Thermal Stability and DNA Binding Activity of a Variant Form of the Sso7d Protein from the Archeon *Sulfolobus solfataricus* Truncated at Leucine 54[†]

Erlet Shehi,[‡] Vincenzo Granata,[§] Pompea Del Vecchio,[§] Guido Barone,[§] Paola Fusi,[‡] Paolo Tortora,[‡] and Giuseppe Graziano*,[||]

Dipartimento di Biotecnologie e Bioscienze, Università di Milano-Bicocca, Piazza della Scienza, 2-20126 Milano, Italy, Dipartimento di Chimica, Università di Napoli Federico II, Via Cinthia, 45-80126 Napoli, Italy, and Dipartimento di Scienze Biologiche ed Ambientali, Università del Sannio, Via Port'Arsa, 11-82100 Benevento, Italy

Received April 2, 2003; Revised Manuscript Received May 23, 2003

ABSTRACT: Sso7d is a 62-residue, basic protein from the hyperthermophilic archaeon *Sulfolobus solfataricus*. Around neutral pH, it exhibits a denaturation temperature close to 100 °C and a non-sequence-specific DNA binding activity. Here, we report the characterization by circular dichroism and fluorescence measurements of a variant form of Sso7d truncated at leucine 54 (L54Δ). It is shown that L54Δ has a folded conformation at neutral pH and that its thermal unfolding is a reversible process, represented well by the two-state N ⇔ D transition model, with a denaturation temperature of 53 °C. Fluorescence titration experiments indicate that L54Δ binds tightly to calf thymus DNA, even though the binding parameters are smaller than those of the wild-type protein. Therefore, the truncation of eight residues at the C-terminus of Sso7d markedly affects the thermal stability of the protein, which nevertheless retains a folded structure and DNA binding activity.

Sso7d[1] is a small, 62-residue, basic (pI 10.2) protein (*1*) from *Sulfolobus solfataricus*, a thermoacidophilic archaebacterium that thrives at 87 °C and acidic pH in volcanic hot springs (*2*). The physiological role of the protein is not entirely clarified because it exhibits a variety of activities. In *in vitro* studies, Sso7d proves to be a non-sequence-specific DNA-binding protein (*3*−*6*), increasing the melting temperature of DNA. Moreover, it promotes the annealing of complementary DNA strands (*7*) and induces negative supercoiling (*8*) and a kink with unwinding of the DNA double helix (*9*, *10*). In addition, Sso7d also possesses a ribonuclease activity (*11*, *12*), and is able to rescue aggregated proteins in an ATP hydrolysis-dependent manner (*13*).

The three-dimensional (3D) structure of Sso7d has been determined in solution by NMR from both natural (PDB entry 1SSO) and recombinant (PDB entry 1JIC) sources, and exhibits a topology similar to that of eukaryotic Src-homology 3 (SH3) domains. The compact globular fold is formed by a double-stranded antiparallel β-sheet (made up of residues 2−7 and 10−15) onto which an orthogonal triple-stranded antiparallel β-sheet is packed (made up of residues 21−25, 28−33 and 41−46). The C-terminus consists of an amphipatic and distorted helical stretch, from Glu47 to Gln61. Specifically, residues 47−50 form a complete turn of a 3_{10}-helix, interrupted by Pro51; residues 52−61 form approximately three turns of α-helix. A cluster of aromatic side chains (Phe5, Tyr7, Phe31, and Tyr33) in herringbone geometry, at the interface of the two β-sheets, constitutes the compact hydrophobic core. Several basic (2 Arg and 13 Lys residues) and acidic (6 Glu and 3 Asp residues) side chains are located on the protein surface; there are no histidines, asparagines, or cysteines. Sso7d is a very stable protein with respect to both pressure and temperature (*14*, *15*); in particular, $T_d \cong 98$ °C and $\Delta_d H(T_d) \cong 270$ kJ/mol over the pH range of 5.0−7.0 (*14*). In previous works, the critical role played by Phe31 in the stability of the folded structure has been elucidated (*15*−*18*).

A variant form of Sso7d truncated at Leu54 (L54Δ) has recently been produced in our labs. The aim is to investigate the role of the C-terminal α-helix for the stability and DNA binding activity of Sso7d by deleting most of the residues constituting such a helical segment. The thermal stability has been investigated by means of circular dichroism (CD) measurements, whereas the DNA binding activity has been measured by recording the intrinsic fluorescence quenching of Trp23 upon association. The experimental data emphasize that (a) the truncated form has a folded structure around neutral pH, (b) it is significantly less stable against temperature than the wild-type protein since the denaturation temperature decreases by 46 °C, and (c) the deletion of the C-terminal α-helix does not impair the DNA binding activity since the association constant for binding of L54Δ to calf

[†] Work supported by grants from the Italian Ministry for Instruction, University and Research (MIUR, Rome, Italy) and from the Italian Space Agency (ASI).

* To whom correspondence should be addressed. Phone: +39/0824/305101. Fax: +39/0824/23013. E-mail: graziano@unisannio.it.

[‡] Università di Milano-Bicocca.

[§] Università di Napoli Federico II.

[||] Università del Sannio.

[1] Abbreviations: Sso7d and Sac7d, 7 kDa DNA-binding proteins from *S. solfataricus* and *Sulfolobus acidocaldarius*, respectively; L54Δ, variant form of Sso7d truncated at leucine 54; CD, circular dichroism; DSC, differential scanning calorimetry; MALDI-TOF/MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; GuHCl, guanidine hydrochloride.

10.1021/bi034520t CCC: $25.00
Published on Web 06/19/2003

thymus DNA (ct-DNA) is similar to that of Sso7d. Such results are discussed on the basis of the structural information available for Sso7d and its complexes with double-stranded oligonucleotides.

MATERIALS AND METHODS

Protein and DNA Preparation. Sso7d from *S. solfataricus* was expressed in *Escherichia coli* and purified as described previously (*19*). The recombinant protein proved to be indistinguishable from its natural counterpart on the basis of its DNA binding and ribonuclease activities and stability against temperature, even though no lysine monomethylation was found (*18*). To produce the variant truncated at Leu54 (L54Δ), the Sso7d-encoding gene was cloned into a pGEM 7Zf(+) plasmid as reported previously (*19*). The truncated mutant was produced by cassette mutagenesis of the wild-type gene. Two partially overlapping oligonucleotides, one carrying the deletion, were extended with the Klenow enzyme. A *Nar*I−*Hin*dIII fragment from the wild-type Sso7d-encoding gene was excised and replaced with the mutagenic cassette. The mutated gene was then sequenced, excised with *Hin*dIII and *Xba*I, and subcloned into plasmid pT7-7. Expression was achieved in *E. coli* strain BL21(DE3)-pLysE grown as described previously (*15*), using isopropyl β-D-thiogalactopyranoside as an inducer. The expressed protein represented ∼15% of the total protein content in crude extracts. Purification of the mutant protein was performed by CM-Sephadex C-25 chromatography as previously reported for wild-type Sso7d (*15*). This procedure allowed us to obtain ∼7.5 mg of protein per liter of cell culture.

The purity of the proteins was confirmed by means of SDS−PAGE and MALDI-TOF mass spectrometry. Before further measurements, sample solutions were dialyzed against the required buffer at 4 °C for 24 h. Dialysis tubes with a cutoff limit of 3500 Da were used. The concentration of both Sso7d- and L54Δ-dialyzed samples was determined spectrophotometrically using an ϵ_{280} of 8300 M^{-1} cm^{-1}, calculated by the method of Gill and von Hippel (*20*), from the absorption of tyrosine (ϵ_{280} = 1400 M^{-1} cm^{-1}) and tryptophan (ϵ_{280} = 5500 M^{-1} cm^{-1}).

Calf thymus DNA was purchased from Sigma (St. Louis, MO) and used without further purification. Lyophilized ct-DNA was dissolved in a solution containing 0.2 M NaCl and sonicated by means of a Vibracell sonicator at 0 °C in 10 min long steps. After each sonication step, the absorbance of the ct-DNA solution at 260 nm showed no changes, indicating that no denaturation of the double helix had occurred during the process. The molecular weight of the obtained fragments corresponded to an average length of 500 base pairs, as measured by intrinsic viscosity measurements (*21*). The ct-DNA samples for fluorescence measurements were prepared by exhaustively dialyzing them against 10 mM sodium phosphate buffer (pH 7.0). The ct-DNA concentration was determined spectrophotometrically based on an ϵ_{260} of 13 200 M^{-1} cm^{-1}, expressed as the molarity of base pairs.

All measurements were taken in 10 mM sodium phosphate buffer. The pH was measured at 25 °C with a Radiometer pHmeter (model PHM93). The phosphate buffer has a low protonation enthalpy (*22*), so the solution pH depends little upon temperature. Doubly deionized water was used throughout. A commercial 8 M solution from Sigma was used for GuHCl.

Circular Dichroism Measurements and Their Analysis. CD spectra were recorded with a Jasco J-715 spectropolarimeter equipped with a Peltier-type temperature control system (model PTC-348WI). The instrument was calibrated with an aqueous solution of *d*-10-(+)-camphorsulfonic acid at 290 nm (*23*). The molar ellipticity per mean residue ([θ] in degrees per square centimeter per decimole) was calculated from the equation $[\theta] = ([\theta]_{obs}\text{mrw})/(10lC)$, where $[\theta]_{obs}$ is the ellipticity measured in degrees, mrw is the mean residue molecular weight (113 Da), C is the protein concentration in grams per liter, and l is the optical path length of the cell in centimeters. Cuvettes with path lengths of 0.2 and 0.5 cm and protein concentrations of 0.2 and 0.8 mg/mL were used in the far-UV and near-UV regions, respectively. CD spectra were recorded with a time constant of 4 s, a 2 nm bandwidth, and a scan rate of 5 nm/min, and were signal-averaged over at least five scans, and baseline corrected by subtracting a buffer spectrum. Thermal unfolding curves were recorded in the temperature mode, over the range of 5−110 °C, with a scan rate of 1.0 °C/min. Samples were rapidly cooled after the first heating run and scanned for a second time to estimate the reversibility of the unfolding transition.

Thermal unfolding transitions were analyzed with the two-state N ⇔ D model whose equilibrium constant is given by (*17*)

$$K_d(T) = \exp\{[-\Delta_d H(T_d)/R](1/T - 1/T_d)\} \qquad (1)$$

where T_d is the denaturation temperature at which $K_d = 1$ and $\Delta_d H(T_d)$ is the denaturation enthalpy change. The denaturation heat capacity change, $\Delta_d C_p$, is considered to equal zero because it cannot reliably be determined from CD measurements (*24*). Correspondingly, the observed molar ellipticity is

$$[\theta] = ([\theta]_N + [\theta]_D K_d)/(1 + K_d) \qquad (2)$$

where $[\theta]_N$ and $[\theta]_D$ are the molar ellipticities of the native and denatured states, respectively, which are assumed to depend linearly on temperature. A nonlinear least-squares regression was carried out to estimate the unknown parameters associated with the unfolding transition, using the Levenberg−Marquardt algorithm, as implemented in the Optimization Toolbox of MATLAB. Since Sso7d has a denaturation temperature close to 100 °C around neutral pH (*14*), and the CD instrument works up to 110 °C, incomplete thermal unfolding transitions have been recorded. The reliability of the thermodynamic parameters obtained from such incomplete transitions was analyzed in detail in a previous article (*17*). For all the examined transition curves, the standard deviation of the calculated points from the experimental ones is very small, indicating the good quality of the fit.

Fluorescence Measurements and Their Analysis. The binding of Sso7d and L54Δ to ct-DNA was studied by determining the quenching of the intrinsic tryptophan fluorescence upon DNA binding. Fluorescence measurements were performed on a Jasco FP777 spectrofluorimeter with excitation at 290 nm (4 nm slit width) to exclude the tyrosine contribution, and emission was monitored at 350 nm (8 nm

slit width). Titration experiments were carried out as reverse titrations by analyzing samples in which the ct-DNA concentration increased and the protein concentration was kept constant (7−8 μM). A sealed quartz cuvette with a path length of 1 cm was used. Measurements were performed in triplicate at 20 °C and pH 7.0 in 10 mM sodium phosphate buffer.

The fractional fluorescence quenching (Q_{obs}) was calculated as $(I_o - I)/I_o$, where I_o and I represent the protein fluorescence intensity at 350 nm observed in the absence and presence of ct-DNA, respectively. The concentration of protein bound to ct-DNA was calculated as

$$[P]_{bound} = [P]_{tot}(Q_{obs}/Q_{max}) \quad (3)$$

where $[P]_{tot}$ is the total protein concentration and Q_{max} is the maximum quenching, i.e., the quenching observed when all the protein in the sample is bound to DNA. This is based on the assumption that the fractional change in fluorescence quenching upon DNA binding is equal to the fraction of bound protein. The validity of this assumption was verified for the binding of the homologous Sac7d to DNA (*25*), by means of a procedure devised by Bujalowski and Lohman (*26*).

The binding of Sso7d and L54Δ to ct-DNA was evaluated by means of the McGhee−von Hippel (*27*) relationship for the association of a large ligand with a linear lattice of overlapping binding sites:

$$[P]_{free} = r/K_b(1 - nr)\{(1 - nr)/[1 - (n - 1)r]\}^{n-1} \quad (4)$$

where $r = [P]_{bound}/[DNA_{bp}]$, K_b is the binding constant, and n is the number of lattice residues covered by one ligand molecule (i.e., the number of DNA base pairs covered by the protein). All the experimental data are well represented by eq 4, so the inclusion of a cooperative interaction parameter was not required. For the analysis of fluorescence data there are two fitting parameters: K_b and n. The latter values were determined by means of a nonlinear least-squares fit of eq 4 to the experimental data, and by simultaneously solving eq 3 for $[P]_{bound}$ with an iterative procedure. The reported binding parameters and errors are the mean values and standard deviations, respectively, for three independent fluorescence titration experiments.

RESULTS

Thermal Stability. The thermal stability of both Sso7d and L54Δ was investigated by means of CD measurements. The far-UV CD spectra of L54Δ are qualitatively similar to those of the wild-type protein over the pH range of 4.0−9.0. The spectra of the two proteins at pH 6.0 are shown in Figure 1. This finding emphasizes that the truncated form has a folded structure in aqueous solution over such a pH range at room temperature. Analysis of the far-UV CD spectra according to the self-consistent method (*28*) has given (a) 11% α-helix and 48% β-sheet in Sso7d, in line with the estimate based on its 3D NMR structure (*3*, *18*), and (b) 4% α-helix and 53% β-sheet in L54Δ. Therefore, the quantitative difference between the spectra of the two proteins is largely due to the decrease in the size of the α-helix contribution as a consequence of the deletion. On the other hand, the polypeptide chain of L54Δ is unfolded both at 5 °C and pH 3.0 and at 5 °C and pH 6.0 in the presence of 1.0 M GuHCl (see Figure 1). These results indicate that (a) L54Δ has a sensitivity to acidic pH significantly higher than that of Sso7d, which is properly folded also at pH 2.5, where $T_d \cong$ 65 °C (*14*, *17*), and (b) it has a poor resistance against the denaturing action of GuHCl.

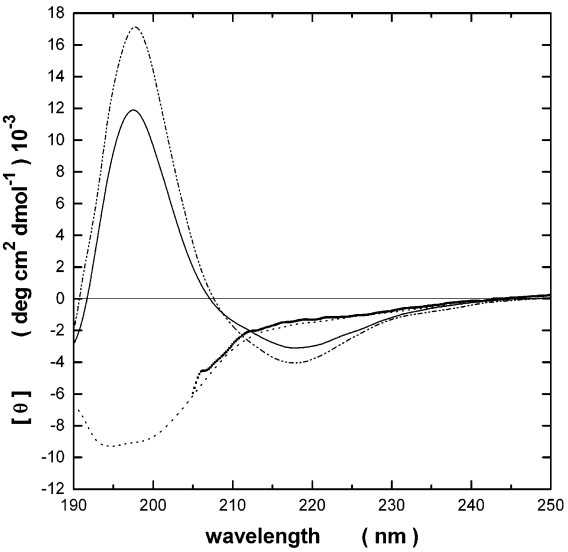

FIGURE 1: Far-UV CD spectra of L54Δ at 5 °C and pH 6.0 (—), pH 3.0 (···), and pH 6.0 and 1.0 M GuHCl (■) and far-UV CD spectrum of Sso7d at 20 °C and pH 6.0 (—··—).

It is worth mentioning that L54Δ is able to properly fold in solution despite its small size and the lack of any disulfide bridge. It should also be noted that a variant form of Sso7d truncated at Glu53 could not be isolated under the same conditions which allowed us to purify L54Δ (unpublished results). This strongly suggests that it failed to attain a properly folded structure, which highlights a crucial role of Leu54 in the folding process. In keeping with this idea, Hard and colleagues emphasized that the side chain of Leu54 is packed well against that of Ala50, anchoring the C-terminal end of the chain to the protein core (*3*). Also, a molecular dynamics simulation confirmed that Leu54 is involved in strong van der Waals interactions with the remaining part of the protein (*16*). In addition, Consonni and colleagues (*18*) pointed out (a) the presence of an H-bond between the carboxyl oxygen of Glu53 and the N-terminal group of Ala1, keeping close to each other the two ends of the chain, and (b) the very slow H−D exchange of the amide proton of Leu54, suggestive of its poor solvent accessibility.

It is worth noting that L54Δ has a limited solubility in aqueous solution, at most 0.9 mg/mL at 25 °C and pH 7.0. The limited solubility is likely to be related to the loss of three net charges (one negative, Glu59, and two positive, Lys60 and Lys62) and the exposure of hydrophobic moieties as a consequence of the deletion of the last eight residues in the Sso7d primary structure. Such limited solubility prevented us from carrying out differential scanning calorimetry (DSC) measurements to investigate the thermal stability of L54Δ, since a protein concentration of at least 2−3 mg/mL would be required (*14*, *17*).

Thermal unfolding of Sso7d and L54Δ was monitored by recording the molar ellipticity at 200 nm, over the range of 5−110 °C, using a heating rate of 1.0 °C/min. The results are presented in Table 1. On the basis of both the observation

Table 1: Thermodynamic Parameters of the Thermal Unfolding of Sso7d and L54Δ, Obtained by Recording the Molar Ellipticity at 200 nm and Different pH Values[a]

	Sso7d		L54Δ	
pH	T_d (°C)	$\Delta_dH(T_d)$ (kJ/mol)	T_d (°C)	$\Delta_dH(T_d)$ (kJ/mol)[1]
3.0	80.0	210	unfolded	–
4.0	95.0	240	44.0	110
5.0	98.5	260	52.0	125
6.0	99.0	270	53.0	125
7.0	98.0	270	52.0	125
8.0	97.5	260	51.0	115
9.0	97.0	250	49.0	95

[a] For all measurements, the buffer was 10 mM phosphate buffer. For each pH, three independent measurements were performed. Each value is the average of the values calculated by the nonlinear regression over the three CD measurements. The uncertainty in the T_d estimates does not exceed 0.5 °C, whereas the uncertainties in the estimates for $\Delta_dH(T_d)$ amount to 10% of the reported values.

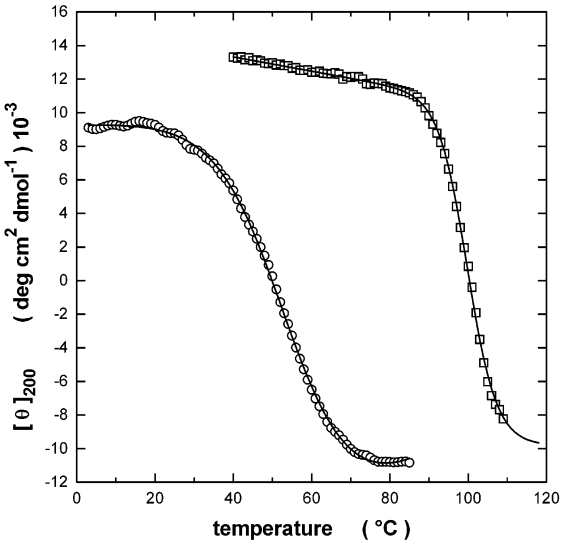

FIGURE 2: Thermal transition curves of L54Δ and Sso7d, obtained by recording the molar ellipticity at 200 nm and pH 6.0 in 10 mM phosphate buffer. Experimental data are shown for L54Δ (○) and for Sso7d (□); fitted values are shown as continuous lines.

that the two proteins consist almost entirely of β-sheets and the notion that in a β-sheet the noncovalent interactions are mainly nonlocal, it follows that molar ellipticity recorded at 200 nm is indicative of changes in both the secondary and tertiary structures. Thermal unfolding proved to be a reversible process, according to the reheating criterion, for both proteins.

The thermal unfolding curves were fit well by eqs 1 and 2, indicating that the two-state N ⇔ D transition model is valid. Both proteins exhibited their maximum denaturation temperatures at pH 6.0, even though the stability depends little on pH over the range of 5.0−8.0. Specifically, at pH 6.0, the thermal unfolding of Sso7d is characterized by a T_d of 99 °C and a $\Delta_dH(T_d)$ of 270 kJ/mol, whereas that of L54Δ is characterized by a T_d of 53 °C and a $\Delta_dH(T_d)$ of 125 kJ/mol. These values indicate that the deletion of the C-terminal α-helix causes a dramatic decrease in both T_d and $\Delta_dH(T_d)$ with respect to the values for the native protein, as shown in Figure 2. In fact, T_d decreased by 46 °C and $\Delta_dH(T_d)$ by ∼150 kJ/mol. It is worth noting that there is good agreement between our determinations of T_d and $\Delta_dH(T_d)$ for Sso7d and those reported by Knapp and colleagues (*14*). They found, however, that, at pH >7.0, the thermal unfolding of Sso7d was an irreversible process characterized by nonsymmetric DSC peaks (*14*). The cause could be the aggregation of unfolded molecules at pH values close to the isoelectric point of Sso7d. We did not observe such irreversibility in far-UV CD measurements, likely because the protein concentration used is 1 order of magnitude smaller than that necessary for DSC experiments (i.e., 0.2 mg/mL vs 2.0 mg/mL).

Since the thermal transition curves were fit well by eqs 1 and 2 that do not consider the existence of Δ_dC_p, the latter quantity has not been determined. Furthermore, for L54Δ, this quantity cannot be estimated as the slope of the plot of $\Delta_dH(T_d)$ versus T_d because the experimental points are clustered together over the pH range where the protein has a folded structure. In any case, we can use a Δ_dC_p of 2.7 kJ K^{-1} mol^{-1}, as determined by Knapp and colleagues (*14*), for both Sso7d and L54Δ to construct the so-called protein stability curve (*29*). At pH 6.0, one obtains a temperature of maximum stability (T_{max}) of ≅285 K for both proteins, whereas the maximum Gibbs energy of stabilization [$\Delta_dG(T_{max})$] equals 33 kJ/mol for Sso7d and 8 kJ/mol for L54Δ. By performing a comparison at a T_d of 53 °C of L54Δ, one obtains the following values: (a) Δ_dH = 146 kJ/mol, Δ_dS = 369 J K^{-1} mol^{-1}, and Δ_dG = 25.6 kJ/mol for Sso7d, and (b) Δ_dH = 125 kJ/mol, Δ_dS = 383 J K^{-1} mol^{-1}, and Δ_dG = 0 for L54Δ. These numbers indicate that the difference in Δ_dG between the two proteins is largely due to enthalpic factors. The dramatic effect on the T_d value caused by the deletion of the last eight residues of Sso7d should be related to the role played by the C-terminal α-helix in capping the small β-barrel and shielding the hydrophobic core from the contact with water. Molecular dynamics simulations on Sso7d and L54Δ are in progress in our labs in an effort to try to provide a microscopic rationalization of the dramatic difference in thermal stability.

It is worth noting that the thermal transition curves of L54Δ are very broad, spanning more than 30 °C. However, the presence of intermediates between the native state and the denatured one can be ruled out. In fact, the thermal transition curves obtained by recording the molar ellipticity at 275 nm in the near-UV region are wholly superimposable with those obtained by recording the molar ellipticity at 200 nm, as shown in Figure 3. Thus, the independence of the wavelength of the thermal unfolding transition shows that the secondary and tertiary structures break down together, which in turn strongly supports the idea that the thermal unfolding of L54Δ is well represented by the two-state N ⇔ D transition model (*30*). Usually, the existence of only two significantly populated states is supported by the presence of isodichroic points in CD spectra recorded at different temperatures. In the case of L54Δ, an isodichroic point occurs at 292 nm in the near-UV region, but only an approximate convergence of the spectra at around 212 nm occurs in the far-UV region (data not shown). This is in line with the results previously obtained for Sso7d (*14*) and Sac7d (*31*).

Therefore, the broadness of the thermal transition in the case of L54Δ just results from a small value of $\Delta_dH(T_d)$. In fact, the T_d and $\Delta_dH(T_d)$ values of L54Δ at pH 6.0 imply that more than 6% of the L54Δ molecules are already unfolded at 35 °C, a temperature significantly lower than

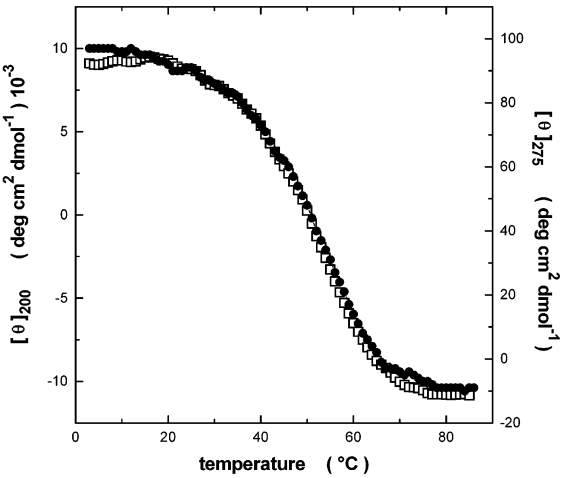

FIGURE 3: Thermal transition curves of L54Δ, obtained by recording the molar ellipticity at 200 (□) and 275 nm (●) at pH 6.0 in 10 mM phosphate buffer.

the T_d of 53 °C. In this respect, it has to be noted that a two-state N ⇔ D transition seems to be sharper when the $\Delta_d H(T_d)$ value is increased and T_d is kept fixed; however, the cooperativity does not change because there are only two thermodynamic states populated, which can be easily verified by simple computations.

An independent approach may also lead to the same conclusions. As originally suggested by Liquori (*32*), a globular protein molecule can be considered a "crystal" molecule, and treated as a thermodynamic system *per se*. In view of the large surface-to-volume ratio, a crystal molecule is characterized by large fluctuations of the thermodynamic quantities around the average values, as pointed out by Cooper (*33*). Clearly, the effect of thermal fluctuations is even more pronounced for a very small protein such as L54Δ, by recognizing that the $\Delta_d H(T_d)$ values are roughly proportional to the number of residues of the protein (*34*). Thus, the broadness of L54Δ thermal unfolding is fully compatible with the two-state N ⇔ D transition model, and has nothing to do with the noncooperative, gradual, and barrier-less thermal unfolding recently identified for the BBL domain by Muñoz and colleagues (*35*).

DNA Binding Activity. The DNA binding activity of L54Δ was studied by measuring the intrinsic fluorescence quenching of Trp23 upon binding, using as a reference Sso7d. Since Hard and colleagues (*3*−*6*) did not find significant differences in the thermodynamics of Sso7d binding to poly(dAdT) and poly(dGdC), we performed measurements using sonicated ct-DNA samples. A blue shift in the emission spectrum of both proteins was observed upon binding to ct-DNA; the fluorescence maximum shifted from 350 nm to approximately 345 nm, indicating a modification of the Trp23 local environment. The quenching of fluorescence intensity at 350 nm amounts, on the average, to 91% in the case of Sso7d and to 78% in the case of L54Δ.

The fluorescence titration curves of Sso7d and L54Δ, constructed at 20 °C and pH 7.0 in 10 mM sodium phosphate buffer from the intensity values at 350 nm, are shown in Figure 4. This figure indicates that the truncated variant binds tightly to ct-DNA. The fitting procedure with respect to eqs 3 and 4 provided the following values: (a) $K_b = (4.4 \pm 0.2) \times 10^6$ M^{-1} and $n = 3.9 \pm 0.1$ base pairs for Sso7d, and

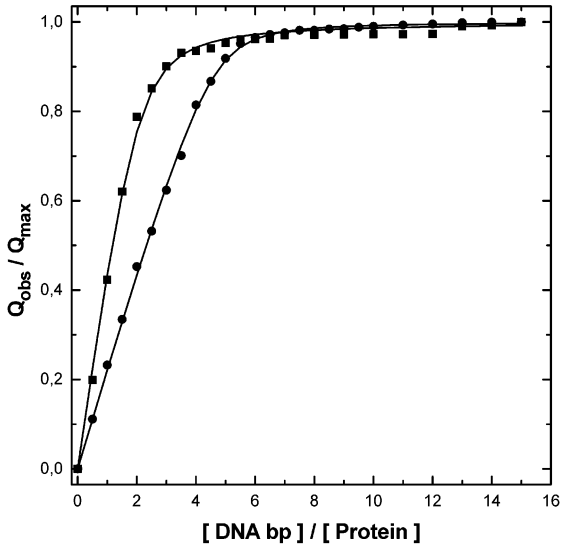

FIGURE 4: Fluorescence titration curves of Sso7d (●) and L54Δ (■) at 20 °C and pH 7.0 in 10 mM phosphate buffer. The concentration of ct-DNA is varied while that of protein is fixed at 7 μM. The solid lines are the best fits of the titration curves performed as described in Materials and Methods.

(b) $K_b = (1.6 \pm 0.2) \times 10^6$ M^{-1} and $n = 1.8 \pm 0.2$ base pairs for L54Δ. The goodness of the fit confirmed that there was no need to consider an additional parameter accounting for the negative or positive cooperativity of the binding. The above figures show that in the case of Sso7d both K_b and n were more than twice as large as those of L54Δ. However, it is worth noting that the binding Gibbs energy change $\Delta_b G$ ($=-RT \ln K_b$) equaled −37.2 kJ/mol for Sso7d and −34.8 kJ/mol for L54Δ, indicating that the energetics of binding to ct-DNA of the two proteins are similar. In fact, both proteins increase the melting temperature of ct-DNA-sonicated samples by more than 40 °C (i.e., from 70 to >110 °C), making it unable to be detected by our CD instrument.

Our experimental data also are in good agreement with those previously reported for Sso7d and Sac7d. Actually, Hard and colleagues (*6*) obtained a K_b of 8.1×10^6 M^{-1} and 4.5 base pairs (n), performing isothermal calorimetry measurements, for the binding of Sso7d to poly(dGdC) at 25 °C and pH 7.1 in 10 mM sodium phosphate and 20 mM NaCl buffer. Edmondson and colleagues (*25*) performed fluorescence measurements for the binding of Sac7d at 25 °C and pH 6.8 in 10 mM potassium phosphate and 50 mM KCl buffer, and obtained the following values: (a) $K_b = 6.5 \times 10^6$ M^{-1} and $n = 4.3$ base pairs for poly(dGdC), and (b) $K_b = 5.4 \times 10^5$ M^{-1} and $n = 4.6$ base pairs for ct-DNA.

The structures of the complexes between Sso7d and double-stranded oligonucleotides have been determined by means of both NMR (*9*), and X-ray crystal diffraction (*10*). In addition, the structures of similar complexes involving the homologous Sac7d have been determined to high resolution (*36*). The protein binds to DNA by placing its triple-stranded antiparallel β-sheet across the DNA minor groove. The external surface of this β-sheet represents a continuous region of positive electrostatic potential, as determined by numerically solving the Poisson−Boltzmann equation (*4*). Beyond electrostatic interactions, the β-sheet is anchored in the DNA minor groove by the formation of specific H-bonds involving side chains protruding into the

groove, and by the intercalation of specific nonpolar side chains (Trp23, Val25, Met28, and Ala44). In particular, the aromatic ring of Trp23 fills up the space between DNA and Sso7d, and its indole NH group is involved in an H-bond. Since the DNA minor groove contains continuous arrays of H-bonding acceptor groups, regardless of the sequence, the topology of the Sso7d−DNA complex clarifies how the binding can be non-sequence-specific but tight. In addition, all the structures that have been determined show that the binding of Sso7d causes a significant bending and unwinding of the standard B-form DNA conformation (*9*, *10*, *36*). Small-angle X-ray scattering measurements showed that regularly repeating bends in a fully saturated complex lead to a zigzag structure with negligible compaction of DNA, suggesting that the primary function of Sso7d is to protect DNA from thermal denaturation at the high growth temperature of *S. solfataricus* (*37*).

In the case of L54Δ, there are no structural data available, but the fluorescence titration measurements suggest that the binding mode should be similar to that of Sso7d. On the other hand, we have no obvious explanation for the much smaller value of the parameter n that amounts to 1.8 ± 0.2 base pairs versus 3.9 ± 0.2 base pairs for the wild-type protein. A possible explanation might be that L54Δ, being a highly fluctuating molecule because of its small size, is able to cover only two base pairs on average. This would also explain the finding that L54Δ does not possess the ribonuclease activity of Sso7d (data not shown). In fact, large fluctuations of surface side chains should not allow the correct geometric disposition of the two putative catalytic residues (i.e., Glu35 and Tyr7) for the in-line cleavage of RNA (*12*, *38*).

DISCUSSION

A major finding of this paper is that the truncated form of Sso7d, L54Δ, has a folded structure in solution around neutral pH. Actually, a polypeptide chain of 54 residues is close to what is considered the minimum length for a stable folded structure. Privalov (*39*), using thermodynamic arguments, suggested that a polypeptide chain should consist of at least 50 residues to be able to adopt a unique and folded 3D structure at room temperature. The stabilization Gibbs energy has to be significantly larger than the random thermal energy RT to prevent unfolded conformations from being significantly populated at room temperature (*40*). If one assumes a threshold $\Delta_d G$ value of ≈10RT (=25 kJ/mol) and that a single residue contributes rarely more than 0.5 kJ/mol at room temperature, on the basis of the experimental data collected so far (*34*, *41*), one readily finds that the minimum length should be around 50 residues. The same estimate can be drawn on the basis of different reasoning. By considering the surface-to-volume ratio of a globule, Dill pointed out that a minimum length of the chain is necessary to build up a well-developed hydrophobic core shielded from the aqueous medium (*42*). Short chain molecules could not fold because the hydrophobic core would be too small to overcome the large negative entropy change associated with the loss of conformational degrees of freedom. In fact, the analysis of a set of 183 nonhomologous proteins, with known 3D structures available in the Protein Data Bank, showed that a polypeptide chain should contain at least 50 residues to bury approximately half of its nonpolar accessible surface area (*43*). The present data on L54Δ and its two-state N ⇔ D thermal unfolding are in line with such theoretical estimates.

Actually, few protein domains with fewer than 50 residues, which adopt a unique structure in the absence of both disulfide bridges and metals and have a two-state N ⇔ D thermal unfolding, have been recently characterized. The 35-residue subdomain from villin headpiece, consisting of three α-helix segments, has a T_d of 70 °C (*44*). The 41-residue peripheral subunit-binding domain, consisting of two parallel α-helices connected by a short loop with a stretch of 3_{10}-helix, has a T_d of 53 °C (*45*, *46*). The structure of the first 39 residues of the ribosomal protein L9, formed by the packing of an α-helix against a triple-stranded antiparallel β-sheet, exhibits a T_d of 65 °C (*47*). The existence of cooperatively folded protein domains that are shorter than the predicted minimum length does not imply that Privalov's and Dill's arguments are wrong. The prediction that a protein must have a minimum length to have a stabilization Gibbs energy significantly larger than the random thermal energy does not contrast with the finding that such very small domains have $\Delta_d G$ values that are only 2−3 times larger than RT at room temperature (*44*−*47*).

According to the far-UV CD spectra, it is plausible to assume that L54Δ has a folded structure close to that of the wild-type protein, except for the deletion of the last eight residues. In this respect, Finkelstein and Ptitsyn (*48*) pointed out that, as a consequence of stereochemical constraints due to the polypeptide chain, the folded structures of globular proteins share the following general features: (a) the compact globule is formed by the packing of regular secondary structure elements that reach from one border of the globule to the other; (b) the α and β segments are connected by turns that are not located inside the globule but lie on the surface and are surrounded by water; and (c) the turns are the fundamental kink points for the folding of the chain. These features can be considered key requirements in constructing a thermodynamically stable globular structure.

It is evident that the supposed structure of L54Δ entirely fulfills the requirements highlighted by Finkelstein and Ptitsyn, as the tight packing of the two β-sheets gives rise to the globular structure. Such a structural motif, i.e., the packing with orthogonal orientation of two β-sheets, allows the formation of a stable globule with chains of minimum length, by taking advantage of the occurrence of hydrophobic residues on one face of the β-sheet and of hydrophilic ones on the other side. This gives rise therefore to a well-packed hydrophobic core, whereas the hydrophilic side chains point to the exterior and interact favorably with water molecules, to ensure a good solubility in aqueous solutions. Finally, the backbone CONH groups in each β-sheet are involved in interstrand H-bonds whose formation more than compensates for the loss of H-bonds with water molecules occurring upon folding (*49*).

The reasoning fits with the observation that the folded structure of Sso7d is very similar to (a) that of eukaryotic SH3 domains, which share a significant stability against temperature (*14*, *31*, *50*−*52*), and (b) that of cold shock proteins from both mesophilic and thermophilic sources (*53*−*55*), after circular permutation (*56*). It seems that such folding topology "has survived in all kingdoms due to its (thermal) stability and because it forms a suitably small and stable

platform for different functions in various organisms" (*3*). This scenario is supported by our data showing that, around neutral pH, the truncated variant L54Δ has a folded structure with a T_d of 53 °C and also possesses DNA binding activity.

ACKNOWLEDGMENT

We thank Dr. Angelo Riccio (Università di Napoli Parthenope, Napoli, Italy) for assistance in writing the programs for the best fit of binding curves.

REFERENCES

1. Choli, T., Henning, P., Wittmann-Lebold, B., and Reinhardt, R. (1988) *Biochim. Biophys. Acta 950*, 193−203.
2. De Rosa, M., Gambacorta, A., Nicolaus, B., Giardina, P., Poerio, E., and Buonocore, V. (1984) *Biochem. J. 224*, 407−414.
3. Baumann, H., Knapp, S., Lundback, T., Ladenstein, R., and Hard, T. (1994) *Nat. Struct. Biol. 1*, 808−809.
4. Baumann, H., Knapp, S., Karshikoff, A., Ladenstein, R., and Hard, T. (1995) *J. Mol. Biol. 247*, 840−846.
5. Lundback, T., and Hard, T. (1996) *J. Phys. Chem. 100*, 17690−17695.
6. Lundback, T., Hansson, H., Knapp, S., Ladenstein, R., and Hard, T. (1998) *J. Mol. Biol. 276*, 775−786.
7. Guagliardi, A., Napoli, A., Rossi, M., and Ciaramella, M. (1997) *J. Mol. Biol. 267*, 841−848.
8. Lopez-Garcia, P., Knapp, S., Ladenstein, R., and Forterre, P. (1998) *Nucleic Acids Res. 26*, 2322−2328.
9. Agback, P., Baumann, H., Knapp, S., Ladenstein, R., and Hard, T. (1998) *Nat. Struct. Biol. 5*, 579−584.
10. Gao, Y. G., Su, S. Y., Robinson, H., Padmanabhan, S., Lim, L., McCrary, B. S., Edmondson, S. P., Shriver, J. W., and Wang, A. H. J. (1998) *Nat. Struct. Biol. 5*, 782−786.
11. Fusi, P., Tedeschi, G., Aliverti, A., Ronchi, S., Tortora, P., and Guerritore, A. (1993) *Eur. J. Biochem. 211*, 305−310.
12. Shehi, E., Serina, S., Fumagalli, G., Vanoni, M., Consonni, R., Zetta, L., Dehò, G., Tortora, P., and Fusi, P. (2001) *FEBS Lett. 497*, 131−136.
13. Guagliardi, A., Cerchia, L., Moracci, M., and Rossi, M. (2000) *J. Biol. Chem. 275*, 31813−31818.
14. Knapp, S., Karshikoff, A., Berndt, K. D., Christova, P., Atanasov, B., and Ladenstein, R. (1996) *J. Mol. Biol. 264*, 1132−1144.
15. Fusi, P., Goossens, K., Consonni, R., Grisa, M., Puricelli, P., Vecchio, G., Vanoni, M., Zetta, L., Heremans, K., and Tortora, P. (1997) *Proteins 29*, 381−390.
16. Mombelli, E., Afshar, M., Fusi, P., Mariani, M., Tortora, P., Connelly, J. P., and Lange, R. (1997) *Biochemistry 36*, 8733−8742.
17. Catanzano, F., Graziano, G., Fusi, P., Tortora, P., and Barone, G. (1998) *Biochemistry 37*, 10493−10498.
18. Consonni, R., Santomo, L., Fusi, P., Tortora, P., and Zetta, L. (1999) *Biochemistry 38*, 12709−12717.
19. Fusi, P., Grisa, M., Mombelli, E., Consonni, R., Tortora, P., and Vanoni, M. (1995) *Gene 154*, 97−102.
20. Gill, S. C., and von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
21. Eigner, J., and Doty, P. (1965) *J. Mol. Biol. 12*, 549−580.
22. Izatt, R. M., and Christensen, J. J. (1976) in *The CRC Handbook of Biochemistry and Molecular Biology, Physical and Chemical Data* (Fasman, G. D., Ed.) 3rd ed., Vol. 1, pp 151−360, CRC Press, Boca Raton, FL.
23. Venyaminov, S. Y., and Yang, J. T. (1996) in *Circular Dichroism and the Conformational Analysis of Biomolecules* (Fasman, G. D., Ed.) pp 69−107, Plenum Press, New York.
24. Del Vecchio, P., Graziano, G., Granata, V., Barone, G., Mandrich, L., Manco, G., and Rossi, M. (2002) *Biochemistry 41*, 1364−1371.
25. McAfee, J. G., Edmondson, S. P., Zegar, I., and Shriver, J. W. (1996) *Biochemistry 35*, 4034−4045.
26. Bujalowski, W., and Lohman, T. M. (1987) *Biochemistry 26*, 3099−3106.
27. McGhee, J. D., and van Hippel, P. H. (1974) *J. Mol. Biol. 86*, 469−480.
28. Sreerama, N., and Woody, R. W. (1993) *Anal. Biochem. 209*, 32−44.
29. Becktel, W. J., and Schellman, J. A. (1987) *Biopolymers 26*, 1862−1877.
30. Lumry, R., Biltonen, R. L., and Brandts, J. F. (1966) *Biopolymers 4*, 917−944.
31. McCrary, B. S., Edmondson, S. P., and Shriver, J. W. (1996) *J. Mol. Biol. 264*, 784−805.
32. Liquori, A. M. (1969) *Q. Rev. Biophys. 2*, 65−92.
33. Cooper, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 2740−2741.
34. Robertson, A. D., and Murphy, K. P. (1997) *Chem. Rev. 97*, 1251−1267.
35. Garcia-Mira, M. M., Sadqi, M., Fischer, N., Sanchez-Ruiz, J. M., and Muñoz, V. (2002) *Science 298*, 2191−2195.
36. Robinson, H., Gao, Y. G., McCrary, B. S., Edmondson, S. P., Shriver, J. W., and Wang, A. H. J. (1998) *Nature 392*, 202−205.
37. Krueger, J. K., McCrary, B. S., Wang, A. H. J., Shriver, J. W., Trewella, J., and Edmondson, S. P. (1999) *Biochemistry 38*, 10247−10255.
38. Consonni, R., Arosio, I., Belloni, B., Focolari, F., Fusi, P., Shehi, E., and Zetta, L. (2003) *Biochemistry 42*, 1421−1429.
39. Privalov, P. L. (1989) *Annu. Rev. Biophys. Biophys. Chem. 18*, 47−69.
40. Karplus, M., and Shakhnovich, E. I. (1992) in *Protein Folding* (Creighton, T. E., Ed.) pp 127−195, Freeman & Company, New York.
41. Makhatadze, G. I., and Privalov, P. L. (1995) *Adv. Protein Chem. 47*, 307−425.
42. Dill, K. A. (1985) *Biochemistry 24*, 1501−1509.
43. Spassov, V. Z., Karshikoff, A. D., and Ladenstein, R. (1995) *Protein Sci. 4*, 1516−1527.
44. McKnight, C. J., Doering, D. S., Matsudaira, P. T., and Kim, P. S. (1996) *J. Mol. Biol. 260*, 126−134.
45. Spector, S., Kuhlman, B., Fairman, R., Wong, E., Boice, J. A., and Raleigh, D. P. (1998) *J. Mol. Biol. 276*, 479−489.
46. Spector, S., Young, P., and Raleigh, D. P. (1999) *Biochemistry 38*, 4128−4136.
47. Horng, J. C., Moroz, V., and Raleigh, D. P. (2003) *J. Mol. Biol. 326*, 1261−1270.
48. Finkelstein, A. V., and Ptitsyn, O. B. (1987) *Prog. Biophys. Mol. Biol. 50*, 171−190.
49. Pace, C. N. (2001) *Biochemistry 40*, 310−313.
50. Lim, W. A., Fox, R. O., and Richards, F. M. (1994) *Protein Sci. 3*, 1261−1266.
51. Knapp, S., Mattson, P. T., Christova, P., Berndt, K. D., Karshikoff, A., Vihinen, M., Smith, C. I. E., and Ladenstein, R. (1998) *Proteins: Struct., Funct., Genet. 31*, 309−319.
52. Filimonov, V. V., Azuaga, A. I., Viguera, A. R., Serrano, L., and Mateo, P. L. (1999) *Biophys. Chem. 77*, 195−208.
53. Schindelin, H., Jiang, W., Inouye, M., and Heinemann, U. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 5119−5123.
54. Wassenberg, D., Welker, C., and Jaenicke, R. (1999) *J. Mol. Biol. 289*, 187−193.
55. Perl, D., Mueller, U., Heinemann, U., and Schmid, F. X. (2000) *Nat. Struct. Biol. 7*, 380−383.
56. Jung, J., and Lee, B. (2001) *Protein Sci. 10*, 1881−1886.

BI034520T