Thermodynamic Analysis of α-spectrin SH3 and Two of Its Circular Permutants with Different Loop Lengths: Discerning the Reasons for Rapid Folding in Proteins † , ‡

ARTICLE *in* BIOCHEMISTRY · FEBRUARY 1999

Impact Factor: 3.02 · DOI: 10.1021/bi981515u · Source: PubMed

7 AUTHORS, INCLUDING:

Jose C. Martinez
University of Granada
40 PUBLICATIONS **1,507** CITATIONS
SEE PROFILE

Rita Berisio
Italian National Research Council
107 PUBLICATIONS **2,289** CITATIONS
SEE PROFILE

Thermodynamic Analysis of α-spectrin SH3 and Two of Its Circular Permutants with Different Loop Lengths: Discerning the Reasons for Rapid Folding in Proteins[†,‡]

Jose C. Martínez,[§,||] Ana R. Viguera,[§] Rita Berisio,[⊥] Matthias Wilmanns,[⊥] Pedro L. Mateo,[||] Vladimir V. Filimonov,*[,||,#,▽] and Luis Serrano*[,§]

EMBL, Meyerhofstrasse 1, 69117 Heidelberg, Germany, Departamento de Quimica Fisica, Facultad de Ciencias, Universidad de Granada, 18071-Granada, Spain, EMBL, DESY, Notkestrasse 85, D-22603 Hamburg, Germany and Institute of Protein Research, Russian Academy of Sciences, Pushino, 142292 Russia

Received June 25, 1998; Revised Manuscript Received October 21, 1998

ABSTRACT: The temperature dependences of the unfolding−refolding reaction of a shorter version of the α-spectrin SH3 domain (PWT) used as a reference and of two circular permutants (with different poly-Gly loop lengths at the newly created fused loop) have been measured by differential scanning microcalorimetry and stopped-flow kinetics, to characterize the thermodynamic nature of the transition and native states. Differential scanning calorimetry results show that all these species do not belong to the same temperature dependency of heat effect. The family of the N47-D48s circular permutant (with 0−6 Gly inserted at the fused-loop) shows a higher enthalpy as happens with the PWT domain. The wild type (WT) and the S19-P20s permutant family have a more similar behavior although the second is far less stable. The crystallographic structure of the PWT shows a hairpin formation in the region corresponding to the unstructured N-terminus tail of the WT, explaining the enthalpic difference. There is a very good correlation between the calorimetric changes and the structural differences between the WT, PWT, and two circular permutants that suggests that their unfolded state cannot be too different. Elongation of the fused loop in the two permutants, taking as a reference the protein with one inserted Gly, results in a small Gibbs energy change of entropic origin as theoretically expected. Eyring plots of the unfolding and refolding semireactions show different behaviors for PWT, S19-P20s, and N47-D48s in agreement with previous studies indicating that they have different transition states. The SH3 transition state is relatively close to the native state with regard to changes in heat capacity and entropy, indicating a high degree of compactness and order. Regarding the differences in thermodynamic parameters, it seems that rapid folding could be achieved in proteins by decreasing the entropic barrier.

Small monomeric proteins with no disulfide bridges or cis-proline bonds follow a two-state folding mechanism in which only the denatured and folded states are significantly populated (*1*−*8*). This facilitates the thermodynamic and kinetic analysis and the interpretation of the results. Although many of these proteins have similar sizes they have very different stabilities and folding kinetics (*9*). These differences are not related to protein size or secondary structure composition and cannot be ascribed to protein function, since similar proteins in different organisms have very different folding rates (*7*). It follows that the window allowed in vivo for protein folding is quite wide, and this property is not necessarily optimized in terms of either speed or stability. Supporting this argument is the fact that it has been found in at least two proteins (the α-spectrin SH3[1] domain and the activation domain of carboxypeptidase A) and that folding speed can be increased by more than 1 order of magnitude through circular permutation (*10*) or by stabilizing α-helices (*9*). Due to the different structures and amino acid compositions of these different two-state small proteins, it is not possible to correlate the kinetic and thermodynamic data with the structural information. More insight into the features that determine these properties can be obtained by analyzing proteins with identical amino acid sequences and different stabilities and folding velocities. Circular permuta-

[†] The work has been supported by the INTAS Grant 93-007-ext, the Grant PB96-1146 from the DGICYT (Spain), and the EU Grants CT96-0013 and CT97-2180. A.R.V. is supported by a fellowship and V.V.F. by a sabbatical from the Spanish Ministry of Education and Science. J.C.M. is supported by an EU-TMR postdoctoral fellowship.

[‡] Atomic coordinates for PWT structure have been deposited in the Protein Data Bank with filename 1pwt.

* To whom correspondence should be adressed. Tel: +49-6221-387320. Fax: +49-6221-387306. E-mail: serrano@embl-heidelberg.de.

[§] EMBL, Heidelberg, Germany.

[||] Universidad de Granada.

[⊥] EMBL, Hamburg, Germany.

[#] Tel: +34-58-243333. Fax: +34-58-272879. E-mail: vladimir@goliat.ugr.es.

[▽] Russian Academy of Sciences.

[1] Abbreviations: SH3, Src homology region 3; WT, wild type; PWT, pseudo wild type; S19-P20s, circular permutation with cut between positions S19 and P20; N47-D48s, circular permutation with cut between positions N47 and D48; 1G, 3G, 5G, number of extra Gly residues inserted in the new loop; CI2, chimotrypsin inhibitor 2; GndHCl, guanidinium hydrochloride; DSC, differential scanning calorimetry; EDTA, ethylenediamine tetraacetic acid; MES, 2-(*N*-morpholino) ethanesulfonic acid; DTT, dithiothreitol.

10.1021/bi981515u CCC: $18.00
Published on Web 12/19/1998

tion represents an excellent example for this comparison. Recently, circular permutants have been done and structurally characterized on the α-spectrin SH3 domain (*10*, *11*).

The α-spectrin SH3 domain (62 residues) folds into an orthogonal β-sandwich that contains three β-hairpins (*12*) and follows a two-state transition as determined by calorimetry and kinetics (*4*). The refolding rate of this domain can be increased seven times by circular permutation when the breaking point is situated in the long irregular RT loop (S19-P20s), but not when the cut is situated in the regular distal β-hairpin (N47-D48s). However, both permutant proteins have similar stabilities as measured by chemical denaturation. The structures of the native and transition states of these two proteins have been studied in detail (*11*). They fold into the same 3D structure as the wild-type SH3 domain except for the engineered loop that fuses the wild-type termini and the cleaved RT loop in the S19-P20s circular permutant that loses nine conserved hydrogen bonds through local hydrogen bond unzipping. No hydrogen bond unzipping occurs in the N47-D48s circular permutant. Contrary to what happens to the native structures, the outcoming transition states seem to be considerably different (*11*). The Gibbs energy of the interactions already present in the transition state is lower for the S19-P20s protein and similar for the WT and N47-D48s, but still all of them have a similar extent of surface buried in the transition state. While the N47-D48s permutant folds a little bit slower than the WT SH3 protein, the S19-P20s permutant folds and unfolds almost 1 order of magnitude faster. The reason for the large differences in folding velocity for the three proteins is not yet understood. The calculation of the entropic and enthalpic components of the transition state could help in finding a common pattern related to the folding velocity.

In this paper, we present a detailed study of the unfolding−refolding reaction of PWT and the two circular permutants with different loop lengths. The idea is to determine the thermodynamic factors behind their different folding behaviors, as well as to corroborate that loop elongation mainly produces an entropic change. Differential scanning microcalorimetry (DSC) has been used to study the equilibrium between the native and unfolded states. Refolding has been studied by pH jump into native conditions, and unfolding was measured in the presence of varying concentrations of GndHCl and then extrapolated to native conditions. Eyring plots of the refolding and unfolding kinetic constants obtained at different temperatures have allowed us to obtain the activation energies and its enthalpic and entropic components in the three species.

EXPERIMENTAL PROCEDURES (MATERIALS AND METHODS)

Protein Expression and Purification. Construction and purification of the four polypeptides used in this study have been described previously (*10*). The sequences are the following: WT, mDETGKELVLALYDYQEKSPREVTMKKGDILTLLNSTNKDWWKVEVNDRQGFVPAAYVKKLD; PWT, m**G**TGKELVLALYDYQEKSPREVTMKKGDILTLLNSTNKDWWKVEVNDRQGFVPAAYVKKLD; S19-P20s, mgPREVTMKKGDILTLLNSTNKDWWKVEVNDRQGFVPAAYVKKLD**SG**TGKELVLALYDYQEKS; N47-D48s, mDRQGFVPAAYVKKLD**SG**GKELVLALYDYQEKSPREVTMKKGDILTLLNSTNKDWWKVEVN.

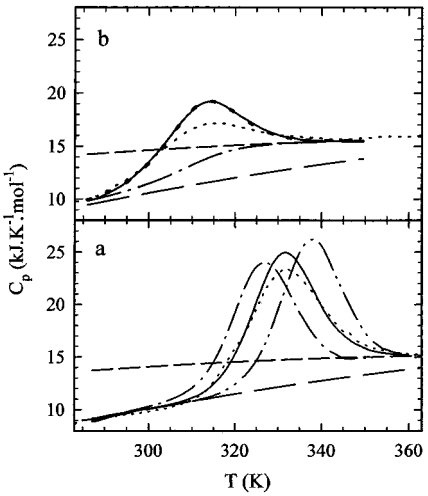

FIGURE 1: Temperature dependencies of the partial molar heat capacities of various mutant of α-spectrin SH3. (a) Dash-dot-dot line, PWT at pH 3.5; solid line, N47-D48s at pH 3.5; short-dash line, S19-P20s1G at pH 3.5; medium-dash-dot line, N47-D48s at pH 7.0. The averaged temperature dependencies of the initial and final heat capacities are also plotted as the long-dash and medium-dash lines, respectively. (b) N47-D48s1G (solid line) and S19-P20s1G (short dash-line) at pH 2.5. The best fit of the N47-D48s1G curve to the two-state model is shown by dots. $C_{p,N}$ is shown by a long-dash line, $C_{p,U}$ is shown by a medium-dash one, and $C_p^{int} = C_{p,N}F_N + C_{p,U}F_U$ is shown by a dash-dot line. F_N and F_U denote the populations of the N and U states, respectively.

The PWT protein contains the same amino acid sequence as the wild-type protein except for the second and third residues which have been substituted by a Gly residue. In the circular permutants the amino acid sequence has been interrupted between amino acids Ser19 and Pro20 and between Asn47 and Asp48, respectively. The previous N and C-termini have been joined by the linker Lys60-Leu61-Asp62-**Ser2-Gly3**-Thr4-Gly5-Lys6, where Ser2 and Gly3 were not present in the WT protein. The "s" in S19-P20s and N47-D48s states for the presence of a Ser residue in the linker to distinguish them from other shorter versions (*10*). In addition, the connecting sequence was elongated with one, three, or five extra-Gly residues, forming the mutants S19-P20s1G, S19-P20s3G, S19-p20s5G, N47-D48s1G, N47-D48s3G, and N47-D48s5G, respectively.

Differential Scanning Calorimetry. The temperature-induced unfolding of the PWT form of the SH3 domain from α-spectrin and its various permuted forms was studied by differential scanning calorimetry (DSC). The computerized version of the DASM-4 instrument (Biopribor, Russia) with platinum cells of 0.47 mL volume has been used at heating rates of 1 and 2 K/min and protein concentrations of 0.8−4 mg/mL. The latter were measured spectrophotometrically using the following extinction coefficients determined by the method of Gill and von Hippel (*13*): 15 370 cm^{-1}·M^{-1} for PWT 15 440 cm^{-1}·M^{-1} for N47-D48s, 15 060 cm^{-1}·M^{-1} for the N47-D48s members with extra-Gly residues, and 15 040 cm^{-1}·M^{-1} for S19-P20s family. The calorimetric records were transformed into the temperature dependencies of the partial molar heat capacity (Figure 1) and analyzed as described elsewhere (*4*) and below. The calorimetric curves recovered about 80% of their initial amplitude upon reheating of the sample, depending on the exposure of the

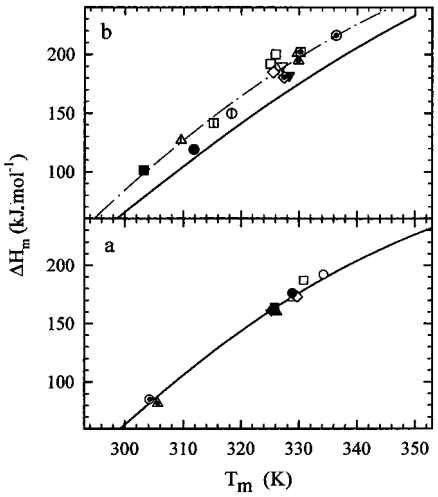

FIGURE 2: The correlations between the unfolding enthalpy, ΔH_m, and transition temperature, T_m. (a) S19-P20s family: S19-P20s, triangles; S19-P20s1G, circles; S19-P20s3G, squares; S19-P20s5G, diamonds; pH 7.0, open symbols; pH 3.5, filled symbols; pH 2.5, dots inside. The solid line corresponds to the best fit of the published WT data *(4)*. (b) N47-D48s family and PWT: PWT, circles; N47-D48s, squares; N47-D48s1G, triangles upward; N47-D48s3G, triangles down; N47-D48s5G, diamonds; pH 2.0, filled symbols; pH 2.5, open symbols with inserted bars; pH 3.5, dotted open symbols; pH 7.0, open symbols. The solid line shows the same as on panel a, while the dash-dot line corresponds to the quadratic regression through the data of this work (see text).

sample to high temperatures. In any case, the extent of refolding was high enough to suggest reversible equilibrium unfolding under all solvent conditions used throughout the DSC study.

Fitting of the DSC Data. To find some reasonable restrictions over the assumed parabolic functions (see Results), we estimated the heat capacity of the unfolded state of PWT by the procedure proposed by Makhatadze and Privalov (*14*). The C_p values calculated for the six fixed temperature values were then fitted to a quadratic function of temperature to get the following empirical expression (see Figure 3):

$$C_{p,U} = -5.95 + 0.116T - 0.00015T^2 \qquad (1)$$

This expression is slightly different from that calculated earlier for the WT protein (*4*), but corresponding best-fit curves are practically coinciding in the temperature range of interest and within the limits of accuracy of C_p estimation. It was found that the experimentally observed temperature dependencies of $C_{p,U}$ (unfortunately available only for temperatures above 338 K) are, on average, smaller than the calculated ones by approximately 0.8 kJ·mol^{-1}, which is comparable with the experimental C_p curve scatter (Figure 3). Thus, in our curve fittings, we could not fix $C_{p,U}$ completely and allowed it to shift on the C_p scale, while the linear and quadratic coefficients remained fixed at the values specified in eq 1.

After performing a preliminary curve fitting under assumption of a linear $C_{p,N}$, we found that a more realistic approximation to the heat capacity of the native state should be close to the following quadratic relation (Figure 3):

$$C_{p,N} = a_N + 0.216T - 0.0002315T^2 \qquad (2)$$

Table 1: Calorimetric Parameters Extracted from DSC Experiments[a]

pH	name	T_m (K)	$\Delta C_{p,U}(T_m)$ (kJ·K^{-1}·mol^{-1})	ΔH_m (kJ·mol^{-1}) exp	ΔH_m (kJ·mol^{-1}) extrap	$\Delta G_U(298)$ (kJ·mol^{-1})
2.0	WT[b]	307	3.7	93		2.3
	PWT	310.5 ± 0.5	4.1	124	128	4.0
	N47-D48s	303.2 ± 0.7	4.4	99	101	1.4
2.5	WT[b]	320	3.4	139		6.9
	PWT	318.7 ± 0.4	3.7	150	159	7.6
	S19-P20s	305.6 ± 0.7	3.9	82	89	1.9
	S19-P20s1G	304.3 ± 0.7	3.9	85	84	1.5
	N47-D48s	315.2 ± 0.5	3.8	142	146	6.0
	N47-D48s1G	309.0 ± 0.5	4.2	127	122	3.4
3.5	WT[b]	336	2.9	188		13.9
	PWT	336.4 ± 0.3	2.8	216	217	16.9
	S19-P20s	326.2 ± 0.4	3.3	160	164	9.7
	S19-P20s1G	328.9 ± 0.3	3.3	176	173	10.9
	S19-P20s3G	325.8 ± 0.4	3.4	164	163	9.5
	S19-P20s5G	325.3 ± 0.4	3.4	161	161	9.3
	N47-D48s	330.3 ± 0.3	3.1	203	199	13.5
	N47-D48s1G	329.8 ± 0.3	3.2	195	198	13.2
	N47-D48s3G	328.3 ± 0.3	3.2	182	193	12.4
	N47-D48s5G	327.5 ± 0.3	3.3	180	190	12.0
7.0	PWT[c]	339.2 ± 0.8	2.7	225		18.5
	S19-P20s	329.2 ± 0.3	3.2	174	174	11.1
	S19-P20s1G	334.3 ± 0.3	3.1	192	190	13.6
	S19-P20s3G	330.8 ± 0.3	3.2	187	179	11.8
	S19-P20s5G	329.7 ± 0.3	3.2	173	176	11.3
	N47-D48s	326.0 ± 0.4	3.3	195	185	11.2
	N47-D48s1G	329.7 ± 0.4	3.2	201	197	13.2
	N47-D48s3G	327.1 ± 0.3	3.3	191	189	11.8
	N47-D48s5G	325.6 ± 0.4	3.4	185	184	11.0

[a] The columns labeled "exp" refer to the experimentally observed ΔH_m values obtained from curve fitting. The columns labeled "extrap" correspond to the extrapolations with temperature-dependent $\Delta C_{p,U}$, respectively. The errors in experimental values of ΔH_m are about 7%, and in the approximated values of $\Delta S_U(298)$ and $\Delta G_U(298)$ they are about 10%. The $\Delta H_U(298)$ values can be easily calculated from eq 3 and are 61.6 kJ·mol^{-1} for PWT, 76.3 kJ·mol^{-1} for the N47-D48s family, and 59.4 kJ·mol^{-1} for S19-P20s family. [c] It was possible to estimate only T_m for PWT at pH 7.0 with reasonable accuracy. [b] Data taken from ref 4.

Thus, a_N became a second adjustable parameter during curve fitting, whereas the other two were ΔH_m and T_m. The results of these final fittings are presented in Figure 2 and Table 1.

Since $\Delta C_{p,U} = C_{p,U} - C_{p,N}$ must also be a quadratic function of temperature, the correlations between ΔH_m and T_m values, obtained by the curve fitting, were approximated by cubic regressions (Figure 2):

$$\Delta H_U(T) = \Delta H_0 + \Delta aT + bT^2 + cT^3 \qquad (3)$$

with ΔH_0, Δa, b, and c being, in general, adjustable parameters. In the case of the PWT protein ΔH_0 is close to −4410, Δa to 27.49, b to −0.05, and c to 2.72×10^{-5}. For the N47-D48s family of proteins ΔH_0 is close to −4300, Δa is about 27.17, b to −0.05, and c to 2.72×10^{-5}. In the case of the S19-P20s family of proteins, their unfolding enthalpy seems to be very similar to that of the WT (Figure 2 b), where ΔH_0 is close to −448, Δa to −5.95, b to 0.0435, and c to -5.98×10^{-5} (*4*).

Then, if we assume (after taking into account that ionization heats are negligible in that pH range) that eq 3 corresponds to the real temperature dependence of the ΔH_U, the heat effect of unfolding, one gets for the heat capacity

increments the following values:

$$\text{PWT} \quad \Delta C_{p,U} = d\Delta H_U/dT = 27.49 - 0.1T + (8.15 \times 10^{-5})T^2 \quad (4)$$

$$\text{S19-P20s} \quad \Delta C_{p,U} = d\Delta H_U/dT = -5.95 + 0.087\,T - (1.795 \times 10^{-4})T^2 \quad (5)$$

$$\text{N47-D48s} \quad \Delta C_{p,U} = d\Delta H_U/dT = 27.17 - 0.1T + (8.15 \times 10^{-5})T^2 \quad (6)$$

The heat capacity increments calculated by substituting experimental T_m values into eq 4−6 are similar to those found directly from the curve fitting (it should be taken into account that usually the latter ones can vary as much as 20% from experiment to experiment due to approximately 10% uncertainties in the positions of the slopes of the instrument baselines). Subtracting these functions from $C_{p,U}$ dependence (eq 1), one gets the $C_{p,N}$ functions which coincide within 10% accuracy with the averaged heat capacity of the initial state (Figure 3). These calculated $C_{p,N}$ functions have the initial slope of about 0.083 kJ·K^{-2}·mol^{-1} (Figure 3), which is very close to that obtained from the fitting of the experimental melting curves.

Now, once the values of ΔH_m, T_m, and $\Delta C_{p,U}$ are known, the Gibbs energy of unfolding at a standard temperature of 298 K can be calculated in two ways by using the following relations:

PWT, N47-D48s

$$\Delta S_U(298) = \Delta H_m/T_m + \Delta a \ln(298/T_m) - 0.1(298 - T_m) + (8.15 \times 10^{-5})(298^2 - T_m{}^2) \quad (7)$$

S19-P20s

$$\Delta S_U(298) = \Delta H_m/T_m + \Delta a \ln(298/T_m) + 0.087(298 - T_m) - (1.795 \times 10^{-4})(298^2 - T_m{}^2) \quad (8)$$

and

$$\Delta G_U(298) = \Delta H_U(298) - (298)\Delta S_U(298) \quad (9)$$

First, the best-fit approximation of $\Delta H_U(T_m)$ can be used to calculate $\Delta H_U(298)$, and, second, the best-fit $\Delta C_{p,U}(T)$ functions (eqs 4, 5, and 6) can be used in combination with the experimentally observed ΔH_m values just for estimating the limits of extrapolation errors introduced into ΔG_U (298).

Temperature Dependence of Refolding and Unfolding Rates. The temperature dependence of the main refolding phase of WT and both circular permutants has been studied over the range 15−42 °C by pH jump experiments. Protein samples adjusted with a concentrated HCl solution to pH 1.8 were mixed 1:10, with 100 mM MES, pH 6.5, in a thermostated stopped-flow. Actual temperature was measured by a thermocouple immerse in one of the syringes. Fitting of the main refolding phase to a monoexponential rendered k values that were used in the Eyring plots [ln(k/T) vs $1/T$]. Unfolding was measured in the presence of varying concentrations of GndHCl from 4.5 to 6.5 M in 100 mM MES, pH 6.5. The unfolding rate in the absence of denaturant was extrapolated by the following quadratic function that takes into account the nonlinear dependence of the natural logarithm of the unfolding rate constant with urea (*15*) or GndHCl (*11*) seen for this domain:

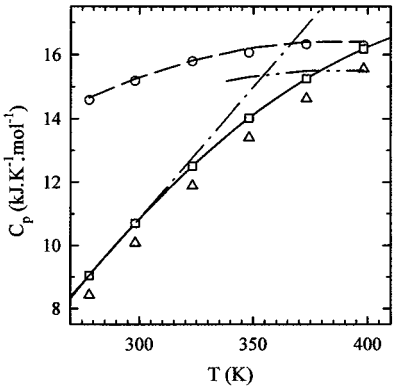

FIGURE 3: The temperature dependencies of the partial molar heat capacity for the unfolded (dash-dot-dot line) and the native states of PWT (open triangles) and N47-D48s family (open squares). The open circles show the $C_{p,U}$ values calculated from the amino acid content by the method of Makhatadze and Privalov (*14*), and the dashed line corresponds to a quadratic regression through the points (see eq 1 in the text). The dash-dot-dot line corresponds to an average of experimental $C_{p,U}(T)$, whereas the heat capacity of the native state was obtained by subtracting from it the $\Delta C_{p,U}(T)$ functions found by fitting the $\Delta H_m(T_m)$ correlations to quadratic polinoms. The dash-dot line drawn through the initial two points of the $C_{p,N}(T)$ has a slope of 0.083 kJ·K^{-2}·mol^{-1} which is close to the average value of the initial slopes observed on the experimental curves, while linear regression through all six points has a slope of 0.059 kJ·K^{-2}·mol^{-1}.

$$\ln k = \ln k(H_2O) + m_{\ddagger-N}[\text{GndHCl}] - 0.033[\text{GndHCl}]^2 \quad (10)$$

the slope $m_{\ddagger-N}$ was found to be dependent on temperature from 0.95 at 16 °C to 0.87 at 40 °C.

For linear Eyring plots the ordinate corresponds to ln(k_B/h) + $\Delta S^{\ddagger}/R$ and the slope to $\Delta H^{\ddagger}/R$, were k_B is the Boltzman constant and h is the Plank constant.

$$\ln(k/T) = \ln(k_B/h) + \Delta S^{\ddagger}/R - \Delta H^{\ddagger}/RT \quad (11)$$

When the change in ΔC_p between the reactant and the transition state is not negligible there is a substantial deviation from linearity, which here is more apparent in refolding. Eyring plots can in this case be fitted to the following equation:

$$\ln(k/T) = A + B(T_0/T) + C \ln(T_0/T) \quad (12)$$

where

$$A = [-\Delta C_p{}^{\ddagger} + \Delta S^{\ddagger}(T_0)]/R - \ln(h/k_B) \quad (13)$$

$$B = [\Delta C_p{}^{\ddagger} - \Delta S^{\ddagger}(T_0)]/R - \Delta G^{\ddagger}(T_0)/RT_0 \quad (14)$$

$$C = -\Delta C_p{}^{\ddagger}/R \quad (15)$$

Thermodynamic parameters corresponding to the transition state can be calculated by the following conversions:

$$\Delta C_p{}^{\ddagger} = -CR \quad (16)$$

$$-T_0\,\Delta S^{\ddagger}(T_0) = -(A - C - 23.76)RT_0 \quad (17)$$

$$\Delta G^{\ddagger}(T_0) = -(A + B - 23.76)RT_0 \quad (18)$$

$$\Delta H^{\ddagger} = -(B + C)RT_0 \quad (19)$$

where $\Delta C_p^{\ddagger}$ is the heat capacity change between denatured or native states and the transition state (kJ·K^{-1}·mol^{-1}), $\Delta S^{\ddagger}$ is the activation entropy (kJ·K^{-1}·mol^{-1}), $\Delta H^{\ddagger}$ is the activation enthalpy (kJ·mol^{-1}), and $\Delta G^{\ddagger}$ is the activation energy of refolding or unfolding (kJ·mol^{-1}). These are the differences of the values in the transition state with respect to the reactant in each case (denatured state in refolding and native in unfolding).

Crystallization and Data Collection. Crystals of PWT were obtained, at room temperature, using the hanging drop method. Drops of 4 μL were prepared from 2 μL of protein solution (6 mg/mL) + 2 μL of reservoir solution (1.1 M ammonium sulfate, 90 mM sodium citrate/citric acid at pH 4.0, 90 mM Bis-Tris propane, 0.9 mM EDTA, 0.9 mM DTT, 0.9 mM sodium azide). Rectangular crystals with maximum dimensions 800 × 200 × 200 μm^3 appeared after 2−3 days. A good cryosolution was obtained by adding 12% [v/v] glycerol to the mother liquor. After 5 min of equilibration in the cryosolution, the crystal was mounted using the loop method (*16*) and flash-frozen in a nitrogen gas stream at 100 K. The data set was collected on a Small Mar imaging plate (radius 90 mm) mounted on a Siemens MacScience MX18 generator. One-hundred and ninety-six images were collected with an oscillation angle $\Delta\varphi = 1°$.

All of the data were processed with DENZO (*17*) and SCALEPACK (*18*). The crystals showed a mosaicity of 0.5° and belonged to the space group $P2_12_12_1$, with unit cell a = 33.01, b = 42.39, and c = 49. 68 Å. The refined value of the mosaicity for this dataset is 0.5°, and the R factor is 5.0%. The dataset was very complete and redundant in all of the resolution shells. A summary of the statistics of the processing is shown in Table 2-1.

Structure Refinement. The starting model for the refinement was the α-spectrin SH3 domain structure (*12*) at 1.8 Å resolution. Its position in the unit cell was found by molecular replacement and rigid body refinement, using the program AMORE (*19*). Reflections in the resolution range 8−3.5 Å were used for this purpose. The R factor after the translation function was 27.7% and 23.5% after the rigid body refinement. The structure was then refined using the maximum likelihood algorithm implemented in the program REFMAC (*20*). Anisotropic scaling of the data and a correction to the disordered solvent were applied.

The solvent model was built automatically using the program ARP (*21*). The program "O" (*22*,*23*) was used for the model building.

The contribution of the hydrogen atoms was included in the computation of the structure factors in the last stage of the refinement; this reduced the R factor and R free by ∼1% for the working set and the test set. The R factor, calculated in the resolution range 15−1.77 Å, was 19.0% and R free 25.0%. The N-terminus of the protein was built by alternating omit map calculation, model building, and refinement. The electron density map was very clear for Lys5 and Met1 (Figure 4b). The loop modeled between these two residues showed high-temperature factors, thus suggesting a high flexibility. The modeling of the N-terminus dropped the R factor down to 18.7% and the R free to 24.7%. The quality of the model was assessed using the programs WHAT_CHECK and PROCHECK (*24*). The final model had 505 protein atoms and 90 water molecules. The refinement statistics are summarized in Table 2, part B.

Table 2: Statistics of the Processing and Refinement Statistics

A. Statistics of Processing	
resolution range (Å)	15−1.77
completeness (overall) (%)	99.8
completeness (last shell) (%)	96.8
R merge overall (%)	5.0
R merge last shell (%)	20.6
no. reflections	43628
no. unique reflections	7163
reflections with $I/\sigma > 5$ (overall) (%)	84.4
reflections with $I/\sigma > 5$ (last shell) (%)	49.0
B. Refinement Statistics	
R factor	18.7
R free	24.7
total no. of protein atoms	505
total no. of water atoms	90
average B (main-chain atoms) (Å)	20.0
average B (side-chain atoms) (Å)	23.3
average B (water) (Å)	33.0
rms bond length (Å)	0.02
rms bond angles (1−3 neighbors) (Å)	0.04
rms ΔB (main-chain bond) (Å^2)	1.0
rms ΔB (side-chain bond) (Å^2)	2.0

RESULTS

Differential Scanning Calorimetry (DSC). The two circular permutants of α-spectrin SH3 domain (S19-P20s and N47-D48s) were constructed by replacing Asp2 and Glu3 in the WT by Ser-Gly, and then joining this new N-terminal sequence (after removing the initial Met) with the C-terminus of the molecule, while opening at the corresponding loops (*10*). Asp2 and Glu3 are located in a nonstructured region of the WT SH3 domain. However they could still interact through long-range electrostatics with the rest of the molecule, as well as influence the conformational distribution of the denatured state. To have an appropriate reference for the mutants that preserves the amino acidic composition, we replaced these two residues by a Gly in the original sequence. Throughout the manuscript we will use the PWT as the reference protein for the circular permutants.

Due to the small size of the domains under study (molecular weight about 7.1−7.4 kDa depending on the loop length), we obtained relatively low molar heats of unfolding resulting in broad and small DSC peaks (Figure 1). This fact complicates the determination of the thermodynamic parameters, in particular at low pH where the permutants and PWT have low T_m values and small heats of unfolding (Figure 1 and Table 1). In a previous work a global curve fitting approach was used to avoid those technical difficulties and to determine the heat effects and other thermodynamic parameters with accuracy, without using high protein concentrations (*4*). In our present study we applied the same approach, although with some modifications. These modifications concern, first of all, the approximation of the initial heat capacity, $C_{p,N}$. When the experimental melting curves were fitted individually to the two-state model, under assumption of the linear dependence of the $C_{p,N}$ on the

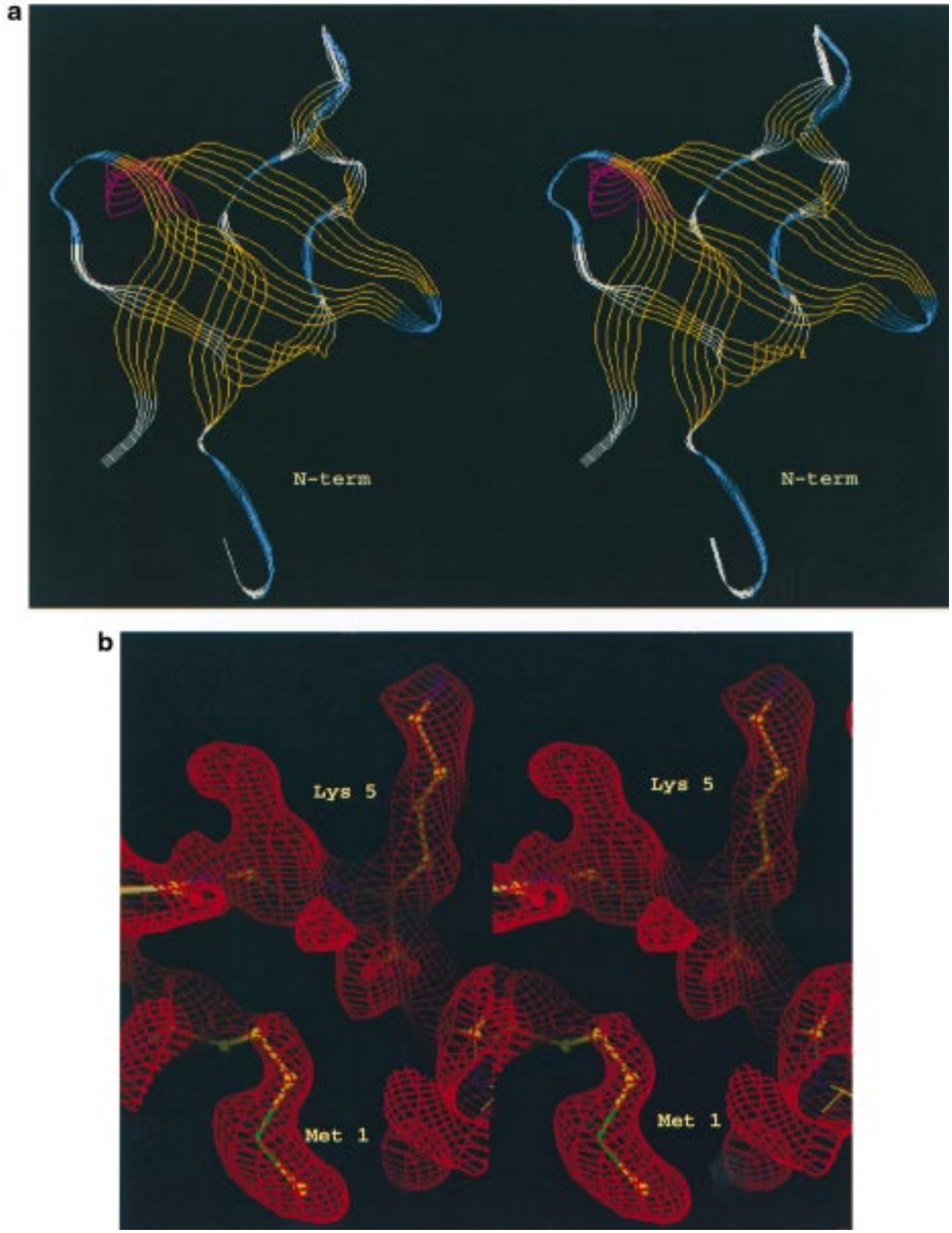

FIGURE 4: (a) Ribbon model of the pseudo wild type (PWT) structure. All of the secondary structure elements are described by colors: yellow, β-strands; pink, helices; blue, turns. (b) Density map of the N-terminal region of PWT structure showing the proximity between residues Met1 and Lys5 in the hairpin.

temperature, it was observed that the initial best-fit slopes usually became higher at lower T_m. In extreme cases, the $C_{p,N}(T)$ function crossed the $C_{p,U}(T)$ at about 353 K; that is, at this temperature the $\Delta C_{p,U}$ became negative (Figure 3), which is highly improbable taking into account all of the information accumulated in the literature (see, for example, ref *25*).

These observations made us look for a more realistic approximation of the initial C_p, and the first logic step was to assume that not only $C_{p,U}$ but also $C_{p,N}$ is a nonlinear

function of temperature. Such nonlinearity of the initial heat capacity is supported by experimental observations that the most curved temperature dependencies of the partial molar heat capacity belong to the hydrophilic amino acids (*14*, *26*). As those are almost completely exposed to the solvent in the native as well as in the unfolded state, there is no reason the unfolded state should have a bent $C_p(T)$ and the folded one not. As a result the $C_{p,N}$ function would have higher slopes at low temperature and lower ones at high temperature (Figure 3). We checked a quadratic approximation of the $C_{p,N}$ (Figure 3, for details see Materials and Methods) and found that its use definitely improves the fitting quality for the low-T_m curves without introducing additional adjustable parameters and decreases remarkably the errors of the best-fit values of ΔH_m and T_m (Figure 2).

Thermodynamic Parameters Obtained from DSC. The corresponding values of the heat effects and Gibbs energies, obtained after fitting the data mentioned above with the equations listed in Materials and Methods, are in Table 1.

(a) PWT. In a previous study we found by urea denaturation that the PWT protein was more stable than the WT, by around 1.7 kJ/mol (*10*). In agreement with this we find here that the heat effects of PWT unfolding measured at a given temperature are generally higher that those of WT (Figure 2 and Table 1). At pH 3.5, we find a difference in ΔG_U between both proteins of 2.1 kJ/mol, which fits very well the data obtained from equilibrium denaturation at pH 7.0 (the WT form cannot be studied by DSC at pH 7.0 due to its low solubility in these conditions). At the same time, the dependence of T_m on pH is very similar. For example, at pH 3.5, WT has T_m at 336 K (*4*), while PWT at 336.4 K, and so on. A priori, no remarkable differences are expected between the thermodynamic functions of WT and PWT since the substitution of Asp2 and Glu3 of the WT protein by a Gly in the PWT molecule is performed in a region of the WT domain that is unstructured in solution (*27*) or in the crystal structure (*12*). However, such differences have been found in all thermodynamic parameters, including unfolding enthalpy and $\Delta C_{p,U}$ which usually are insensitive to surface mutations. Nevertheless, neither heat effects nor $\Delta C_{p,U}$ versus T functions of PWT are shifted with respect to WT (Table 1).

(b) N47-D48s Family. These circular permutants have an enthalpy change similar to that of PWT, although their stability is significantly reduced. Moreover, it appears that they have higher stability at pH 3.5 than at pH 7.0, in contrast to both WT and PWT domains, which are more stable at pH 7.0 than at pH 3.5. This could be caused by local electrostatic interactions of the newly made N- and C-terminal groups which usually have the p*K*'s in the pH range between 7.0 and 3.5. Increasing the length of the newly formed fused loop by one Gly residue results in a slight stabilization at pH 7.0, suggesting that it alleviates some conformational restraint (Table 1). In fact, in a previous study we found that, by increasing the length of the loop by inserting a Ser residue (S19-P20 to S19-P20s and N47-D48 to N47-D48s), we stabilized the permutants (*10*). A further increase of loop size, however, destabilizes the protein. Although we could not detect any changes in the unfolding heat effect, we found an entropic change related to the length of the newly formed loop. This is in agreement with the assumption that loop elongation mainly produces an entropic destabilization (*28*). More importantly, these results essentially enable us to obtain correlations of ΔH_m versus T_m in a wider temperature range, which increases the reliability of DSC parameters.

(c) Unlike PWT and N47-D48s family, S19-P20s permuted forms have heat effects of unfolding very similar to those observed for WT, as the experimentally observed values of ΔH_m and T_m belong to the best-fit correlation published for WT. These S19-P20s domains, however, are less stable than WT, but unlike N47-D48s forms, they are more stable at pH 7.0 than at pH 3.5. Therefore, a cut introduced between positions S19 and P20 is unfavorable in terms of both the heat effect and entropy change. Part of the unfavorable entropy change should come from the removal of a proline residue from the middle of the polypeptide chain, which must increase the entropy of the unfolded state (*29*) and, hence, decrease the stability (in terms of ΔG_U) of the native state. The reason for the decrease in heat effect from the level of PWT back to the values of WT is most probably due to the loss of nine conserved hydrogen bonds as have been seen in the crystal structure (*11*). As happened in the N47-D48s family, at neutral pH and at pH 3.5, the introduction of a Gly in the fused loop stabilizes the protein while further elongation destabilizes the protein. The destabilization seems to be, within the experimental error, entropic in nature as we explained before (Table 1).

Crystal Structure of PWT. When comparing WT to PWT, there is an increase in heat effect observed for the latter, without a simultaneous raise in its melting temperature. The sequences of the two proteins differ only at the N-terminus. In fact, we have changed MDETG for MGTG, thus resulting in a shorter sequence with two missing acidic groups. Since the mutation was introduced in an unstructured area, as revealed by NMR (*27*) and X-ray (*12*), it may be suggested that the N-terminal end of PWT becomes more structured within the native state, rather than the denatured ensemble becoming more unfolded. Thus one might expect that the N-terminal lacking of two charged acidic groups sticks to the protein globule either to increase the number of van der Waals contacts (responsible for heat effect increase), and/or to protect some exposed hydrophobic groups and to fix the positions of the five amino acids (this will decrease both the heat capacity and entropy of the native state giving rise to lower $\Delta C_{p,U}$ and ΔS_U values).

To confirm this possibility we have crystallized the PWT under the same conditions already published for the WT protein (*12*). The crystallographic statistics are summarized in Table 2, and Figure 4 shows the final structure calculated for PWT and a detail of the N-terminal part. Regarding the N-terminus, the density is very clear on Lys5 and Met1 and a bit discontinuous on Gly2. In agreement with our hypothesis, the protein forms an extra hairpin at the N-terminus, which is stabilized by two hydrogen bonds, one formed between the carbonyl oxygen of Met1 and the amido nitrogen of Gly4 and the other one between the carbonyl oxygen of Gly4 and the amino nitrogen of Met1. These new contacts explain the enthalpic increase we obtain by calorimetry.

Temperature Dependence of the Kinetic Constants. Refolding of PWT, S19-P20s, and N47-D48s from acid-denatured forms at neutral pH has been followed by stopped-flow. Temperatures from 15 to 40 °C were used (Figure 5a). At higher temperatures the protein does not recover com-

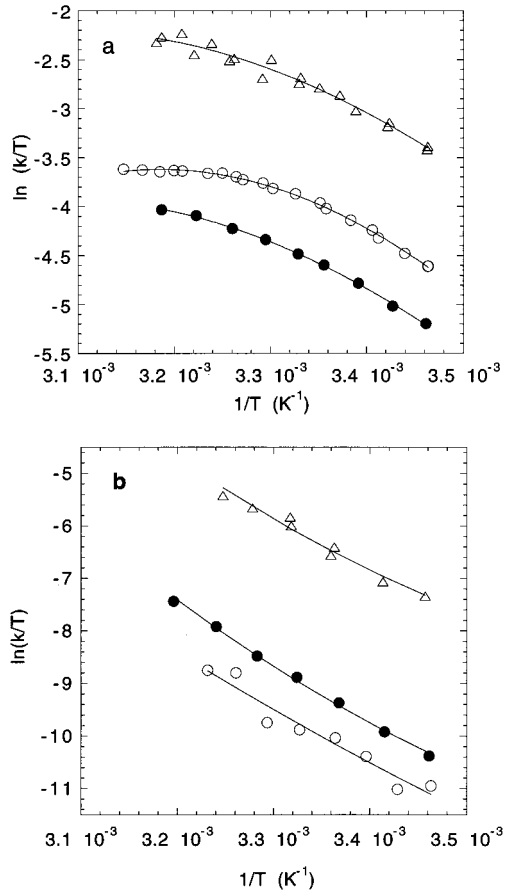

FIGURE 5: (a) Eyring plots of the temperature dependence of the refolding rate constant in water obtained by pH jump from acid-denatured forms to pH 6.5. Continuos lines represent best fits to eq 12 in Materials and Methods. (b) Eyring plots for the extrapolated values of the rate constants obtained in GndHCl to native conditions at pH 6.5. Best fits to eq 12 in Materials and Methods are also plotted; linear fitting to eq 11 rendered very similar values of the activation parameters: PWT (open circles), S19-P20s (open triangles), and N47-D48s (closed circles).

pletely the native state at the final pH used and deviations could appear as the observed kinetic constant is the addition of the two microscopic ones. For the unfolding reaction (Figure 5b), four different final GndHCl concentrations were used and the unfolding rate constant at the particular temperature investigated was obtained by nonlinear extrapolation to water (data not shown; eq 10 in Materials and Methods).

The nonlinear regression fitting of the data in the Eyring plots [ln(k/T) vs $1/T$] to eq 12 in Materials and Methods renders the activation parameters for unfolding and refolding of PWT, S19-P20s, and N47-D48s shown in Table 3. By subtraction of the values of both semireactions, the equilibrium values can be extracted. We find a reasonable qualitative correlation between the Gibbs energy values and those obtained by calorimetry or equilibrium chemical denaturation. The discrepancies with the equilibrium values for the Gibbs energies are within the experimental error, while they are larger for ΔH and ΔS. These discrepancies could be attributed to the fact that we have used a constant ΔC_p value for the calculation of the activation parameters, which is an approximation, since we do not know the dependence of $\Delta C_{p,\ddagger-N}$ and $\Delta C_{p,\ddagger-U}$ with temperature.

Table 3: Activation Parameters for Unfolding and Fast Refolding Calculated from the Eyring Plot According to Equation 12 in Materials and Methods with $T_0 = 298$ K

			$\Delta C_p^{\ddagger}$ (kJ·K^{-1}·mol^{-1})	$\Delta S^{\ddagger}$ (J·K^{-1}·mol^{-1})	$\Delta H^{\ddagger}$ (kJ·mol^{-1})	$\Delta G^{\ddagger}$ (kJ·mol^{-1})
refolding	PWT	(‡-U)	-2.3 ± 0.1	-108 ± 2	37 ± 1	69 ± 1
	S19-P20s	(‡-U)	-1.5 ± 0.4	-94 ± 6	38 ± 2	66 ± 4
	N47-D48s	(‡-U)	-1.6 ± 0.1	-100 ± 2	41 ± 1	70 ± 1
unfolding	PWT	(‡-N)	0.7	0 ± 17	84 ± 8	84 ± 10
	S19-P20s	(‡-N)	1.8	19 ± 15	80 ± 4	75 ± 6
	N47-D48s	(‡-N)	1.7	22 ± 7	88 ± 2	81 ± 4

Table 4: Thermodynamic Parameters for the Different Forms of α-Spectrin SH3 under Native Conditions, at pH 6.5[a]

		N	TS	U
ΔC_p	PWT	−3.0	−2.3	0
	S19-P20s	−3.3	−1.5	0
	N47-D48s	−3.3	−1.6	0
ΔH	PWT	−73	37	0
	S19-P20s	−59	38	0
	N47-D48s	−74	41	0
ΔS	PWT	−185	−108	0
	S19-P20s	−162	−94	0
	N47-D48s	−210	−100	0
ΔG	PWT	−15	69	0
	S19-P20s	−9	66	0
	N47-D48s	−11	70	0

[a] The reaction proceeds from the native state "N", through the transition state "TS", to the unfolded state "U". The unfolded state is arbitrarily taken as the reference point.

In any case the results obtained from the refolding reaction (Table 3) are totally reliable for two reasons: on one hand they are obtained directly and not from extrapolation and on the other they exhibit a significant curvature which allows one to estimate $\Delta C_{p,\ddagger-N}$. Therefore they were used, together with the DSC equilibrium results, to analyze the transition-state energetics of all of the proteins. This guarantees that the differences observed between them (Table 4) are real and not the result of experimental uncertainties.

DISCUSSION.

Comparison of WT and PWT Forms of the Protein. A priori no remarkable differences between the unfolding enthalpies of the WT and PWT should arise from the substitution of Asp2 and Glu3 by a Gly residue, into the *unstructured* part of the WT domain. However, such differences have been found in *all* thermodynamic parameters, including unfolding enthalpy and heat capacity change, which usually are insensitive to surface mutations. The only alternatives to the intramolecular stabilization concern the possibility of a change in the unfolded state of the protein or an increased tendency for dimerization (the elevated heat effects might be attributed to the heats of dimerization; 20 kJ·mol^{-1} is a typical value for this kind of process). However, we did not find any concentration dependency on T_m, which should discard this possibility. Thus, the increase in stability manifested in the higher values of $\Delta G_U(298)$, already found by chemical denaturation (*10*), has a rather big enthalpic contribution, with an unfolding enthalpy about 20 kJ·mol^{-1} higher than that of WT at all studied pH values. The most simple structural explanation of this unusual enthalpic stabilization might come from the possibility that, after

replacing two charged groups by a Gly, the unstructured N-terminal tail could be organized, thus decreasing the enthalpy of the native conformation and increasing ΔH_U. The crystallographic results support our ideas since the N-terminal tail is structured in the PWT mutant and two or three new hydrogen bonds are formed.

Stability of PWT and Permuted SH3 Domains and Comparison with the X-ray Structures. Calorimetry results suggest that there are two tendencies among the different domains analyzed here. WT and S19-P20s have very similar heat effects of unfolding, and the experimentally observed values of ΔH_m and T_m could be fitted to the same correlation. However, the unfolding enthalpies of PWT and N47-D48s are about 20 kJ·mol^{-1} higher at all studied pH values, despite the melting temperatures of N47-D48s being smaller than those of WT and PWT at any given pH. A second important feature is the T_m dependence upon pH for the N47-D48s protein. Unlike the WT and PWT, this protein is more stable at pH 3.5 than at pH 7.0 and this difference reaches as much as 4 K.

The above results can be explained on the basis of the X-ray structure differences of these proteins (ref *11* and this manuscript). In the N47-D48s cleavage of this loop does not disrupt the structure and the newly created N- and C-termini residues are almost superimposable with the equivalent residues in the uncleaved PWT protein. On the other hand, the newly created N-terminal hairpin produced by replacement of Asp2 and Glu3 by a Gly cannot be made here, since we have covalently linked the PWT N- and C-termini. However, in this protein a new hydrogen bond between Ser2 and Glu3 is formed in the fused loop. The regions equivalent to the N- and C-terminus of the PWT protein are more organized. Therefore, we do not expect any important enthalpic difference between this permutant and the PWT protein. The different pH dependence of the stability in this mutant could be due to the proximal difference of the newly formed N- and C-terminal charged groups (in the PWT and WT, the terminal charged groups are not located so close to each other) and their possible interactions with charged neighbors. Opposite to this, in the S19-P20s permutant the disrupted RT loop is hydrogen bond unzipped and rather unordered (the new termini are not contacting, as in the case of PWT). Consequently, the enthalpic gain produced by the formation of the fused loop is more than compensated by the loss of nine hydrogen bonds at the cleaved loop. This unzipping also results in a gain in entropy that compensates for the enthalpic loss. The good correlation found here between the thermodynamic and structural properties of the four proteins indicates that their unfolded states cannot be very different from a thermodynamic point of view.

Importance of Loop Length on Thermostability of the Permutant Proteins. In addition to the "basic" permuted forms of the protein, we constructed elongated forms where one, three, and five extra-Gly residues were added to the new formed loop to increase its flexibility and size, in order to describe the influence of the loop length in the stability of the protein. This is an important subject since theoretical analysis of the effect of loop length on protein folding and stability is based on the assumption that the changes must be of entropic origin. Addition of the first Gly residue seems to stabilize the two permutant proteins, suggesting that there is some conformational restraint in the new loop that is alleviated by inserting one amino acid (Table 1). Therefore, this indicates that this mutant should be used as a reference and not the WT protein as was explained previously (*28*). In both families, the stability decreases upon increasing the loop size in a very similar way (on average no more than 1 kJ·mol^{-1} by extra-Gly added). This tendency seems to be purely entropic in nature, since we cannot find changes in the unfolding heat effects or in the heat capacity change of unfolding on changing the loop length. However, we cannot discard small compensatory enthalpic changes. The entropic effect mentioned above should arise from an increasing difficulty to close the loop on the formation of the antiparallel β-strand. The most interesting result is that insertion of a long Gly loop in the two permutant proteins produces a very small destabilization effect (*28*). Similar results have been found in the Chymotrypsin inhibitor protein CI-2 when 10 Gly, or Gln, residues have been inserted in the long reactive loop (*30*). The opposite of this insertion of up to 10 Gly residues in a short loop in ROP protein, produces a very large destabilizing effect (up to 2.5 kcal·mol^{-1}) (*31*). The main difference between the ROP experiment and those done on CI-2 and SH3 domains is that the insertion region in the first case is short and therefore rather rigid, while in the other two cases it takes place in a longer flexible loop. These results suggest that loop length is not necessarily a major thermodynamic problem in protein stability as long as its flexibility is not conformationally constrained.

Characterization of the Transition State of Folding and Folding Velocity. The activation heat capacity, enthalpy, and entropy can be calculated from the analysis of the Eyring plots for unfolding and refolding. Curvature in the Eyring plot is considered to represent changes in heat capacity on going from the denatured to the transition state and/or from the native to the transition state. However, it is possible that the presence of residual structure in the unfolded state that is destabilized at higher temperatures would provoke also a curvature. This does not seem to be the case in the α-spectrin SH3 domain since we observe curvature in the refolding Eyring plots, when the experiments were performed in the presence of different final urea concentrations (urea should destabilize any residual structure present at low pH and temperature) (data not shown). As it was found for CI-2 (*1*, *2*), we observed a significant curvature in the refolding plots for the PWT and the two permutants, indicating the burial of a significant hydrophobic surface in the refolding transition state. In the unfolding reaction there is an apparent small curvature for the three proteins. Fitting of the data using the calorimetric results shows that, in the case of the S19-P20s permutant, this curvature represents the same change in heat capacity as in the refolding reaction (Table 4). In the case of the N47-D48s permutant we find the same behavior. The PWT, on the other hand, behaves like CI-2 (*2*): the change in heat capacity in the refolding reaction is three times larger than in the unfolding direction. Regarding the WT protein, although we do not discuss the data here since it is not our reference protein, we found the same overall behavior as for the PWT protein (data not shown). These results indicate that the transition state of the permutants is more solvent accessible than that of the PWT protein.

The activation enthalpy, entropy, and Gibbs energy changes, as we explain below, fit reasonably well with previous structural and kinetic data. The S19-P20s mutant

has a lower Gibbs energy activation barrier in both directions (Table 4), in agreement with the kinetic data that showed faster unfolding and refolding rates than in the two other proteins (ref *11* and this work). This lower activation barrier seems to be due to a lower entropy barrier in the refolding semireaction and to a lower enthalpic barrier in the unfolding semireaction (Table 4). The X-ray structure of this permutant showed a local unzipping of the cleaved loop, resulting in the loss of 9 hydrogen bonds. This explains the lower enthalpic unfolding activation barrier. The analysis of the transition state of this protein shows that the region around Thr24 in this loop is more disorganized than in the N47-D48s and WT proteins (*11*), explaining the lower entropic refolding barrier.

The Eyring plot analysis of the N47-D48s permutant shows a different entropic and enthalpic balance in both semireactions compared to the PWT protein, which are difficult to interpret. However, it is clear that the overall entropic cost is larger than in the PWT protein as found by calorimetry and in agreement with the fact that the cleaved β-turn (N47-D48) is as structured as in the PWT protein. The above data agree with other evidences that the transition states of the three species are different (*11*).

Comparison of the data obtained for the three proteins with that obtained with another two-state transition protein, CI-2 (*2*), shows some similarities and discrepancies. In both cases the main enthalpic change occurs when going from the native to the transition state, indicating the breaking of numerous interactions. However, while in CI-2 this change in enthalpy was accompanied by a significant change in entropy, in our three proteins the entropic change in the transition between the native and the transition state is rather small, which agrees with the fact that their unfolding kinetic rates are larger. However, refolding is going faster in CI2 where the entropy change is clearly lower (*2*). The entropy includes two components: the entropy component associated with the increase of the configurational freedom of the polypeptide chain (which is usually positive and its value upon unfolding is 15−20 J·K^{-1} per mole of amino acid residue at 298 K, as derived from theoretical studies (*32*−*34*)) and the other with the hydration of groups that become exposed on unfolding (this change is negative, even if one takes into account that which concerns the hydration of polar and nonpolar groups (*35*)). With these considerations it is clear that the lower values obtained for the entropy change of the transition state of these proteins should arise from the fact that this state appears to be more ordered than that of CI-2, but also more solvent-exposed.

In summary, the transition state of α-spectrin SH3 domain seems to be a high-energy form of the native state in which some stabilizing interactions, present in the native structure, have been lost. Loss of such interactions seems to be accompanied by a partial opening of the hydrophobic core, which may be interpreted as an instance of Hammond behavior. This situation is similar to the ones found for a mutant of T4 lysozime (*36*), chimotrysin inhibitor 2 (*1*, *2*, *37*), and barnase (*38*−*40*). The small destabilization observed by inserting up to 6 Gly residues in a loop seems to be entropic in nature as expected. Rapid folding seems to be achieved by having a very unstructured transition state, and small entropic barriers appear to be the reason that modulates folding velocities.

REFERENCES

1. Jackson, S. E., and Fersht A. R. (1991) *Biochemistry 30*, 10428−10435.
2. Jackson, S. E., and Fersht, A. R. (1991) *Biochemistry 30*, 10436−10443.
3. Alexander, P., Orban, J., and Bryan, P. (1992) *Biochemistry 31*, 7243−7248.
4. Viguera, A. R., Martínez, J. C., Filimonov, V. V., Mateo, P. L., and Serrano, L (1994) *Biochemistry 33*, 2142−2150.
5. Huang, G. S., and Oas, T. G. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 6878−6882.
6. Schindler, T., Herrler, M., Marahiel, M. A., and Schmid, F. X. (1995) *Nat. Struct. Biol. 2*, 663−673.
7. Kragelund, B. B., Robinson, C. V., Knudesn, J., Dobson, C. M., and Poulsen, F. M. (1995) *Biochemistry 34*, 7217−7224.
8. Villegas, V., Azuaga, A. I., Catasus, Ll., Reverter, D., Mateo, P. L., Aviles, F. X., and Serrano, L. (1995) *Biochemistry 34*, 15105−15110.
9. Viguera, A. R., Villegas, V., Aviles, F. X., and Serrano, L. (1996) *Folding Des. 2*, 23−33.
10. Viguera, A. R., Blanco, F. J., and Serrano, L (1995) *J. Mol. Biol. 247*, 670−681.
11. Viguera, A. R., Wilmanns, M., and Serrano, L. (1996) *Nat. Struct. Biol. 3*, 874−880.
12. Musacchio, A., Noble, M. E. M., Pautit, R., Wierenga, R., and Saraste, M. (1992) *Nature 359*, 851−855.
13. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
14. Makhatadze, G. I., and Privalov, P. L. (1990) *J. Mol. Biol. 213*, 375−384.
15. Prieto, J., Wilmanns, M., Jimenez, M. A., Rico, M., and Serrano, L. (1997) *J. Mol. Biol. 268*, 760−778.
16. Teng, T. Y. (1990) *J. Appl. Crystallogr. 23*, 387−391.
17. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
18. Gerwirth, D. (1986) *The HKL manual: a description of the programs Denzo*, XDisplayF, Scalepack.
19. Navaza, J. (1994) *Acta Crystallogr., Sect. A 50*, 157−163.
20. Murshudov, G. N., Vagin, A., and Dodson, E. J. (1996) *Acta Crystallogr., Sect. D 53*, 240−255.
21. Lamzin, V. S., and Wilson, K. S. (1993) *Acta Crystallogr., Sect. D 49*, 129−147.
22. Jones, T. A., and Kjeldgaard, M. (1992) *O-The manual*, Uppsala, Sweden.
23. Kleywegt, G. J., and Jones, T. A. (1994) *From first map to final model* (Bayley, S., Hubbard, R., and Waller, D. A., Eds.) pp 59−66.
24. Wilson, K. S., Butterworth, S., Dauter, Z., Lamzin, V. S., Walsh, M., Wodak, S., Pontius, J., Richelle, J., Vaguine, A., Sander, C., Hooft, R. W. W., Vriend, G., Thornton, J. M., Laskowsky, R. A., MacArthur, M. W., Dodson, E. J., Murshudov, G., Oldfield, T. J., Kaptein, R., and Rullman, J. A. C. (1998) *J. Mol. Biol. 276*, 417−436.
25. Fu, L., and Freire, E. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9335−9338.
26. Vogl, T., Hinz, H.-J., and Hedwig, G. R. (1995) *Biophys. Chem. 54*, 261−269.
27. Blanco, F. J., Ortiz, A. R., and Serrano, L. (1996) *J. Biomol. NMR 9*, 347−357.
28. Viguera, A. R., and Serrano, L. (1997) *Nat. Struct. Biol. 4*, 939−946.
29. Matthews, B. W., Nicholson, H., and Becktel, W. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6663−6667.
30. Ladurner, A. G., and Fersht, A. R. (1997) *J. Mol. Biol. 273*, 330−337.
31. Nagi, A. D., and Regan, L. (1997) *Folding Des. 2*, 67−75.
32. Nemethy, G., Leach, S. J., and Scheraga, H. A. (1966) *J. Phys. Chem. 70*, 998−1004.
33. Bryngelson, J. D., and Wolynes, P. G. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7524−7528.
34. Lee, K. H., Xie, D., Freire, E., and Amzel, L. M. (1994) *Proteins: Struct., Funct., Genet. 20*, 68−84.
35. Makhatadze, G. I., and Privalov, P. L. (1996) *Protein Sci. 5*, 507−510.

36. Chen, B., Baase, W. A., and Shellman, J. A. (1989) *Biochemistry 28*, 691−699.
37. Itzhaki, L. S., Otzen, D. E., and Fersht, A. R (1995) *J. Mol. Biol. 254*, 260−288.
38. Matouscheck, A., and Fersht, A. R. (1993) *Proc. Natl. Acad. Sci. U.S.A 90*, 7814−7818.
39. Mattews, J. M., and Fersht, A. R. (1995) *Biochemistry 34*, 6805−6814.
40. Matouscheck, A., Otzen, D. E., Itzhaki, L. S., Jackson, S. E., and Fersht, A. R. (1995) *Biochemistry 34*, 13656−13662.

BI981515U