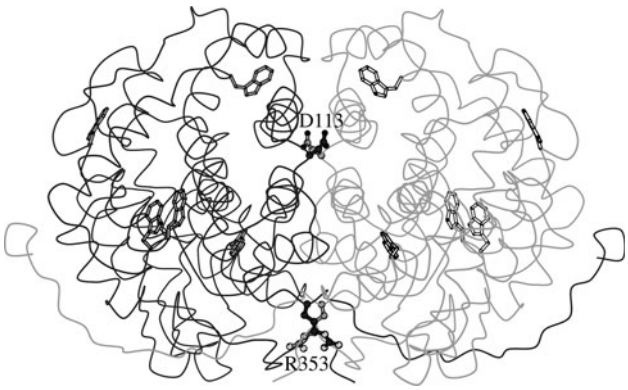

Fig. 1 The Cα trace of *Pf*CS, with the two monomers drawn in *light* and *dark grey*, and individual residues added in *ball-and-stick* representation. The active sites are indicated by R353, which crosses from one monomer to complete the active site of the other, but are separate from the ion pair network in the dimer interface, one part of which is D113, the site of the D113S and D113A mutations. All the tryptophan residues, drawn in white, are distant from the dimer interface. Figure produced using MOLSCRIPT (Kraulis 1991)

In the current work, we show that for the dimeric *Pf*CS, the first stage in denaturation is the reversible separation of the monomers, forming an inactive but folded monomer conformation. Having identified the first stage in unfolding for this enzyme system, in our subsequent computational analysis we compare the energy of the dimer with that of the two folded individual monomers, thereby avoiding the difficulties inherent in estimating the structure of an unfolded conformation.

To improve correlation with experiment, the parameters used in the experimental studies were incorporated into the computational analyses, permitting a direct comparison of

the two approaches in the study of the contribution of the inter-subunit ionic interactions to the thermostability of *Pf*CS. The method was tested by calculating dimerisation energies for the wild-type enzyme and also for two mutated *Pf*CS proteins, where one of the residues responsible for the inter-subunit electrostatic interactions yet distant from the active site, Asp113 (Fig. 1), was changed to a residue capable only of hydrogen bonding (D113S) and to one incapable of any polar interaction (D113A). Agreement between the experimentally measured and computationally derived stabilities of these proteins shows that this is a robust method for the prediction of thermostabilising mutations in *Pf*CS.

Materials and methods

Expression and purification of citrate synthase

Citrate synthases from *P. furiosus* wild-type, and mutants D113S and D113A, were produced as described by Arnott et al. (2000). Briefly, the wild-type and mutant CS genes were inserted into the *Escherichia coli* expression vector pREC7/*Nde*I and expressed under the control of the recA promoter by induction with nalidixic acid. Soluble, active expressed proteins were purified from cell extracts as described by James et al. (1994) by a combination of heat treatment and affinity chromatography using a Matrex Red Gel A affinity column.

Citrate synthase assay

Citrate synthase activity was measured spectrophotometrically at 55°C by the method of Srere et al. (1963). The standard reaction mixture contained 0.2 mM oxaloacetic acid, 0.15 mM acetyl-coenzyme A, and 0.2 mM 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB) in 50 mM EPPS buffer, pH 8 (at 55°C), containing 100 mM KCl and 2 mM EDTA.

Guanidine hydrochloride-induced unfolding

Stock solutions of 8 M guanidine hydrochloride (GdnHCl) were prepared as described by Pace and Scholtz (1998) and the concentration was determined by refractive index measurements. Twenty to thirty different denaturant concentrations were used to define each denaturation curve and were prepared by adding a standard volume of protein solution to a mixture of denaturant stock solution and buffer (50 mM sodium phosphate, pH 7, 2 mM EDTA) to obtain a protein concentration of 10 μg/ml and the required concentrations of denaturant. The solutions were then incubated at 50°C for up to 18 h to allow equilibrium to be reached.

Fluorescence measurements were carried out at 50°C using a Perkin Elmer 50B spectrofluorimeter with excitation wavelength 278 nm and emission wavelength 340 nm. Catalytic activity measurements were carried out at 55°C using the standard assay described above.

Analytical ultracentrifugation

Sedimentation equilibrium was performed in a Beckman Optima XL-A analytical ultracentrifuge (Palo Alto, USA) with an An-60 Ti 4-hole rotor. *Pf*CS stock solution was diluted in buffer to give nine samples for sedimentation equilibrium analysis. Four samples were diluted in 50 mM phosphate buffer, pH 7, 2 mM EDTA, and the remaining five samples were diluted in the phosphate buffer containing 2.4 M GdnHCl. The samples were loaded into 12 mm path length, charcoal-filled epon six-channel centrepieces and run at speeds of 18,000 and 24,000 rpm at 4°C. A series of scans was taken 3 h apart until equilibrium was ascertained using the program WINMATCH (http://www.rasmb.bbri.org/rasmb/windows/uconn_uaf/). The final data were the average of 10 scans acquired at a wavelength of 280 nm with a radial step size of 0.001 cm.

The sedimentation coefficient of the protein in the presence of 0 and 2.4 M GdnHCl was determined by sedimentation velocity. Protein samples in the phosphate buffer at 0 and 2.4 M GdnHCl were loaded into 12 mm path length, charcoal-filled epon double sector centrepieces and centrifuged at 4°C in a Beckman Coulter Optima XL-I analytical ultracentrifuge (Palo Alto, USA) using an An-50 Ti 8-hole rotor. Interference optics were used; the laser delay was adjusted prior to the run to obtain high-quality interference fringes. A series of 200 scans, 2.5 min apart, was acquired at 50,000 rpm.

The partial specific volume ($\bar{v}$) of *Pf*CS, calculated from its amino acid composition using the program SEDNTERP (Laue et al. 1992; http://www.jphilo.mailway.com/), was 0.742 ml/g in the phosphate buffer at 4°C and 0.730 ml/g in the presence of 2.4 M GdnHCl. SEDNTERP was also used to compute buffer densities [1.00481 g/ml (phosphate buffer); 1.06506 g/ml (phosphate buffer with 2.4 M GdnHCl)] and viscosities [0.01599 Poise (phosphate buffer); 0.01794 Poise (phosphate buffer with 2.4 M GdnHCl)] at 4°C.

Computational methods

The wild-type *Pf*CS structure was obtained from the PDB (Berman et al. 2000) (PDB ID: 1AJ8). The WHAT IF web interface (Rodriguez et al. 1998) was used to add hydrogen atoms and to assign the protonation state of the histidine residues based on their potential to form hydrogen bonds (Nielsen et al. 1999; Nielsen and Vriend 2001).

The calculations were performed with the UHBD package (Madura et al. 1995), following the method of Elcock and McCammon (1997), which combines a solvent-accessible surface area (SASA) term to describe hydration of non-polar groups with continuum electrostatics to describe the polar groups. The PARSE parameter files of atomic radii and partial charges were adjusted to incorporate increasing atomic radii with temperature, having separate files for calculations at 25, 50, 75 and 100°C (the files were a generous gift from A.H. Elcock, University of Iowa, IA, USA). Similarly, the value of the solvent dielectric was decreased with temperature, from 78.8 at 25°C, to 69.9 at 50°C, 62.3 at 75°C and 55.6 at 100°C (Weast 1973). This was necessary to reproduce the increased contribution of electrostatic interactions to stability at higher temperatures. An accurate value for the protein dielectric constant is difficult to obtain; those quoted in the literature vary between 4 (interior) and 20 (surface) (Warshel et al. 2006), so for these whole protein calculations the protein dielectric constant was set at 12 across all temperatures. The proteins were placed within a 200^3 grid with a grid spacing of 0.5 Å. The ionic strength was chosen to be 50 mM, comparable to that used in experimental measurements of $\Delta\Delta G^*$; changing this parameter changed the magnitude of the energies obtained, but did not change their order. The remaining parameters were: maxits 500 (number of iterations), nmap 1.4 (probe radius for accessible surface), nsph 300 (number of surface points per atom for surface calculation), rion 2 (ion exclusion radius), and bcfl 2.0 (boundary condition 2).

To calculate the hydrophobic contribution to the dimerisation energy, the SASA of each atom was determined in UHBD, multiplied by a different temperature-dependent radius scaling factor (RSF) depending on whether the atom was polar, aliphatic or aromatic, and then finally summed to give a total value for each structure (Elcock and McCammon 1997).

Mutant structures were generated using O (Jones et al. 1991), with the changed amino acids being modelled in all available rotamer conformations, for each of which a separate energy calculation was carried out. The final dimerisation energy of the protein was calculated as the Boltzmann-weighted average of the contributions from all the individual rotamers.

Results

Guanidine hydrochloride-induced changes in fluorescence and activity

Initial studies of the GdnHCl-induced unfolding of wild-type *Pf*CS, monitored by changes in fluorescence emission at 340 nm, revealed a single-transition unfolding curve (Fig. 2), consistent with a two-state unfolding mechanism. However, activity measurements indicated that the enzyme was completely inactivated at low concentrations of GdnHCl prior to the onset of a significant change in the fluorescence emission (Fig. 2).

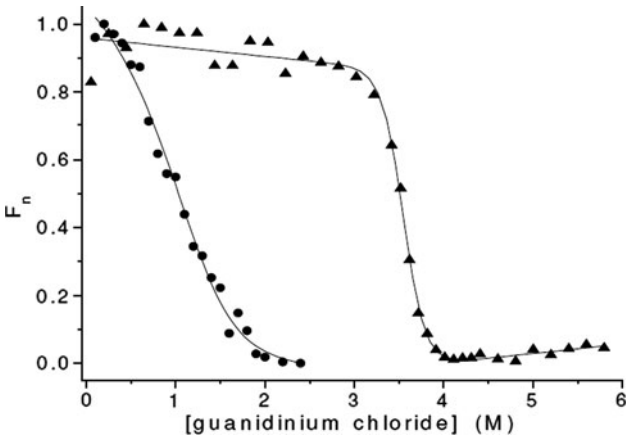

Fig. 2 GdnHCl-induced unfolding of *Pf*CS at 50°C as determined by changes in relative activity (*circle*) and fluorescence emission (*triangle*). The enzyme was equilibrated in 50 mM sodium phosphate, pH 7, 2 mM EDTA, containing varying concentrations of GdnHCl, for up to 18 h at 50°C prior to measurements. Fluorescence emission at 340 nm and citrate synthase activity were measured at 50°C and the values normalised to the fraction of the value obtained with native, active protein (F_n). Data were fitted by non-linear least square analysis using Origin software

Reversibility of the unfolding transitions was examined by a 20-fold dilution in phosphate buffer of the *Pf*CS that had been fully unfolded in 6 M GdnHCl. Complete recovery of activity was observed after incubation for 1 h at 50°C, whereas the fluorescence signal (F_{340}) reached 70% of the initial value. The incomplete recovery in the fluorescence signal is probably due to the need to use higher concentrations of protein (50 μg/ml) in order to detect a signal following dilution.

Analytical ultracentrifugation analysis

The effects of GndHCl on the wild-type *Pf*CS were further analysed by analytical ultracentrifugation. Sedimentation equilibrium data were fitted with a monomer–dimer self-association model (encoded in the Beckman Coulter software module within the Origin data analysis programme). The monomer mass was fixed at 43,290 Da (computed from the amino acid composition). Fits (to the 25,000 rpm data) are shown in Fig. 3. In the phosphate buffer *Pf*CS dimerises with $K_d = 17$ nM, whereas this is dramatically elevated to 131 μM upon addition of 2.4 M GdnHCl. Therefore, at the concentrations used throughout this study, *Pf*CS is a dimer in phosphate buffer and a monomer in the presence of GdnHCl.

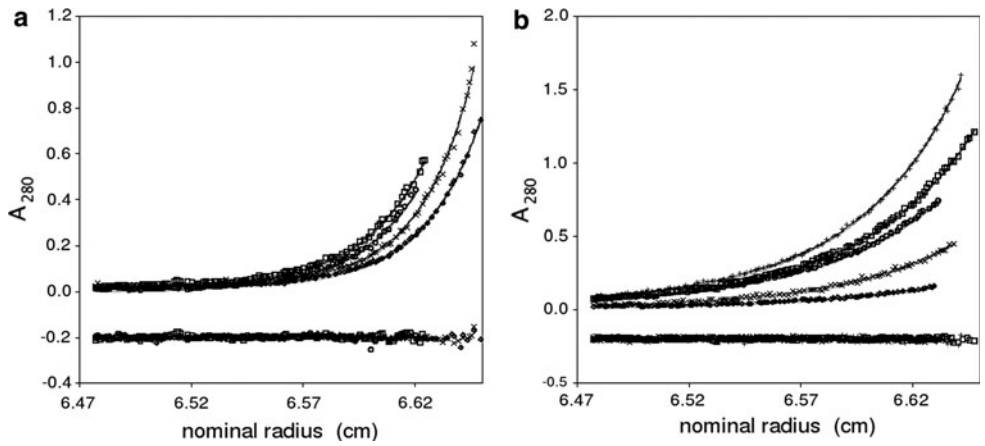

Fig. 3 Sedimentation equilibrium data fitted with a monomer–dimer self-association model for **a** *Pf*CS in 50 mM phosphate buffer, pH 7, 2 mM EDTA (K_d = 17 nM); loading concentrations were: 1.86 μM (*diamond symbol*), 3.01 μM (*times symbol*), 3.94 μM (*circle*), 5.34 μM (*square*), and **b** 50 mM phosphate buffer, pH 7, 2 mM EDTA containing 2.4 M GdnHCl (K_d = 131 μM); loading concentrations were: 0.70 μM (*diamond symbol*), 1.86 μM (*times symbol*), 3.25 μM (*circle*), 4.64 μM (*square*), 5.92 μM (*plus symbol*). The data have been normalised in the radial dimension to permit ease of display. Residuals of the fits to the data are shown beneath the exponential distributions

Sedimentation velocity data were analysed using the program SEDFIT (Schuck 2000; Schuck et al. 2002). Sedimentation boundaries were initially modelled as numerical finite element solutions of the Lamm equation using the *c(s)* analysis (Fig. 4). Apparent sedimentation coefficients were further obtained via the non-interacting discrete species model that employs finite element analysis. The apparent sedimentation coefficients were then corrected to standard conditions of temperature and solvent before extrapolation to infinite dilution to obtain sedimentation coefficients independent of concentration, $s^0_{20,w}$. In phosphate buffer, between 82 and 91% of *Pf*CS (depending on loading concentration) sediments with $s^0_{20,w}$ = 5.41 S; upon addition of 2.4 M GdnHCl the major species (between 70 and 100%, again dependent on loading concentration) has $s^0_{20,w}$ = 3.10 S. The sedimentation coefficients of the dimer ($s^0_{20,w}$ = 5.36 S) and monomer ($s^0_{20,w}$ = 3.16 S) of *Pf*CS were calculated from the crystal structure (Berman et al. 2000, PDB ID: 1AJ8) using the SOMO module (Rai et al. 2005) of the program UltraScan II (Demeler et al. 2005). This supports the hypothesis that in phosphate buffer the major species is a dimer; in 2.4 M GdnHCl *Pf*CS the major species is a monomer, whereas the minor peak/shoulder at a lower sedimentation coefficient is most probably partially denatured monomer.

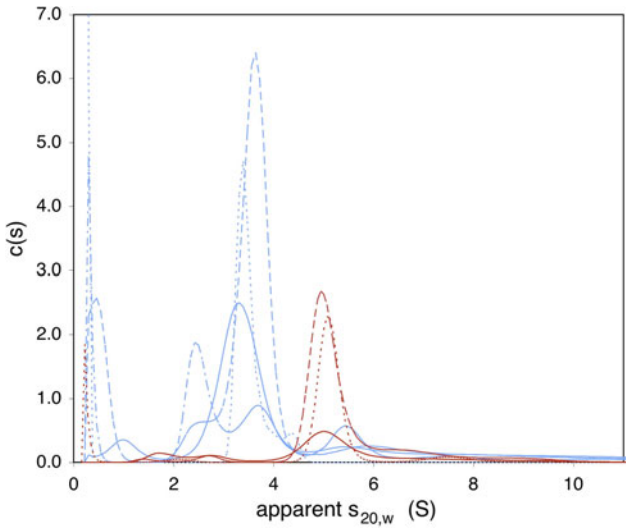

Fig. 4 *c(s)* distributions derived via SEDFIT from SV interference data for *Pf*CS in (*red*) 50 mM phosphate buffer, pH 7, 2 mM EDTA at loading concentrations of 13.4 μM (*dashed*), 6.7 μM (*dotted*) and 4.2 μM (*solid*) and (*blue*) 50 mM phosphate buffer, pH 7, 2 mM EDTA containing 2.4 M GdnHCl at loading concentrations of 21.0 μM (*dashed*), 12.6 μM (*dotted*), 10.6 μM (*solid*) and 9.0 μM (*dash-dotted*) The peak positions were used to guide subsequent non-interacting discrete species model analysis of the raw data from which apparent sedimentation coefficients were derived (colour figure online)

Unfolding model for *P. furiosus* citrate synthase

Citrate synthase requires a completely associated dimer for catalytic activity, as the active site spans the dimer interface, R353 from one monomer being required to bind co-factor in the other active site (Fig. 1). However, the tryptophan residues of *Pf*CS, responsible for the bulk of the fluorescence observed, are buried within the monomer structure, not on the dimer interface, as can also be seen in Fig. 1; therefore, fluorescence changes happen only once the monomers unfold. Thus the intrinsic fluorescence, activity and ultracentrifugation data for *Pf*CS are consistent with a two-step sequential unfolding mechanism involving an initial dimer-to-monomer dissociation at low

concentrations of denaturant (step I), followed by complete unfolding of the monomer at higher denaturant concentrations (step II). This model can be summarised as:

$$N_2 \overset{I}{\rightleftharpoons} 2N \overset{II}{\rightleftharpoons} 2U$$

Consequently, the unfolding data in Fig. 2 were analysed using a concentration-dependent dimer/monomer model for step I and a concentration-independent monomer model for step II, as described by Neet and Timm (1994) and Ásgeirsson et al. (2000). For example, the fluoresence data (step II) were analysed using a non-linear, least squares programme to fit the entire unfolding curve to the equation (Pace, 1990):

$$F_{obs} = \frac{\left[(F_f + m_f[D]) + (F_u + m_u[D])e^{-\left(\frac{\Delta G_u}{RT} - \frac{m[D]}{RT}\right)}\right]}{1 + e^{-\left(\frac{\Delta G_u}{RT} - \frac{m[D]}{RT}\right)}} \quad (1)$$

where F_f and m_f, and F_u and m_u, are the slope and intercept of the pre- and post-transition baselines, respectively, [*D*] is the concentration of GndHCl, ΔG_u the free energy of unfolding of the monomer, and m is a measure of the dependence of ΔG_u on denaturant concentration. The activity data (step I) were similarly analysed, using a modified version of Eq. 1 as fully described in Ásgeirsson et al. (2000). The resulting lines of best fit are shown in Fig. 2 and the corresponding ΔG and *m* values in Table 1. The ΔG value for the wild-type enzyme corresponds reasonably with that calculated from the K_d of 17 nM determined by ultracentrifugal analysis ($\Delta G = 9.8$ kcal/mol at 4°C).

The effect of GdnHCl on the activity and intrinsic fluorescence of the *Pf*CS mutants D113A and D113S was also investigated. The data were analysed as described above and the corresponding ΔG and *m* values are also shown in Table 1. Meyers et al. (1995) provide evidence that the amount of surface exposed to solvent upon unfolding is a main structural determinant of an *m* value, and the observation that the *m* values for the N–U transition are larger than those for N_2–2N is consistent with this. However, the significantly larger m_u value of *Pf*CS D113S than that of *Pf*CS D113A was unexpected and is unexplained, although others have noted similar effects on single amino acid changes in a protein; for instance, mutations in *Staphylococcus aureus* nuclease were shown to cause variation in *m* values over a threefold range (Shortle 1995). In this study, Shortle ascribes these changes in m to changes in the relative stabilities of the native and denatured states and acknowledges that little information is available on the range of structures taken up by denatured proteins. To avoid the necessity of modelling the denatured states in analysis of the N–U transition, this work has concentrated on the preceding N_2–2N dissociation.

Table 1 Comparison of the effects of GdnHCl on the catalytic activity and fluorescence emission of wildtype *Pf*CS and mutants D113A and D113S

	ΔG (kcal mol^{-1})	*m* (kcal mol^{-1} Ml^{-1})
ΔG_d (N2 ⇌ 2N)		
*Pf*CS wt	14.7 ± 1.0	−2.55 ± 1.43
*Pf*CS D113A	12.4 ± 1.5	−2.42 ± 1.11
*Pf*CS D113S	12.3 ± 1.4	−2.68 ± 0.88
ΔG_u (N ⇌ U)		
*Pf*CS wt	18.7 ± 2.2	−4.18 ± 0.65
*Pf*CS D113A	17.5 ± 5.2	−4.76 ± 1.41
*Pf*CS D113S	29.9 ± 5.9	−8.30 ± 1.62

Data were analysed by non-linear least square analysis using Origin software, to derive the free energy of dissociation of the dimer (ΔG_d), the free energy of unfolding of the monomer (ΔG_u), the associated *m* values, and standard deviations from the curve-fitting procedure

Computation analysis of thermostability

The fluorescence, activity and analytical ultracentrifugation results showed that the *Pf*CS dimer separated into two folded monomers before polypeptide unfolding. Thus the thermostability of *Pf*CS is critically dependent on interactions that stabilise the dimer, and this is reflected in the energy of dimerisation, which is the difference between the energy of the folded dimer and that of the two folded monomers separately. Knowing that this first stage in unfolding is reversible meant that monomer structures derived from the crystal structure of the dimer could be used in the energy calculations rather than the unfolded conformations. In developing the computational method, initial calculations focussed on the wild-type *Pf*CS, calculating the electrostatic and hydrophobic contributions to the dimerisation energy across a range of temperatures from 25 to 100°C, mesophilic to hyperthermophilic. Results of these calculations are given in Table 2. The calculated values of the electrostatic and hydrophobic contributions to the dimerisation energy of the wild-type and mutant proteins at 100°C are given in Table 3, together with the experimentally determined $\Delta\Delta G^*$ values for thermal inactivation measured at the same temperature. The dimerisation energy values for the D113S *Pf*CS mutant are a Boltzmann-weighted average of those obtained for all three possible rotamers (a, b, c) of a serine residue.

Discussion

Oligomer stabilisation in hyperthermophilic proteins

In this work we show through comparison of fluorescence and ultracentrifugation analysis that citrate synthase is one

Table 2 Electrostatic and hydrophobic contributions to the dimerisation energy of *Pf*CS

Temperature (°C)	Electrostatic contribution (kcal mol^{-1})	Hydrophobic contribution (kcal mol^{-1})	Total dimerisation energy (kcal mol^{-1})
25	8.5	−52.6	−44.1
50	1.9	−62.3	−60.4
75	−7.3	−68.6	−75.9
100	−12.6	−71.6	−84.2

Values over a range of temperatures were calculated by subtracting the energy values, calculated in UHBD, for the two *Pf*CS monomers from the energy of the dimer

Table 3 Electrostatic and hydrophobic contributions to the dimerisation energy of wild-type and mutant *Pf*CSs

	Electrostatic contribution (kcal mol^{-1})	Hydrophobic contribution (kcal mol^{-1})	Total dimerisation energy (kcal mol^{-1})	Difference in dimerisation energy (kcal mol^{-1})	$\Delta\Delta G^*$ (kcal mol^{-1})
*Pf*CS	−12.6	−71.6	−84.2	0	0
D113S	−12.0	−71.5	−83.5	0.70	0.79
D113Sa	−9.4	−71.4			
D113Sb	−7.7	−71.4			
D113Sc	−12.1	−71.6			
D113A	−9.8	−70.8	−80.6	3.6	1.1

Values given are the electrostatic and hydrophobic dimerisation energies, calculated by subtracting the energy values, calculated in UHBD, for the two *Pf*CS monomers from the energy of the dimer. Values for D113S are given by the Boltzmann-weighted average of those for the individual rotamers D113Sa, D113Sb and D113Sc. All calculations were performed at a temperature of 100°C. The differences in dimerisation energy were calculated by subtraction of the wild-type dimerisation energy from those of the mutants. $\Delta\Delta G^*$ values ($\Delta G^*_{\text{mutant}} - \Delta G^*_{\text{wild-type}}$) for irreversible thermal inactivation, included for comparison, are from Arnott et al. (2000)

of the significant proportion of oligomeric proteins for which denaturation begins with reversible dissociation of monomers before the final irreversible unfolding step. A recent review (Rumfeldt et al. 2008) suggests that this is a common result for dimeric proteins, although examples of three-state unfolding involving dimeric intermediates and two-state unfolding have been described. In an earlier study of structures of citrate synthases from organisms with a range of growth temperature optima from psychrophilic to hyperthermophilic, the extent of electrostatic interactions across the dimerisation interface was shown to correlate with increased thermostability (Bell et al. 2002). A recent comparison of 123 dimeric thermo- and hyperthermo-stable enzyme structures with mesophilic counterparts has shown an increase in ion pairs in the thermophilic interfaces, although no correlation was found between the number of interfacial ion pairs and the temperature optima (Maugini et al. 2009). However, within the citrate synthase family of structures this is the case; the hyperthermophilic *Pf*CS has the largest electrostatic network, one of the key residues of which is D113. Mutation of this residue has already been shown to reduce thermostability as measured by resistance to irreversible thermal inactivation (Arnott et al. 2000).

The $\Delta G_{\text{d}}(\mathbf{N_2} \rightleftharpoons \mathbf{2N})$ values obtained here for the D113S and D113A mutated proteins are lower than that of wild type *Pf*CS (Table 1), corroborating the importance of electrostatic interactions in the inter-subunit bonding. Thus, for thermophilic oligomeric structures such as the citrate synthase dimer, electrostatic interactions are an important factor in thermostability.

Calculation of the electrostatic contribution of individual amino acids to thermostability

Calculation of the free energies of salt bridges in interfaces has been an area of active research, particularly in the area of protein–protein complexes, whether transient or permanent. Albeck et al. (2000) used the continuum electrostatic approach as implemented in DELPHI to calculate the difference in energy between a protein–protein complex and the two isolated proteins, investigating two salt bridges and two hydrogen bonds. The residues were individually mutated to Ala experimentally and computationally but, although the calculations of the effect of removal of the salt bridges correlated well with experimental results, the hydrogen bond removal was less satisfactory.

Since then there has been an explosion of new computational methods developed to predict the pathway and the strength of protein–protein interactions (reviewed in Potapov et al. 2009), many of which are web-based and freely available. These algorithms are knowledge-based, deriving

parameters from a large training set of interactions that are then validated by applying them to an even larger test set. According to Potapov et al. (2009), there are significant drawbacks to all these methods. Although they predict the correct trend of mutations, the values obtained underestimate experimental results, giving, of necessity, an averaged value. Also, although predictions of the effect of mutation to alanine were shown to have a correlation coefficient of 0.5 with experiment, mutation to other residues is significantly less accurate, the correlation falling to 0.38. Thus they conclude that there is still a need for more computationally intensive but more realistic energy calculations for individual interactions. Another important drawback to the use of these methods for calculation of the effect of mutation on free energy at elevated temperatures, is that none have been trained on thermostable or hyperthermostable structures.

Similarly, there have been many descriptions of the use of UHBD and other programs such as DELPHI to calculate the contribution of electrostatic interactions to thermal stability, some of which were followed by comparisons with experiment (e.g. Torrez et al. 2003; Spector et al. 2000), but these calculations were carried out for temperatures around 25°C, while thermal denaturation takes place experimentally at a significantly higher temperature. For instance, one such study, an earlier computational study of the same CS enzyme (Kumar and Nussinov 2004), used room temperature and zero ionic strength as parameters, conditions that are not appropriate for thermophilic or hyperthermophilic enzymes.

Our work uses the method of Elcock and McCammon (1997) to scale the parameters such as solvent dielectric constant and atomic radii to values more appropriate for 100°C conditions. Consequently, our results for the *Pf*CS wild-type (Table 2) calculated that the hydrophobic contribution to dimerisation is stabilising at all temperatures, but that the electrostatic contribution is destabilising at mesophilic (25°C) and moderately thermophilic (50°C) temperatures, only becoming stabilising at hyperthermophilic temperatures. This is in agreement with current theory that stabilisation through salt bridges is important only in proteins from hyperthermophiles (Elcock 1998; Karshikoff and Ladenstein 2001), and indicates that the use of scaled parameters aids accurate prediction at high temperatures.

The absolute values obtained by this method for the dimerisation energy were inversely correlated with the protein dielectric constant used in the calculation (data not shown). Using a protein dielectric of 4, the electrostatic dimerisation energies were 100 kcal/mol higher than those using 12, and with a value of 20 they were 20 kcal/mol more negative. However, whichever value was used, the trends within the energy values obtained were always in the same direction, e.g. with increasing temperature the electrostatic contribution became more negative. Using a protein dielectric constant of 12 gave the results shown in Table 2, where electrostatic interactions become stabilising at higher temperatures, and this value was used in all calculations presented. However this uncertainty in the protein dielectric constant means that the absolute values of dimerisation energy obtained cannot be significant; rather it is the relative values that should be compared to those in the experimental results. Thus the absolute value of the experimental wild-type ΔG_d given in Table 1 measured at 50°C is not the same as the dimerisation energy at that temperature in Table 2, although an agreement could have been achieved by using a non-standard value for the protein dielectric constant.

Trends in computed values compare favourably with other experimental data when this model was used to calculate p*K*a in enzyme active sites (Nielsen and McCammon 2003) and the free energies of individual amino acids at elevated temperature (Elcock and McCammon 1997). Potapov et al. (2009) surveyed six methods for calculations of $\Delta\Delta G$ on mutation and obtained at best a correlation of 0.59 between experimental and computational values, although all methods showed a correct trend in predictions. This work shows that prediction of trends in $\Delta\Delta G$ values can be extended to high temperatures.

A major improvement with the method reported here is removal of the necessity of estimating an unfolded conformation of the protein to obtain meaningful energies of dissociation. The fluorescence data show that the *Pf*CS monomers have essentially the native fold, which is known from the crystal structure of the dimer, so the comparison of the energies of the dimer and monomer conformations is free from the uncertainty inherent in modelling an unfolded protein. The calculated energy values shown in Table 3 are obtained by subtracting the energy of the two monomers from that of the dimer, and therefore are a measure of the difference in energy of these two states with the most negative being the most stable. A similar approach has been used in hetero-protein complex calculations where the energy of the two proteins in isolation is subtracted from that of the complex (Albeck et al. 2000). By presenting their energies with respect to the wild-type enzyme, the values for the mutants can be shown to correlate with the published $\Delta\Delta G^*$ values for thermal inactivation in Table 3. The experimentally determined ΔG_d ($N_2 \rightleftharpoons 2N$) values measure the energy barrier to monomer separation, so the more stable species has a larger positive value, and the results in Table 1 show the mutants to be less stable than the wild-type by an amount similar in magnitude to those seen in Table 3. In making this comparison, it should be noted that irreversible thermal inactivation determines differences in activation energies ($\Delta\Delta G^*$) between wild-

type and mutant enzymes, whereas $\Delta\Delta G_d$ values are free energy differences between dimer and monomeric forms. However, mutational alteration of the stability of the native folded enzyme will affect the values of both $\Delta\Delta G^*$ and $\Delta\Delta G_d$ in a similar manner.

To calculate the energy of mutated proteins, the conformation of the altered amino acid must be estimated. Earlier studies, such as that of Kumar and Nussinov (2004) on CS, calculated the electrostatic contribution with respect to hydrophobic isosteres, which are not experimentally accessible, so cannot be readily compared with experimental measurements. This may have been due to the difficulty of estimating the relaxation of structures following mutation, and a drawback of grid-based methods such as UHBD and DELPHI is that they are critically dependent on the positions of atoms relative to the grid. For point mutations, such as those introduced here, we have made the assumption that the overall structure is unaltered and that the mutated amino acid takes up one of the well-characterised rotamer conformations (Lovell et al. 2000), and can thus be modelled into the structure of the wild-type protein. Although this may not be valid for multiple mutations, for the single mutations used here a good correlation is observed with experiment (Table 3).

In Table 3, the serine rotamers are presented in order of descending frequency of occurrence in protein structures (Lovell et al. 2000), showing that the third, least common form is in fact the most stable. This is probably not the one that would have been chosen automatically since it is neither the most common in other structures (that is rotamer a), nor follows the direction of the D113 side chain (approximately in the direction of the b rotamer), but is the only one of the three that is able to form a hydrogen bond with H93 on the other monomer and thus restore some of the stabilisation across the dimer interface. Thus, an unexpected outcome of this method was in conformer prediction; if the energies of all possible rotamers are calculated, the method can make a clear prediction of the most stable conformation of an altered amino acid, despite this not being the most common rotamer.

Conclusion

We have shown that the hyperthermostable *Pf*CS dimer dissociates reversibly into two folded monomers prior to irreversible unfolding, and knowing this we have developed a robust method for calculating the contribution of individual amino acids to protein thermostability at high temperatures. Using this method, with parameters appropriate for experimental conditions, we have calculated dimerisation energies of the hyperthermostable, homodimeric *Pf*CS that correlate both with published experimental $\Delta\Delta G^*$ values for thermal inactivation (Table 3, Arnott et al. 2000), and with the ΔG_d ($N_2 \rightleftharpoons 2N$) measurements calculated from guanidine-induced inactivation (Table 1). Both computational and experimental approaches show the mutants to be less stable than the wild-type, the results being of the same order of magnitude and in the same order. This approach, based on knowledge of the underlying mechanism of unfolding and the structure of the molecule, now needs to be tested on other systems to determine its wider applicability to the many other thermophilic or hyperthermophilic oligomeric proteins shown to unfold after dissociation of monomers.

Acknowledgments AYK and VM thank the Biotechnology & Biological Sciences Research Council, UK, for their studentships. The PARSE parameter files were a generous gift from Dr Adrian H. Elcock (University of Iowa, IA, USA). Generous support and encouragement from Dr Adrian Mulholland (University of Bristol, UK) and Dr Adrian H. Elcock are gratefully acknowledged.

References

Albeck S, Unger R, Schreiber G (2000) Evaluation of direct and cooperative contributions towards the strength of buried hydrogen bonds and salt bridges. J Mol Biol 298:503–520

Arnott MA, Michael RA, Thompson CR, Hough DW, Danson MJ (2000) Thermostability and thermoactivity of citrate synthase from the thermophilic and hyperthermophilic Archaea, *Thermoplasma acidophilum* and *Pyrococcus furiosus*. J Mol Biol 304:657–668

Ásgeirsson B, Hauksson JB, Gunnarsson GH (2000) Dissociation and unfolding of cold-active alkaline phosphatase from Atlantic cod in the presence of guanidinium chloride. Eur J Biochem 67:6403–6412

Baker NA (2004) Poisson–Boltzmann methods for biomolecular electrostatics. Meth Enz 383:94–118

Baker NA (2005) Improving implicit solvent simulations: a Poisson-centric view. Curr Opin Struct Biol 15:137–143

Baker NA, McCammon JA (2003) Electrostatic interactions. In: Bourne PE, Weissig H (eds) Structural bioinformatics. Wiley-Liss, New York, pp 427–440

Bell GS, Russell RJM, Connaris H, Hough DW, Danson MJ, Taylor GL (2002) Stepwise adaptations of citrate synthase to survival at life's extremes. Eur J Biochem 269:6250–6260

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. Nucl Acids Res 28:235–242

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

Demeler B (2005) UltraScan, a comprehensive data analysis software package for analytical ultracentrifugation experiments. In: Scott DJ, Harding SE, Rowe AJ (eds) Modern analytical ultracentrifugation: techniques and methods. Royal Society of Chemistry (UK), pp 210–229

Eijsink VGH, Bjørk A, Gåseidnes S, Sirevåg R, Synstad B, van den Burg, Vreind G et al (2004) Rational engineering of enzyme stability. J Biotechnol 113:105–120

Elcock AH (1998) The stability of salt bridges at high temperatures: implications for hyperthermophilic proteins. J Mol Biol 284:489–502

Elcock AH, McCammon JA (1997) Continuum solvation model for studing protein hydration thermodynamics at high temperatures. J Phys Chem B 101:9624–9634

Gribenko AV, Patel MM, Lui J, McCallum SA, Wang C, Makhatadze GI (2009) Rational stabilization of enzymes by computational redesign of surface charge-charge interactions. Proc Natl Acad Sci 106:2601–2606

James KD, Russell RJM, Parker L, Daniel RM, Hough DW, Danson MJ (1994) Citrate synthases from the Archaea: development of a bio-specific affinity column purification procedure. FEMS Microbiol Lett 119:181–186

Jones TA, Zou J-Y, Cowan SW, Kjeldgaard M (1991) Improved methods for the building of protein models in electron density maps and the location of errors in these models. Acta Cryst A47:110–119

Karshikoff A, Ladenstein R (2001) Ion pairs and the thermotolerance of proteins from hyperthermophiles: a 'traffic rule' for hot roads. Trends Biochem Sci 26:550–556

Koehl P (2005) Electrostatics calculations: latest methodological advances. Curr Opin Struct Biol 16:142–151

Kraulis P (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of proteins. J Appl Crystallogr 24:946–950

Kumar S, Nussinov R (2004) Different roles of electrostatics in heat and cold: adaptation by citate synthase. ChemBioChem 5:280–290

Laue TM, Shah BD, Ridgeway TM, Pelletier SL (1992) Computer-aided interpretation of analytical sedimentation data for proteins. In: Harding SE, Rowe AJ, Horton JC (eds) Analytical ultracentrifugation in biochemistry and polymer science. Royal Society for Chemistry, London, pp 90–125

Lovell SC, Word JM, Richardson JS, Richardson DC (2000) The penultimate rotamer library. Proteins 40:389–408

Madura JD, Briggs JM, Wade RC, Davis ME, Luty BA, Ilin A, Antosiewicz J, Gilson MK, Bagheri B, Scott LR, McCammon JA (1995) Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian Dynamics program. Comp Phys Commun 91:57–95

Matsui I, Harata K (2007) Implication for buried polar contacts and ion pairs in hyperthermostable enzymes. FEBS J 274:4012–4022

Maugini E, Tronelli D, Bossa F, Pascarella S (2009) Structural adaptation of the subunit interface of oligomeric thermophilic and hyperthermophilic enzymes. Comp Biol Chem 33:137–148

Meyers JK, Pace CN, Scholtz JM (1995) Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. Protein Sci 4:2138–2148

Neet KK, Timm DE (1994) Conformational stability of dimeric proteins: quantitative studies by equilibrium denaturation. Protein Sci 3:2167–2174

Nielsen JE, McCammon JA (2003) Calculating pKa values in enzyme active sites. Prot Sci 12:1894–1901

Nielsen JE, Vriend G (2001) Optimizing the hydrogen-bond network in Poisson–Boltzmann equation-based pKa calculations. Proteins 43:403–412

Nielsen JE, Andersen KV, Honig B, Hooft RW, Kelbe G, Vriend G, Wade RC (1999) Improving macromolecular electrostatics calculations. Prot Eng 12:657–662

Pace CN (1990) Measuring and increasing protein stability. Trends Biotech 8:93–98

Pace CN, Scholtz JM (1998) Measuring the conformational stability of a protein. In: Creighton TE (ed) Protein structure: a practical approach, 2nd edn. IRL Press at Oxford University Press, Oxford, pp 299–321

Potapov V, Cohen M, Schreiber G (2009) Assessing computational methods for predicting protein stability upon mutation: good on average but not in the details. Prot. Eng Des Select 22:553–560

Rai N, Nöllmann M, Spotorno B, Tassara G, Byron O, Rocco M (2005) SOMO (SOlution MOdeler): differences between X-Ray- and NMR-derived bead models suggest a role for side chain flexibility in protein hydrodynamics. Structure 13:723–734

Robinson-Rechavi M, Alibés A, Godzik A (2006) Contribution of electrostatic interactions, compactness and quaternary structure to protein thermostability: lessons from structural genomics of *Thermotoga maritima*. J Mol Biol 356:547–557

Rodriguez R, Chinea G, Lopez N, Pons T, Vriend G (1998) Homology modelling, model and software evaluation: three related resources. CABIOS 14:523–528

Rumfeldt JAO, Galvagnion C, Vassall KA, Meiering EM (2008) Conformational stability and folding mechanisms of dimeric proteins. Prog Biophys Mol Biol 98:61–84

Russell RJ, Ferguson JM, Hough DW, Danson MJ, Taylor GL (1997) The crystal structure of citrate synthase from the hyperthermophilic archaeon *Pyrococcus furiosus* at 1.9 Å resolution. Biochemistry 36:9983–9994

Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. Biophys J 78:1606–1619

Schuck P, Perugini MA, Gonzales NR, Howlett GJ, Schubert D (2002) Size-distribution analysis of proteins by analytical ultracentrifugation: strategies and application to model systems. Biophys J 82:1096–1111

Shortle D (1995) Staphylococcal nuclease: a showcase of *m*-value effects. Adv Prot Chem 46:217–247

Spector S, Wang M, Carp SA, Robblee J, Hendsch ZS, Fairman R, Tidor B, Raleigh DP (2000) Rational modification of protein stability by the mutation of charged surface residues. Biochemistry 39:872–879

Srere PA, Brazil H, Gonen L (1963) The citrate condensing enzyme of pigeon breast muscle and moth flight muscle. Acta Chem Scand 17:S129–S134

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

Strickler SS, Gribenko AV, Gribenko AV, Keiffer TR, Tomlinson J, Reihle T, Loladze VV, Makhatadze GI (2006) Protein stability and surface electrostatics: a charged relationship. Biochemistry 45:2761–2766

Torrez M, Schultehenrich M, Livesay DR (2003) Conferring thermostability to mesophilic proteins through optimised electrostatic surfaces. Biophys J 85:2845–2853

Vielle C, Zeikus G (2001) Hyperthermophilic enzymes: Sources, uses and molecular mechanisms for thermostability. Microbiol Mol Biol Rev 65:1–43

Warshel A, Sharma PK, Kato M, Parson WW (2006) Modelling electrostatic effects in proteins. Biochim Biophys Acta 1764: 1647–1676

Weast RC (ed) (1973) CRC press handbook of chemistry and physics 53rd edition. Chemical Rubber Company, OH

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