Folding Dynamics of the src SH3 Domain[†]

Viara P. Grantcharova and David Baker*

Department of Biochemistry, University of Washington, Seattle, Washington 98195

Received July 22, 1997; Revised Manuscript Received September 10, 1997[⊗]

ABSTRACT: The thermodynamics and kinetics of folding of the chicken src SH3 domain were characterized using equilibrium and stopped-flow fluorescence, circular dichroism (CD), and nuclear magnetic resonance (NMR) hydrogen exchange experiments. As found for other SH3 domains, guanidinium chloride (GdmCl) denaturation melts followed by both fluorescence and circular dichroism were nearly superimposable, indicating the concerted formation of secondary and tertiary structure. Kinetic studies confirmed the two-state character of the folding reaction. Except for a very slow refolding phase due to proline isomerization, both folding and unfolding traces fit well to single exponentials over a wide range of GdmCl concentrations, and no burst phase in amplitude was observed during the dead time of the stopped-flow instrument. The entropy, enthalpy, and heat capacity changes upon unfolding were determined by global fitting of temperature melts at varying GdmCl concentrations (0.4−3.7 M). Estimates of the free energy of unfolding, $\Delta G_U{}^{H_2O}$, from guanidine denaturation, thermal denaturation, and kinetic experiments were in good agreement. To complement these data on the global characteristics of src SH3 folding, individual hydrogen−deuterium (HD) exchange rates were measured for approximately half of the backbone amides in 0 and 0.7 M GdmCl. The calculated free energies of the opening reaction leading to exchange (ΔG_{HD}) indicated that unfolding is highly cooperative—slowly exchanging protons were distributed throughout the core of the protein. The slowly exchanging protons exhibited ΔG_{HD} values higher than the global $\Delta G_U{}^{H_2O}$ by ∼1 kcal/mol, suggesting that the denatured state might be somewhat compact under native conditions. Comparison of the src SH3 with homologous SH3 domains as well as with other small well-characterized β-sheet proteins provides insights into the determinants of folding kinetics and protein stability.

SH3 domains have emerged in the last decade as intergal parts of many signal transduction and cytoskeletal proteins and have been shown to mediate a myriad of protein−protein interactions (*1*). Beside their biological importance as protein adapters, their ability to fold independently, modest size, and lack of disulfide bonds and bound cofactors make the SH3 domains an attractive model system for understanding the principles of protein folding at their simplest level. The abundance of both crystal and solution structures for many members of the SH3 family (*2−10*) further facilitates detailed investigations of their folding. Except for some variability in the loop regions, all structures display the distinctive SH3 fold (*11*): two 3-stranded β-sheets packed orthogonally against each other to form a single hydrophobic core. Several SH3 domains have been the subject of detailed thermodynamic and kinetic studies to complement this structural information. Experiments on the spectrin SH3 domain (*12*) point to a moderately stable protein with simple two-state folding kinetics. A search for putative folding initiation sites failed to detect any peptide hairpins with nativelike structure in solution, suggesting the importance of tertiary interactions in stabilizing SH3 domains (*13*). Studies on the *Drosophila* drkN SH3 domain took advantage of the low stability of this domain to examine directly the unfolded state (*14*). The existence of a dynamic equilibrium

between folded and unfolded protein in H_2O enabled the application of innovative NMR[1] experiments for the detection of residual structure in the unfolded state under native conditions and in the presence of GdmCl (*15*, *16*). Other SH3 domains studied include Sem5 SH3 (*17*), PI3K SH3 (J. I. Guijarro, personal communication), and human fyn SH3 (K. W. Plaxco, personal communication).

The present study was motivated by two objectives. First, our laboratory has identified src SH3 mutants with highly simplified sequences that still fold properly (*18*). The wild-type SH3 kinetic and thermodynamic parameters reported here provide an important point of reference for the biophysical analysis of such mutants. Second, a systematic comparison between homologous SH3 domains should make apparent the spectrum of stabilities and folding rates available to this particular fold and reveal how changes in the sequence affect the parameters of folding. Such information can later be extended to comparisons with other, SH3-like proteins, which are functionally unrelated to the original members of the family: Sso7d (*19*), PsaE (*20*), and HIV-integrase DNA binding domain (*21*). Our results on the src SH3 domain confirm the two-state folding model for small proteins and for the first time provide residue-specific information on SH3 domain stability. Significant differences in the folding parameters of src SH3 and the other homologous proteins point to the conclusion that the determinants of the protein fold are distinct from the factors influencing stability and

---

[1] Abbreviations: CD, circular dichroism; ESI, electrospray ionization; HD, hydrogen−deuterium; HSQC, heteronuclear single quantum coherence; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; GdmCl, guanidinium chloride.

folding kinetics. The src SH3 data presented here further illuminates the folding of all β-sheet proteins (22−27).

EXPERIMENTAL PROCEDURES

Chemicals. $^{15}NH_4Cl$ for growing labeled src SH3 and D_2O (99.9% isotope pure) were purchased from Isotec, Inc. GdmCl used in all denaturation experiments was ultrapure grade from U.S. Biochemical Corp. Deuterated GdmCl was prepared by repeated lyophilization from D_2O. The precise concentration of GdmCl solutions was determined using refractometry (28). Solutions for spectroscopic measurements were made with double-distilled deionized water.

Protein Expression and Purification. The chicken *src* SH3 gene (for sequence see Figure 5) was cloned in plasmid pET_15b (Novagen) and expressed in the BL21(DE3)pLysS strain of *Escherichia coli*. A histidine tag at the N-terminus of the protein contained a thrombin cleavage site (MSG-SSHHHHHSSGLVPR↓GSHM) to facilitate its removal after purification. After a single-step purification on a Zn^{2+} affinity column (Pharmacia Biotech), the histidine tag was removed by an overnight incubation with biotinylated thrombin (Novagen). Thrombin was then sequestered with streptavidin−agarose (Novagen) and the preparation was run over the Zn^{2+} column again to retain the cleaved histidine tag. The final yield of protein from 1 L of cells grown in LB medium was ∼30 mg. The purity of the prep was verified by ESI mass spectroscopy to be >98%. Thereafter, the protein was stored in 50 mM sodium phosphate, pH 6, conditions used for all subsequent experiments. Protein concentration was determined by UV absorbance at 280 nm using an extinction coefficient of 16 500 M^{-1} cm^{-1} (29).

^{15}N-labeled SH3 for NMR spectroscopy was grown in M9 minimal medium with ^{15}N-labeled NH_4Cl as the only nitrogen source. Purification and His tag removal was carried out as described above for the unlabeled protein.

Equilibrium Fluorescence and CD Measurements. Equilibrium fluorescence data was collected on a Spex Fluorolog2 spectrofluorometer. Excitation was at 280 nm and emission was recorded at 225 nm. src SH3 (3 μM) in 50 mM sodium phosphate, pH 6, was prepared at varying concentrations of GdmCl. The same experiment was performed in deuterated buffer to measure isotope effects on protein stability. CD data was collected on an Aviv 62A DS spectrometer. SH3 (50 μM) in 50 mM sodium phosphate, pH 6, was mixed in a Microlab titrator (Hamilton) with varying amounts of SH3 (50 μM) in 7M GdmCl, pH 6, to yield the desired final GdmCl concentrations. Changes in ellipticity were followed at 235 nm. Temperature was maintained at 295 K with a Peltier device. To obtain a value for $\Delta G_U^{H_2O}$, the denaturation curves were fitted by nonlinear least-squares analysis using the linear extrapolation model as applied by Santoro and Bolen (30).

Equilibrium Temperature−GdmCl Denaturation. Temperature melts in deuterated buffer were conducted at 16 different GdmCl concentrations between 0.4 and 3.7 M. Changes in ellipticity were followed on an Aviv 62A CD spectrometer at 235 nm. The instrument was placed in kinetic mode, and data were collected every degree for 30 s. Temperature was equilibrated for 1 min between measurements. The temperature range was from 273 to 371 K for low GdmCl concentrations and from 268 to 378 K for the higher concentrations. All denaturation curves were fitted

simultaneously to eq 1 (31) to generate a three-dimensional guanidine−temperature denaturation surface:

$$\text{signal} = N_0 + C_{nt}T + C_{ng}D + \frac{(D_0 + D_{nt}T + D_{ng}D)e^{-\Delta G/RT}}{1 + e^{-\Delta G/RT}} \quad (1)$$

where

$$\Delta G_U = \Delta H_g + mD - T\Delta S_g \Delta C_p[T_g - T + T \ln (T/T_g)] \quad (2)$$

The signal of the native protein at a particular temperature and denaturant concentration is given by $(N_0 + C_{nt}T + C_{ng}D)$, and the signal of the denatured protein by $(D_0 + D_{nt}T + D_{ng}D)$. ΔH_g and ΔS_g are the enthalpy and entropy changes upon unfolding at T_g. ΔC_p, the change in heat capacity, and m, the dependence of ΔG_U on [GdmCl], are both assumed to be independent of temperature and [GdmCl]. D is the GdmCl concentration in (molar), T_g is the reference temperature (295 K), and T is the absolute temperature. A second fit to the data included individual dependences of ΔH, ΔS, and ΔC_p on GdmCl concentration (26). The program S-PLUS (Statistical Science, Inc.) was used for the fitting and error analysis. With the estimates of m, ΔH, ΔS, and ΔC_p, the free energy of unfolding, ΔG_U, can be calculated at any temperature and GdmCl concentration (eq 2).

Kinetic Measurements. Kinetics of folding and unfolding were followed by stopped-flow fluorescence as described previously (32). The final src SH3 concentration was 5 μM. Folding and unfolding were initiated by 10-fold dilutions of the protein from 4 and 0 M GdmCl, respectively, to the appropriate final GdmCl concentrations. Signal sampling was every 0.5 ms with a calibrated instrument dead time of 1.7 ms. The temperature was held constant at 295 K with a circulating water bath. At each GdmCl concentration, 5−7 experiments were performed sequentially and the kinetic traces were averaged. The exact GdmCl concentration of the multiple shots was measured by refractive index. Folding and unfolding rates were determined from single-exponential fits to the kinetic traces using Biologic software. The rates in 0 M GdmCl and the kinetic m values were obtained by fitting kinetic data to:

$$\ln k_{obs} = \ln [k_f^{H_2O} \exp(-m_f[\text{GdmCl}]/RT) + k_u^{H_2O} \exp(m_u[\text{GdmCl}]/RT)] \quad (3)$$

A slow phase ($k = 0.017 \pm 0.002$ s^{-1}) due to proline isomerization was followed by sampling every 100 ms for 400 s. Because of the large difference in time scale, the rates of the two phases could be measured independently.

NMR Spectroscopy. HD exchange experiments were done with 1.7 mM ^{15}N-labeled src SH3, in 50 mM potassium phosphate, pH 6, at two GdmCl concentrations. The 0 M GdmCl experiment was performed on a 750 MHz Bruker DMX spectrometer, while the 0.7 M GdmCl experiment was performed on a 500 MHz Bruker DMX instrument. In both cases, HD exchange was initiated by rapidly dissolving protein lyophilized from NH_4HCO_3 in deuterated buffer containing either 0 or 0.7 M GdmCl. Multiple sensitivity-enhanced HSQC spectra were collected at 295 K as described
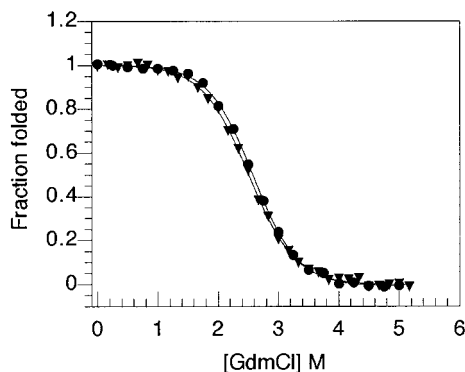
FIGURE 1: Equilibrium GdmCl denaturation followed by fluorescence (●) and CD (▼) at 295 K. Fluorescence was recorded at 225 nm and CD at 235 nm. Solid lines are the fits to the experimental data using the linear extrapolation model. Baselines were corrected for sloping (*30*) before conversion to fraction folded.

by Yi and Baker (*33*) to follow the exchange. The exact pH was measured after the experiment, and the correction pD = pH + 0.4 was applied. Processing of the HSQC spectra as well as peak volume intergration was performed using the program FELIX 2.30 (Biosym Technologies, San Diego, CA). In the analysis of the spectra, we made use of the previously published src SH3 assignments (*34*). As our construct differed by five residues at the N-terminus and one at the C-terminus, ambiguous peaks were assigned by a NOESY-HSQC experiment (*35*) with a mixing time $\tau = 90$ ms. Observed exchange rates (k_{obs}) for each amide were determined by fitting a single exponential to the data with the program Kaleidagraph (Abeldeck). Random-coil exchange rates (k_{rc}) were calculated by the method of Bai et al. (*36*), which corrects for the steric and inductive effects of neighboring side chains. We used low-salt exchange rate constants for the 0 M GdmCl experiment and high-salt rate constants for the 0.7 M GdmCl experiment to account for the effect of ionic strength on exchange rates.

RESULTS

Equilibrium Denaturation. Guanidine denaturation of the src SH3 domain was followed by both CD and fluorescence to monitor changes in secondary and tertiary structure, respectively. src SH3 has two partially buried tryptophans whose fluorescence decreases significantly upon unfolding (*18*). Taking into account sloping of the folded and unfolded baselines, the denaturation curves were fitted using the linear extrapolation model (*30*) to estimate the free energy of unfolding in the absence of GdmCl, $\Delta G_U^{H_2O}$. The unfolding transitions detected by the two probes were nearly coincident (Figure 1), suggesting that disruption of secondary and tertiary structure occurred concurrently. The estimates of the free energies of unfolding and the *m* values obtained using the two methods were within experimental error (Table 1). Chemical denaturation was also performed in deuterated buffer to assess isotope effects on stability and for comparison of $\Delta G_U^{D_2O}$ with the ΔG_{HD} values determined for slowly exchanging amide protons (see below). As observed for cytochrome *c* (*37*), the src SH3 domain is ~0.5 kcal/mol more stable in D_2O than in H_2O (Table 1).

To obtain thermodynamic parameters for the unfolding of the src SH3 domain, we performed temperature denaturations followed by CD at 16 different GdmCl concentrations and over a wide temperature range (Figure 2). This method

Table 1: Summary of Kinetic and Thermodynamic Parameters in H_2O and D_2O at 295 K[a]

	in H_2O		in D_2O	
	GdmCl denaturation	kinetic experiments	GdmCl denaturation	temperature denaturations[b]
$\Delta G_U^{H_2O}$ (kcal/mol)	$4.1^c \pm 0.1$ $3.8^d + 0.1$	$3.7^d \pm 0.06$	$4.7^c \pm 0.17$	4.7 ± 0.22
m [kcal/(mol·M)]	$1.6^c \pm 0.03$ $1.5^e \pm 0.03$	$1.4^f \pm 0.03$	$1.7^c \pm 0.06$	1.7 ± 0.07
ΔH (kcal/mol)				6.9 ± 0.1^g
ΔS [cal/(K·mol)]				7.7 ± 0.4^g
ΔC_p [kcal/(K·mol)]				0.70 ± 0.02

[a] All experiments were done in 50 mM phosphate, pH (pD) 6. Standard errors are those estimated from the least-squares fit. [b] Data were obtained from the global fitting of the temperature−GdmCl denaturation surface. [c] Determined by fluorescence. [d] Value obtained from the extrapolated rates of folding ($k_f = 56.7 \pm 3.0$ s^{-1}) and unfolding ($k_u = 0.1 \pm 0.01$ s^{-1}) in 0 M GdmCl. [e] Determined by CD. [f] $m = m_f + m_u$; $m_f = 0.99 \pm 0.02$ and $m_u = 0.45 \pm 0.01$. [g] Because of the high correlation between the estimates of ΔH and ΔS, the actual uncertainties in these parameters are likely to be considerably larger than the errors obtained from the least-squares fitting given in the table.
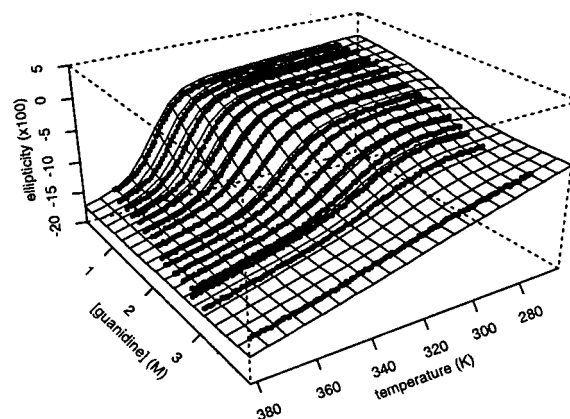

FIGURE 2: Three-dimensional guanidine−temperature denaturation surface. Closed circles represent experimentally determined values from the individual temperature melts at different GdmCl concentrations. Continuous net is the best fit of eq 1 to the experimental data.

reduces the uncertainties in estimating quantities in 0 M denaturant since the extrapolations are based on more data compared to a GdmCl denaturation experiment at a single temperature. The temperature−GdmCl denaturation surface was fit using two models in which the denaturant dependence of the thermodynamic parameters was described either with (1) a single overall *m* value (eq 1) or with (2) explicit denaturant dependences of ΔH, ΔS, and ΔC_p. ΔC_p was assumed to be temperature-independent in both cases (*38*). The ΔG_U estimates from the two models were similar (4.7 $\pm$ 0.22 and 4.44 $\pm$ 0.2 kcal/mol), however, the ΔH and ΔS values differed considerably. Given the high correlation between estimates of ΔH and ΔS (0.998) in both models, independent treatment of their GdmCl dependences, as in model 2, is probably not warranted. We report ΔH, ΔS, ΔC_p, and $\Delta G_U^{H_2O}$ values at 295 K (Table 1) for the first model because of the fewer free parameters involved. The temperature melts were done under conditions identical to the NMR experiments described below to allow meaningful comparison of the free energy estimates. The simultaneous fit of all temperature denaturation curves yielded $\Delta G_U^{D_2O}$ and *m* values at 295 K very close to those obtained from the individual GdmCl denaturation experiments (Table 1). Ap-
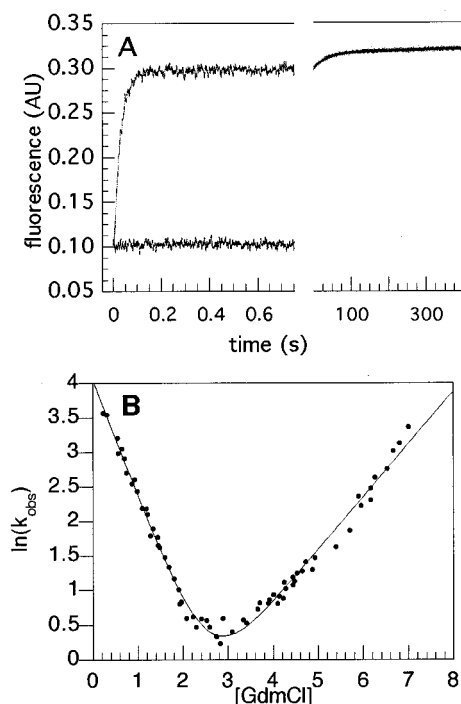
FIGURE 3: (A) Refolding kinetics of src SH3 in 0.4 M GdmCl at 295 K followed by stopped-flow fluorescence. The slow phase due to proline isomerization accounts for ∼10% of the total change in signal. The baseline for the denatured protein in 0.4 M GdmCl is assumed to be the same as that in 4 M GdmCl since equilibrium GdmCl denaturation did not reveal any dependence of the fluorescence signal of the unfolded protein on GdmCl. (B) GdmCl dependence of ln k_{obs} at 295 K. Closed circles are averages of the rates obtained from 5−7 folding or unfolding experiments at a particular GdmCl concentration. Solid line is the best fit of eq 2 to the experimental data.

parently, the assumption of the linear extrapolation model that ΔG_U is linearly dependent on [GdmCl] over a broad concentration range holds true for this protein.

Stopped-Flow Kinetic Measurements. The kinetics of folding were examined at 295 K over a wide range of denaturant concentrations (0.4−7 M GdmCl). Under all conditions both the folding and unfolding traces followed by fluorescence fit well to single exponentials, indicating a lack of intermediates in the folding reaction. There appear to be no early burst-phase events that significantly alter the environment of the tryptophans. The signal immediately after the instrument dead time (1.7 ms) approaches very closely the unfolded baseline; the small difference could be accounted for by extrapolating the trace to the zero time point (Figure 3A). The src SH3 domain contains two *trans*-prolines in its native state, and we detect a slow proline isomerization phase with a rate constant of $k = 0.017 \pm 0.002$ s^{-1} (Figure 3A). The observed contribution of proline isomerization to the total change in amplitude during folding (∼10%) is somewhat smaller than expected from an assumed *cis/trans* equilibrium ratio of 1:4 in solution. Similar results were reported previously for the spectrin SH3 domain (*12*) and may be explained by local sequence effects favoring the *trans*-proline conformation in the unfolded state (K. W. Plaxco, personal communication). Figure 4 shows the combined kinetic data from all experiments. The dependence of ln k_{obs} on [GdmCl] has the typical v-curve shape, with the left arm of the curve dominated by the folding rate and the right arm by the unfolding rate. The data were fit using a simple model in which the logarithms of the folding and
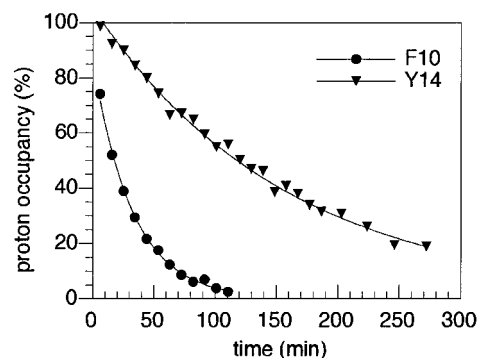

FIGURE 4: Time course of HD exchange for two amide protons in 0.7 M GdmCl at 295 K: F10 (●) is a fast-exchanging proton, while Y14 (▼) is a slowly exchanging one. Proton occupancy was determined by integrating peak intensities of amide protons from HSQC spectra at different time points. The extrapolated intensity of the slowest exchanging proton at time zero was taken as 100% proton occupancy for all protons. The solid lines are the best single-exponential fits to the experimental data.

unfolding rate constants depend linearly on denaturant with slopes of m_f and m_u, respectively (see Experimental Procedures, eq 3). The ratio of the transition and equilibrium m values, m_f/m, is one measure of the position of the transition state along the reaction coordinate. The value of 0.69 for the src SH3 suggests that ∼70% of the buried surface area of the folded protein is excluded from solvent in the transition state.

Equilibrium HD Exchange Experiments. Equilibrium HD exchange of amide protons with solvent is a probe for local structural stability. Under EX2 conditions ($k_f >> k_{rc}$) (*39*), the free energy for the opening reaction that leads to exchange of individual protons can be obtained from the measured rates of exchange (k_{obs}) and an estimate of the random-coil exchange rate (k_{rc}):

$$\Delta G_{HD} = -RT \ln (k_{obs}/k_{rc}) \qquad (4)$$

In a simple model (*37*), both denaturant-independent local fluctuations and denaturant-dependent global unfolding contribute to exchange. With increasing GdmCl concentrations, global unfolding becomes dominant and a convergence of ΔG_{HD} values to the global ΔG_U is expected. We were able to measure HD exchange rates for 33 nonoverlapping amide protons by following their peak intensities in ^{1}H−^{15}N HSQC spectra. Representative time courses of HD exchange are shown for two protons (Figure 4). The observed rates, k_{obs}, were obtained from single-exponential fits to the data. Random-coil exchange rates were calculated as described in the Experimental Procedures section. Even though the EX2 exchange mechanism was not tested directly, the assumption is very likely to be correct for both experimental conditions, on the basis of our determination of k_f in 0 and 0.7 M GdmCl and estimates of the fastest k_{rc} (in 0 M GdmCl, $k_f = 56.7$ s^{-1} and $k_{rc} = 4$ s^{-1}; in 0.7 M GdmCl, $k_f = 14.9$ s^{-1} and $k_{rc} = 1$ s^{-1}). Since the measured k_fs reflect global folding, protons exchanging through local fluctuations should have even higher closing rates. Panels A and B of Figure 5 show the ΔG_{HD} values for the amide protons in 0 and 0.7 M GdmCl, respectively, calculated from eq 4. The secondary structure elements in which the amides participate are indicated. In Figure 5A the cutoff line at 4.67 kcal/mol is the $\Delta G_U^{D_2O}$ value from equilibrium GdmCl denaturation and represents the expected ΔG_{HD} for protons
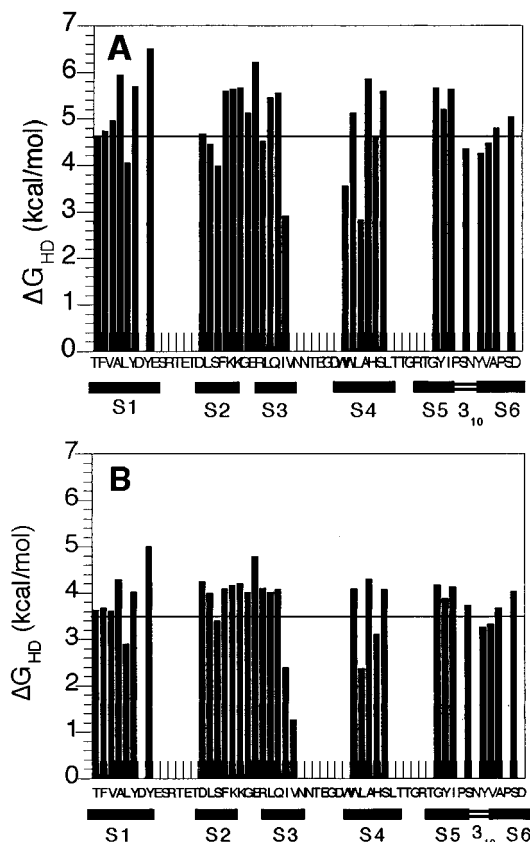
FIGURE 5: ΔG_{HD} values at 295 K for amide protons in 0 M GdmCl (A) and 0.7 M GdmCl (B). Solid lines at 4.67 kcal/mol (A) and 3.47 kcal/mol (B) indicate the free energy of unfolding in D_2O, $\Delta G_U^{D_2O}$, as determined from the temperature–GdmCl denaturations (Figure 2). Residues are shown within the context of the secondary structure in which they participate according to the NMR structure solved by Yu et al. (*34*): S, strand; 3_{10}, 3_{10} helix.

FIGURE 6: Summary of HD exchange data for the src SH3 domain. Residues are colored according to local stability: blue, slow exchange ($\Delta G_{HD} \geq 5$ kcal/mol); yellow, intermediate exchange (4 kcal/mol $< \Delta G_{HD} < 5$ kcal/mol); red, fast exchange or not measurable ($\Delta G_{HD} \leq 4$ kcal/mol). The NMR solution structure (*34*) is shown.

Thermodynamic Parameters for src SH3. The equilibrium thermodynamic parameters ΔH, ΔS, and ΔC_p determine the stability of proteins at various temperatures. The values obtained for the src SH3 domain from global fitting of the temperature melts (Table 1) are the first reported for an SH3 domain near neutral pH. After correction for the size of the protein, the enthalpy and entropy measurements for src SH3 [$\Delta H = 1.2$ cal/g; $\Delta S = 0.002$ cal/(K·g)] fall within the ranges [for ΔH, −0.5 to 5 cal/g; for ΔS, −0.005 to 0.013 cal/(K·g)] previously determined for other single-domain proteins (*41–44*). The only other SH3 domain for which the corresponding parameters have been determined is spectrin. Although these studies were done at low pH (2–4), the ΔC_p values agree quite well. The ΔC_p values for a wide variety of proteins have been shown to correlate with changes in buried surface area and polar and nonpolar solvent-accessible surface area upon unfolding (*38*). For the src SH3 domain the calculated ΔC_p is higher than the experimentally determined value: 1.14 (calculated) vs 0.69 (experimental) kcal/(mol·K). Since the theoretical model assumes complete unfolding and hydration of the denatured state, the discrepancy between the two values might point to a slightly compact unfolded state.

Backbone Dynamics of src SH3. We have examined the equilibrium backbone dynamics of the src SH3 by monitoring the time course of HD exchange by 2D NMR. The experiment was prompted by the suggestion (*34*) that the protein is unusually flexible since only three slowly exchanging protons were observed. We chose to determine amide proton ΔG_{HD} values at only two GdmCl concentrations because at 0 M GdmCl the slowly exchanging protons were interspersed throughout the structure of the protein and it did not seem that a more extensive characterization of denaturant dependence, as done for cytochrome *c* and RNaseH (*45, 46*), would be revealing in this case. When evaluating the HD exchange results, we considered the value of ΔG_{HD} in 0 M GdmCl to be a measure of local stability and the dependence of ΔG_{HD} on [GdmCl], to be an estimate of the contribution of local fluctuations toward exchange. High ΔG_{HD} values in 0 M GdmCl correlate very well with a large dependence of ΔG_{HD} on [GdmCl], as expected,

if they are exchanging through global unfolding only. The corresponding value in 0.7 M GdmCl is 3.49 kcal/mol (Figure 5B). Comparison of Figure 5 panels A and B shows the convergence of free energies in 0.7 M GdmCl as predicted by the simple model for denaturant dependence of HD exchange. Also apparent from the plots is that a number of protons have ΔG_{HD} values higher than the equilibrium value. Allowing for ~0.5 kcal/mol variations in ΔG_{HD} due to uncertainties in the k_{rc} estimates (*40*), 13 protons in 0 M GdmCl and 5 protons in 0.7 M GdmCl rise significantly above the global unfolding threshold. The structural context of these outliers (Figure 6) as well as the implications of this discrepancy in free energies for the denatured state of src SH3 are discussed below.

DISCUSSION

Two-State Folding Process. Both equilibrium and kinetic data on the folding of the src SH3 are well fit by a two-state model. Several tests diagnostic for the absence of stable intermediates provide convincing evidence: (1) nearly superimposable equilibrium denaturation curves followed by fluorescence and CD; (2) single-exponential folding and unfolding kinetics; and (3) agreement between kinetically and thermodynamically determined ΔG_U and *m* values. This characterization of the src SH3 domain adds it to the list of single-domain proteins that fold without detectable populations of partially folded intermediates.

because protons in stable regions exchange primarily through global unfolding. The following residue-specific map of src SH3 stability and dynamics emerged. The slowest exchanging protons are located in two regions of the protein—at the base of the so-called RT loop (the second half of strand 1 and the diverging turn between strands 2 and 3) and in strands 4 and 5, which form a hairpin (Figure 6). Since these amides also have the largest dependence of ΔG_{HD} on [GdmCl], exchange from these elements probably requires disruption of the hydrophobic core. Inspection of the src SH3 structure suggests that such a motion would be equivalent to opening of the β-sandwich (global unfolding). Strands 1 and 6, on the other hand, occur at the N- and C-termini, respectively, and judging from their intermediate ΔG_{HD} values and moderate dependence of ΔG_{HD} on [GdmCl], they may be free to dissociate from the rest of the structure without substantial exposure of hydrophobic surface. The same is true also for the second half of strand 3, which bridges the two β-sheets and is slightly distorted. The third set of amide protons occur at the ends of strands or in loops. They exhibit the lowest ΔG_{HD} values and the lowest dependence of ΔG_{HD} on [GdmCl], indicative of exchange mainly through local unfolding events. The three sets of protons do not occur contiguously along structural elements (Figure 6), as is expected for β-sheet proteins where a single strand can contain amides with alternating high and low protection depending on the solvent accessibility and the strength of hydrogen bonding on either side of the strand.

The HD exchange data suggest that, unlike cytochrome *c* and RNase H, src SH3 does not undergo large concerted motions of subdomains. Instead, local unfolding events are limited to loops and the terminal strands. This qualitative difference probably results from the larger size ($\sim$100 aa) of cyt *c* and RNase H, which makes the stabilization of partial structures possible. In contrast, src SH3 is a small β-sheet protein for which independent modules are most likely unstable when outside the context of the rest of the protein. Protein L, another small protein (60 aa) studied in our lab (*47*), and chymotrypsin inhibitor 2 (*48*) also display similar fully cooperative unfolding transitions.

Discrepancy between ΔG_{HD} and ΔG_U. A number of the src SH3 amide protons have ΔG_{HD} values significantly above the ΔG_U for global unfolding. Since ΔG_{HD} is obtained directly in 0 M GdmCl, while $\Delta G_U^{H_2O}$ is inferred from measurements in the transition region, such discrepancy has often been attributed to a breakdown of the linear extrapolation model at low GdmCl concentrations (*47, 49–51*). An upward curvature in the dependence of ΔG_U on GdmCl would lead to an underestimation of $\Delta G_U^{H_2O}$ upon linear extrapolation. However, the close match between the values for $\Delta G_U^{H_2O}$ obtained from two different methods (GdmCl denaturation and global fitting of the temperature melts) gives us good confidence in its accuracy. Bai et al. (*37*) suggest that the presence of prolines results in larger free energies being measured by HD exchange than by equilibrium denaturation because proline isomerization does not occur on the time scale of k_{rc}. Considering the small contribution proline isomerization makes to the overall folding reaction of the src SH3 domain ($\sim$10%), the correction would be very small ($\sim$0.06 kcal/mol). A more likely explanation is that the random-coil exchange rates based on model peptides (*36*) do not represent accurately exchange from the unfolded state. A 3–4-fold overestimation of k_{rc} would account for the $\sim$1

kcal/mol difference in free energies in 0 M GdmCl. Residual hydrogen bonds, steric exclusion, or electrostatic repulsion of the OD^- ions that catalyze exchange (*52, 53*) could all potentially slow down exchange from the denatured state. In 0.7 M GdmCl, the discrepancy between ΔG_{HD} and ΔG_U is reduced (the average difference between ΔG_{HD} and ΔG_U is $\sim$0.7 kcal/mol). This result is consistent with the observation by Zhang and Forman-Kay (*15*) that under denaturing conditions, residual structure in the unfolded state is destabilized.

There are precedents for residual structure in the denatured state. Most relevant for our case are the detailed NMR studies on the N-terminal SH3 domain of *Drosophila* drk (*15, 16*), which revealed that backbone parameters (^{15}N chemical shifts and *J* coupling constants) for the unfolded state deviate significantly from random coil values. Medium-range NOEs suggest that a number of regions sample turnlike conformations, and the overall dynamics of the backbone pointed to a compact unfolded state. In light of these investigations, the protection from exchange in the unfolded state of the src SH3 could result from either global compactness of the chain or the persistence of specific interactions (either native or nonnative). As mentioned above, one region of the src SH3 with unusually high ΔG_{HD} values is the diverging turn between strands 2 and 3 and the first part of strand 1. Interestingly, the corresponding region in the drkN SH3 domain exhibits line broadening, which is indicative of slow interconversion between different structures. This fragment was also identified in a library of sites with propensity for local interactions (*54*). Peptide studies on the spectrin SH3 domain (*13*) did not find any hairpins that form independently in aqueous solution. These experiments, however, did not include the diverging turn, and it is also conceivable that several of the residual elements are required to stabilize each other. We are currently investigating the role of this turn in the folding process by mutagenesis and peptide studies.

Comparison with Other β-Sheet Proteins. The characterization of src SH3 enables the comparison of its kinetic and thermodynamic properties with those of other SH3 domains. The stability and equilibrium *m* value for spectrin (30% sequence identity to src SH3) at neutral pH (*55*) closely match those for src SH3 under identical conditions, suggesting that the two-state folding of both proteins is accompanied by similar burial of surface area and formation of favorable interactions. At comparable free energies of unfolding, however, both the rates of folding and unfolding are more than 10-fold slower for spectrin: k_f is 2.83 s^{-1} vs 56.7 s^{-1} for the src SH3, and k_u is 0.007 s^{-1} vs 0.1 s^{-1}. The same trend is observed when the two proteins were characterized at pH 3.5 to examine the dependence of SH3 domain stability on pH. Interestingly, for both proteins the rate of unfolding is affected much more strongly by pH than the rate of folding, pointing to the importance of electrostatic interactions in stabilizing the native fold, but not so much the transition state. The fyn SH3 (69% sequence identity to src), on the other hand, is more stable than src SH3 at neutral pH and has a slightly higher folding rate (94 s^{-1}) and a significantly lower unfolding rate (0.001 s^{-1}) (K. W. Plaxco, personal communication). As it is the unfolding rate that differs by 100-fold, the fyn SH3 must be making more favorable interactions in its folded state; the denatured and transition states likely remain the same as judged by the similar

equilibrium *m* value and m_f/m ratio (0.69 and 0.68 for src and fyn SH3, respectively). On the other side of the spectrum is drk SH3, which is only marginally stable (∼0 kcal/mol) in water and has an estimated folding rate of 0.9 s^{-1} (*16*). The SH3 fold can thus accomodate a variety of stabilities and kinetic rates of folding. The variations in kinetic and thermodynamic parameters may be related to the distinct functional requirements on the members of this diverse class of proteins.

A survey of the folding kinetics of β-sheet proteins studied so far reveals a broad range of rate constants. Large multidomain β-sheet proteins such as interleukin 1β (*24*) fold relatively slowly, while small single-domain β-sheet proteins, like the src SH3, fold on time scales similar to those of small helical proteins (*32*). CspB (67 aa), whose folding kinetics and thermodynamics have been extensively characterized (*26*), is an extreme example of a rapidly folding β-sheet protein (k_f = 1070 s^{-1}). The similar folding rates of small β-sheet and α-helical proteins contrast with the different ability of theoreticians to predict their folds: the success rate of secondary structure prediction programs is ∼ 85% for helices versus 60% for strands. This dichotomy can be understood by noting that helices are stabilized primarily by local interactions and strands by nonlocal ones. This fact complicates structure prediction of β-sheets because (1) a limited sequence window around a given residue is unlikely to contain all of its interacting partners and (2) there is considerable variation in strand conformation due to differences in protein context, resulting in weak local sequence signatures. Many *ab initio* tertiary structure prediction methods assemble proteins in a hierarchical fashion (*56*, *57*) and thus are also hampered by the importance of nonlocal interactions in sheet formation. The fact that in nature β-sheet proteins are not at a disadvantage in finding their native state suggests that early restriction of conformational freedom does occur in β-sheet proteins. A better understanding of the stages in β-sheet formation should improve the prediction of the structure of this important class of proteins.

ACKNOWLEDGMENT

SUPPORTING INFORMATION AVAILABLE

One table, showing observed HD exchange rates of src SH3 domain in 0 and 0.7 M GdmCl (1 page). Ordering information is given on any current masthead page.

REFERENCES

1. Pawson, T. (1995) *Science 373*, 573−580.
2. Kuriyan, J., & Cowburn, D. (1993) *Curr. Opin. Struct. Biol. 3*, 828.
3. Borchert, T. V., Mathieu, M., Zeelen, J. P., Courtneidge, S. A., & Wierenga, R. K. (1994) *FEBS Lett. 341*, 79−85.
4. Kohda, D., Terasawa, H., Ichikawa, S., Orura, K., Hatanaka, H., Mandiyan, V., Ullrich, A., Schlessinger, J., & Inagaki, F. (1994) *Structure 2*, 1029−1040.
5. Wittekind, M., Mapelli, C., Farmer, B. T., II, Suen, K.-L., Goldfarb, V., Tsao, J., Lavoie, T., Barbacid, M., Meyers, C. A., & Müller, L. (1994) *Biochemistry 33*, 13531−13539.
6. Wu, X., Knudsen, B., Feller, S. M., Zheng, J., Sali, A., Cowburn, D., Hanafusa, H., & Kuriyan, J. (1995) *Structure 3*, 215−226.
7. Guruprasad, L., Dhanaraj, V., Timm, D., Blundell, T. L., Gout, I., & Waterfield, M. D. (1995) *J. Mol. Biol. 248*, 856−866.
8. Hiroaki, H., Klaus, W., & Senn, H. (1996) *J. Biomol. NMR 8*, 105−122.
9. Xu, W., Harrison, S. C., & Eck, M. J. (1997) *Nature 385*, 595−602.
10. Sicheri, F., Moarefi, I., & Kuriyan, J. (1997) *Nature 385*, 602−605.
11. Efimov, A. V. (1994) *Structure 2*, 999−1002.
12. Viguera, A. R., Martinez, J. C., Filimonov, V. V., Mateo, P. L., & Serrano, L. (1994) *Biochemistry 33*, 2142−2150.
13. Viguera, A. R., Jiménez, M. A., Rico, M., & Serrano, L. (1996) *J. Mol. Biol. 255*, 507−521.
14. Zhang, O., & Forman-Kay, J. D. (1995) *Biochemistry 34*, 6784−6794.
15. Zhang, O., & Forman-Kay, J. D. (1997) *Biochemistry 36*, 3959−3970.
16. Farrow, N. A., Zhang, O., Forman-Kay, J. D., & Kay, L. E. (1997) *Biochemistry 36*, 2390−2402.
17. Lim, W., Fox, R. O., & Richards, F. M. (1994) *Protein Sci. 3*, 1261.
18. Riddle, D. S., Santiago, J., Grantcharova, V. P., & Baker, D. (1997) *Nat. Struct. Biol.* (in press).
19. Bauman, H., Knapp, S., Lundbäck, T., Ladenstein, R., & Härd, T. (1994) *Nat. Struct. Biol. 1*, 808−819.
20. Falzone, C. J., Kao, Y. H., Zhao, J., Bryant, D. A., & Lecomte, J. T. J. (1994) *Biochemistry 33*, 6052−6062.
21. Eijkelenboom, A. P. A. M., Puras Lutzke, R. A., Boelens, R., Plasterk, R. H. A., Kaptein, R., & Hard, K. (1995) *Nat. Struct. Biol. 2*, 807−810.
22. Ropson, I. J., Gordon, J. I., & Frieden, C. (1990) *Biochemistry 29*, 9591−9599.
23. Koide, S., Dyson, H. J., & Wright, P. E. (1993) *Biochemistry 32*, 12299−12310.
24. Varley, P., Gronenborn, A. M., Christensen, H., Wingfield, P. T., Pain, R. H., & Clore, G. M. (1993) *Science 260*, 1110−1113.
25. Liu, Z. P., Rizo, J., & Gierasch, L. (1994) *Biochemistry 33*, 134−142.
26. Schindler, T., & Schmid, F. X. (1996) *Biochemistry 35*, 16833−16842.
27. Plaxco, K. W., Spitzfaden, C., Campbell, I. D., & Dobson, C. M. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 10703−10706.
28. Nozaki, Y. (1972) *Methods Enzymol. 26*, 43−51.
29. Gill, S. C., & von Hippel, P. H. (1989) *Anal. Biochem. 182*, 319−326.
30. Santoro, M. M., & Bolen, D. W. (1988) *Biochemistry 27*, 8063−8068.
31. Huang, G. S., & Oas, T. G. (1996) *Biochemistry 35*, 6173−6180.
32. Scalley, M. L., Yi, Q., Gu, H., McCormack, A., Yates, J. R., III, & Baker, D. (1996) *Biochemistry 36*, 3373−3382.
33. Yi, Q., & Baker, D. (1996) *Protein Sci. 5*, 1060−1066.
34. Yu, H., Rosen, M. K., & Schreiber, S. L. (1993) *FEBS Lett. 324*, 87−92.
35. Zhang, O., Kay, L. E., Olivier, J. P., & Forman-Kay, J. D. (1994) *J. Biomol. NMR 4*, 845−858.
36. Bai, Y., Milne, J. S., Mayne, L., & Englander, S. W. (1993) *Proteins: Struct., Funct., Genet. 17*, 75−86.
37. Bai, Y., Milne, J. S., Mayne, L., & Englander, S. W. (1994) *Proteins: Struct., Funct., Genet. 20*, 4−14.
38. Gómez, J., Hisler, V. J., Xie, D., & Friere, E. (1995) *Proteins: Struct., Funct., Genet. 22*, 404−412.
39. Hvidt, A. (1964) *C. R. Trav. Lab. Carlsberg 34*, 299−317.
40. Mori, S., van Zijl, P. C. M., Shortle, D. (1997) *Proteins: Struct., Funct., Genet. 28*, 325−332.
41. Privalov, P. L. (1979) *Adv. Protein Chem. 33*, 167−241.

42. Makhatadze, G. I., Kim, K. S., Woodward, C., & Privalov, P. (1993) *Protein Sci. 2*, 2028−2036.
43. Makhatadze, G. I., Clore, G. M., Gronenborn, A. M., & Privalov, P. L. (1994) *Biochemistry 33*, 9327−9332.
44. Wintrode, P. L., Makhatadze, G. I., & Privalov, P. L. (1994) *Proteins: Struct., Funct., Genet. 18*, 246−253.
45. Bai, Y., Sosnick, T. R., Mayne, L., & Englander, S. W. (1995) *Science 269*, 192−197.
46. Chamberlain, A. K., Handel, T. M., & Marqusee, S. (1996) *Nature Struct. Biology 3*, 782−787.
47. Yi, Q., Scalley, M., Simons, K., & Baker, D. (1997) *Folding Des. 2*, 271−280.
48. Itzhaki, L. S., Neira, J. L., & Fersht, A. R. (1997) *J. Mol. Biol. 270*, 1−10.
49. Johnson, C. M., & Fersht, A. R. (1995) *Biochemistry 34*, 6795−604.
50. Orban, J., Alexander, P., Bryan, P., & Khare, D. (1995) *Biochemistry 34*, 15291−15300.
51. Mayo, S., & Baldwin, R. (1993) *Science 262*, 873−876.
52. Kim, P. S., & Baldwin, R. L. (1982) *Biochemistry 21*, 1−5.
53. Matthew, J. B., & Richards, F. M. (1983) *J. Biol. Chem. 258*, 3039−3044.
54. Bystroff, C., & Baker, D. (1997) (submitted for publication).
55. Viguera, A. R., Blanco, F. J., & Serrano, L. (1995) *J. Mol. Biol. 247*, 670−681.
56. Srinivasan, R., & Rose, G. D. (1995) *Proteins: Struct., Funct., Genet. 22*, 81−99.
57. Simons, K. T., Kooperberg, C., Huang, E., & Baker, D. (1997) *J. Mol. Biol. 268*, 1−17.