ELSEVIER

Biophysical Chemistry 77 (1999) 195–208

Biophysical Chemistry

A thermodynamic analysis of a family of small globular proteins: SH3 domains

Vladimir V. Filimonov[a,b], Ana I. Azuaga[a], Ana R. Viguera[c], Luis Serrano[c], Pedro L. Mateo[a,*]

[a] *Department of Physical Chemistry and Institute of Biotechnology, Faculty of Sciences, University of Granada, 18071 Granada, Spain*
[b] *Institute of Protein Research of the Russian Academy of Sciences, Puschino, 142292 Moscow region, Russia*
[c] *EMBL, Meierhofstrasse 1, 69117 Heidelberg, Germany*

Received 28 December 1998; received in revised form 18 February 1999; accepted 18 February 1999

Abstract

The stability and folding thermodynamics of two SH3-domains, belonging to *Fyn* and *Abl* proteins, have been studied by scanning calorimetry and urea-induced unfolding. They undergo an essentially two-state unfolding with parameters similar to those of the previously studied α-spectrin SH3 domain. The correlations between the thermodynamic parameters (heat capacity increment, $\Delta C_{p,U}$, the proportionality factor, $\boldsymbol{m}$, and the Gibbs energy, ΔG_w^{298}) of unfolding and some integral structural parameters, such as polar and non-polar areas exposed upon domain denaturation, have been analyzed. The experimental data on $\Delta C_{p,U}$ and the $\boldsymbol{m}$-factor of the linear extrapolation model (LEM) obey the simple empirical correlations deduced elsewhere. The Gibbs energies calculated from the DSC data were compared with those found by fitting urea-unfolding curves to the LEM and the denaturant-binding model (DBM). The ΔG_w^{298} values found with DBM correlate better with the DSC data, while those obtained with LEM are systematically smaller. The systematic difference between the parameters calculated with LEM and DBM are explained by an inherent imperfection of the LEM.

Keywords: SH3 domains; Urea unfolding; Scanning calorimetry; Thermodynamic analysis; Folding; Stability

* Corresponding author. Tel.: +34-958-243333, fax: +34-958-272879, e-mail: pmateo@goliat.ugr.es

0301-4622/99/$ - see front matter
PII: S0301-4622(99)00025-3

1. Introduction

Spectrin homology 3 (SH3) domains are found in a large variety of proteins participating in cellular signal transduction [1–4]. These domains do not occupy fixed positions within their host proteins and can be expressed and purified independently [5]. Most of the SH3 domains recognized so far in various proteins have an overall length of approximately 65 amino acids, but some of them may consist of as many as 85 amino acids thanks to inserts into the loops connecting the main conservative elements of the secondary structure [6]. Nevertheless, despite the differences in length and amino-acid sequences (Fig. 1), the SH3 domains are structurally fairly homologous, especially in terms of their overall fold pattern [2,6,7]. It has also been proposed that all SH3 domains recognize proline-rich sequences folded into a PPII polyproline helix [8–11].

In addition to many structural and functional studies SH3 domains have also been characterized physico-chemically. Their smallness, their simple fold pattern, made up of β-structure elements, and their lack of either disulfide bridges or strongly bound ligands all make the SH3 domains very suitable for folding studies. In a previous work we have shown that Spc-SH3 essentially exhibits a two-state folding–unfolding transition in both thermodynamic and kinetic aspects [12].

In recent years SH3-domains have become the subject of considerable interest and the three-dimensional structure of several members of this family has already been determined. This, together with the high reversibility of their unfolding and the diversity of their primary structures, affords good opportunities for establishing correlations between protein structure and the energetics of its folding. We have extended our thermodynamic studies of Spc-SH3 to other two members of the family to show that both of them obey a two-state unfolding mechanism with thermodynamic parameters very similar to those of the spectrin domain. An attempt has also been made to provide a structural interpretation of the observed thermodynamic parameters within the framework of the existing thermodynamic models of protein folding.

2. Materials and methods

The Abl-SH3, Fyn-SH3 and Spc-SH3 proteins were cloned, expressed and purified as described elsewhere [9,13]. Fyn-SH3 contained 63 ($M_r = 7.0$ kDa) and Abl-SH3 64 ($M_r = 7.1$ kDa) residues. The concentrations were measured spectrophotometrically using the extinction coefficients reported elsewhere [9].

2.1. Differential scanning calorimetry

DSC was done as described elsewhere [12] using a DASM-4 microcalorimeter (Biopribor, Russia) with 0.5 ml cell volumes, at heating rates of 1

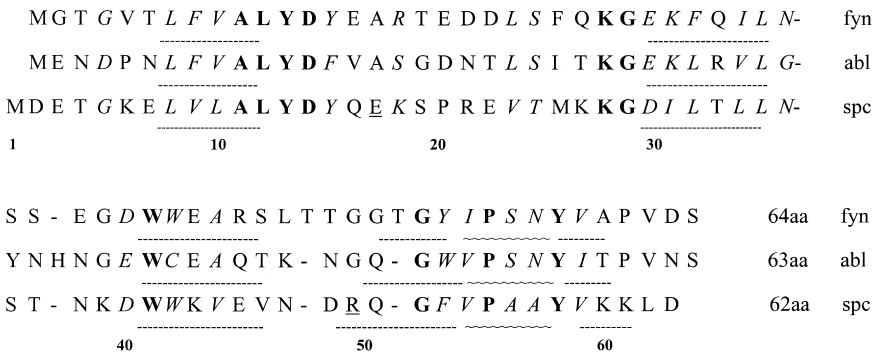

Fig. 1. Sequence and structure-based alignment of the SH3 domains studied. Solid letters and italics refer to the fully and semi-conserved positions. Dashed lines show the fragments involved in the conserved β-structure elements, whilst the wavy lines show the conserved loops. The Glu17 and Arg49 residues forming a salt-bridge in Spc-SH3 are underlined.

or 2 $K \cdot min^{-1}$ and protein concentrations of 2–4 $mg \cdot ml^{-1}$. Before the calorimetric experiments the samples were thoroughly dialyzed against 10 mM glycine, PIPES or phosphate to maintain the pH within the regions 2–3.5, 7 or 10.5.

DSC data were analyzed as described elsewhere [12] with the following modification. Unlike Spc-SH3, the stability of both Fyn-SH3 and Abl-SH3 was only slightly dependent upon pH and the unfolding of these domains always occurred above 320 K, where the heat capacity of the unfolded state is well approximated by a linear temperature dependence. Therefore, the heat capacities of the native and denatured states, $C_{p,N}$ and $C_{p,U}$, were represented by linear temperature functions, which implies that the $\Delta C_{p,U}$ values found by curve fitting were also linearly-dependent upon temperature. Nevertheless, the Gibbs energies of unfolding at 298 K, ΔG_w, were calculated using the quadratic dependence of $\Delta C_{p,U}$ on temperature found for Spc-SH3 [12], whereas constant $\Delta C_{p,U}$ values, deduced from the slopes of linear regressions through the ΔH_m vs. T_m, data were used elsewhere. As was pointed out by Viguera et al. [12] these slopes correspond to the average values of $\Delta C_{p,U}$ over the temperature range 310–350 K, i.e. to $\Delta C_{p,U}$ at approximately 330 K.

2.2. Urea denaturation

Urea solutions were prepared gravimetrically in volumetric flasks. For each data point 100 μl of the protein in sodium phosphate, pH 7.0, was mixed with 750 μl of a given denaturant solution, resulting in a final buffer concentration of 50 mM. The mixtures were left to equilibrate for at least 1 h (no changes in sample fluorescence were observed between 1 and 12 h of incubation). Fluorescence emission spectra were measured in an Aminco Bowman Series 2 luminiscence spectrometer. Excitation was at 290 nm with a 2 nm slit, while fluorescence was detected through an 8 nm slit. The measurements were carried out at 298 K at a protein concentration of 2.2 μM.

The normalized experimental dependencies of fluorescence upon urea molarity were analyzed by a non-linear curve fitting to the three models most frequently used for a two-state unfolding: LEM (linear extrapolation model), DBM (denaturant binding model) and BEM (binomial extrapolation model) [14–16]. The LEM is the simplest and most frequently used approach, corresponding to the equation

$$\Delta G_U = \Delta G_w + \boldsymbol{m}[c_{\mathrm{den}}] \tag{1}$$

where $[c_{\mathrm{den}}]$ is the molar concentration of the denaturant (urea, in our case), $\boldsymbol{m}$ is the proportionality factor and ΔG_w is the free-energy of unfolding at zero denaturant concentration.

The second approach, DBM, corresponds to the simplified identical- and independent-site model of denaturant binding

$$\Delta G_U = \Delta G_w + \Delta n_b \, RT \, \ln(1 + K_b a_{\mathrm{den}}) \tag{2}$$

Here a_{den} is the denaturant activity, which in the case of urea might be calculated from its molarity $[c_{\mathrm{den}}]$, by the empirical formula [14]

$$a_{\mathrm{den}} = 0.9815\,[c_{\mathrm{den}}] - 0.02978[c_{\mathrm{den}}]^2 + 0.00308[c_{\mathrm{den}}]^3 \tag{3}$$

and K_b and Δn_b, are the average denaturant binding constant and an increment in the effective number of binding sites at protein unfolding, respectively. This model provides more flexibility for curve fitting since it contains one additional adjustable parameter compared to LEM.

The third model, BEM, takes into consideration recent experimental results showing a binomial dependence of the unfolding rate constant upon urea, which turned out to be the same for barnase [15] and Spc-SH3 [16]. Taking into account this non-linearity the corrected equation for the Gibbs energy change is

$$\Delta G_U = \Delta G_w + \boldsymbol{m}[c_{\mathrm{den}}] - 0.008295[c_{\mathrm{den}}]^2 \tag{4}$$

In all three models the baselines were approximated by linear functions

$$Fl_N = a_N + b_N[c_{\mathrm{den}}] \tag{5}$$

$$Fl_U = a_U + b_U[c_{\mathrm{den}}] \tag{6}$$

Thus, the overall dependence of the normalized fluorescence on urea concentration is

$$Fl = [Fl_N + Fl_U K_U]/(1 + K_U) \quad (7)$$

where K_U is a two-state equilibrium constant

$$K_U = \exp(-\Delta G_U/RT) \quad (8)$$

2.3. Calculation of solvent-protected areas

The areas exposed to the solvent in the native and unfolded states were calculated using NACCESS software kindly provided by Dr S Hubbard (Biocomputing Group, EMBL). The program is based on the method of Lee and Richards [17]; a probe size of 1.4 Å, slice size of 0.05 Å and van der Waals radii suggested by Chothia [18] were used. The native structures were taken from the PDB files *pdb1shg.ent, pdb1abl.ent* and *pdb1shf.ent*. The unfolded states of the proteins were approximated by the extended β-structure, simulated with the Biopolymer module of INSIGHTII software (Biosym, California, USA) on INDIGO Workstation (Silicon Graphics, USA). The WHATIF program, kindly provided by Dr.G.Vriend (EMBL, Heidelberg), was also used for some structural analyses.

To estimate the m and $\Delta C_{p,U}$ from ΔASA_{tot} we used formulas deduced elsewhere [19]

$$\boldsymbol{m}_{\text{urea}} = (0.14 \pm 0.01) \quad \Delta ASA_{\text{tot}} \quad (9)$$

$$\Delta C_{p,U} = (0.19 \pm 0.01.) \quad \Delta ASA_{\text{tot}} \quad (10)$$

2.4. Curve simulation

The normalized fluorescence curves were simulated assuming that at zero urea concentration the unfolded state of barnase has a fluorescence intensity of approximately 20% of that of the native form [20]. The baselines and the unfolding curves were simulated with the equations described in the Materials and Methods using the parameters shown in the second line of Table 1. The curves were simulated for several ΔG_w values starting with the experimental value of 43.9 $\text{kJ}\cdot\text{mol}^{-1}$. In order to avoid any ambiguity in the native-state baselines the values of ΔG_w lower of than 25 $\text{kJ}\cdot\text{mol}^{-1}$ were not used. The LEM procedure was applied to the simulated curves without imposing any restriction over the baseline parameters, $\boldsymbol{m}$ or ΔG_w, i.e. all six parameters were adjusted independently for each curve.

Table 1
The results obtained by fitting to LEM the data simulated with the *DBM* (italics) using the parameters (solid letters) reported for barnase at 25°C (line A with corresponding references)

	$\Delta G_{w,DBM}$ $\text{kJ}\cdot\text{mol}^{-1}$	$\Delta G_{w,LBM}$ $\text{kJ}\cdot\text{mol}^{-1}$	$\boldsymbol{m}$ kJ/ mol·M	Δn_b	K_b M^{-1}	a_N	b_N M^{-1}	a_U	b_U M^{-1}
A	**43.9 ± .3** [31]	**37.3** [31]	**8.03** [31]	**53 ± 6** [31]	**0.088 ± 0.012** [31]	**1.0** [20]	**0.016** [23]	**0.2** [20]	**0.021** [23]
B	–	–	–	*53*	*0.09*	*1.0*	*0.02*	*0.2*	*0.02*
1	*44.0*	37.42	7.69	–	–	0.9996	0.0204	0.2018	0.0198
2	*35.0*	29.76	8.01	–	–	0.9991	0.0212	0.2021	0.0197
3	*30.0*	25.67	8.27	–	–	0.9998	0.0222	0.2020	0.0198
4	*25.0*	21.61	8.60	–	–	0.9986	0.0241	0.2017	0.0198
4a	*25.0*	21.90	8.70	–	–	1.0020	0.02_{fix}	0.2024	0.0197

Line B contains six parameters fixed during simulations, while the column $\Delta G_{w,DBM}$ corresponds to the assumed changes in the Gibbs energy. The lines 1–4 show the results of the best fits to the LEM model with all six parameters being adjustable, while the line 4a corresponds to the case with b_N fixed at 0.02 M^{-1}.

3. Results

3.1. DSC data

As with Spc-SH3, the temperature unfolding of the Abl and Fyn-SH3 domains under standard conditions used for DSC experiments (neutral pH, concentration 2–4 $mg \cdot ml^{-1}$, heating rate 2 $K \cdot min^{-1}$) is a reversible process. The second heatings of the samples after their cooling inside the calorimeter cell usually resulted in a less than 15% loss in the amplitude of the heat-absorbance peak when the heating was terminated immediately after completing denaturation (Fig. 2). In addition to protein aggregation, the irreversibility of the temperature-induced unfolding may be the result of other processes, such as the chemical modification of various protein groups. The slow development of these processes is accelerated by temperature and when accompanied by exothermic heat effects can result in a step-wise decline in the apparent heat capacity of the unfolded state [21]. This effect appears most clearly in the unfolding curves of Abl-SH3 recorded at alkaline pH (Fig. 3), where its unfolding, unlike that of Fyn-SH3 (Fig. 2), is completely irreversible. The second heating of Abl-SH3 at pH 10.5 does not reveal any trace of the transition peak (data not shown), while the third heating results in a slightly concave curve (Fig. 3), typical for the heat capacity of the unfolded state [22]. Since during DSC experiments no visible traces of aggregation were detected at any pH, the complete irreversibility of unfolding observed for Abl-SH3 at alkaline conditions should be attributed to the oxidation of the single SH-group of Cys40 present only in this domain. This cysteine occupies the place of the tryptophane residue conserved in other SH3 domains and imbedded in the hydrophobic core, so that the free SH-group, according to static surface accessibility calculations, is inaccessible to the solvent in the native state, but should become accessible upon denaturation. The importance of this sulfhydril group for the correct folding of Abl-SH3 has been confirmed in urea-unfolding experiments. We found that the omission of DTT from the Abl-SH3 solutions resulted in 0.4–1.4 $kJ \cdot mol^{-1}$ decrease in ΔG_w depending on the model used (Table 2). Unfortunately, however, the addition of thiol-reducing agents, such as DTT into solutions produces large thermal effects on heating and therefore they were not added to the DSC samples. Nevertheless, the unfolding of Abl-SH3 was fairly reversible at neutral and, more so, at acidic pH when the heating was stopped immediately after the transition. It seems, therefore, that under these conditions oxidation of the SH-group occurs very slowly and does not affect the DSC unfolding curves to any great extent.

The heat capacity functions of the native state, $C_{p,N}(T)$, found for the *Abl* and *Fyn* domains are very similar to that of Spc-SH3. The $\Delta C_{p,m}$ of Fyn-SH3 at $T_m = 343.6$ K (pH 7 and 2) has been found to be equal to 3.3 ± 0.4 $kJ \cdot K^{-1} \cdot mol^{-1}$, with a decrease, $\delta \Delta C_{p,U}$, close to -0.06 $kJ \cdot K^{-2} \cdot mol^{-1}$. This sloping down of $\Delta C_{p,U}$ completely compensates the increase in $C_{p,N}$ so that the apparent $C_{p,U}(T)$ at temperatures above 75°C appears with a zero or even a slightly negative slope. We could not determine $\Delta C_{p,m}$ precisely from the unfolding curves of Abl-SH3 due to the poorly reproducible sloping down of the apparent $C_{p,U}$ mentioned above. The $\Delta C_{p,U}$ can also be estimated from the slope of the $\Delta H_m(T_m)$ dependence. To do this with reasonable accuracy it is necessary to change T_m over a wide temperature range, which was not possible for Fyn-SH3 and Abl-SH3. It appears though that, within the limits of experimental error, the Fyn-SH3 data fit into the $\Delta H_m(T_m)$ correlation observed previously for Spc-SH3, while the Abl-SH3 data lie some 15 $kJ \cdot mol^{-1}$ below that line (Fig. 4). The latter shift is statistically reliable, although it is approximately equal to the absolute error of the ΔH_m determination. The $\Delta C_{p,U}$ value, estimated for Abl-SH3 from its ΔH_m vs. T_m correlation, is approximately 3.3 ± 0.4 $kJ \cdot K^{-1} \cdot mol^{-1}$, which is also close to that of spectrin SH3 (Table 2).

3.2. Unfolding by urea

The normalized urea-unfolding curves followed by fluorescence change at pH 7.0 are shown in Fig. 5. These curves have the same drawback as the temperature unfolding functions in that they

Table 2
The thermodynamic parameters of unfolding for various SH3-domains at 25°C

SH3-domains	LEM model		BEM model		DBM model			DSC			
	m, kJ/mol M	ΔG_w, kJ/mol	m, kJ/mol M	ΔG_w, kJ/mol	Δn_b	$K_b \times 27310^3$, M^{-1}	ΔG_w, kJ/mol	T_m, K	ΔH_m, kJ/mol	ΔG_w, kJ/mol	$\Delta C_{p,m}$, kJ/Kmol
SPC, pH 7.0	2.6 ± 0.1	13.6 ± 0.8	3.0 ± 0.1	15.0 ± 0.8	17.7 ± 1.0	86 ± 4	15.4 ± 0.8	339.0 ± 0.2	197 ± 10	15.6 ± 1.2	3.4 ± 0.1
SPC, pH 3.5	2.9 ± 0.1	11.6 ± 0.8	3.3 ± 0.1	13.0 ± 0.8	18.1 ± 1.2	91 ± 4	13.5 ± 0.8	336.0 ± 0.2	188 ± 10	13.9 ± 1.2	3.4 ± 0.1
ABL, pH 7.0											
−DTT	2.3 ± 0.1	14.7 ± 0.8	2.8 ± 0.1	16.2 ± 0.8	17.2 ± 1.2	86 ± 4	16.9 ± 0.8	341.5 ± 0.2	194 ± 10	15.0 ± 1.5	3.3 ± 0.4
+DTT	2.3 ± 0.1	15.1 ± 0.8	2.8 ± 0.1	17.6 ± 0.8	17.2 ± 1.2	86 ± 4	18.2 ± 0.8				
FYN, pH 7.0	2.7 ± 0.2	21.1 ± 1.5	3.3 ± 0.1	23.4 ± 1.0	18.9 ± 1.5	95 ± 5	24.0 ± 1.5	343.6 ± 0.2	233 ± 15	20.5 ± 2.0	3.3 ± 0.4[a]

The heat capacity increments correspond to the slopes of the linear regressions through ΔH_m points, while the $\Delta C_{p,U}$ of Fyn-SH3 (labeled with [a]) was found directly from the calorimetric curves. The Gibbs energies were calculated from DSC data using $\Delta C_{p,U}(T)$ function taken from Viguera et al. [12].

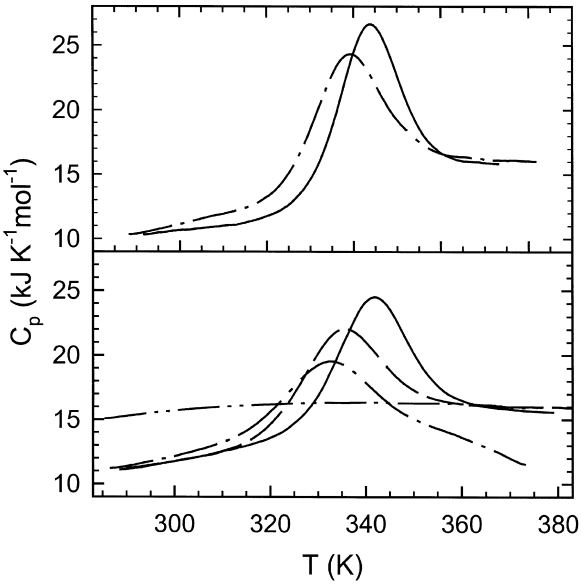

Fig. 2. Temperature dependencies of the apparent partial molar heat capacities of Fyn-SH3 (*upper panel*) and Abl-SH3 (*lower panel*) under various conditions. The data were obtained at pH 7 (solid lines), pH 10.5 (dash–dot lines) and pH 2.0 (dashed line). The dash–dot–dot line in the lower panel corresponds to the third heating of the Abl-SH3 sample at pH 10.5.

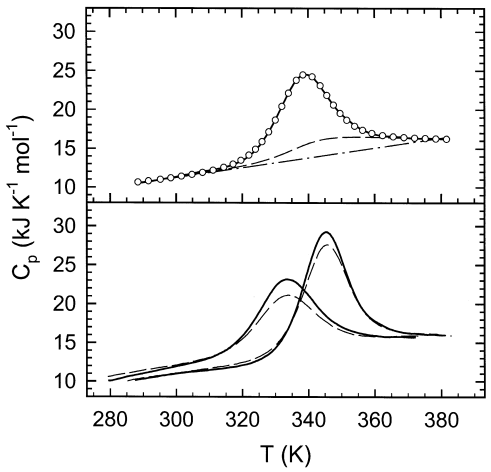

Fig. 3. Temperature dependencies of the partial molar heat capacity for Fyn- and Abl-SH3 domains: *upper panel*, *unfolding* curve of Fyn-SH3 in 10 mM glycine buffer, pH 10.5 (solid line) and its best fitting to the two-state model (open circles); the dash–dot line shows the heat capacity of the native state, as calculated by fitting, and the dashed line corresponds to the temperature dependence of the partial internal, or chemical, heat capacity, C_p^{int}, of the protein; *lower panel,* the heat capacities of Abl-SH3 at pH 2.0 (curves on the left) and Fyn-SH3 at pH 7.0 (curves on the right); solid lines show the unfolding profiles of the first heatings whilst dashed lines correspond to the reheatings.

are very broad and in some cases do not contain important information about the baselines. This complicates their analysis even within the simplest versions of a two-state model and results in increased errors in the thermodynamic parameters independently of the model used to fit the data. In such a situation some reasonable restrictions must be sought to impose upon the base lines. According to a statistical analysis made in [23] over a very large number of urea-unfolding curves of barnase and its mutants (with three tryptophane residues compared to two in SH3 domains) the average values for the slopes, b_N and b_U, of the normalized curves were 0.016 and 0.02 M^{-1}, respectively. Statistical analysis of the unfolding curves accumulated for our SH3 domains (in particular for Spc-SH3 and its modifications) indicates that the baselines for SH3 domains have on average positive slopes similar to those of barnase. It seems thus that the baseline slopes reflect mostly the dependence of the tryptophane fluorescence upon urea and when different proteins have equal amounts of the chromophores situated in a very similar surrounding the normalized baselines should be virtually the same. This admits the possibility of fitting several unfolding curves simultaneously and thus decreasing the number of independent adjustable parameters.

We have assumed that the slope, b_U, of the unfolded-state base lines is similar for all SH3-proteins since they have the same number of tryptophane residues, while the slope b_N is an adjustable parameter since the disposition of chromophores in the native structure of Abl-SH3 is not the same as in the others. The exact value of the slope b_U has been obtained in experiments with unstable Spc-SH3 mutants (unpublished results). The data on urea-unfolding of SH3-domains are shown in Table 2.

4. Discussion

The data presented above clearly show that the

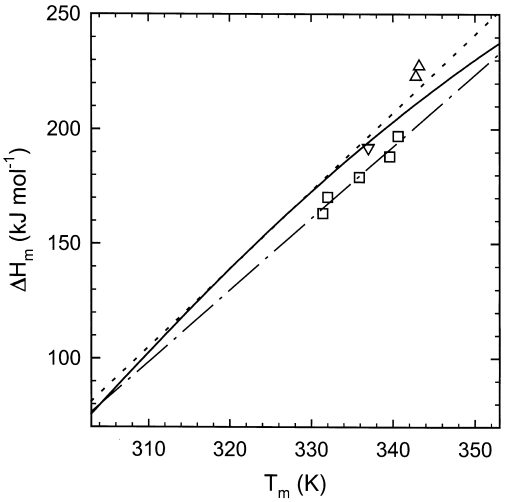

Fig. 4. The dependence of the unfolding enthalpy, ΔH_m, on the unfolding temperature, T_m, for the SH3 domains studied so far. The solid line is the best fitting of the experimental enthalpy values of Spc-SH3 to a third-order function corresponding to $\Delta C_{p,U} = -5.95 + 0.087T - 1.8 \cdot 10^{-4} T^2$ [12], while the dotted line reflects the linear regression through these data with the slope of 3.4 $kJ \cdot K^{-1} \cdot mol^{-1}$. The open squares show the data obtained for Abl-SH3 at pH 7.0, 3.5, 2.5, 2.1 and 2.0 (in the sequence of decreasing stability with the dash–dot line representing the linear regression through these points with the slope of 3.3 $kJ \cdot K^{-1} \cdot mol^{-1}$). The upright triangles show the data obtained for Fyn-SH3 at pH 7.0 and 2.0, while the upside-down is at pH 10.5.

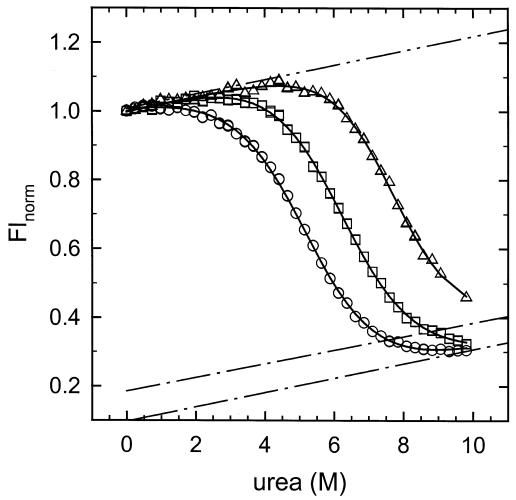

Fig. 5. The changes in the relative fluorescence induced by urea at pH 7.0 in Spc-SH3 (circles), Abl-SH3 (squares) and Fyn-SH3 (triangles). The three data sets were fitted simultaneously to the LEM model (solid lines); the dash–dot–dot line, with a slope of 0.020, shows the common best-fit base line for the native states of the three domains, while dash–dot lines correspond to the unfolded states of Fyn-SH3 (upper line), and of Spc-SH3 and Abl-SH3 domains (lower line), respectively.

native structures of three SH3 domains with very similar folds have remarkably different stabilities (ΔG_U) under identical solvent conditions, despite that such parameters as unfolding enthalpies, heat capacity increments, or ***m***-values are very similar (Table 2). In addition to the Gibbs energy at pH 7.0 there is another important difference between both Abl-SH3 and Fyn-SH3, and Spc-SH3: the stability of the two former depend only slightly on pH. Since a sharp dependence of the stability on pH is usually attributed to some charged side-chains with abnormal pK values, Spc-SH3 should have in its native structure at least one acidic group either protected from the solvent or forming a buried salt-bridge, whereas the other two domains should not. In fact, there are two salt-bridges within the Spc-SH3 structure, Glu7-Arg60 and Glu17-Arg49, while the COOH group of Glu22 is partially protected from the solvent. A visual inspection of the Fyn-SH3 structure, however, does not reveal any salt-bridge, perhaps because all four aminoacids forming salt-bridges within Spc-SH3 are replaced by uncharged groups both in Fyn- and Abl-SH3. As the stability of protein structure depends on multiple factors and still cannot be predicted, it is not clear why Fyn-SH3 has so much higher stability than, for example, Abl-SH3. Although the unfolding enthalpy is much less dependent on some factors, such as electrostatic interactions, than the entropy, and, therefore, ΔH_U might be estimated from the structural parameters with a higher accuracy, it is not clear either why Abl-SH3 has a lower unfolding enthalpy at the same T_m than the other two domains.

One of the technical problems encountered in this study concerned the fitting of the wide urea-unfolding curves, particularly in the case of Fyn-SH3. The applicability of the fixed-slope procedure used here might be questioned, if the environments of the Trp residues in the denatured states of the different SH3 domains are dissimilar

and therefore respond differently to urea concentration. Nevertheless, $\Delta G_w = 25.1 \pm 2.5$ kJ·mol^{-1} has been recently found by Plaxco et al. [24] from the guanidinium hydrochloride unfolding of Fyn-SH3, which is in a good agreement with the values obtained here, thus supporting our assumption.

4.1. Correlation of the free energy of unfolding with empirically derived solvent-accessible parameters

A number of empirical correlations between the thermodynamic parameters of protein unfolding and the general structural characteristics of globular proteins have been proposed [19,25–28]. Two major difficulties were encountered when we tried to use these algorithms for a structural interpretation of our experimental results. Firstly, despite the fact that for calculating the accessible polar (ASA_{pol}) and non-polar (ASA_{np}) areas in globular proteins most of the authors use the same algorithm, ACCESS, based on the theoretical model of Lee and Richards [17], there are remarkable differences between the ASA values reported by different authors for the same protein. For example, according to Spolar et al. [25] the polar and non-polar surfaces exposed upon the unfolding of ribonuclease A are 5815 and 4475 Å^2, whilst, according to Myers et al. [19], they are 6819 and 3638 Å^2, and more recent calculations by Makhatadze and Privalov [29] have resulted in 5273 and 4141 Å^2. Freire et al. did not publish their ASA values, but Myers et al. [19] pointed out that the difference between their calculations and those of Murphy and Freire [26] could be as high as 25%. Two groups have reported ΔASA parameters for Spc-SH3. According to Makhatadze and Privalov [29], the ΔASA_{np} and $\Delta ASA_{\mathrm{pol}}$ for this domain are equal to 2760 and 1444 Å^2, while Myers et al. [19] have published values of 3439 and 1121 Å^2, respectively. The reasons for such a large discrepancy in this case are not clear, since it appears that both groups used the same software as well as the same extended conformation model for the unfolded state. It would seem that all these inconsistencies cannot be eliminated until clear standards are introduced into exposed-area calculations. Meanwhile we have used the simplest empirical formulas (9–10), published by Myers et al. [19]. It must be emphasized that formula Eq. (9) defines a correlation between the $\Delta ASA_{\mathrm{tot}}$ and $\boldsymbol{m}$ values found by different authors using the LEM model. According to these formulas, the $\Delta ASA_{\mathrm{tot}} = 4560$ Å^2 reported by Myers et al. [19] for Spc-SH3 should correspond to $\Delta C_{p,U} = 3.6$ kJ·K^{-1}·mol^{-1} and $\boldsymbol{m} = 2.67$ kJ·mol^{-1}·M^{-1}, whereas the experimental values are 3.4 ± 0.1 kJ·K^{-1}·mol^{-1} ([12], this work) and 2.6 ± 0.1 kJ·mol^{-1}·M^{-1} (Table 2). According to our calculations $\Delta ASA_{\mathrm{tot}} = 4171$ Å^2 for Abl-SH3, which should correspond to $\boldsymbol{m} = 2.44$ kJ·mol^{-1}·M^{-1} (experimental 2.3 ± 0.1) and $\Delta C_{p,U} = 3.3$ kJ·K^{-1}·mol^{-1} (experimental 3.3 ± 0.4). Finally, for Fyn-SH3 we found $\Delta ASA_{\mathrm{tot}} = 4260$ Å^2, which should correspond to $\boldsymbol{m} = 2.5$ kJ·mol^{-1}·M^{-1} (experimental 2.7 ± 0.1) and $\Delta C_{p,U} = 3.4$ kJ·K^{-1}·mol^{-1} (experimental 3.3 ± 0.4). Therefore for this family of proteins the empirical approximation describes the observed experimental parameters reasonably well.

4.2. Calculating the free energy of unfolding from extrapolation of the experimental data

A very important issue, which has been widely discussed in the literature, concerns the accuracy of the LEM model and its applicability for calculating ΔG_w from the denaturant unfolding curves. Although very small globular proteins, such as SH3-domains are not the best subjects for answering this question because of their very broad unfolding curves, some observations can still be made. The best choice among the SH3-domains studied so far is Spc-SH3, for which the most complete set of DSC and urea-unfolding data has been accumulated ([9,12,30], present work). The ΔG_w values calculated for Spc-SH3 and its mutants with the LEM model appear to be somewhat lower (10–15%) than the calorimetric ones, while DBM and BEM give Gibbs energy values coinciding with the DSC data (Table 2). This observation is in line with the conclusions drawn in other publications [15,16,31]. In their study into barnase unfolding Johnson and Fersht [31] have found that the $\boldsymbol{m}$-factor of the LEM is not constant but appear to depend on urea concentration

in accordance with an empirical equation

$$\boldsymbol{m} = (2.65 \pm 0.05) + (0.08 \pm 0.01)[c_{\text{urea}}] \qquad (11)$$

and that the application of the DBM model to the barnase data gives much closer correspondence between the free energy values. For pH 6.3 they report the following values: $\Delta G_w = 44$ kJ·mol^{-1}, $\Delta n_b = 56 \pm 6$, $K_b = 0.088 \pm 0.012$ M^{-1}. This latter K_b agrees well with our average value found for the SH3-domains (Table 2) and coincides with the independent estimations of the affinity constant for a binding-unit consisting of two peptide groups [14]. Therefore, it appears that for the majority of typical globular proteins the number of adjustable parameters within the DBM model might be reduced to two by fixing K_b at 0.09 M^{-1}. If this average value of K_b is used for SH3 domains the best-fit value of Δn_b for all of them is approximately 18 ± 1. When K_b is fixed the DBM model has the same number of adjustable parameters as LEM, but the former has the advantage of estimating ΔG_w more closely to the DSC data. In addition, Δn_b has a clearer structural interpretation than $\boldsymbol{m}$ and might be effectively used for characterizing urea-induced protein unfolding. The BEM model is based on the empirical kinetic observation that the rate constants of barnase [15] and Spc-SH3 [16] show a non-linear dependence upon urea. This method gives better approximations of Gibbs energy than the LEM model for the reasons explained below.

The dependence of the $\boldsymbol{m}$-factor on urea concentration, revealed and analyzed by Johnson and Fersht [31] might reflect either an intrinsic drawback of the LEM, or some real differences in denaturant-binding to the native or unfolded state of proteins with different stability. This latter hypothesis has been supported by a number of authors who suggested that the reversed correlation between the mid-point of unfolding transition, $[\boldsymbol{c}_{\text{den}}]_{1/2}$, and the $\boldsymbol{m}$-factor might arise from a correlation between the compactness of the unfolded state and overall protein stability [32,33]. To discriminate between these two options we made an analysis: urea-unfolding curves simulated with DBM model were subsequently fitted using LEM. The parameters reported by Fersht's group for barnase were chosen for curve simulating because the data for barnase are most precise and complete. Furthermore, the unfolding curves of barnase are much sharper than those of SH3 domains, which helps to resolve the problem of the baselines. The results of this modeling are set out in Fig. 6 and Table 1, where it can be seen that the LEM model fits the simulated data perfectly and what is more with small errors in the baseline parameters.

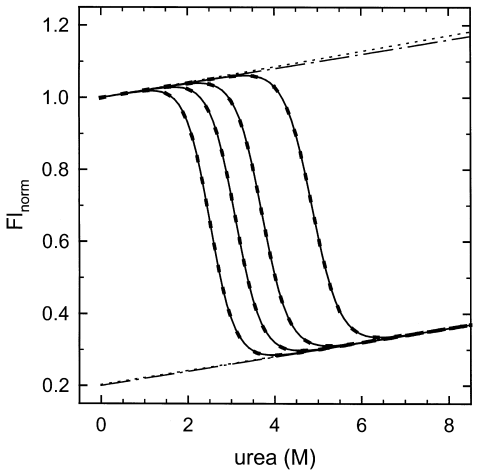

Fig. 6. The barnase unfolding curves simulated with the DBM (solid lines) using the parameters specified in Table 1, and their best fits to the LEM (solid dots). The dash–dot lines show the simulated baselines, while the dotted lines correspond to the best-fit baselines for the case of $\Delta G_w = 35$ kJ mol^{-1}.

The best-fit values of ΔG_w are, however, remarkably smaller than those assumed for DBM simulations. The relative deviation in ΔG_w is as high as 13% and almost independent of overall protein stability, $[\boldsymbol{c}_{\text{den}}]_{1/2}$. On the other hand, our ΔG_w and $\boldsymbol{m}$ values, calculated with non-corrected LEM, correspond well with those reported by Fersht's group for barnase (Table 1). This coincidence is hardly accidental and proves the applicability of our simulation/fitting analysis.

Table 1 shows that the $\boldsymbol{m}$-factor found by fitting the simulated curves decreases as ΔG_w increases. Since within LEM the ΔG_w is directly proportional to $[\boldsymbol{c}_{\text{urea}}]_{1/2}$ we plotted the correlation between $\boldsymbol{m}$ and the Gibbs energy together

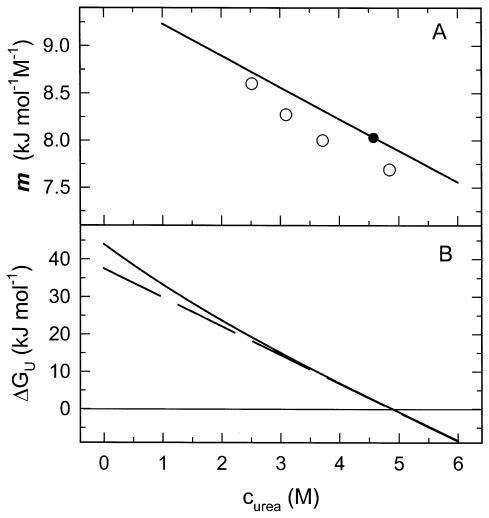

Fig. 7. Panel A: The correlation between $[c_{urea}]_{1/2}$ and ***m***-factor. The filled circle shows the data of Johnson and Fersht, [31]. The solid line corresponds to the equation 1a and the open circles to our best-fit parameters specified in Table 1. Panel B: The dependencies of the ΔG_U on urea concentration calculated for barnase at maximum stability Table 1 with LEM (dashed line) and DBM (solid line).

with the data of Johnson and Fersht [31]. It can be seen in Fig. 7 that the ***m***-factors found by modeling conform well with the empirical dependence reported by Johnson and Fersht. The conclusion is clear: fitting the same experimental data with LEM and DBM must give different results and the former model should always give smaller values of ΔG_w. The reason for this lies in the replacement of the term $RT\Delta n_b ln(1 + K_b a_{den})$ in DBM (Eq. (2)) by the term $\boldsymbol{m}c_{den}$ in LEM (Eq. (1), see Materials and Methods). Since a two-state model with the thermodynamic parameters of each state independent of urea has been assumed in all simulations the ***m***-factor of LEM should depend upon denaturant concentration but not upon protein stability or other thermodynamic parameters of the native and unfolded states. This drawback of LEM might be reduced by introducing a second-order term as suggested by Johnson and Fersht, but this correction must depend upon the Δn_b and K_b values, which are specific for each protein.

The conclusion about the low accuracy of the LEM is based on the comparison of the LEM-derived Gibbs energies with those determined from the DSC data by long extrapolations. Our analysis is, however, supported by an independent evidence arising from the hydrogen-exchange experiments monitored by NMR. Among other parameters, this technique allows to measure the Gibbs energies of the protein unfolding in the region of denaturant concentration invisible to classical unfolding experiments [34,35]. The unfolding Gibbs energy measured by hydrogen exchange is a non-linear function of the denaturant concentration, clearly deviating from the linear extrapolation at low $[c_{den}]$. For cytochrome *c* the ΔG_w measured by hydrogen exchange is approximately 20% higher than the LEM-extrapolated value [34], while the shape of the ΔG_U dependence on GuHCl concentration closely resembles the function simulated by us for barnase with DBM (Fig. 7B).

This means that neither LEM nor its corrected version BEM have any advantage over DBM. The former because it gives incorrect Gibbs energies and the latter because it has the same number (three) of adjustable parameters as DBM. Furthermore, whereas both K_b and Δn_b of DBM are open to clear structural and thermodynamic interpretation, the physical meaning of the ***m***-factor and of its correction are not so obvious.

Another complication concerns the apparent pH-dependence of the ***m***-factors calculated either with the LEM or BEM models for Spc-SH3 (Table 2). This dependence was found by Pace et al. [36,37] with barnase and interpreted in terms of an increasing charge repulsion expanding the unfolded state and thus facilitating the denaturant binding, which must be reflected at least in the increased Δn_b values. Our analysis of the Spc-SH3 data with the DBM model does not, however, show such significant changes with pH in either Δn_b, or K_b. This observation leads to the question as to whether other factors, such as a possible non-linearity of the baselines or their dependence upon pH might be of greater importance than has been assumed until now.The free-energy values calculated for Fyn-SH3 and Abl-SH3 with LEM seem to agree with the calorimetric ones better than do those for Spc-SH3. It must be emphasized, however, that both calorimetric and

urea-unfolding data for the former two domains are less complete and accurate than those for Spc-SH3. The exact temperature dependencies of the heat capacity for Fyn-SH3 and Abl-SH3 are unknown, which, when bearing in mind the high stability of these domains, must increase the error of extrapolating the Gibbs energy to 298 K.

5. Conclusions

Our calorimetric and urea-unfolding data of the SH3 domains agree with each other reasonably well and also correlate with some structural parameters. Nevertheless, in the light of our analysis it might be concluded that further progress in establishing a better correlation between structural and thermodynamic parameters depends both on increasing the accuracy of the experimental data and improving the calculation algorithms, primarily, by introducing clear and realistic standards. The important factors to be taken into consideration when using the simple empirical approximations are the solvent conditions (pH, ionic strength, etc.) under which the experimental data have been obtained.

Unfortunately, increasing the accuracy of DSC data on protein unfolding has various constraints. Even in the case of complete reversibility of unfolding the best scanning calorimeters do not allow us to determine the $\Delta C_{p,U}$ with an error less than 10%. This uncertainty together with temperature dependence of the $\Delta C_{p,U}$ make the margin of errors of ΔG_w determination from DSC data even higher than 10%. Furthermore, not only the SH3 domains but many other globular proteins, which are the most widely used subjects for thermodynamic studies, have very strong correlation between $\Delta ASA_{\mathrm{pol}}$ and ΔASA_{np}, each of them being simply proportional to the molecular weight. These factors suggest that isothermal calorimetry, applied to the studies of protein–protein or protein–ligand interactions and accompanied by an adequate structural analysis, can provide more precise and detailed correlations between structural and thermodynamic parameters.

6. Nomenclature

- Spc-SH3, Fyn-SH3 and Abl-SH3: the SH3-domains from the α-spectrin, *Fyn* and *Abl* proteins, respectively;
- DSC; differential scanning calorimetry;
- $\Delta C_{p,U}$: increment of the partial molar heat capacity upon unfolding;
- ΔH_m and $\Delta C_{p,m}$: increments of the enthalpy and heat capacity at the transition point T_m;
- ΔG_U and ΔG_w: the Gibbs energy change upon unfolding and its change at 298 K in water;
- GuHCl: guanidinium hydrochloride;
- DTT: dithiotreitol;
- LEM: linear extrapolation model;
- DBM: denaturant-binding model;
- BEM: binomial extrapolation model, which takes into account the non-linear dependence of the protein unfolding rate with urea;
- $\Delta ASA_{\mathrm{tot}}$: change in solvent-accessible surface area;
- ΔASA_{np} and $\Delta ASA_{\mathrm{pol}}$: the non-polar and polar contributions to $\Delta ASA_{\mathrm{tot}}$; and
- $\boldsymbol{m}$: the proportionality factor of LEM.

Acknowledgements

This work has been supported by INTAS Grant 93-007, DGICYT Grant PB96-1443 (Spain) and EU Grants CT96-0013 and CT97-2180. A.I.A and A.R.V. acknowledge a predoctoral fellowship from Junta de Andalucia (Spain) and a postdoctoral fellowship from the European Union, respectively. VVF acknowledges a sabbatical fellowship from the Spanish Government. The authors are thankful to Dr John Trout for revising the English text.

References

[1] B.J. Mayer, M. Hamaguchi, H. Hanafusa, A novel viral oncogene with structural similarity to phospholipase C. Nature 332 (1988) 272–275.
[2] A. Musacchio, T. Gibson, V.P. Lehto, M. Saraste, SH3 — An abundant protein domain in search of a function. FEBS Lett. 307 (1992) 55–61.

[3] V.P. Lehto, V.M. Wassenius, P. Salven, M. Saraste, Transforming and membrane proteins. Nature 334 (1988) 388.

[4] Y. Pawson, G.D. Gish, SH2 and SH3 domains: from structure to function. Cell 71 (1992) 359–362.

[5] A. Mussacchio, M.E.M. Noble, R. Pautit, R. Wierenga, M. Saraste, Crystal structure of a Src-homology 3 (SH3) domain. Nature 359 (1992) 851–855.

[6] S. Koyama, H. Yu, D.C. Dalgarno, T.B. Shin, L.D. Zydowsky, S.L. Schreiber, Structure of the P13K SH3 domain and analysis of SH3 family. Cell 72 (1993) 945–952.

[7] D. Kohda, H. Terasawa, S. Ichikawa, K. Ogura, H. Hatanaka, V. Mandiyan, A. Ullrich, J. Schlessinger, F. Inagaki, Solution structure and ligand-binding site of the carboxy-terminal SH3 domain GRB2. Structure 2 (1994) 1029–1040.

[8] R. Ren, B.J. Mayer, P. Cicchetti, D. Baltimore, Identification of a 10-aminoacid proline-rich SH3 binding site. Science 259 (1993) 1157–1161.

[9] A.R. Viguera, J.R.L. Arrondo, A. Musacchio, M. Saraste, L. Serrano, Characterization of the interaction of natural proline-rich peptides with five different SH3 domains. Biochemistry 33 (1994) 10925–10933.

[10] W.A. Lim, R.O. Fox, F.M. Richards, Stability and peptidic binding affinity of an SH3 domain from the *Caenorhabiditis elegans* signaling protein Sem-5. Protein Sci. 3 (1994) 1261–1266.

[11] M. Pisabarro, A. Ortiz, A.R. Viguera, F. Gago, L. Serrano, Molecular modeling of the interaction of polyproline-based peptides with the Abl-SH3 domain: rational modification of the interaction. Protein Eng. 7 (1994) 1455–1462.

[12] A.R. Viguera, J.C. Martinez, V.V. Filimonov, P.L. Mateo, L. Serrano, Thermodynamic and kinetic analysis of the SH3 domain of spectrin shows a two-state folding transition. Biochemistry 33 (1994) 2142–2150.

[13] W.A. Lim, F.M. Richards, R.O. Fox, Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. Nature 372 (1994) 375–379.

[14] C.N. Pace, Determination and analysis of urea and guanidine hydrochloride denaturation curves. Methods Enzymol. 131 (1986) 266–280.

[15] A. Matouschek, J.M. Matthews, C.M. Johnson, A.R. Fersht, Extrapolation to water of kinetic and equilibrium data for the unfolding of barnase in urea solutions. Protein Eng. 7 (1994) 1089–1095.

[16] A.R. Viguera, M. Wilmanns, L. Serrano, Different folding transition states could result in the same native structure. Nat. Struct. Biol. 3 (1996) 874–880.

[17] B.K. Lee, F.M. Richards, The interpretation of protein structures: estimation of static accessibility. J. Mol. Biol. 55 (1971) 379–400.

[18] C. Chothia, The nature of the accessible and buried surfaces in proteins. J. Mol. Biol. 105 (1976) 1–14.

[19] J.K. Myers, C.N. Pace, J.M. Scholtz, Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. Prot. Sci. 4 (1995) 2138–2148.

[20] J.T. Kellis, Jr., K. Nyberg, A.R. Fersht, Energetics of complementary side-chain packing in a protein hydrophobic core. Biochemistry 28 (1989) 4914–4922.

[21] J.C. Martinez, V.V. Filimonov, P.L. Mateo, G. Schreiber, A.R. Fersht, A calorimetric study of the thermal stability of barstar and its interaction with barnase. Biochemistry 34 (1995) 5224–5233.

[22] P.L. Privalov, E.I. Tiktopulo, S.Yu. Venyaminov, Yu.V. Griko, G.I. Makhatadze, N.N. Khechinashvili, Heat capacity and conformation of proteins in the denatured state. J. Mol. Biol. 205 (1989) 737–750.

[23] A. Horovitz, J.M. Matthews, A.R. Fersht, α-helix stability in proteins. Factors that influence stability at an internal position. J. Mol. Biol. 227 (1992) 560–568.

[24] K.W. Plaxco, J.I. Giujarro, C.J. Morton, M. Pitkeathly, I.D. Campbell, C.M. Dobson, The folding kinetics and thermodynamics of the Fyn-SH3 domain. Biochemistry 37 (1998) 2529–2537.

[25] R.S. Spolar, J.R. Livingstone, M.T. Record, Jr., Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. Biochemistry 31 (1992) 3947–3955.

[26] K.P. Murphy, E. Freire, Thermodynamics of structural stability and cooperative folding behavior in proteins. Adv. Protein Chem. 43 (1992) 313–361.

[27] G.I. Makhatadze, P.L. Privalov, Contribution of hydration to protein folding thermodynamics. I. The enthalpy of hydration. J. Mol. Biol. 232 (1993) 639–659.

[28] J. Gomez, E. Freire, Thermodynamic mapping of the inhibitor site of the aspartic protease endothiapepsin. J. Mol. Biol. 252 (1995) 337–350.

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

[30] A.R. Viguera, F.J. Blanco, L. Serrano, The order of secondary structure elements does not determine the structure of a protein but does affect its folding kinetics. J. Mol. Biol. 247 (1995) 670–681.

[31] C. Johnson, A.R. Fersht, Protein stability as a function of denaturant concentration: the thermal stability of barnase in the presence of urea. Biochemistry 34 (1995) 6795–6804.

[32] C.K. Smith, Z. Bu, K.S. Anderson, J.M. Sturtevant, D.M. Engelman, L. Regan, Surface point mutations that significantly alter the structure and stability of a protein's denatured state. Protein Sci. 5 (1996) 2009–2019.

[33] E.M. Nicholson, J.M. Scholtz, Conformational stability of the *Escherichia coli* HPr protein: test of the linear extrapolation method and a thermodynamic characterization of cold denaturation. Biochemistry 35 (1996) 11369–11378.

[34] S.W. Englander, L. Mayne, Y. Bai, T.R. Sosnick, Hydrogen exchange: the modern legacy of Liderstrøm-Lang. Protein Sci. 6 (1997) 1101–1109.
[35] Y. Bai, J.S. Milne, L. Mayne, S.W. Englander, Protein stability parameters measured by hydrogen exchange. Proteins: Struct. Funct. Genet. 20 (1994) 4–14.
[36] C.N. Pace, D.V. Laurents, J.A. Thomson, pH dependence of the urea and guanidine hydrochloride denaturation of ribonuclease A and ribonuclease T1. Biochemistry 29 (1990) 2564–2572.
[37] C.N. Pace, D.V. Laurents, R.E. Erickson, Urea denaturation of barnase: pH dependence and characterization of the unfolded state. Biochemistry 31 (1992) 2728–2734.