Eur. J. Biochem. **263**, 246–253 (1999) © FEBS 1999

A thermodynamic study of the 434-repressor N-terminal domain and of its covalently linked dimers

Javier Ruiz-Sanz[1], András Simoncsits[2], Imre Törö[2], Sandor Pongor[2], Pedro L. Mateo[1] and Vladimir V. Filimonov[1,3]

[1]*Department of Physical Chemistry and Institute of Biotechnology, Faculty of Sciences, University of Granada, Spain;* [2]*International Centre for Genetic Engineering and Biotechnology (ICGEB), Trieste, Italy;* [3]*Institute of Protein Research of the Russian Academy of Sciences, Puschino, Moscow Region, Russia*

The isolated N-terminal 1–69 domain of the 434-phage repressor, R69, and its covalently linked (head-to-tail and tail-to-tail) dimers have been studied by differential scanning microcalorimetry (DSC) and CD. At neutral solvent conditions the R69 domain maintains its native structure, both in isolated form and within the dimers. The stability of the domain depends highly upon pH within the acidic range, thus at pH 2 and low ionic strength R69 is already partially unfolded at room temperature. The thermodynamic parameters of unfolding calculated from the DSC data are typical for small globular proteins. At neutral pH and moderate ionic strength, the domains of the dimers behave as two independent units with unfolding parameters similar to those of the isolated domain, which means that linking two R69 domains, either by a long peptide linker or by a designed C-terminal disulfide bridge, does not induce any cooperation between them.

Keywords: thermal stability; differential scanning microcalorimetry; circular dichroism; domain stability; interdomain interactions.

Many regulatory DNA-binding proteins, including the repressor from the 434-phage, are homodimers consisting of two globular DNA-binding domains and polypeptide fragments responsible for dimerization [1]. Although within the 434-repressor–DNA complex the DNA-binding domains are neighbours and some contacts between them are revealed by structural analysis [2], the strength and functional importance of these contacts remains in question.

Combining protein domains via recombinant or synthetic linkers is a common strategy for designing artificial proteins [3,4]. The rationale underlying such experiments is that the structure of the individual domains would remain intact upon transposing them from a given protein to a completely different framework, but the functional properties of such a chimera could differ from those of the native structure. Among other things, this strategy allows us to estimate the strength of the inherent interdomain interactions and, because the domains can be combined in different ways providing symmetric (tail-to-tail) or asymmetric (head-to-tail) constructs with linkers of varying lengths, it might also be possible to find out how important the symmetry of the native structure is to its function.

In a previous study we designed an artificial protein framework based on two, covalently linked DNA-binding domains R69 [5]. This single-chain architecture allowed us to develop DNA-binding proteins with altered specificity, with a DNA-binding affinity comparable to or even exceeding that of the natural phage-434 repressor [6,7]. As both the 3D structure [8,9] and folding properties [10] of the constituent R69 domain are known, in this work we used this system as a model for comparing the folding and stability properties of free and linker-connected protein domains. In one of the tandem constructs, RR69, a linker of 20 amino acids connects the N- and C-terminal residues of the two constituent R69 domains [5,6]. In another construct, $[R69C]_2$, the two R69 domains are linked by a disulfide bridge formed between extra cysteine residues added to the C-termini of both domains. Here we show, using differential scanning microcalorimetry (DSC) and CD spectroscopy, that the R69 domains in the covalent dimers fold independently under close-to physiological conditions and have thermodynamic properties similar to those of the isolated domains, which unfold via a two-state mechanism. When conditions are far from physiological ones (acidic pH and low ionic strength), the positively charged domains within the dimers destabilize each other, this repulsive effect being stronger in $[R69C]_2$, probably due to its short interdomain linker.

MATERIALS AND METHODS

The N-terminal DNA-binding domain (amino acids 1–69) of the 434-phage repressor, R69 and its covalently linked head-to-tail tandem, RR69, were expressed and purified as described elsewhere [5]. The tail-to-tail cross-linked dimer $[R69C]_2$ was obtained by expressing gene *r69c*. The new

Correspondence to P. L. Mateo, Department of Physical Chemistry, Faculty of Sciences, University of Granada, 18071 Granada, Spain.
Fax: + 34 958 272879, Tel.: + 34 958 243333,
E-mail: pmateo@goliat.ugr.es

Abbreviations: DSC, differential scanning microcalorimetry; ASA, solvent-accessible surface area in $Å^2$; ΔASA_{tot}, total change in solvent-accessible surface upon unfolding; ΔASA_{npol} and ΔASA_{pol}, its nonpolar and polar components; R69, isolated 1–69 N-terminal fragment of the 434-phage repressor; RR69, single-chain, head-to-tail tandem of R69 linked by a 20-amino-acid natural linker; $[R69C]_2$, the tail-to-tail cross-linked symmetric tandem of R69, linked via a disulfide bridge between cysteine residues added to the C-termini.

(Received 1 February 1999, revised 12 April 1999, accepted 20 April 1999)

gene was constructed from the pSETR69 template [7] by PCR using the T7 promoter primer (TAATACGACTCACTATAGGG, Novagen) and the mutagenic primer 5′-TCTTGGATCCTCAGCATCTAACATTCGAATCAGAGG-3′, where the underlined residues are complementary to the extra cysteine codon. The PCR product was cleaved with *Bam*HI and *Xba*I and cloned into the pSET5a vector [7]. Expression was performed as previously described [7] and the spontaneously cross-linked dimer product was purified by FPLC on a Mono-SHR column followed by reverse phase-HPLC. Protein samples were checked for homogeneity by SDS/PAGE and gel filtration on a Superdex75 analytical column and stored as a lyophilized, refrigerated powder. Before all experiments the samples were dialysed overnight against appropriate buffers at 20 mM concentration: Pipes (pH 7.0), sodium acetate (pH 4.5 and 4.0) and glycine (pH 3.0 and 2.0). To increase the ionic strength, 200 mM NaCl was added to the buffers.

The protein concentration of the dialysed samples was measured spectrophotometrically at 25 °C and 280 nm using the following extinction coefficients: 5900, R69; 12 980, RR69, and 11 910 $M^{-1}\cdot cm^{-1}$, $[R69C]_2$, as determined by the method of Gill and Von Hippel [11].

DSC was performed on a DASM-4 microcalorimeter (BioPribor, Russia) at heating rates of 1 and 2 $K\cdot min^{-1}$ as described in detail elsewhere [12] at protein concentrations in the range of 1.5–7.0 $mg\cdot mL^{-1}$. Some measurements were carried out on a MicroCal VP-DSC calorimeter at heating rates of 1 and 1.5 $K\cdot min^{-1}$. The partial molar heat capacity was calculated assuming 0.73 $mL\cdot g^{-1}$ for the partial specific volume of the proteins, while values of 7.55 kDa, 17.25 kDa and 15.30 kDa were taken for the molecular masses of R69, RR69 and $[R69C]_2$, respectively. The DSC traces transformed into the temperature dependencies of the partial molar heat capacity were subject to single and multiple fitting procedures using ORIGIN 4.1 software from MicroCal. User-defined procedures based on the equations for single and multiple analysis [13,14] were used instead of the standard routines provided by MicroCal.

The areas exposed to the solvent in the native and unfolded states of R69 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 [15]; a probe size of 1.4 Å, slice size of 0.05 Å and van der Waals radii suggested by Chothia [16] were used. The native structure coordinates were taken from the PDB file *PRA1.ent*. The unfolded state of R69 was approximated by the extended β-structure, simulated with the Biopolymer module of INSIGHT II software (Biosym, California, USA) on an INDIGO Workstation (Silicon Graphics, USA).

CD spectra were recorded at 25 °C with a Jasco-750 spectropolarimeter using cells with 2 or 0.2-mm path lengths and protein concentrations of about 0.1 or 0.8 $mg\cdot mL^{-1}$, respectively.

RESULTS

R69 domain

The unfolding of the R69 domain has been studied at neutral and acidic pH (from pH 7 to pH 2) at two ionic strengths, 'low' (20 mM buffer) and 'high' (+200 mM sodium chloride). Under all conditions the heat-induced unfolding was highly reversible and independent of both the scanning rate and protein concentration. A typical calorimetric curve for the R69 monomer is shown in Fig. 1 together with its two-state fitting.

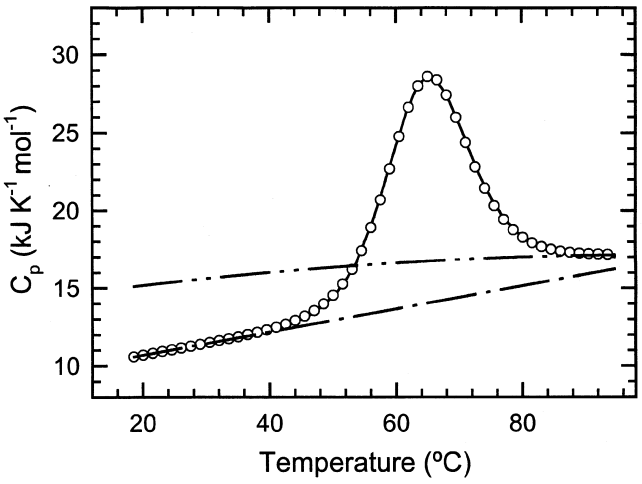

Fig. 1. Temperature dependence of the partial molar heat capacity for the R69 monomer at pH 4.5 (solid line). The open circles show the best fitting to the monomolecular two-state model; the dash-dot lines correspond to the heat capacities of the native (lower) and unfolded (upper) states.

Any treatment of DSC curves (independently of the assumed transition model) should include a reasonable approximation of the baselines, i.e. the temperature dependencies of the heat capacity of the initial and the final conformations. As shown by Privalov and coworkers [17], the heat capacity of the unfolded state is a nonlinear function of temperature, which bends considerably below 40 °C. When the transition is narrow and consequently the heat absorbance peak is high, this curvature of the $C_{p,U}(T)$ is negligible and a linear approximation of $C_{p,U}(T)$ would be accurate enough. With R69, however, the transitions are rather broad and hence the curvature of $C_{p,U}(T)$ should be

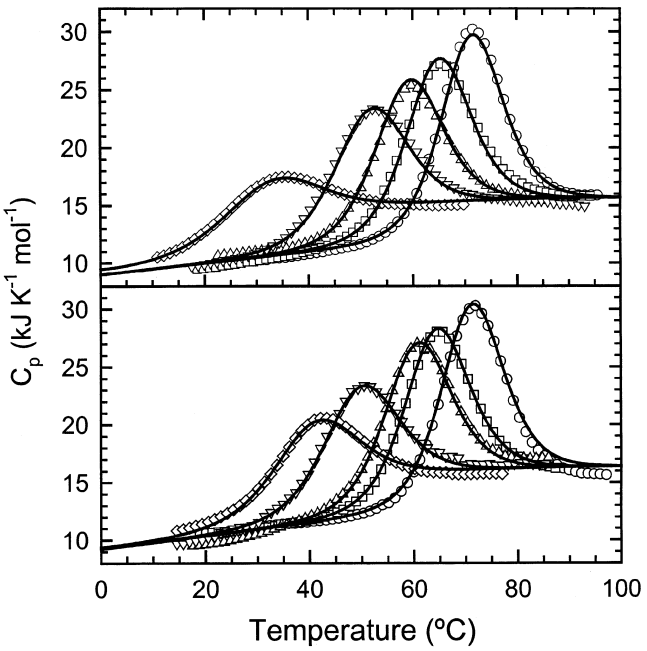

Fig. 2. The results of simultaneous fittings of the C_p functions of the monomer at low (upper panel) and high (lower panel) ionic strength. The symbols correspond to the experimental data at pH values: 2.0 (◇); 3.0 (▽); 4.0 (△); 4.5 (□) and 7.0 (○), while solid lines refer to the best-fit curves.

Table 1. The thermodynamic parameters of the heat-induced unfolding of the R69 and its dimers, RR69 and [R69C]$_2$, found by the single (in parentheses) and multiple DSC curve fittings under various conditions. The letters L and H refer to low and high ionic strength. The parameters are calculated per mole of the R69 domain.

	T_m (°C)			ΔH_m (kJ·mol^{-1})			ΔG_U^{298} (kJ·mol^{-1})
Conditions	R69	RR69	[R69S]$_2$	R69	RR69	[R69S]$_2$	R69
pH 7.0							
L	70.7 ± 0.5	64.8 ± 0.7	67.2 ± 0.5	247 ± 10	210 ± 15	226 ± 10	22.4 ± 1.5
	(70.5 ± 0.5)	(65.7 ± 0.7)	(68.1 ± 0.5)	(250 ± 10)	(234 ± 15)	(234 ± 10)	
H	70.7 ± 0.5	66.6 ± 0.6	67.6 ± 0.6	245 ± 10	241 ± 10	243 ± 10	22.3 ± 1.5
	(71.1 ± 0.5)	(67.8 ± 0.6)	(68.7 ± 0.6)	(251 ± 10)	(244 ± 10)	(251 ± 10)	
pH 4.5							
L	64.1 ± 0.5	60.1 ± 0.7	58.8 ± 0.5	222 ± 10	194 ± 8	194 ± 8	18.4 ± 1.0
	(64.2 ± 0.5)	(62.0 ± 0.7)	(58.8 ± 0.5)	(227 ± 10)	(197 ± 8)	(201 ± 8)	
H	63.7 ± 0.5	63.5 ± 0.5	60.2 ± 0.6	226 ± 10	229 ± 10	218 ± 10	18.1 ± 1.0
	(63.7 ± 0.5)	(63.8 ± 0.5)	(59.7 ± 0.6)	(226 ± 10)	(236 ± 10)	(217 ± 10)	
pH 4.0							
L	58.2 ± 0.5	58.2 ± 0.5	53.6 ± 0.5	209 ± 8	191 ± 8	185 ± 8	15.0 ± 0.8
	(58.4 ± 0.5)	(59.0 ± 0.5)	(53.5 ± 0.5)	(204 ± 8)	(186 ± 8)	(182 ± 8)	
H	59.8 ± 0.5	57.8 ± 0.7	57.2 ± 0.5	214 ± 10	209 ± 8	208 ± 8	15.8 ± 0.9
	(59.8 ± 0.5)	(58.4 ± 0.7)	(57.5 ± 0.5)	(210 ± 10)	(203 ± 8)	(213 ± 8)	
pH 3.0							
L	50.5 ± 0.8	42.3 ± 1.0	37.2 ± 1.3	183 ± 8	144 ± 8	134 ± 8	10.8 ± 0.8
H	48.2 ± 0.8	46.9 ± 1.0	42.8 ± 1.0	176 ± 8	169 ± 8	163 ± 8	9.5 ± 0.8
pH 2.0							
L	27.7 ± 2.0	32.2 ± 1.5	24.4 ± 2.0	101 ± 10	115 ± 8	93 ± 10	0.9 ± 0.5
H	38.3 ± 2.0	39.4 ± 1.2	35.6 ± 1.5	140 ± 10	142 ± 10	143 ± 8	4.9 ± 0.6

taken into account [13]. Thus, the heat capacity of the unfolded state was assumed to be a quadratic function of temperature, with the first and second-order coefficients found from the nonlinear quadratic regression through the C_p values calculated from the amino-acid content, as described elsewhere [14,18]. As seen from Fig. 1, the slightly concave $C_{p,U}(T)$ agrees well with the post-transitional heat capacity changes observed in our DSC experiments. The transition enthalpies obtained from single fittings are shown in Table 1.

Figure 2 demonstrates two sets of C_p curves obtained for the R69 monomer at low and high ionic strengths together with the multiple best fittings. One of the largest uncertainties in DSC measurements, which gives rise to errors in the absolute heat capacity values, derives from a random vertical displacement of the instrumental baseline. This problem can be solved during a simultaneous curve fitting in two ways: (a) the parameter that defines the displacement of the $C_{p,N}$, can be left adjustable for each curve, or (b) the individual curves aligned so that their absolute heat capacity at a selected temperature coincides with an average value. We chose the second option and, before performing a simultaneous fitting, the experimental curves were displaced vertically to coincide at 373 K with the average value of 15.8 kJ·K^{-1}·mol^{-1} measured for the C_p of the monomer at this temperature. As the stability of R69 depends not only upon pH but also upon the ionic strength, the multiple-fitting analysis was performed separately for the two C_p curve sets obtained at both ionic strengths. It is widely accepted

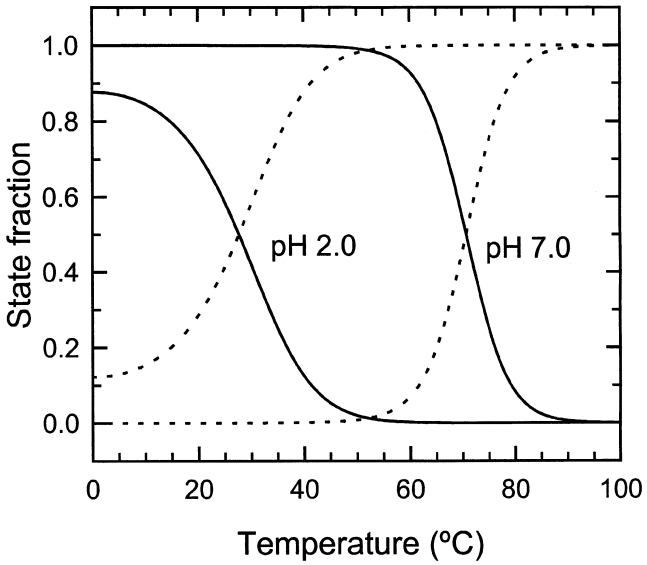

Fig. 3. Temperature dependencies of the populations of the native (solid lines) and unfolded (dashed lines) states of the R69 monomer at pH 2.0 (left) and 7.0 (right) at low ionic strength. The line cross-points correspond to the transition midpoints, T_m (27.7 °C at pH 2.0 and 70.7 °C at pH 7.0).

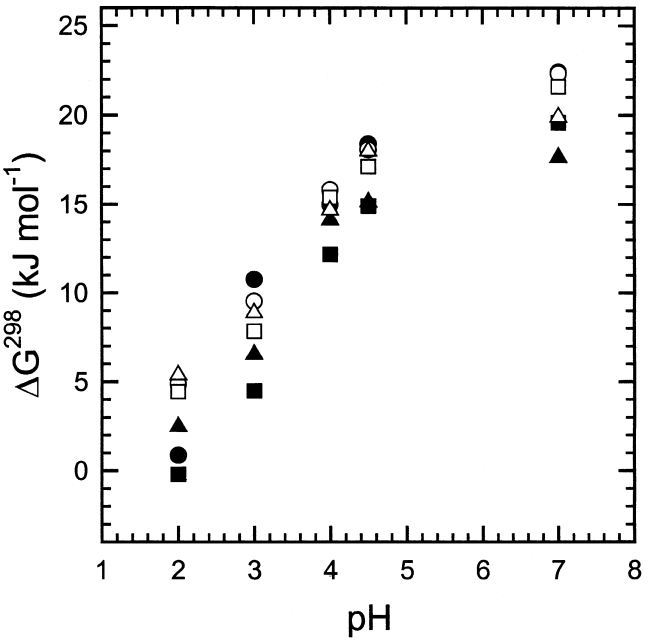

Fig. 4. The dependence of the standard Gibbs energy of unfolding (per mole of R69) on pH for R69 (circles), RR69 (triangles) and [R69C]$_2$ (squares). Filled symbols: low ionic strength; open symbols: high ionic strength.

[19,20] that the effect of pH (in the acidic range) and ionic strength on protein stability is mostly of entropic origin, i.e. neither of these factors changes the enthalpy of the native state unfolding to any great extent. It was also found that the influence of these two solvent factors on $\Delta C_{p,U}$ is also small, which means that there is one single $\Delta H_U(T)$ function, independent of both ionic strength and pH. This assumption allows us to decrease the number of adjustable parameters and to arrive at some realistic approximations for $C_{p,N}(T)$ and $C_{p,U}(T)$ in those cases when one of them is unknown, as happens with $C_{p,N}(T)$ at pH 2, for example. As can be seen from Table 1, the values obtained from individual fittings of the curves recorded at neutral pH, where an independent approximation of the initial heat capacity is reliable, agree well with those found by multiple analysis. At pH values below 4, even at low temperatures, the unfolded state is remarkably populated (Fig. 3) and thus a reliable approximation of $C_{p,N}$ for any single curve is impossible as the individual fittings result in large uncertainties in the thermodynamic parameters. For this reason the results of the individual fittings for the pH range below 4 are not shown in Table 1.

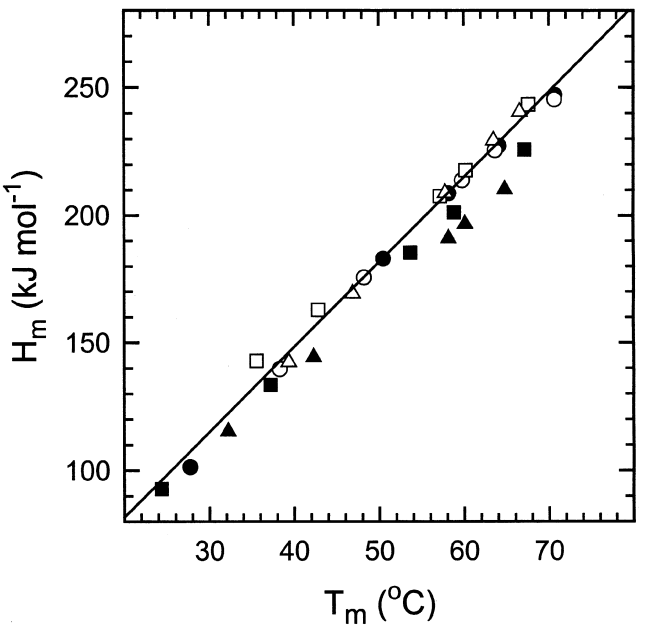

Fig. 5. The correlation between the transition midpoint and the unfolding enthalpy for R69 (circles), RR69 (triangles) and [R69C]$_2$ (squares) found by simultaneous DSC curve fittings. Filled symbols, low ionic strength; open symbols, high ionic strength.

RR69 and [R69S]$_2$ dimers

A general unfolding scheme for a two-domain protein should include interdomain interactions, which might depend on various structural and thermodynamic factors, such as the structure of the linker, the conformation of the partner, etc. If the domain interaction is considered as an additional two-state unit the general unfolding scheme would involve $2^3 = 8$ states (three units with two accessible states). Such a scheme, which looks like a cube, has never been considered in detail because in the majority of real cases many of these states are not populated. One of the simplest variants, which will be analysed below, has been introduced by Filimonov and coworkers [12,21,22] and later developed by Ramsay & Freire [23]. This simple model does not include a breakdown of interdomain contacts as an initial step preceding the unfolding, but takes into consideration the interdomain interaction within a simple four-state model:

$$\begin{array}{ccc} A_N B_N & \overset{K_{A1}}{\Leftrightarrow} & A_U B_N \\ {\scriptstyle K_{B1}}\,\Updownarrow & & \Updownarrow\,{\scriptstyle K_{B2}} \\ A_N B_U & \underset{K_{A2}}{\Leftrightarrow} & A_U B_U \end{array} \qquad (1)$$

Here the subscripts N and U refer to the folded and unfolded conformations of the domains. The parallel edges of the square correspond to the equilibrium unfolding of the same unit, while the subscripts of the equilibrium constants show whether this unit unfolds first or second. In general:

$$K_{A1} \neq K_{A2} \quad \text{and} \quad K_{B1} \neq K_{B2} \qquad (2)$$

But it must hold that:

$$K_{A1}K_{B2} = K_{B1}K_{A2} \qquad (3)$$

Taking A_NB_N as the reference state, the partition function of the system can be written as:

$$Q_{AB} = 1 + K_{A1} + K_{B1} + K_{A1}K_{B2} \qquad (4)$$

and the excess heat capacity is the sum of three asymmetric peaks of heat absorption. The condition shown in Eqn (2) means that the unfolding parameters of each domain depend upon the macroscopic state of its partner and this interdependence might be either a stabilizing or destabilizing one. Let us introduce K_{in}, an effective equilibrium constant corresponding to the interdomain interaction, so that $K_{A1} = K_{in}K_{A2}$. Hence $K_{B1} = K_{in}K_{B2}$. If $K_{in} > 1$ then $K_{A1} > K_{A2}$ and domain A unfolds more easily when domain B is folded; the same is true for domain B, hence the interaction between the domains is a destabilizing one. Correspondingly, if $K_{in} < 1$ the interaction is stabilizing. When $K_{in} = 1$ ($K_{A1} = K_{A2} = K_A$ and

Table 2. The α-helix content of R69 and of its covalently linked dimers as calculated from the R69 structure [8,9] and CD spectra recorded at 25 °C under various solvent conditions. The data in the column 'Structure' were calculated under the assumption that the domains within the chimeras have the same conformation as the isolated R69 and that the linkers are not in the helical conformation.

	α-Helicity (%)		Unfolded state fraction	
Construct/conditions	Structure	CD	DSC	CD
R69				
pH 7.0	59	56 ± 3	< 0.01	< 0.01
pH 2.0, no NaCl	–	42 ± 5	0.4 ± 0.05	0.3 ± 0.1
pH 2.0, 200 mM NaCl	–	49 ± 5	0.15 ± 0.03	0.13 ± 0.05
RR69				
pH 7.0	52	51 ± 3	< 0.01	< 0.02
pH 2.0, no NaCl	–	36 ± 5	0.27	0.3 ± 0.1
pH 2.0, 200 mM NaCl	–	44 ± 5	0.10	0.15 ± 0.05
$[R69C]_2$				
pH 7.0	58	57 ± 3	< 0.01	< 0.02
pH 2.0, no NaCl	–	42 ± 5	0.52	0.44
pH 2.0, 200 mM NaCl	–	50 ± 5	0.14	0.15

$K_{B1} = K_{B2} = K_B$) the domains do not 'feel' the conformation of the partner and the partition function shown in Eqn (4) reduces to:

$$Q_{AB} = (1 + K_A)(1 + K_B) \quad (5)$$

which reflects that the system has converted into a set of two independent subsystems and the excess heat capacity will be equal to the sum of two independent peaks:

$$C_p^{exc} = (K_A/R)[\Delta h_A/T(1 + K_A)]^2 + (K_B/R)[\Delta h_B/T(1 + K_B)]^2 \quad (6)$$

where the unfolding enthalpies Δh_A and Δh_B coincide with the intrinsic unfolding enthalpies of each domain. Nevertheless, even in this simple case, neither of the two constants needs to be equal to the unfolding constants of the isolated domains, $K_{X,iso}$, because the linker or other structural factors might and very often do change the stability of the domains within the dimer without inducing any cooperativity. When the domains are identical, Eqn (6) becomes:

$$C_p^{exc} = 2(K/R)[\Delta h/T(1 + K)]^2 \quad (7)$$

which differs from the expression for the isolated domain only by a factor 2.

The model first introduced by Rowe *et al.* [24], and later developed by Brandts *et al.* [25], includes the domain separation step and so subsequent domain unfolding is independent. This model converts into Eqn (1) when the interdomain interactions break down well before unfolding ($K_{dis} \gg 1$) or, on the contrary, when cooperation between the folded domains is very strong ($K_{dis} \ll 1$).

Eqn (7) has been used throughout this study to fit the DSC data of our constructs under the assumption that the domains within the dimers are identical and unfold independently. This simplifying assumption is realistic because the isolated R69 does not form stable dimers in solution even under the conditions of NMR experiments, i.e. at about 1 mM concentration [9], and therefore the interaction between the domains within the dimers might be expected to be rather weak, which has been proved to be correct during data analysis.

Under all the conditions studied so far (identical to those used for the R69 monomer), the unfolding of both chimeras was highly reversible and independent of the scanning rate and the protein concentration. The thermodynamic parameters were found either by single or multiple fittings of the C_p functions expressed per mole of R69 in order that they should be comparable with those of the isolated monomer (Table 1). At high ionic strength the positions and shapes of the RR69 and $[R69C]_2$ melting curves are very similar to those of the monomer, with only slight differences in the T_m values (Table 1). The fittings of the melting curves to a two-state model are reasonably good, which shows that at high ionic strength the domains within RR69 and $[R69C]_2$ unfold independently. Furthermore, the linkers do not change the intrinsic stability of the domains within the tandems as both the

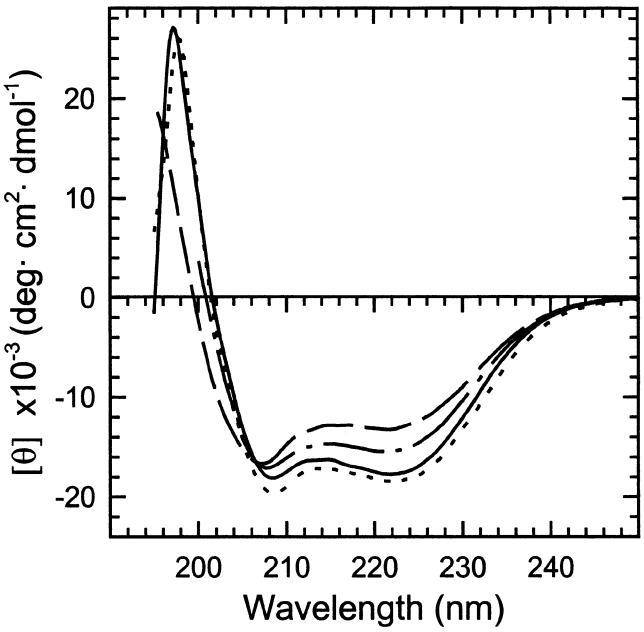

Fig. 6. The CD spectra of the R69 domain at 25 °C and various solvent conditions. Solid line, pH 7.0 and low ionic strength; dashed line, pH 2.0 and low ionic strength; dash-dot line, pH 2.0 and high ionic strength. The CD-spectrum of the RR69 tandem recalculated per the amino-acid content of a single domain (without taking into account the linker) is shown by a dotted line.

T_m and ΔG_U^{298} values are similar to those of R69 alone (Fig. 4 and Table 1) and the ΔH_m versus T_m data do not deviate from the linear regression, with a slope of 3.33 ± 0.07 kJ·K^{-1}·mol^{-1}, found for R69 (Fig. 5).

At low ionic strength, however, the temperature-induced transitions in the dimers differ from those of the monomer. As a rule, the T_m values for the dimers are lower, while the unfolding enthalpies lie below the linear regression of the monomer (Fig. 5). An analysis of the fitting errors shows that the melting curves of the dimers recorded at low ionic strength seem to be a little wider than those obtained at high ionic strength, which means that the simplified two-state model is not precisely applicable to the low-ionic-strength data, probably because $K_{in} \neq 1$, as will be discussed below.

CD spectra

The DSC curve fitting allows us to calculate, among other parameters, the temperature dependencies of state fractions for each solvent condition, as shown in Fig. 3. The results of our analyses indicate that the R69 domain has a stable, folded conformation at 25 °C at pH above 4.0, both at low and high ionic strengths, but starts to unfold below this pH threshold. If the unfolding scheme has been chosen correctly, the populations found from the DSC data should agree with the estimations made from other experimental data because within the two-state model, any optical parameter can be defined through the populations of the states as:

$$\langle X(T)\rangle = X_N F_N + X_U F_U \qquad (8)$$

where $X_N(T)$ and $X_U(T)$ are the temperature dependencies of the parameter for the native and unfolded states correspondingly. To check the correctness of our analysis and to obtain some additional structural information we have recorded the CD spectra of the domain and its dimers under various conditions (Fig. 6). At pH 7.0, the far-ultraviolet spectra of both the isolated R69 and the chimeras have the characteristic features of a structure with a high α-helix content. The estimations of the α-helicity from the CD spectrum (Table 2) give the value of about 56% for R69, which corresponds closely with the helicity calculated directly from the published 3D structure [8,9]. This result might be expected as, according to NMR data, the domain is folded in solution at neutral pH. Moreover, a comparison of the CD spectra of the monomer and dimers reveals that at pH 7.0 the linkers within both chimeras add nothing to the internal helicity of the R69 domain, as the CD-spectra calculated per mole of the R69 residue practically coincide for all three protein variants.

The CD spectra recorded at pH 2.0 have lower amplitudes indicating that the samples may have a lower α-helicity under these solvent conditions. The apparent decrease in α-helix content at pH 2.0 might, however, be explained by a partial unfolding of the domains at room temperature, as suggested by the DSC data analysis, which predicts that at pH 2.0 and low ionic strength as much as 40% of the monomer molecules are unfolded at 25 °C. At high ionic strength, however, the unfolded fraction is much smaller (about 15%) due to the clear stabilizing effect caused by salt. These proportions are slightly different for the isolated domain and the dimers, which means that linking two domains into a chimera might influence their intrinsic stabilities.

DISCUSSION

The CD spectra and the unfolding parameters of isolated R69 (including its partial denaturation at pH 2.0) are not unusual for small globular proteins of this size. The strong pH-dependence of the intrinsic stability of the R69 structure can be put down to one or two acidic side-chains with anomalous pK values. To account for the observed destabilizing effect the neighbourhood of such anomalous groups should impede its protonation within the native conformation below pH 4. Thus, the good candidates should be imbedded within positively charged clusters or else form solvent-protected salt-bridges, as, for example, Glu35, which is reported to form a buried salt-bridge with Arg10 [10]. Due to its small size, R69 has a relatively small Gibbs energy of unfolding, even under the conditions of maximum stability. Hence, even a single acidic group with its pK shifted 2 units below its normal value might account for an almost complete reduction of the unfolding energy at pH 2.

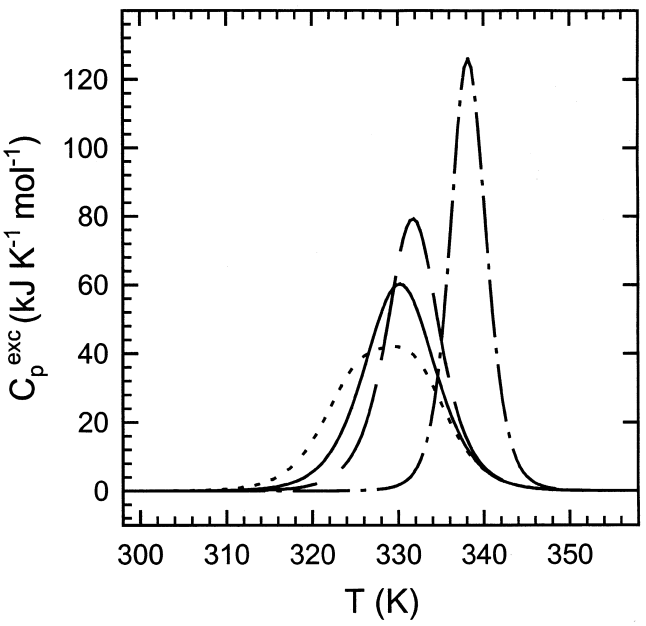

Fig. 7. Excess heat capacity functions simulated for the unfolding. Using the scheme shown in Eqn (1) under the assumption that $K_{A2} = K_{B2} = K_{iso}$ and $K_{A1} = K_{B1} = K_{in}K_{iso}$. K_{in} is defined by Eqn (11) with ΔS_{in} independent of temperature. The parameters of the isolated domain unfolding, which define K_{iso}, were: $\Delta h_U = \Delta h_A = \Delta h_B = 330$ kJ·mol^{-1}; $T_m = 330$ K; $\Delta C_{p,U} = 3$ kJ·K^{-1}·mol^{-1}; $\Delta G_U^{298} = 27$ kJ·mol^{-1}. The ΔS_{in} (kJ·mol^{-1}) was set to: 0, solid line; 0.01($K_{in} = 3.33$; $\Delta\Delta G_U^{298} \cong -3$ kJ mol^{-1}), dotted line; −0.01($K_{in} = 0.3$; $\Delta\Delta G_U^{298} \cong 3$ kJ·mol^{-1}), dashed line; −0.05 ($K_{in} = 2.45 \times 10^{-3}$; $\Delta\Delta G_U^{298} \cong 15$ kJ·mol^{-1}), dash-dot line.

As the effect of pH and ionic strength on domain stability is entropic, the slope of the regression line through the ΔH_m versus T_m points corresponds to an average $\Delta C_{p,U}$, the heat capacity difference between the unfolded and native states. A number of empirical relations correlating some integral structural characteristics, such as solvent-protected areas, with the thermodynamic parameters of protein unfolding has been suggested in the literature [26–28]. Taking into account the solvent-accessible surface area (ASA) upon R69 folding ($\Delta ASA_{tot} = 4586$ Å^2; $\Delta ASA_{pol} = 1424$ Å^2; $\Delta ASA_{npol} = 3162$ Å^2) the expressions for $\Delta C_{p,U}$ (kJ·K^{-1}·mol^{-1}) suggested by Gomez & Freire [26] might be transformed into:

$$\Delta C_{p,calc} = -20.16 + 0.171T - 2.988 \times 10^{-4}T^2 \qquad (9)$$

from which $\Delta C_{p,calc}(298\ K) = 4.24$ kJ·K^{-1}·mol^{-1} and the average value over the temperature range 303–363 K, $\Delta C_{p,calc,av}$ =

3.5 $kJ \cdot K^{-1} \cdot mol^{-1}$. This latter estimation agrees well with the above-mentioned experimental value. The more simple relationship between ΔASA_{tot} and the heat capacity increment suggested by Myers *et al*. [29] also gives a rather good estimation of the average $\Delta C_{p,U}$:

$$\Delta C_{p,av} = 0.19 \times \Delta ASA_{tot} = 3.64\,kJ \cdot K^{-1} \cdot mol^{-1} \quad (10)$$

showing that R69 is a typical globular protein with a standard distribution of hydrophobic and hydrophilic groups between the globule interior and its surface.

As mentioned above, charged groups ought to play a very important role not only in the interaction of R69 with DNA, but also in modulating the domain stability and unfolding mechanism within the covalently linked dimers, in particular at low pH and ionic strength. Under these extreme conditions, where the domains acquire a large noncompensated charge, the DSC melting curves not only shift to a lower temperature but also begin to deviate from the two-state model. The consequences of such a domain interdependence might be illustrated by simulations of the excess heat capacity for a hypothetical case of two equivalent domains. Let us assume that there is an electrostatic interaction between two domains accompanied by a constant entropy change, ΔS_{in}, and thus:

$$K_{in} = e^{(\Delta S_{in}/R)} \quad (11)$$

The results of computer simulations of the heat capacity excess functions, C_p^{exc}, are illustrated in Fig. 7. The solid line corresponds to an independent unfolding of two domains in accordance with Eqn (7). For the three remaining curves, the unfolding constant of the second transition was kept equal to K_{iso}, while the first one was changed to $K_{iso}K_{in}$. When $\Delta S_{in} > 0$, $K_{in} > 1$ and the interaction between the two folded domains is destabilizing. Even a small positive ΔS_{in} value, changing the ΔG_U only by 10%, brings about a remarkable distortion in the C_p^{exc} shape: the melting curve becomes broader, smaller and asymmetrical. When ΔS_{in} is negative the melting curves become sharper and shift to the right, reflecting an increase in the degree of cooperation between the domains. At a K_{in} of about 2.5×10^{-3} the heat capacity peak becomes very sharp and an analysis of the state populations shows that the unfolding transition is practically a two-state one, as the domains merge into a single cooperative unit. These simulations show that the heat capacity curve is very sensitive even to a small (in terms of Gibbs energy) positive or negative interaction between the domains.

Nevertheless, in our DSC data analyses we never observed any sharpening of the melting curves compared to those of R69, which means that under our solvent conditions there was no positive co-operation between the domains. Furthermore, the stability of the isolated domain was always slightly higher than within the dimers. A certain broadening of the dimer melting curves, probably caused by a domain repulsion under extreme conditions, might be responsible for an apparent decrease in the unfolding enthalpy values found by the two-state fitting. This may happen because the two-state fitting algorithm looks for a compromise between the calorimetric enthalpy (the area under the peak) and the van't Hoff enthalpy (peak amplitude and half-width).

Both stabilizing and destabilizing interactions between the domains within a dimer would depend upon its organization because even 'nonstructured' tails or linkers might influence the stability of the native structure [20,30–32], through electrostatic interactions or solvent redistribution around the molecule, for example. It should be noted that the behaviour of $[R69C]_2$ and RR69 might be different, not only because their linkers are different in length, but also because $[R69C]_2$ is fully symmetrical whilst RR69 is not.

As can be seen from Table 1, under identical solvent conditions (except for pH 7.0) the domains within $[R69C]_2$ are usually less stable than isolated R69 or the domains within RR69. Hence, the mutual destabilizing induced by linking two domains with a short disulfide bridge is stronger than the effect of the long, flexible linker of RR69. At neutral pH, the domains are equally stable in all the constructs. Therefore, a large part of the destabilizing effect must be put down to mutual electrostatic repulsion, which gets stronger with acidic pH values, but which should decrease after the unfolding of any domain due to the higher flexibility of the unfolded chains. This latter factor might be responsible for the observed broadening of the DSC curves.

However, under natural conditions DSC does not reveal any interaction between the domains in the RR69 and $[R69S]_2$ dimers. Robinson and Sauer [3] found a correlation between linker length and the stability of head-to-tail recombinant dimers of the P22 Arc repressor monomer (53 residues) and identified an optimum linker length for their construct. The Arc repressor is, however, quite different from the compact R69 domain, not only because it binds to DNA as a tetramer, but also because the Arc domains are not independent folding units [4]. Jana *et al*. [32] compared covalent dimers of the λ Cro repressor, the structure of which is similar to that of R69 [8]. They found that varying the linker length in recombinant head-to-tail dimers has no apparent effect on the stability and DNA-binding affinity of the dimers. They also found, however, that the Cro-dimer connected by a disulfide bridge in the middle of the monomer/monomer interface, is more stable but has lower DNA-binding affinity than the other constructs. In our opinion, the DNA-binding affinity might be lower due to the perturbation of the dimer structure caused by the disulfide bridge. These results, together with ours, confirm that independently folding domains can be included in artificial constructs with relatively small limitations on linker design. The thermodynamic equivalence of the two domains of the head-to-tail RR69 is also remarkable because in this construct the natural N-terminal domain is repeated in tandem and so the second domain is in a nonnatural, C-terminal position without a free N-terminus. This equivalence indicates that independently folding domains can indeed be freely shuffled in artificial constructs.

The fact that cooperativity between the R69 domains could not be detected in this work does not exclude the formation of important interdomain contacts upon the binding of DNA. A few specific contacts might contribute significantly to the DNA-binding affinity and specificity, even if they are weak in comparison with the entropic gain of covalently linking two DNA-binding domains.

ACKNOWLEDGEMENTS

This work has been supported by PB96-1446 grant from the Spanish Government, INTAS-93-007 Grant, and ERBBMH1CT931454 Project from the European Union and its ERBCIPDCT940298 extension granted to V. V. F. through PECO. V. V. F. also acknowledges sabbatical grants from the Spanish Ministry of Science and Education. We also thank Dr John Trout for revising the English text.

REFERENCES

1. Freemont, P.S., Lane, A.N. & Sanderson, M.R. (1991) Structural aspects of protein-DNA recognition. *Biochem. J.* **278**, 1–23.

2. Aggarwal, A.K., Rodgers, D.W., Drottar, M., Ptashne, M. & Harrison, S.C. (1988) Recognition of a DNA operator by the repressor of phage 434: a view at high resolution. *Science* **242**, 899–907.
3. Robinson, C.R. & Sauer, R.T. (1998) Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. *Proc. Natl Acad. Sci. USA.* **95**, 5929–5934.
4. Robinson, C.R., Rentzeperis, D., Silva, J.L. & Sauer, R.T. (1997) Formation of denatured dimer limits the thermal stability of Arc repressor. *J. Mol. Biol.* **31**, 692–700.
5. Percipalle, P., Simoncsits, A., Zakhariev, S., Guarnaccia, C., Sanchez, R. & Pongor, S. (1995) Rationally designed helix-turn-helix proteins and their conformational changes upon DNA-binding. *EMBO J.* **14**, 3200–3205.
6. Chen, J., Pongor, S. & Simoncsits, A. (1997) Recognition of DNA by single-chain derivatives of the phage 434 repressor: high affinity binding depends on both the contacted and non-contacted base pairs. *Nucleic Acids Res.* **25**, 2047–2054.
7. Simoncsits, A., Chen, J., Percipalle, P., Wang, S., Töro, I. & Pongor, S. (1997) Single-chain repressors containing engineered DNA-binding domains of the phage 434 repressor recognise symmetric or asymmetric DNA operators. *J. Mol. Biol.* **267**, 118–131.
8. Mondragon, A., Subbiah, S., Almo, S.C., Drottar, M. & Harrison, S.C. (1989) Structure of the amino-terminal domain of phage 434 repressor at 2.0 Å resolution. *J. Mol. Biol.* **205**, 189–201.
9. Neri, D., Billeter, M. & Wüthrich, K. (1992) Determination of the nuclear magnetic resonance solution structure of the DNA-binding domain (residues 1–69) of the 434 repressor and comparison with X-ray crystal structure. *J. Mol. Biol.* **223**, 743–767.
10. Pervushin, K., Billeter, M., Siegal, G. & Wüthrich, K. (1996) Structural role of a buried salt bridge in the 434 repressor DNA-binding domain. *J. Mol. Biol.* **264**, 1002–1012.
11. Gill, S.C. & von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* **182**, 319–326.
12. Privalov, P.L. & Potekhin, S.A. (1986) Scanning microcalorimetry in studying temperature-induced changes in proteins. *Methods Enzymol.* **131**, 4–51.
13. Martinez, J.C., Viguera, A.R., Serrano, L., Filimonov, V.V. & Mateo, P.L. (1998) The DSC data analysis for small, single-domain proteins. Application to the SH3 domain. *React. Func. Polymers* **36**, 221–225.
14. Viguera, A.R., Martinez, J.C., Filimonov, V.V., Mateo, P.L. & Serrano, L. (1994) Thermodynamic and kinetic analysis of the SH3 domain of spectrin shows a two-state folding transition. *Biochemistry* **33**, 2142–2150.
15. Lee, B.K. & Richards, F.M. (1971) The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55**, 379–400.
16. Chothia, C. (1976) The nature of the accessible and buried surfaces in proteins. *J. Mol. Biol.* **105**, 1–14.
17. Privalov, P.L., Tiktopulo, E.I., Venyaminov, S.Y., Griko, Y.V., Makhatadze, G.I. & Khechinashvili, N.N. (1989) Heat capacity and conformation of proteins in the denatured state. *J. Mol. Biol.* **205**, 737–750.
18. Makhatadze, G.I. & Privalov, P.L. (1990) Heat capacity of proteins. I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect. *J. Mol. Biol.* **213**, 375–384.
19. Privalov, P.L. (1979) Stability of proteins. Small globular proteins. *Adv. Prot. Chem.* **33**, 167–241.
20. Martinez, J.C., Viguera, A.R., Bariso, R., Willmanns, M., Mateo, P.L., Filimonov, V.V. & Serrano, L. (1999) Thermodynamic analysis of α-spectrin SH3 and two its circular permutants with different loop lengths: discerning the reasons for rapid folding in proteins. *Biochemistry* **38**, 549–559.
21. Filimonov, V.V., Potekhin, S.A., Matveyev, S.V. & Privalov, P.L. (1982) Thermodynamic analysis of scanning calorimetry data. Algorithms for the deconvolution of complex heat absorbance curves. *Mol. Biol. USSR* **16**, 435–444.
22. Matveyev, S.V., Filimonov, V.V. & Privalov, P.L. (1982) A thermodynamic approach to the investigation of structure organisation of the 5S RNA from *E. coli* ribosomes. I. DSC data analysis. *Mol. Biol. USSR* **16**, 1234–1244.
23. Ramsay, G. & Freire, E. (1990) Linked thermal and solute perturbation analysis of co-operative domain interactions in proteins. Structural stability of diphtheria toxin. *Biochemistry* **29**, 8677–8683.
24. Rowe, E.S. (1976) Dissociation and denaturation equilibria, and kinetics of a homogenious human immunoglobulin Fab fragment. *Biochemistry* **15**, 905–916.
25. Brandts, J.F., Hu, C.Q. & Lin, L.N. (1989) A simple model for proteins with interacting domains. Application to scanning calorimetry data. *Biochemistry* **28**, 8588–8596.
26. Gomez, J. & Freire, E. (1995) Thermodynamic mapping of the inhibitor site of the aspartic protease endothiapepsin. *J. Mol. Biol.* **252**, 337–350.
27. Makhatadze, G.I. & Privalov, P.L. (1995) Energetics of protein structure. *Adv. Prot. Chem.* **47**, 307–425.
28. Spolar, R.S., Livingstone, J.R. & Record, M.T. Jr (1992) Use of liquid and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from the water. *Biochemistry* **31**, 3947–3955.
29. Myers, J.K., Pace, C.N. & Scholtz, J.M. (1995) Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–2148.
30. Hamil, S.J., Meekhof, S.E. & Clarke, J. (1998) The effect of boundary selection on the stability and folding of the third fibronectin type III domain from human tenascin. *Biochemistry* **37**, 8071–8079.
31. Pfuhl, M., Improta, S., Politou, A.S. & Pastore, A. (1997) When a module is also a domain. The role of the N-terminus in the stability and the dynamics of immunoglobulin domains from titin. *J. Mol. Biol.* **265**, 242–256.
32. Jana, R., Hazbun, T.R., Fields, J.D. & Mossing, M.C. (1998) Single-chain lambda Cro repressors confirm high intrinsic dimer-DNA affinity. *Biochemistry* **37**, 6446–6455.