Structural and kinetic characterization of the simplified SH3 domain FP1

QIAN YI,[1] PONNI RAJAGOPAL,[1] RACHEL E. KLEVIT,[1] AND DAVID BAKER[1,2]

[1]Department of Biochemistry and [2]Howard Hughes Medical Institute, University of Washington, Seattle, Washington 98195, USA

(RECEIVED October 11, 2002; FINAL REVISION January 15, 2003; ACCEPTED January 15, 2003)

Abstract

The simplified SH3 domain sequence, FP1, obtained in phage display selection experiments has an amino acid composition that is 95% Ile, Lys, Glu, Ala, Gly. Here we use NMR to investigate the tertiary structure of FP1. We find that the overall topology of FP1 resembles that of the src SH3 domain, the hydrogen-deuterium exchange and chemical shift perturbation profiles are similar to those of naturally occurring SH3 domains, and the ^{15}N relaxation rates are in the range of naturally occurring small proteins. Guided by the structure, we further simplify the FP1 sequence and compare the effects on folding kinetics of point mutations in FP1 and the wild-type src SH3 domain. The results suggest that the folding transition state of FP1 is similar to but somewhat less polarized than that of the wild-type src SH3 domain.

Keywords: FP1; simplified SH3; folding kinetics; NMR structure

Naturally occurring proteins are made from the 20 amino acids encoded by the modern genetic code. However, it has been argued (Osawa et al. 1992) that early in the history of life, the modern genetic code evolved from a simpler code representing only a handful of amino acids. Is this possible? Recent success in protein design has demonstrated that helical proteins can be built from simplified amino acid alphabets. Regan and coworkers (Munson et al. 1994, 1996) produced variants of the helix-bundle RNA-binding protein with highly simplified hydrophobic cores. The structure of a designed four-helix bundle protein, DHP, built from seven amino acids, was determined to 2.9 Å resolution and found to be nearly as well packed as native proteins (Schafmeister et al. 1997). T4 lysozyme with 10 core residues replaced with methionine folds to a partially active and thus presumably very native-like conformation (Gassner et al. 1996). The phage 434 Cro protein with 11 out of 13 core residues replaced with leucine exhibits the folding cooperativity and nuclear magnetic resonance (NMR) dispersion expected for a fully folded protein (Desjarlais and Handel 1995).

Reprint requests to: David Baker, Department of Biochemistry, Box 357350, University of Washington, Seattle, WA 98195, USA; e-mail: dabaker@u.washington.edu; fax: (206) 685-1792.

Article and publication are at http://www.proteinscience.org/cgi/doi/10.1110/ps.0238603.

Here we characterize the tertiary structure of the simplified 57-residue β-sheet protein FP1 previously obtained from a combinatorial phage-display library based on the src-SH3 domain (Riddle et al. 1997; see Table 1 for sequence comparison between FP1 and src-SH3). The FP1 scaffold that supports the proline-rich peptide binding site of the SH3 domain is 95% Ile, Lys, Glu, Ala, and Gly. Previous qualitative structural characterization by gel filtration chromatography, circular dichroism, and NMR indicated that FP1 is folded. Compared to the wild-type src SH3, it is slightly destabilized, folds approximately twice as fast, and unfolds five times as fast. It binds to a proline-rich peptide ligand 20 times more weakly than the wild-type src SH3. In this study, we used NMR techniques to determine the structure of FP1, and we found that the topology of FP1 is very similar to that of the wild-type SH3. Further simplification of the FP1 sequence was achieved by simplifying the binding residues that were kept constant in the previous design strategy (Riddle et al. 1997), and the folding of the variants is compared to that of corresponding variants of the wild-type src-SH3 domain.

Results

NMR structure determination

The FP1 structural characterization was carried out in 50mM sodium phosphate, 90% H_2O/10%D_2O, pH 6.0 with

Table 1. *Sequence alignment between the simplified protein FP1 and src-SH3*

```
     9           19          29          39          49          59
SH3 TFVALYDYES RTETDLSFKK GERLQIVNNT EGDWWLAHSL STGRTGYIPS NYVAPSD
FP1 EFIAIYDYKA ETEEDLTIKK GEKLEIIEK. EGDWWKAKAI GSGEIGYIPA NYIAAAE
```

The numbering system for FP1 starts at 9. The residues in blue were binding residues that were kept unchanged during the combinatorial phage-display experiment (Riddle et al.1997); the other residues were allowed to be one of Ile, Glu, Lys, Gly, or Ala in the original oligonucleotide design. The residues in FP1 that are Ile, Glu, Lys, Gly, and Ala are highlighted in red. The "." indicates a deletion in FP1.

400 mM sodium sulfate. The initial comparison of the ^{15}N-HSQC spectra of src-SH3 and FP1 (Fig. 1) showed that the chemical shift dispersion of the amide protons of FP1 (from 7.0 ppm to 10.0 ppm), is comparable to that of src SH3, suggesting that FP1 is folded.

Hydrogen-deuterium (H-D) exchange rates for the backbone amides of FP1 were measured to determine the accessibility of the backbone to solvent. Figure 2A shows the free energy for H-D exchange for the protected amides in FP1. With the exception of the last several residues at the C-terminus, the H-D exchange pattern is very similar in FP1 and the wild-type src SH3 (Fig. 2B; Grantcharova and Baker 1997).

To determine the tertiary structure of FP1, the resonance assignments of $^{13}C_\alpha$, $^{13}C_\beta$, ^{15}N, and ^{1}H-N for the simplified protein FP1 in solution were carried out using standard 3D-triple-resonance NMR experiments (HNCA, HNCOCA, HNCACB, and CBCACONH). The sidechain proton assignments were obtained from 3D-HCCH-TOCSY, and 3D-HCCH-COSY experiments in combination with 3D-^{15}N-edited NOESY-HSQC, 3D-^{13}C-edited NOESY-HSQC experiments (about 95% of proton resonances were assigned).

$^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts are predominately determined by the backbone conformation, and the chemical shift differences from the random coil values, $\Delta^{13}C_\alpha$ and $\Delta^{13}C_\beta$ (Ikura et al. 1991; Wishart and Sykes 1994), are indicators of secondary structure in folded proteins. In general, $^{13}C_\alpha$ resonances are shifted downfield by an average of 2.6 ppm for α-helices, and shifted upfield by 1.7 ppm for β-sheets. The correlation of $\Delta^{13}C_\alpha$ and $\Delta^{13}C_\beta$ values with secondary structure is enhanced by calculating ($\Delta^{13}C_\alpha$ – $\Delta^{13}C_\beta$) for each residue (Metzler et al. 1993; Constantine et al. 1997). Also, whereas the individual $\Delta^{13}C_\alpha$ and $\Delta^{13}C_\beta$ values depend on the exact ^{13}C chemical shift referencing (Wishart and Sykes 1994), ($\Delta^{13}C_\alpha$ – $\Delta^{13}C_\beta$) does not, provided that the $^{13}C_\alpha$ and $^{13}C_\beta$ chemical shifts are taken from spectra that are similarly referenced. The ($\Delta^{13}C_\alpha$ – $\Delta^{13}C_\beta$) values for FP1 are shown in Figure 3A. Because the $\Delta^{13}C_\alpha$ and $\Delta^{13}C_\beta$ chemical shifts of src-SH3 have not been reported, for comparison we show in Figure 3B the ($\Delta^{13}C_\alpha$ – $\Delta^{13}C_\beta$) for the structurally related drk SH3 domain (Zhang and Forman-Kay 1995), generously provided by Dr. Julie Forman-Kay. Most values of $\Delta^{13}C_\alpha$ – $\Delta^{13}C_\beta$ are negative, indicating that FP1 contains primarily extended local structure with little helical content as in drk-SH3. The magnitude of ($\Delta^{13}C_\alpha$ – $\Delta^{13}C_\beta$) of the FP1 in general is somewhat smaller than that of the drk-SH3, and no three consecutive residues in FP1 have ($\Delta^{13}C_\alpha$ – $\Delta^{13}C_\beta$) values above 3ppm,

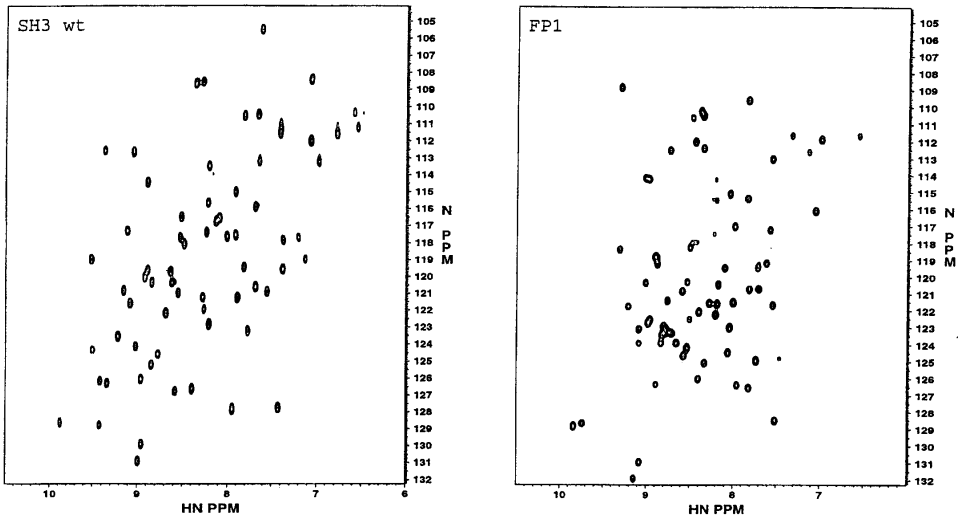

Figure 1. Comparison of ^{15}N-HSQC of src-SH3 in 50 mM sodium phosphate (pH 6.0) and FP1 in 50 mM sodium phosphate (pH 6.0) and 400 mM sodium sulfate.

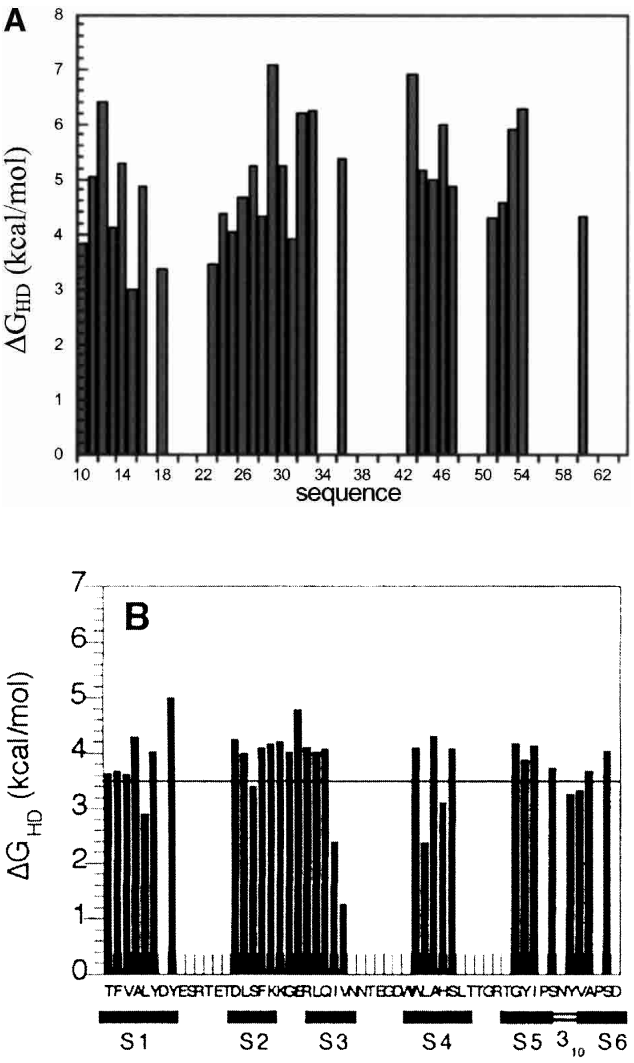

Figure 2. Comparison of amide proton H-D exchange in src SH3 and FP1. ΔG_{HD} was calculated using $\Delta G_{HD} = -RT\ln(k_{obs}/k_{rc})$ assuming the exchange is under EX2 condition (Bai et al. 1994). (*A*) FP1 in 50 mM sodium phosphate (pH 6.0) and 400 mM sodium sulfate. (*B*) The wild-type src-SH3 in 50 mM sodium phosphate (pH 6.0).

suggesting some structural averaging. This is also supported by the measurement of three-bond coupling constants of $J^{HN\alpha}$: No three consecutive residues in FP1 have $J^{HN\alpha}$ values all greater than 9.0 for the β-sheet conformation or all smaller than 5.0 for the α-helix conformation (data not shown). However, the overall pattern of $(\Delta^{13}C_\alpha - \Delta^{13}C_\beta)$ is remarkably similar in FP1 and the drk SH3 domain, suggesting a similar distribution of secondary structure propensity.

The tertiary structure of FP1 was investigated by collecting a total of 826 distance constraints for residues Glu9–Glu64 from a 120-msec 3D ^1H-^{15}N NOESY-HSQC and a 120-msec 3D ^1H-^{13}C NOESY-HMQC, and a total of 40 dihedral-angle constraints from a J-modulated HSQC experiment and a 3D HNHB experiment. X-PLOR was used to calculate 100 initial structures based on the distance and dihedral-angle constraints provided by the NMR experiments. Fifty-two of the 100 structures satisfied the selection criteria of no NOE constraint violations > 0.3 Å, no dihedral angle constraint violations > 3°, no deviations from ideal bond length > 0.03 Å, and no deviations from ideal geometry > 3°. The precision and quality of the structures, and the root mean square deviations (RMSD) are summarized in Table 2. An overlay of 10 final structures of FP1 is shown in Figure 4. The average RMSD of the 52 final calculated structures (residues 9–64) from the averaged structure coordinates are 1.52 Å for the backbone atoms and 2.22 Å for the heavy atoms. Residues 14–28 are much less well defined compared to the rest of the structures because of a very low density of medium- and long-range NOEs. The hydrophobic residues Phe10, Ile11, A12, Leu32, Ile34, Ile35, Trp42,

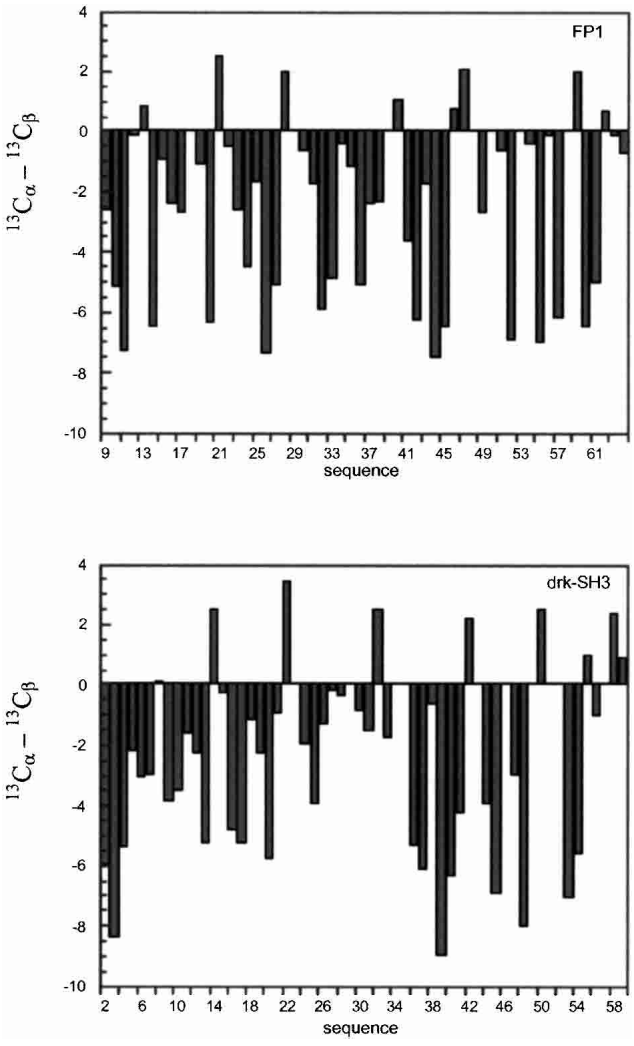

Figure 3. Comparison of the chemical shift deviations of $(\Delta^{13}C_\alpha - \Delta^{13}C_\beta)$ between FP1 and wild-type drk-SH3. (*A*) FP1 in 50 mM sodium phosphate (pH 6.0) and 400 mM sodium sulfate. (*B*) The wild-type drk-SH3.

Table 2. *Deviation from experimental constraints of the calculated FP1 structures*[a]

	# Constraints	RMSD
All noes	826	0.00849 ± 0.00216
Intraresidue	270	0.00295 ± 0.00259
Interresidue sequential ($\|i - j\| = 1$)	248	0.00953 ± 0.00443
Interresidue medium range ($1 < \|i - j\| < 5$)	138	0.0153 ± 0.0051
Interresidue long range ($\|i - j\| \geq 5$)	170	0.0107 ± 0.0045
RMSD from dihedral angle constraint	40	0.402 ± 0.091
RMSD from idealized covalent geometry[b]		
Angles (°)		0.595 ± 0.031
Bonds (A°)		0.00577 ± 0.00023
Improper (°)		0.259 ± 0.022
Average energies[c]		
E_{noe} (kcal mole^{-1})		3.990 ± 2.344
E_{dih} (kcal mole^{-1})		0.433 ± 0.204
E_{repel} (kcal mole^{-1})		3.482 ± 1.118
E_{LJ} (kcal mole^{-1})[c]		−78.84 ± 23.74
Percentage of residues (non-Gly, non-Pro) in most-favored and additional allowed regions of the Ramachandran plot[d]		89%
Coordinate precision, average pairwise RMSD (A°)		
Residues 9–64 (C′, C_α, N)		2.11 ± 0.55
Residues 9–64 (heavy atoms)		3.17 ± 0.55
Residues 10–13 (C′, C_α, N)		0.46 ± 0.20
Residues 10–13 (heavy atoms)		4.77 ± 1.36
Residues 30–63 (C′, C_α, N)		1.09 ± 0.25
Residues 30–63 (heavy atoms)		3.50 ± 0.71

[a] Comparison of the 50 accepted structures from XPLOR simulation. All structures have no NOE constraint violations >0.3 A°, no dihedral angle constraint violations >3°, no deviations from ideal bone length >0.03 A°, and no deviations from ideal geometry >3°.
[b] Force constant used in the XPLOR target function: covalent bonds, 100 kcal (mole * A°)$^{-1}$; angles, 500 kcal (mole * rad)$^{-1}$; improper, 500 kcal (mole * rad)$^{-1}$; NOE constraints with 50 kcal (mole * A°)$^{-1}$ and dihedral angle constraints 200 kcal (mole * rad)$^{-1}$. Square-well potentials were used for NOE and dihedral constraints.
[c] Calculated using the param19x.pro CHARMM energy function. This term is used only for calculation of the Lennard-Jones energy term and not during the simulated annealing refinement.
[d] Ramachandran plots calculated using PROCHECK-NMR.

Ala44, Ala46, Ile55, Pro56, Ala57, and Ile60 are packed in the core of the protein. No regular secondary structure elements could be consistently assigned within the accepted FP1 structures, consistent with the chemical shift perturbation results described above. As shown in Figure 5, the overall topology of FP1 is very similar to the c-src SH3 NMR structure (Yu et al. 1992).

^{15}N NMR relaxation measurements were also performed to assess the backbone dynamics. The T_1 and T_2 of the backbone ^{15}N in FP1 are 0.58 ± 0.03 (sec) and 0.084 ± 0.013 (sec) respectively. The ^{15}N-^{1}H heteronuclear NOEs are 0.69 ± 0.08. These results are comparable to those of naturally occurring proteins with similar sizes. In summary, the overall topology, hydrogen bonding pattern, and chemical shift perturbation profile of FP1 are remarkably similar to those of the naturally occurring SH3 domains, and the relaxation data suggest similar levels of backbone motions.

Further simplification

The previous simplification of SH3 was achieved restricting nonfunctional residues to a five-letter alphabet: the nonpolar residue isoleucine (I), the polar residues lysine (K) and glutamate (E), alanine (A) for core positions, with not enough space for isoleucine and glycine (G) for conformational flexibility in turns. Is it possible to further simplify the sequence and retain the ability to fold? Four single-point mutations were made in FP1, Y14A, D23A, L32I, and S49A, to investigate this possibility. Y14 and D23 are involved in binding to the proline-rich peptide, and were kept unchanged in the previous simplification strategy, which relied on the peptide binding activity. The isoleucine substitution for leucine at position 32 is aimed to further simplify the hydrophobic composition of FP1. S49 may play an important role in folding of FP1, because T50 in SH3 plays an important role in the folding of the wt. src-SH3 domain (Riddle et al. 1999). The point mutations are all highlighted in the FP1 averaged structure (Fig. 6A).

Y14A and D23A were expressed to a significant level, whereas S49A was poorly expressed. They are all less stable than the original FP1, as shown in Table 33, but S49A is more dramatically destabilized than Y14A and D23A. L32I was so poorly expressed that a reasonable amount of protein

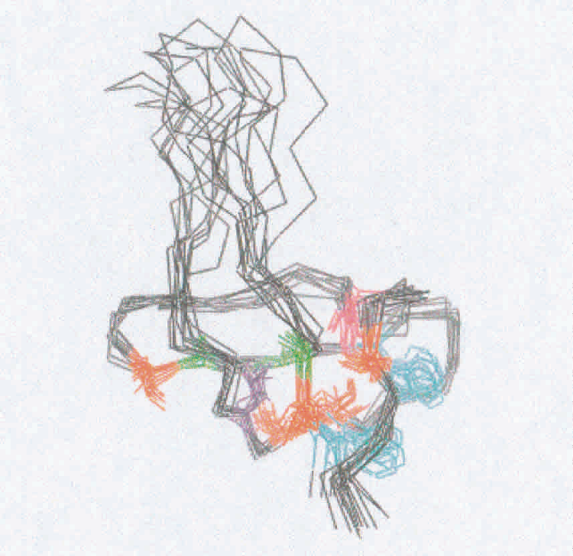

Figure 4. A superposition of the backbone C_α of 10 final refined FP1 structures. The sidechains of core hydrophobic residues are shown in color.

could not be obtained for thermodynamic and kinetic study. These results suggest that the sequence of FP1 can be further simplified at exposed positions, but not in the protein core.

Folding kinetics

Previous studies (Riddle et al. 1999) demonstrated that the distal β-hairpin is formed in the transition state and is critical for the folding of the SH3 topology. Residues A45 and G51 play important roles in the folding transition of src SH3, and the corresponding mutations were made in FP1 (A44G and G50A) to investigate whether it has a folding mechanism similar to that of the wild-type SH3. Figure 6B shows the folding profiles of Y14A, D23A, A44G, and G50A in comparison to that of FP1, and Table 3 lists the kinetic parameters for the mutants. Y14 in FP1, like Y14 in SH3, does not have much impact on either folding or unfolding. As in the wild-type SH3, the D23A mutant in FP1 does not have a significant effect on folding. However, the D23A mutant increases the unfolding of FP1 by 4.5-fold, whereas the unfolding of src SH3 was only increased twofold by the corresponding substitution. A44G in FP1 slows down the folding by threefold and has little effect on the unfolding, whereas A45G in SH3 slows down the folding by sixfold and has little effect on the unfolding. S49A in FP1 slows down the folding by threefold, but accelerates the unfolding by threefold, whereas T50A in SH3 slows down the folding by 13-fold, but has a minor effect on the unfolding. G50A in FP1 decreases the folding rate by twofold and increases the unfolding rate by twofold, whereas G51A in SH3 decreases the folding rate by eightfold and has no effect on the unfolding rate. The lower ϕ values associated with the A44G, S49G, and G40A mutations in FP1 compared to src SH3 (Table 3) suggest that FP1 has a less polarized folding transition state than that of src SH3 (Riddle et al. 1999).

Discussion

NMR characterization of the tertiary structure of FP1 reveals that the protein has the overall SH3 domain fold. The hydrogen bonding pattern and the chemical shift perturbation profiles are remarkably similar to those of naturally occurring SH3 domains, and the ^{15}N relaxation rates are in the range of those of small native proteins, suggesting that FP1 is nearly as well ordered as naturally occurring proteins. The fact that we were able to determine an NMR solution structure for FP1 sets it apart from most designed all-β proteins, which for the most part have had very poorly resolved NMR spectra. (Our recent results with computational protein design highlight the difficulties with designing all-β proteins: Of nine α, αβ, and all-β proteins completely redesigned using the method described by Kuhlman and Baker (2000), the only protein to be unfolded under physiological conditions was a completely redesigned SH3 domain). The relatively well defined NMR spectrum of FP1 compared to other designed all-β proteins may reflect the selection for function used to obtain FP1 initially—protein function is likely to put greater constraints on the mobility of the structure than overall folding.

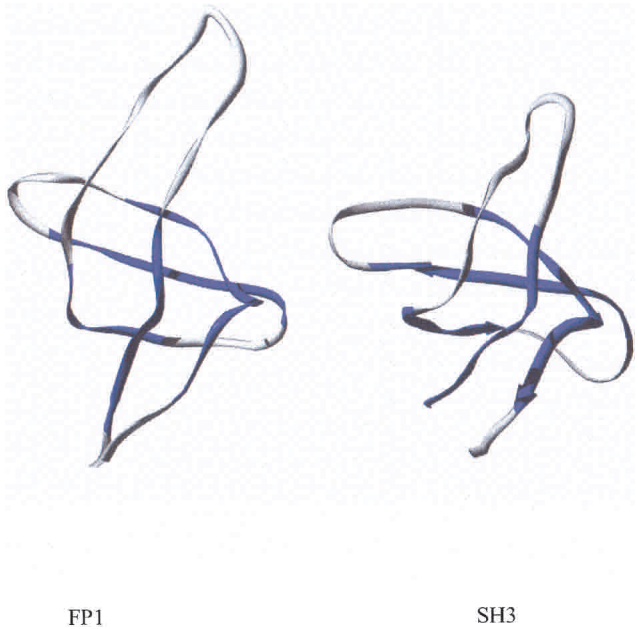

Figure 5. Comparison of the overall topology of FP1 and the wild-type src SH3 domain (Yu et al. 1992). The secondary structural elements of src SH3 and the corresponding regions of FP1 are highlighted in blue.

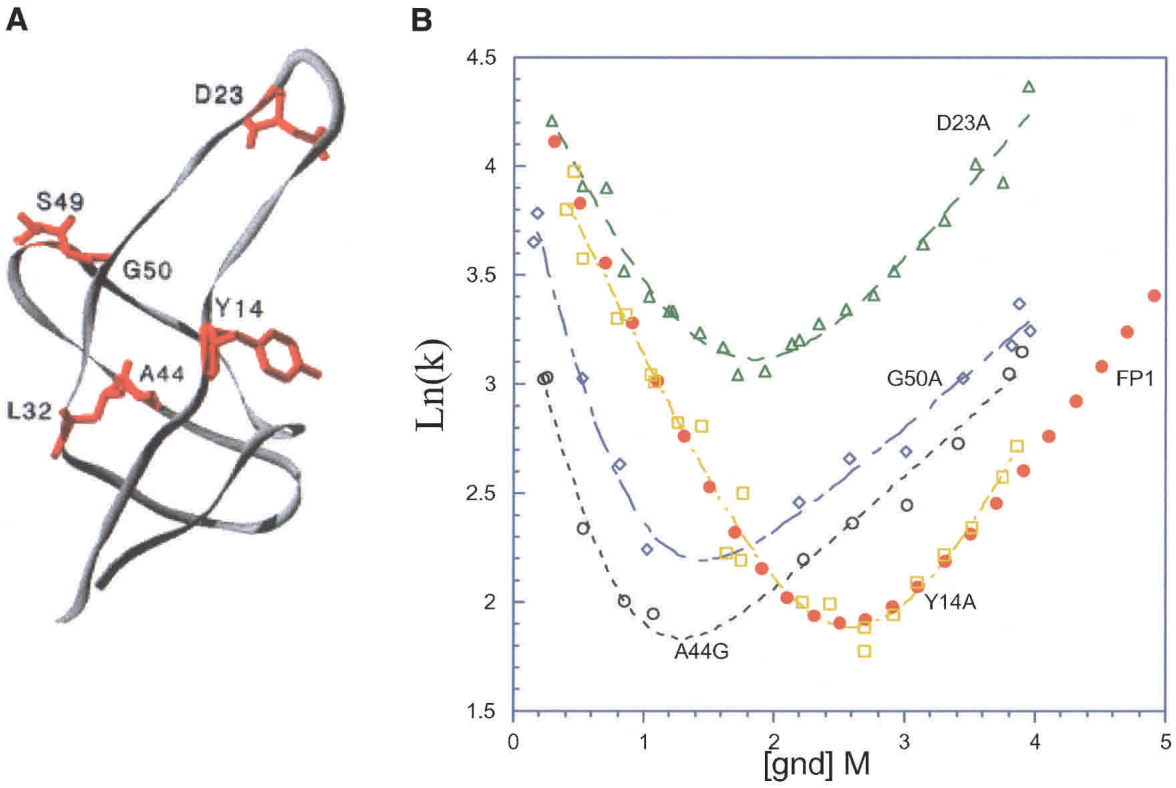

Figure 6. (*A*) The mutants of FP1 whose folding kinetics were characterized are highlighted in red in the average FP1 structure. (*B*) Dependence of the folding and unfolding rates on the denaturant concentration for FP1 and the FP1 mutants indicated in (A).

A striking feature of the folding mechanism of the src, spectrin (Martinez and Serrano 1999), and Fyn SH3 domains (Northey et al. 2002) is the polarization of the folding transition state: Part of the structure is largely formed and the rest, largely disrupted at the rate-limiting step in folding. The limited number of ϕ values we were able to obtain for FP1 suggest that this polarization is somewhat decreased in the simplified protein. In src SH3, Y14A has a ϕ value close to zero, whereas A45G, T50A and G51A have ϕ values close to 1.0. In FP1, the ϕ value of Y14A is increased to 0.20, and the ϕ values of the latter three mutations are all below 0.8. In the wild-type protein, favorable specific interactions between the distal β hairpin and the diverging turn may be sufficient to overcome the entropy of ordering additional residues, whereas partial disruption of some of these interactions in the course of simplification of FP1 may necessitate the formation of a large fraction of the protein structure to overcome the entropic barrier to folding.

Materials and methods

Sample preparation

FP1 was cloned into pET 15b (Novagen) and transformed into BL21 (DE3/plysS) cells, and protein was overexpressed. The construct contains an N-terminal 6 X His-tag to facilitate purification. The His-tag was not removed in this study. All mutants were made using the Quick Change site-directed mutagenesis kit (Stratagene). All mutants were sequenced to ensure that the mutagenesis was successful, and the purified proteins were verified by mass spectrometry.

The ^{15}N- single-labeled and the ^{13}C, ^{15}N- double-labeled protein samples were made by growing the transformed *E. coli* cells in M9 minimal medium using 99.9% $^{15}NH_4Cl$ as sole nitrogen source and 99.9% ^{13}C-glucose as sole carbon source. The level of labeling was estimated to be greater than 95% based on the result from mass spectrometry. The NMR sample conditions were 1.2 mM protein, 90% H_2O/10% D_2O 50 mM sodium phosphate (pH 6.0), and 400 mM sodium sulfate.

NMR experiments

All spectra were collected at 22°C on a four-channel Bruker DMX 500 MHz spectrometer equipped with a triple resonance, triple axis gradient probe. The backbone assignments were accomplished by straight-forward 3D-triple resonance experiments of HNCA, HNCOCA, HNCACB, and CBCACONH. The sidechain assignments were obtained from a 3D- HCCH-TOCSY, 3D- HCCH-COSY, and 120-msec 3D- ^{15}N-edited NOESY-HSQC and a 120-msec 3D- ^{13}C-edited NOESY-HSQC experiment. A series of J-modulated ^{1}H, ^{15}N- HSQC experiments were collected to obtain the $J^{HN\alpha}$ coupling constants in order to derive the backbone ϕ angle constraints for structural determination. A 3D- HNHB experiment was also carried out in order to gain some χ1 angle

Table 3. *Comparison of kinetic parameters of wild-type SH3 and simplified FP1 variants*

Protein	$k_f^{0.3M}$ (s^{-1})	$k_u^{3.8M}$ (s^{-1})	$\Delta\Delta G_u$[a] (cal $mole^{-1}$)	Φ_f[b]
SH3[c]	34.8	2.92		
Y14A[c]	36.2	5.36	−0.31	−0.08
D23A[c]	30.9	6.75	−0.56	0.13
A45G[c]	5.53	2.36	−0.92	1.20
T50A[c]	2.69	4.90	−1.79	0.86
G51A[c]	4.01	2.77	−1.21	1.06
FP1[d]	66.7	13.1		
Y14A	48.4	15.2	−0.12	0.20
D23A	53.5	65.4	−1.05	0.09
A44G	20.3	18.7	−0.87	0.77
S49A	18.2	34.1	−1.31	0.58
G50A	36.2	28.2	−0.79	0.44

[a] $\Delta\Delta G_u$ was calculated using the following equation:

$$\Delta\Delta G_u = -RT(\ln((k_{0.3M}{}^{fp1}/k_{0.3M}{}^{mut}) + (\ln(k_{3.8M}{}^{mut}/k_{3.8M}{}^{fp1}))$$

[b] Φ_f was calculated using the following equation:

$$\Phi_f = -RT\ \ln(k_{0.2M}{}^{fp1}/k_{0.3M}{}^{mut})/\Delta\Delta G_u$$

[c] Kinetic parameters are from Riddle et al. (1999).
[d] Kinetic parameters are from Riddle et al. (1997).

constraints for structural determination. Spectra widths were typically 7000 Hz in the ^1H dimension, 2000 Hz in the ^{15}N dimension, and 12, 000 Hz in the ^{13}C dimension at 500 MHz field strength. Sixteen transients were typically collected with a relaxation of 1.2 sec.

NMR data processing

NMR data were processed with NMRPipe software (Delaglio et al. 1995). Spectra were typically linearly predicted in the indirectly detected dimensions, apodized with a $\pi/3$-shifted sine bell square in both dimensions, and zero-filled. The NOESY spectra used in the structure determination were first multiplied by a $\pi/2$-shifted sine bell squared function and zero-filled to double the number of data points before Fourier transformation. The program PIPP (Garrett et al. 1991) was used for NOE crosspeak assignments and measurement of peak intensities.

Hydrogen-deuterium exchange

H-D exchange experiments were carried out by acquiring a series of ^{15}N-HSQC spectra immediately after dissolving lyophilized FP1 protein in deuterated buffer. The exchange rates were estimated as described (Yi et al. 1996).

Structure calculation

Two NOESY spectra of 3D-^{15}N-edited NOESY-HSQC and a 120-msec 3D- ^{13}C-edited NOESY-HSQC were assigned and translated to 823 distance constraints classified as strong (1.80–3.0 Å), medium (1.80–4.5 Å), weak (1.80–5.5 Å), or very weak (1.80–6.0 Å). Twenty-seven ϕ angle constraints and 13 $\chi 1$ angle constraints were obtained from the J-modulated HSQC series and 3D- HNHB experiment, respectively. Most angle constraints were allowed to vary by ±30°, and some were allowed to vary by ±50°. Structure calculations were carried out with X-PLOR using the standard substructure-embedding and simulated-annealing protocols of dg_sub_embed.inp, dgsa.inp, and refine.inp (Nigles et al. 1988; Brunger 1993). Some minor modification was made to dgsa.inp and refine.inp protocols with the starting temperature at 2500 K and the 5000 cooling steps. In addition, the chirality of nonstereospecifically assigned methyl and methylene protons was evaluated using the X-PLOR protocol swap15v.inp (Nigles et al. 1988; Folmer et al. 1997). Distance constraints involving the degenerate resonances of methyl and methylene protons were averaged.

Biophysical characterization

Biophysical characterization of all the mutants was performed at 22°C in 50 mM sodium phosphate, pH 6.0. The kinetics of folding and unfolding were determined by stop-flow fluorescence using a Bio-Logic SFM-4 stopped-flow instrument. All kinetic data were analyzed using a two-state model as described (Scalley et al. 1997). The $\Delta\Delta G$ and Φ values listed in Table 2 for the mutants were calculated from the folding rates at 0.23 M guanidine, and the unfolding rates at 3.86 M guanidine without extrapolation to 0 M guanidine, as described (Riddle et al. 1999).

Acknowledgments

This work was supported by a grant from the NIH. We thank Dr. Julie Forman-Kay for generously providing the chemical shift data of the drk-SH3.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Bai, Y., Milne, J.S., Mayne, L., and Englander, S.W. 1994. Protein stability parameters measured by hydrogen exchange. *Proteins.* **20:** 4–14.

Brunger, A.T. 1993. *X-PLOR manual version 3.1: A system for crystallography and NMR.*Yale University Press, New Haven, CT.

Constantine, K.L., Mueller, L., Goldfarb, V., Wittekind, M., Metzler, W.J., Yanchunas, Jr., J., Robertson, J.G., Malley, M.F., Friedrichs, M.S., and Farmer II, B.T. 1997. Characterization of $NADP^+$ binding to perdeuterated MurB: Backbone atom NMR assignments and chemical-shift changes. *J. Mol. Biol.* **267:** 1223–1246.

Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. 1995. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6:** 277–293.

Desjarlais, J.R. and Handel, T.M. 1995. De novo design of the hydrophobic cores of proteins. *Protein Sci.* **4:** 2006–2048.

Folmer, R.H., Hilbers, C.W., Konings, R.N., and Higles, M. 1997. Floating stereospecific assignment revisited: Application to an 18kD protein and comparison with J-coupling data. *J. Biomol. NMR* **9:** 245–258.

Garrett, D.S., Powers, R., Gronenborn, A.M., and Clore, G.M. 1991. A common sense approach to peakpicking in two-, three-, and four-dimensional spectra using automatic computer analysis of contour diagrams. *J. Magn. Reson.* **95:** 214–220.

Gassner, N.C., Basse, W.A., and Mathews, B.W. 1996. A test of the jigsaw puzzle model for protein folding by multiple methionine substitutions within the core of T4 lysozyme. *Proc. Natl. Acad. Sci.* **93:** 12155–12158.

Grantcharova, V.P. and Baker, D. 1997. Folding dynamics of the src SH_3 domain. *Biochemistry* **36:** 15685–15692.

Ikura, M., Spera, S., Barbato, G., Kay, L.E., Krinks, M., and Bax, A. 1991. Secondary structure and side-chain 1H and 13C resonance assignments of calmodulin in solution by heteronuclear multidimensional NMR spectroscopy. *Biochemistry* **24:** 9216–9228.

Kuhlman, B. and Baker, D. 2000. Native protein sequences are close to optimal for their structures. *PNAS* **97:** 10383–10388.

Martinez, J.C., and Serrano, L. 1999. The folding transition state between SH3

domains is conformationally restricted and evolutionarily conserved. *Nat. Struct. Biol.* **6:** 1010–1016.

Metzler, W.J., Constantine, K.L., Friedrichs, M.S., Bell, A.J., Ernst, E.I., Lavoie, T.B., and Mueller, L. 1993. Characterization of the three-dimensional solution structure of human profiling: 1H, 13C, and 15N NMR assignments and global folding pattern. *Biochemistry* **32:** 13818–13829.

Munson, M., O'Brien, R., Sturtevant, J.M., and Regan, L. 1994. Redesigning the hydrophobic core of a four-helix-bundle protein. *Protein Sci.* **3:** 2015–2022.

Munson, M., Balasubramanian, S., Fleming, K.G., Nagi, A.D., O'Brien, R., Sturtevant, J.M., and Regan, L. 1996. What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties. *Protein Sci.* **5:** 1584–1593.

Nigles, M., Clore, G.M., and Gronenborn, A.M. 1988. Determination of the three dimensional structure of proteins from interproton distance data by hybrid distance geometry-dynamical simulated-annealing calculations. *FEBS Lett.* **229:** 317–324.

Northey, J.G., Maxwell, K.L., and Davidson, A.R. 2002. Protein folding kinetics beyond the ϕ value: Using multiple amino acid substitutions to investigate the structure of the SH3 domain folding transition state. *J. Mol. Biol.* **5:** 389–402.

Osawa, S., Jukes, T.H., Watanabe, K., and Muto, A. 1992. Recent evidence for evolution of the genetic code. *Microbiol. Rev.* **56:** 229–264.

Riddle, D.S., Santiago, J.V., Bray-Hall, S.T., Doshi, N., Grantcharova, V.P., Yi, Q., and Baker, D. 1997. Functional rapidly folding proteins from simplified amino acid sequences. *Nat. Struct. Biol.* **4:** 805–809.

Riddle, D.S., Grantcharova, V.P., Santiago, J.V., Alm, E., Ruczinski, I., and Baker, D. 1999. Experiment and theory highlight role of native state topology in SH3 folding. *Nat. Struct. Biol.* **6:** 1016–1024.

Scalley, M.L., Yi, Q., Gu, H., McCormack, A., Yates III, J.R., and Baker, D. 1997. Kinetics of folding of the IgG binding domain of peptostreptococcal protein L. *Biochemistry* **36:** 3373–3382.

Schafmeister, C.E., LaPorte, S.L., Miercke, L.J.W., and Stroud, R.W. 1997. A designed four helix bundle protein with native-like structure. *Nat. Struct. Biol.* **4:** 1039–1046.

Wishart, D.S. and Sykes, B.D. 1994. Chemical shifts as a tool for structure determination. *Methods Enzymol.* **239:** 363–393.

Yi, Q. and Baker, D. 1996. Direct evidence for a two-state protein folding transition from hydrogen-deuterium exchange, mass spectrometry, and NMR. *Protein Science* **5:** 1060–1066.

Yu, H., Rosen, M.K., Shin, T.B., Seidel-Dugan, C., Brigge, J.S., and Schreiber, S.L. 1992. Solution structure of the SH3 domain of Src and identification of its ligand binding site. *Science* **258:** 1665–1668.

Zhang, O. and Forman-Kay, J.D. 1995. Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer. *Biochemistry* **34:** 6784–6794.