doi:10.1016/j.jmb.2006.08.047 J. Mol. Biol. (2006) 363, 958–976

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Identification of a Collapsed Intermediate with Non-native Long-range Interactions on the Folding Pathway of a Pair of Fyn SH3 Domain Mutants by NMR Relaxation Dispersion Spectroscopy

Philipp Neudecker[1,2], Arash Zarrine-Afsar[1], Wing-Yiu Choy[1,2] D. Ranjith Muhandiram[1,2], Alan R. Davidson[1] and Lewis E. Kay[1,2]*

[1]*Departments of Medical Genetics and Biochemistry University of Toronto, Toronto ON, Canada M5S 1A8*

[2]*Department of Chemistry University of Toronto, Toronto ON, Canada M5S 3H6*

**Corresponding author*

Recent ^{15}N and ^{13}C spin-relaxation dispersion studies of fast-folding mutants of the Fyn SH3 domain have established that folding proceeds through a low-populated on-pathway intermediate (I) where the central β-sheet is at least partially formed, but without interactions between the NH_2- and COOH-terminal β-strands that exist in the folded state (F). Initial studies focused on mutants where Gly48 is replaced; in an effort to establish whether this intermediate is a general feature of Fyn SH3 folding a series of ^{15}N relaxation experiments monitoring the folding of Fyn SH3 mutants N53P/V55L and A39V/N53P/V55L are reported here. For these mutants as well, folding proceeds through an on-pathway intermediate with similar features to those observed for G48M and G48V Fyn SH3 domains. However, the ^{15}N chemical shifts extracted for the intermediate indicate pronounced non-native contacts between the NH_2 and COOH-terminal regions not observed previously. The kinetic parameters extracted for the folding of A39V/N53P/V55L Fyn SH3 from the three-state folding model F↔I↔U are in good agreement with folding and unfolding rates extrapolated to zero denaturant obtained from stopped-flow experiments analyzed in terms of a simplified two-state folding reaction. The folding of the triple mutant was studied over a wide range of temperatures, establishing that there is no difference in heat capacities between F and I states. This confirms a compact folding intermediate structure, which is supported by the ^{15}N chemical shifts of the I state extracted from the dispersion data. The temperature-dependent relaxation data simplifies data analysis because at low temperatures (<25 °C) the unfolded state (U) is negligibly populated relative to I and F. A comparison between parameters extracted at low temperatures where the F↔I exchange model is appropriate with those from the more complex, three-state model at higher temperatures has been used to validate the protocol for analysis of three-site exchange relaxation data.

Keywords: chemical exchange; CPMG NMR relaxation dispersion; Fyn SH3 domain; protein folding; stopped-flow kinetics

Present address: W.-Y. Choy, Department of Biochemistry, University of Western Ontario, London, ON, Canada N6A 5C1.

Abbreviations used: CPMG, Carr-Purcell-Meiboom-Gill; HSQC/HMQC, heteronuclear single/multiple quantum coherence; RMSD, root-mean-square deviation; SH3, Src homology 3; TPPI, time-proportional phase incrementation; NOESY, nuclear Overhauser enhancement spectroscopy; TOCSY, total correlated spectroscopy.

E-mail address of the corresponding author: kay@pound.med.utoronto.ca

Introduction

An understanding of the process(es) by which a protein folds into its native conformation requires knowledge of the various states that populate its folding pathway(s), their relative stabilities, the kinetics of interconversion between them, as well as their structural properties. In practice such information can be difficult to obtain. Many, if not all, intermediates are populated at very low levels and are transient so that state-of-the-art biophysical techniques that provide detailed information about folded proteins, for example, become less useful. As a result information about these "elusive" excited states is often obtained through indirect studies that invoke changes in variables such as temperature, pressure, denaturant concentration or perturbations in the primary amino acid sequence.[1] One group of proteins whose folding properties have been studied in detail *via* the approaches summarized above is the SH3 domain family.[2–11] SH3 domains are small modules comprised of approximately 60 residues that fold into a five-strand β-sandwich structure[12] (Figure 1). As in the case of many other small protein moieties, folding studies of SH3 domains involving calorimetric, equilibrium and kinetic folding/unfolding experiments are consistent with a two-state folding transition, with little evidence for the formation of partially folded intermediates (reviewed by Capaldi & Radford).[13]

Traditionally, many of the techniques that have been used to study protein folding are based on the use of a single probe or report a single average property that involves many or all residues in the molecule. It is thus difficult to obtain site-specific information over a wide range of sites that can be necessary to detect intermediates in the first place or that is necessary to obtain a complete description of

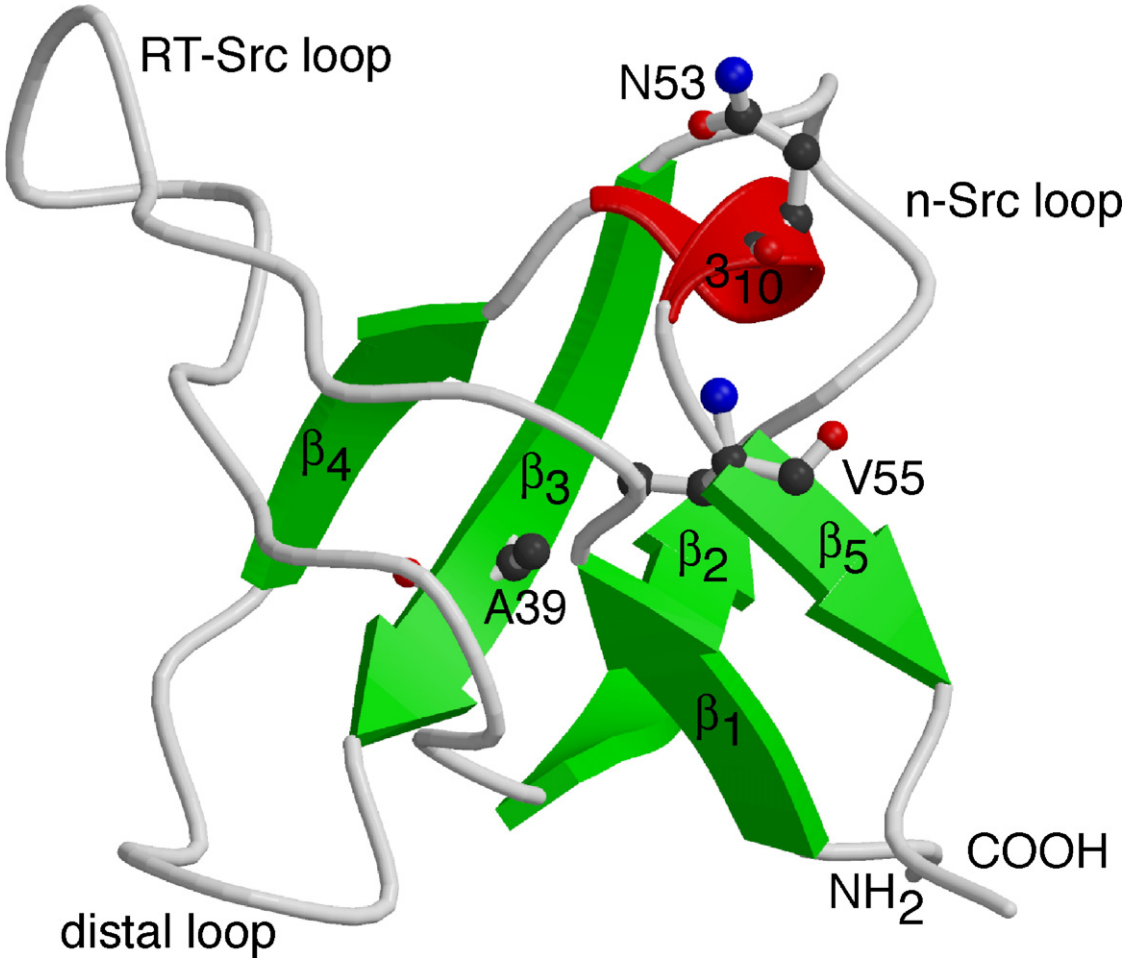

Figure 1. Schematic representation of the secondary structure of the major conformation (SH3-1) of the wild-type *Homo sapiens* Fyn SH3 domain,[12] featuring the characteristic SH3 domain β-sandwich fold formed by the terminal (strands β_1 from Leu3 to Ala6 and β_5 from Val55 to Pro57) and the approximately orthogonal central (strands β_2 from Asp25 to Asn30, β_3 from Trp36 to Ser41, and β_4 from Thr47 to Ile50) β-sheets, along with a helical turn with 3_{10} geometry from Pro51 to Tyr54. The second conformation (SH3-2), which differs in the structure of the n-Src loop region (Leu29 to Asp35), was not found to be significantly populated in solution.[14] The only sequence differences between the WT *H. sapiens* SH3 domain and the WT *Gallus gallus* Fyn SH3 domains studied here are a valine to glutamate (*G. gallus*) substitution at position 5. The residues Ala39, Asn53, and Val55 mutated in this study are shown in ball-and-stick representation. The Figure was drawn with MolScript 2.1.2[51] and rendered with Raster3D 2.7.[52]

such intermediates when they do exist. In principle, Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion NMR spectroscopy provides a complementary approach to other biophysical probes of folding because the technique is extremely sensitive to processes that involve interconversion between ground and excited states of molecules, as long as the excited states are populated at levels of 0.5% or higher and the exchange kinetics are on the millisecond time-scale, because site-specific information is available over a wide range of sites simultaneously. In general, the analysis of dispersion profiles leads to the extraction of the rates of interconversion between ground and excited states (kinetics), their populations (thermodynamics), and structural information in the form of chemical shift differences between the exchanging states. We have recently reported a number of studies of fast-folding mutants of the SH3 domain from the Fyn tyrosine kinase from *Gallus gallus* (G48M and G48V substitutions) using CPMG relaxation dispersion experiments.[7–11] Notably, on-pathway folding intermediates were observed for both mutants with fractional populations on the order of 1–2% in the absence of denaturant. One-dimensional energy landscapes of the folding reaction could be constructed for both the G48M and G48V domains and ensembles of structures that model the intermediate state determined. These studies established that NMR relaxation dispersion measurements could be used to probe intermediates that were otherwise "invisible" to other techniques. However, a number of important questions emerge relating to the generality of the results. First, are the observed intermediates in the G48M and G48V Fyn SH3 domain folding pathways a consequence of mutation of position 48 or a general property of the Fyn SH3 folding pathway and second, are the structural features of SH3 domain folding intermediates similar in general?

Here we address these questions by studying the folding of a pair of additional mutants from the Fyn SH3 domain, N53P/V55L and A39V/N53P/V55L, but where glycine at position 48 is retained. We show that the ^{15}N relaxation dispersion data recorded on both mutants are consistent with a folding model involving an on-pathway intermediate, I, and present a detailed analysis of the kinetics and thermodynamics of folding for both of these mutants as well as an analysis of the structures of the I states based on chemical shift information that is obtained directly from the dispersion data. Many of the structural features of the I state previously observed for the G48M and G48V mutants are maintained in the mutants studied here. However, the N and C-terminal β-strands of the domain that are largely unstructured in the I states of G48M/V Fyn SH3 form non-native interactions in the folding intermediate states of both N53P/V55L and A39V/N53P/V55L mutants. The results of the present work suggest that the observed Fyn SH3 domain folding intermediates observed previously are not "artifacts" of Gly48 mutations and that similarly structured intermediates may well be common to folding of SH3 domains in general.

Results and Discussion

Sequence-specific NMR resonance assignments

[^{1}H, ^{15}N] HSQC spectra of both WT Fyn SH3[14] and G48M Fyn SH3[7] modules have been assigned previously. Unfortunately, correlation spectra of N53P/V55L and A39V/N53P/V55L Fyn SH3 studied here are not sufficiently similar to those assigned previously so that sequence-specific resonance assignments could not be readily transferred. As a result ^{15}N-edited total correlated spectroscopy (TOCSY) and nuclear Overhauser enhancement spectroscopy (NOESY) experiments were recorded, leading to the unambiguous assignment of all ^{1}H and ^{15}N backbone amide resonances from Leu3 (N53P/V55L Fyn SH3, black) or Thr2 (A39V/N53P/V55L Fyn SH3, red) to Asp59 (Figure 2), with the exception of Ser31 in the n-Src loop for which correlations were not observed. It is noteworthy that the amide resonance of Ser31 was found to be extensively broadened in the WT Fyn SH3 domain that had been investigated previously.[14] In addition, the $^{15}N^{\varepsilon 1}$-$^{1}H^{\varepsilon 1}$ indole resonances of the two tryptophan residues in the protein (Trp36 and Trp37), as well as the side-chain NH_2 groups of Gln27 and Asn30 could also be assigned unambiguously. Although the backbone amide NOE patterns verify that the β-sandwich fold of the WT Fyn SH3 domain[12,14] is conserved in both mutants investigated here, Figure 2 shows that introduction of the mutation A39V into the N53P/V55L Fyn SH3 sequence causes extensive backbone amide chemical shift changes, presumably *via* subtle repacking of the hydrophobic core which is unusually rich in aromatic residues, so that resonance positions are strongly influenced by ring-current effects.[15]

Stopped-flow folding kinetics

The folding kinetics of A39V/N53P/V55L Fyn SH3 were initially studied by stopped-flow guanidinium hydrochloride denaturation monitored by tryptophan fluorescence experiments performed at room temperature. At guanidinium hydrochloride concentrations below 1.0 M the folding rate is much faster than 100/s and cannot be quantified reliably because the tryptophan fluorescence relaxes within the dead-time of the stopped-flow device (≈12 ms). At denaturant concentrations c_{denat} between 1.0 M and 6.8 M the fluorescence relaxation detected after the dead-time is effectively monoexponential. The obtained relaxation rates $k_{ex}(c_{denat})$ are well fit by a two-state folding model, F↔U, according to the chevron equation (9), yielding folding and unfolding rates extrapolated to zero denaturant concentration of $k_f = 791(\pm 102)/s$ and $k_u = 1.46(\pm 0.19)/s$, respectively (Figure 3). These values of k_f and k_u

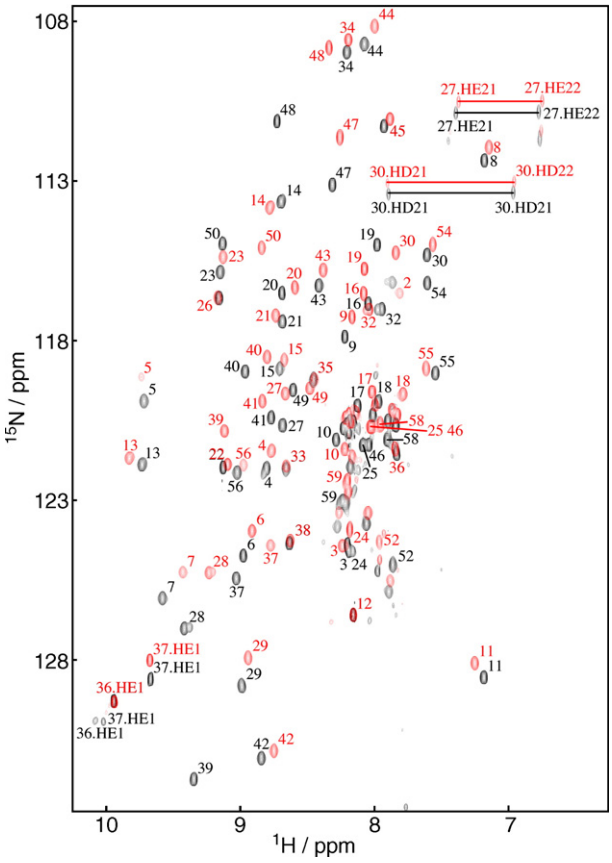

Figure 2. Overlay of [^{1}H, ^{15}N] HSQC spectra of uniformly ^{15}N-labeled N53P/V55L (black) and A39V/N53P/V55L (red) Fyn SH3 domains at 40 °C. Amide proton resonances are labeled according to their residue numbers, starting from Thr2 (Thr84 of full-length *G. gallus* Fyn). The constructs used contain additional residues N and C-terminal to the SH3 domain itself, which give rise to several unassigned resonances in the random-coil region. The n–Src loop region of N53P/V55L Fyn SH3 and especially the indole resonances of Trp36 and Trp37 show a minor conformation populated to less than 20%, which exchanges with the major conformation on a time-scale too slow for observation in the ^{15}N relaxation dispersion experiments.

predict an equilibrium population of the unfolded state of 0.18% at 23 °C.

CPMG relaxation dispersion studies of folding kinetics

N53P/V55L Fyn SH3

CPMG relaxation dispersion experiments are particularly sensitive reporters of exchange between multiple conformers, with optimal sensitivity to processes with exchange rates on the order of 1000/s and populations of excited states in excess of 0.5%. The folding of N53P/V55L Fyn SH3 is slower than optimal for this class of experiment (≈100/s at 45 °C). Thus, for the vast majority of resonances sizable relaxation dispersion profiles are only observed at elevated temperatures (i.e. 40 °C or higher, Figure 4(a)) where the unfolded state is relatively highly populated. At 45 °C exchange contributions, R_{ex}, of at least 2.5/s to ^{15}N relaxation rates were obtained for 37 backbone amide resonances along with the indole resonance of Trp37 and the residues giving rise to these correlations were selected for further analysis. Relaxation dispersion profiles recorded for these residues were initially fit to a two-site exchange model, with best fit global exchange parameters, $k_{ex,FE} = k_{FE} + k_{EF} = 107(\pm 14)$/s, and an excited state population, $p_E = 4.0(\pm 0.5)\%$ (Table 1). The strong correlation of chemical shift differences between excited and folded states, $|\Delta\varpi_{F\rightarrow E}| = |\varpi_E - \varpi_F|$, generated for each of the fitted residues and the ^{15}N chemical shift differences be-

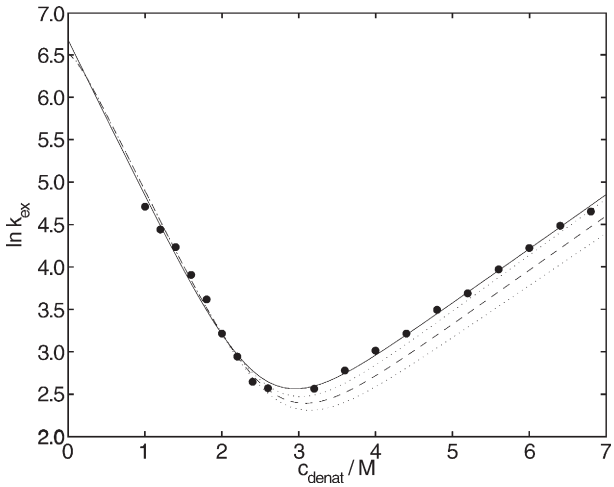

Figure 3. Chevron plot of the stopped-flow folding kinetics of A39V/N53P/V55L Fyn SH3 as a function of guanidinium hydrochloride concentration c_{denat} (filled circles). Values of $k_f = 791(\pm 102)$/s, $m_f = -1.85(\pm 0.08)$/M, $k_u = 1.46(\pm 0.19)$/s, and $m_u = 0.639(\pm 0.024)$/M are determined on the basis of a least-squares fitting of a model of two-site folding given by equation (9) to the data (continuous line). The broken line shows the calculated chevron plot, ln $\lambda_-(c_{denat})$ *versus* c_{denat} (equation (5)), based on kinetic parameters obtained from fits of the ^{15}N relaxation dispersion data (rates at 23 °C of $k_{FI} = 18.4(\pm 0.1)$/s, $k_{IF} = 822(\pm 5)$/s, $k_{IU} = 55.1(\pm 2.4)$/s, $k_{UI} = 901(\pm 37)$/s), stopped-flow *m*-values $m_{U\rightarrow IU} = -1.85(\pm 0.08)$/M and $m_{I\rightarrow IU} = 0.639(\pm 0.024)$/M, and the assumption that the *m*-values of the I and F states are identical, as are the *m*-values of both transition states. The vertical shift of the unfolding arm caused by the ≈20% discrepancy between the predicted unfolding rate of 1.21(±0.05)/s and the unfolding rate extrapolated to zero denaturant $k_u = 1.46(\pm 0.19)$/s is not significant (that these two rates are different cannot be proved to greater than 80% confidence). The dotted lines correspond to chevron plots calculated with $k_{IU} = 57.5$/s and $m_{I\rightarrow IU} = 0.663$/M (each one standard error higher than the best fit values, upper dotted line) and $k_{IU} = 52.7$/s, $m_{I\rightarrow IU} = 0.615$/M (one standard error lower, lower dotted line). The folding rate below denaturant concentrations of 1.0 M is too fast to be quantified given the stopped-flow dead-time of 12 ms, precluding observation of the slight "rollover" caused by the transient population of the intermediate predicted by the CPMG kinetics.

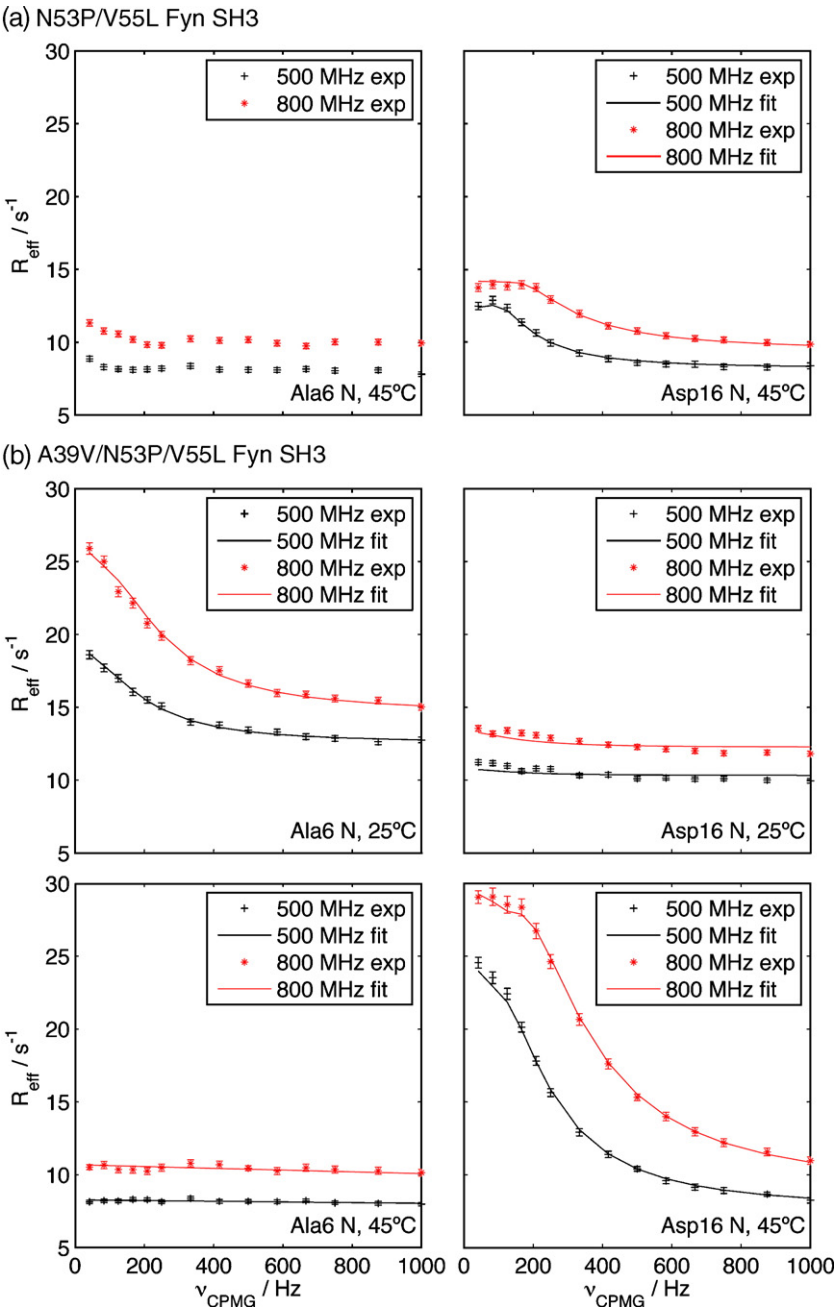

Figure 4. ^{15}N single-quantum relaxation rates R_{eff} as a function of CPMG frequency ν_{CPMG} for Ala6 and Asp16 of (a) N53P/V55L Fyn SH3 at 45 °C and (b) A39V/N53P/V55L Fyn SH3 at 25 °C (top) and 45 °C (bottom). Continuous lines are the relaxation dispersion profiles predicted using parameters generated from fitting the three-site folding model described in the text to ^{15}N relaxation dispersion data at all temperatures together (data for Ala6 N, N53P/V55L Fyn SH3 were not used in the fits).

tween random coil (RC) values[16] and F, $|\varpi_{RC} - \varpi_F|$, (RMSD 1.21 ppm) verifies that the excited state corresponds to an unfolded protein (data not shown). However, the relatively high reduced χ^2 value obtained for the global data fit (2.07) indicates that the observed exchange process is likely more complex than two-state. We considered, therefore, a folding scheme of the form, F↔I↔U, corresponding to an on-pathway folding intermediate. A model with exchange rates $k_{ex,FI}$ = 1915(±208)/s and $k_{ex,IU}$ = 820(±44)/s, along with excited state populations p_I = 0.67(±0.04)% and p_U = 5.5(±0.6)% reproduces the experimental relaxation dispersion data significantly better (reduced χ^2 of 1.00). The three-site model can also be used to fit dispersion profiles recorded at 40 °C, 45 °C, and 50 °C simultaneously, assuming that the chemical shift differences $\Delta\varpi_{F\rightarrow I}$ and $\Delta\varpi_{F\rightarrow U}$ are constant in the temperature range considered.[17] The assumption of temperature independent $\Delta\varpi$ values leads to improved convergence in fits and improves the accuracy of extracted chemical shift values.[18] The data were well fit by a model that makes no assumptions about the temperature dependence of the exchange rates and populations (reduced χ^2 of 1.12); however, the assumption that the kinetics of interconversion between states obeys transition-state theory, as implemented in a previous study of the temperature dependence of the folding of an SH3 domain,[7] appears to be valid in this temperature range as well (reduced χ^2 of 1.23; Table 2). Dispersion profiles recorded at 55 °C, however, could not be well fit, indicating that additional states might become populated when

Table 1. Exchange parameters extracted from fitting a two-site model of chemical exchange to CPMG relaxation dispersion data recorded on N53P/V55L and A39V/N53P/V55L Fyn SH3

Mutant	T (°C)	k_{FI} (s^{-1})	k_{IF} (s^{-1})	$k_{ex,FI}$ (s^{-1})	k_{FU} (s^{-1})	k_{UF} (s^{-1})	$k_{ex,FU}$ (s^{-1})	p_I (%)	p_U (%)	red. χ^2
N53P/V55L	45.0				4.31±0.05	103±14	107±14		4.03±0.51	2.07
A39V/N53P/V55L	10.0	7.5±0.2	363±20	371±21				2.02±0.08		1.37
	15.0	9.6±0.1	351±10	360±10				2.66±0.06		0.88
	20.0	14.0±0.1	559±8	573±9				2.44±0.03		0.75
	25.0	21.9±0.2	1013±12	1034±12				2.11±0.02		0.95
	30.0	31.6±0.6	1593±25	1625±26				1.95±0.03		1.96
	35.0				17.5±0.2	1456±31	1474±31		1.19±0.02	5.70
	40.0				14.3±0.2	921±18	935±18		1.53±0.02	4.16
	45.0				24.7±0.2	776±9	801±9		3.09±0.02	2.06
	50.0				45.6±0.3	653±6	699±6		6.53±0.04	1.57

Data at each temperature are fitted individually, assuming exchange models of the form, F↔I or F↔U. Note that fits at many of the temperatures have high reduced (red.) χ^2 values indicating that the two-site model is a poor approximation; parameters extracted in these cases must be interpreted with caution.

approaching thermal denaturation. In addition to the folding model described above, we also considered linear three-site exchange schemes with an off-pathway intermediate, I↔F↔U. However, the data were not as well reproduced (reduced χ^2 of 1.19 or 1.35 assuming transition-state theory, as compared to 1.12 or 1.23, respectively, see above) as with the on-pathway model. Finally, solutions to the F↔U↔I model produced unrealistic ^{15}N chemical shifts (i.e. differences that exceed the ^{15}N chemical shift range in proteins). These results provide strong evidence that folding of N53P/V55L Fyn SH3 proceeds predominantly *via* an on-pathway intermediate under native folding conditions.

A39V/N53P/V55L Fyn SH3

As described above, the folding of N53P/V55L Fyn is slower than optimal for studies by relaxation dispersion NMR spectroscopy. Since the mutation A39V is known to speed up folding by almost an order of magnitude,[6] a triple mutant, A39V/N53P/V55L, was constructed as well in the hope that larger dispersion profiles could be obtained for the extraction of more accurate exchange parameters. Although the unfolded state of this mutant is populated to only 0.18% at 23 °C according to stopped-flow kinetics (see above), surprisingly large relaxation dispersion profiles were observed over a temperature range from 5 °C to 50 °C (Figure 4(b)). Interestingly, some of the dispersion profiles that are large at 25 °C become small at 45 °C, as can be seen in Figure 4(b) for Ala6, while the opposite occurs for profiles from other residues such as Asp16, also shown.

Initially dispersion data sets were analyzed at each temperature separately using a two-state model of exchange; however, the data recorded at 5 °C were excluded from analysis due to rather poor spectral quality (broad lines causing resonance overlap and low signal/noise ratios). Reduced χ^2 values obtained from two-site fits of dispersion profiles at each of the temperatures separately suggests that such a model is appropriate for temperatures of 25 °C and below (Table 1). Notably, $\Delta\varpi_{F\rightarrow E}$ values extracted independently at each of these temperatures are in good agreement (for

Table 2. Exchange parameters extracted from fitting a linear folding model with an on-pathway intermediate F↔I↔U to CPMG relaxation data at multiple temperatures

Mutant	T (°C)	k_{FI} (s^{-1})	k_{IF} (s^{-1})	$k_{ex,FI}$ (s^{-1})	k_{IU} (s^{-1})	k_{UI} (s^{-1})	$k_{ex,IU}$ (s^{-1})	p_I (%)	p_U (%)	red. χ^2
N53P/V55L	40.0	13.0±0.7	2123±126	2136±127	313±13	96±14	409±19	0.60±0.01	1.96±0.27	1.23
	45.0	15.9±0.6	2438±126	2454±127	832±24	150±13	981±30	0.63±0.01	3.48±0.27	
	50.0	19.4±0.8	2789±158	2809±159	2142±103	230±17	2372±114	0.65±0.02	6.02±0.33	
A39V/N53P/V55L	10.0	7.1±0.1	258±11	265±11	n.d.	n.d.	n.d.	2.67±0.09	n.d.	1.02
	15.0	9.9±0.1	392±9	402±9	n.d.	n.d.	n.d.	2.47±0.04	n.d.	
	20.0	14.5±0.1	615±9	630±9	n.d.	n.d.	n.d.	2.31±0.02	n.d.	
	25.0	21.8±0.2	994±10	1016±11	n.d.	n.d.	n.d.	2.15±0.01	n.d.	
	30.0	34.3±0.5	1765±22	1799±22	161±4	839±28	1001±34	1.90±0.01	0.36±0.01	
	35.0	51.0±1.6	2951±55	3002±56	389±9	924±15	1313±24	1.69±0.02	0.71±0.01	
	40.0	98.9±4.7	5352±122	5451±127	739±26	909±10	1648±36	1.79±0.05	1.45±0.01	
	45.0	214.9±23.9	10866±422	11081±443	1348±126	845±31	2193±141	1.88±0.15	3.00±0.01	

All temperature dependent data are fit simultaneously. The temperature dependence of the folding kinetics of N53P/V55L Fyn SH3 was assumed to follow simple transition-state theory without heat capacity differences between the states; dispersion data were fit directly with these assumptions. The assumption of $\Delta C_P(X\rightarrow Y)=0$ no longer holds for A39V/N53P/V55L Fyn SH3, since a wide temperature range was probed. In fits of dispersion data for this mutant no assumptions were made regarding the temperature dependence of the folding/unfolding rates. Exchange parameters involving the unfolded state of A39V/N53P/V55L Fyn SH3 could not be determined (n.d.) below 25 °C because this state is not sufficiently populated to contribute to relaxation dispersion profiles.

example, RMSD between values at 15 °C and 25 °C of 0.26 ppm), suggesting that the assumption of invariance of $\Delta\varpi$ over the full temperature range for both N53P/V55L and A39V/N53P/V55L Fyn SH3 is reasonable. The two-site model that assumes temperature invariance of ^{15}N chemical shift differences $\Delta\varpi_{F\rightarrow E}$ can be used to simultaneously fit CPMG dispersion data recorded at 10 °C, 15 °C, 20 °C, and 25 °C derived from the 37 backbone amide resonances and the indole resonance of Trp37 that give rise to dispersion profiles of at least 2.5/s either without (reduced χ^2 of 1.08) or with the assumption that the exchange rates obey transition-state theory (reduced χ^2 of 1.19). By contrast, the two-site model fails to reproduce dispersion data recorded at higher temperatures with reduced χ^2 values as large as 5.7 obtained in fits of profiles recorded at 35 °C, suggesting that a more complex exchange process is occurring at higher temperatures. Indeed, at 35 °C folding rates extracted on a per-residue basis using a model that assumes a two-state folding reaction, differ by as much as an order of magnitude, while unfolding rates differ by even more.[11] These results are also inconsistent with expectations based on a two-site exchange model according to which exchange rates are invariant with residue.

Insights into the nature of the exchange processes can be obtained from a preliminary analysis of the ^{15}N chemical shifts obtained from the two-state fits of the data (i.e. analysis assuming only one excited state). As described above, stopped-flow studies of the folding of this domain at room temperature indicate that the unfolded state is populated to an amount (0.2%) that essentially renders this state invisible to our methodology. Indeed, a comparison between ^{15}N chemical shift differences between the excited state and the folded state ($\Delta\varpi_{F\rightarrow E}=\varpi_E-\varpi_F$) measured at 25 °C shows a poor correlation with $\varpi_{RC}-\varpi_F$, where ϖ_{RC} are tabulated random coil chemical shifts[16] (Figure 5(a)). In addition, a poor correlation is noted between ^{15}N excited state chemical shifts extracted from dispersion profiles at 25 and 45 °C (inset to Figure 5(a)). These comparisons strongly suggest that the excited state at lower temperature does not correspond to an unfolded polypeptide chain. In contrast, a much better correlation is noted between the chemical shifts of the excited state at 45 °C and those from a random coil (Figure 5(b)), consistent with the dominant excited state at higher temperature being largely unfolded. As we describe below the dispersion data for A39V/N53P/V55L Fyn SH3 is well fit by a model of the form F↔I↔U, with the excited state predominating at low temperature corresponding to a largely structured intermediate, I, while the major excited state at high temperature is the unfolded state, U. The temperature dependence of the dispersion profiles illustrated in Figure 4(b) can now be understood. At 25 °C dispersion profiles reflect largely the F↔I equilibrium, since the population of the I excited state (p_I) is much higher than that of the U excited state (p_U), while at 45 °C

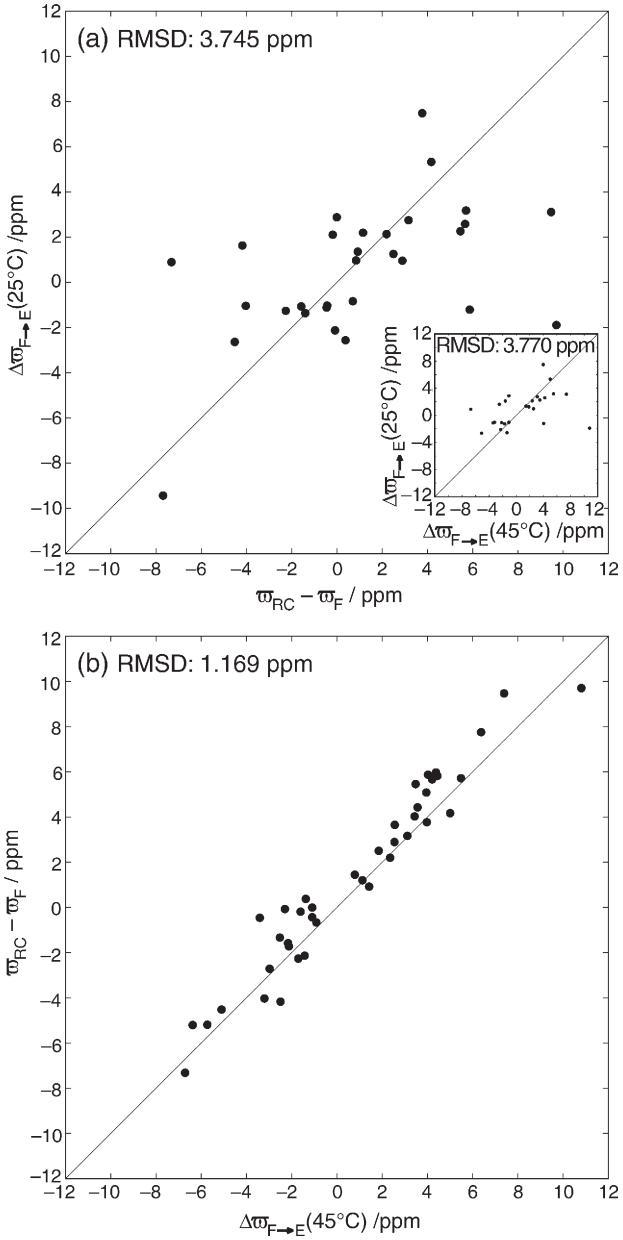

Figure 5. (a) Comparison of chemical shift differences between excited and folded states of A39V/N53P/V55L Fyn SH3 extracted from fits of CPMG data at 25 °C using a two-site model of exchange F↔E, $\Delta\varpi_{F\rightarrow E}=\varpi_E-\varpi_F$, with differences between random coil, ϖ_{RC},[16] and folded, ϖ_F, ^{15}N shift values. A correlation between $\Delta\varpi_{F\rightarrow E}$(25 °C) and $\Delta\varpi_{F\rightarrow E}$(45 °C) is provided in the inset. The signs of the chemical shift differences are obtained as described in the text. These plots establish that the exited state at 25 °C is not a random coil ensemble. (b) The excellent correlation of ^{15}N chemical shifts extracted for the excited state at 45 °C with random coil values ϖ_{RC}[16] identifies this state as unfolded.

the opposite prevails. Since for a given residue the size of the dispersion curve depends on the chemical shift difference between ground and excited conformations and since $|\Delta\varpi_{F\rightarrow I}| \gg |\Delta\varpi_{F\rightarrow U}|$ for Ala6 (see Table 3) ^{15}N profiles for this residue

Table 3. ^{15}N chemical shifts of the folded (F), intermediate (I), and unfolded (U) states of N53P/V55L and A39V/N53P/V55L Fyn SH3 extracted from fits using the three-site model of exchange F↔I↔U to ^{15}N relaxation dispersion data sets (exchange parameters listed in Table 2)

Mutant		N53P/V55L						A39V/N53P/V55L						
Residue	Atom	ϖ_F (ppm)	$\Delta\varpi_{F\to I}$ (ppm)	ϖ_I (ppm)	$\Delta\varpi_{F\to U}$ (ppm)	ϖ_U (ppm)	ϖ_{RC} (ppm)	ϖ_F (ppm)	±	$\Delta\varpi_{F\to I}$ (ppm)	ϖ_I (ppm)	$\Delta\varpi_{F\to U}$ (ppm)	ϖ_U (ppm)	ϖ_{RC} (ppm)
T2	N							116.90	−	**−1.39**	**115.51**	−0.78	116.12	115.50
L3	N							124.63	−	1.08	125.71	−1.19	123.44	124.20
F4	N	121.69	−2.47	119.22	−1.87	119.82	121.30	121.49	+	2.12	123.61	−1.77	119.72	121.30
E5	N	119.58	6.43	126.01	1.86	121.44	122.60	118.83	+	7.82	126.65	2.57	121.40	122.60
A6	N							123.85	**+**	**2.21**	**126.06**	−0.16	123.69	125.00
L7	N	125.76	−2.56	123.20	−5.55	120.21	121.00	125.52	−	−2.56	122.96	−5.03	120.49	121.00
Y8	N	112.05	2.49	114.54	7.47	119.52	121.30	111.83	+	3.03	114.86	7.67	119.50	121.30
D9	N	117.57	1.26	118.83	4.09	121.66	123.20	117.33		−1.18	116.15	4.40	121.73	123.20
Y10	N							121.11	+	2.78	123.89	−1.21	119.90	121.10
E11	N	128.23	−0.89	127.34	−6.77	121.46	123.00	128.21		−0.41	127.80	−6.67	121.54	123.00
A12	N	126.28	−0.94	125.34	−1.87	124.41	125.00	126.72		0.45	127.17	−2.20	124.52	125.00
R13	N	121.58	−1.04	120.54	−1.27	120.31	119.70	121.84		0.43	122.27	−1.48	120.36	119.70
T14	N	113.33	1.69	115.02	1.62	114.95	115.00	114.08		1.34	115.42	1.32	115.40	115.00
E15	N	118.57	1.83	120.40	3.49	122.06	122.60	118.57		0.53	119.10	3.67	122.24	122.60
D16	N	116.53	1.59	118.12	4.12	120.65	121.60	116.52		0.50	117.02	4.04	120.56	121.60
D17	N	119.73	0.75	120.48	0.81	120.54	121.20	119.75		−0.41	119.34	0.83	120.58	121.20
L18	N	119.60	1.88	121.48	2.74	122.34	122.60	119.71	+	0.84	120.55	2.58	122.29	122.60
S19	N							116.00		−0.86	115.14	−0.46	115.54	116.70
F20	N	116.20	−1.36	114.84	4.61	120.81	122.20	116.23		−0.64	115.59	4.76	120.99	122.20
H21	N	117.09	−1.55	115.54	2.31	119.40	120.60	116.95		0.55	117.50	2.67	119.62	120.60
G23	N	115.54	1.39	116.93	−7.09	108.45	110.40	115.59		−0.54	115.05	−6.04	109.55	110.40
E24	N							124.23		−0.90	123.33	−3.23	121.00	120.20
F26	N	116.36	1.54	117.90	3.65	120.01	121.90	116.44	+	2.20	118.64	3.36	119.80	121.90
Q27	N	120.34	−0.92	119.42	1.11	121.45	122.20	119.70	+	1.20	120.90	1.77	121.47	122.20
I28	N	126.70	−0.40	126.30	−5.14	121.56	121.20	125.38	+	1.63	127.01	−2.66	122.72	121.20
L29	N	128.50	−0.70	127.80	−2.94	125.56	126.00	128.27	−	1.27	129.54	−1.81	126.46	126.00
N30	N	115.00	−1.33	113.67	4.21	119.21	119.70	115.28		−0.57	114.71	3.84	119.12	119.70
S32	N							117.21		0.52	117.73	0.54	117.75	117.60
E33	N							122.19		0.55	122.74	−0.28	121.91	122.10
D35	N	118.91	1.56	120.47	1.79	120.70	120.40	119.20		−0.36	118.84	1.18	120.38	120.40
W36	N							121.72	−	−2.63	119.09	−1.18	120.54	122.10
W37	N	125.13	−1.22	123.91	−4.23	120.90	124.10	124.56	−	1.09	125.65	−3.75	120.81	124.10
W37	NE1	128.29	0.70	128.99	1.52	129.81	129.60	128.22		−0.70	127.52	1.40	129.62	129.60
E38	N	124.04	−0.88	123.16	−2.73	121.31	123.00	124.34		0.49	124.83	−2.51	121.83	123.00
A/V39	N	131.42	−1.25	130.17	−8.64	122.78	120.40	121.06		0.37	121.43	−0.95	120.11	120.40
R40	N							118.58		−0.43	118.15	4.78	123.36	124.40
S41	N	120.10	1.55	121.65	−4.46	115.64	117.10	119.82		0.66	120.48	−3.17	116.65	117.10
L42	N	130.77	1.11	131.88	−8.45	122.32	123.70	131.02		0.83	131.85	−6.63	124.39	123.70
T43	N	115.97	1.25	117.22	−1.93	114.04	114.60	116.18	−	−0.99	115.19	−2.17	114.01	114.60
T44	N	108.40	−1.41	106.99	7.01	115.41	116.00	108.25		−0.63	107.62	6.18	114.43	116.00
G45	N							111.23		0.51	111.74	−0.41	110.82	111.20
T47	N	112.81	−1.53	111.28	1.32	114.13	114.80	111.63	+	2.73	114.36	2.94	114.57	114.80
G48	N							109.01	+	2.11	111.12	2.21	111.22	111.20
Y49	N	119.23	1.10	120.33	1.02	120.25	120.30	119.46	+	−1.00	118.46	0.59	120.05	120.30
I50	N	114.65	1.01	115.66	10.11	124.76	124.50	114.80		−1.95	112.85	9.99	124.79	124.50
S52	N	124.71	−8.37	116.34	−8.26	116.45	117.00	124.69	−	−9.14	115.55	*−7.56±5.24*	*117.13±5.24*	117.00
Y54	N	115.89	1.77	117.66	3.26	119.15	120.70	115.03	+	2.51	117.54	4.01	119.04	120.70
L55	N	118.71	3.20	121.91	5.75	124.46	124.60	118.89	+	3.11	122.00	5.38	124.27	124.60
A56	N	121.82	4.86	126.68	4.46	126.28	126.00	121.83	+	5.41	127.24	4.39	126.22	126.00
V58	N	120.79	−2.61	118.18	−2.25	118.54	119.60	120.98	−	−2.12	118.86	−2.14	118.84	120.90

Native state chemical shifts, ϖ_F, are reported at 40 °C and 25 °C for N53P/V55L and A39V/N53P/V55L, respectively, while $\Delta\varpi_{F\to I} = \varpi_I - \varpi_F$ and $\Delta\varpi_{F\to U} = \varpi_U - \varpi_F$ were assumed to be temperature-invariant. For N53P/V55L Fyn SH3, the unknown absolute signs of $\Delta\varpi_{F\to U}$ values were inferred from the random coil chemical shifts ϖ_{RC}[16] that fix the signs of $\Delta\varpi_{F\to I}$ *via* the relative sign information obtained from the fit.[7,8] The absolute signs of $\Delta\varpi_{F\to U}$ for A39V/N53P/V55L Fyn SH3 were determined experimentally for almost all resonances using the method of Skrynnikov *et al.* [19] from which the signs of $\Delta\varpi_{F\to I}$ could be inferred. In the case of Thr2 and Ala6 (boldface) the experimentally determined absolute signs of $\Delta\varpi_{F\to I}$ were used to obtain the signs of $\Delta\varpi_{F\to U}$. None of the absolute signs could be determined experimentally for Ser32, Glu33, and Thr44, so the sign of $\Delta\varpi_{F\to U}$ was inferred from the random coil chemical shift value, ϖ_{RC}. The experimentally determined signs of $\Delta\varpi_{F\to I}$ (for A39V/N53P/V55L Fyn SH3) are reported in the column preceding $\Delta\varpi_{F\to I}$ (and those for $\Delta\varpi_{F\to U}$ are listed in the Table). Monte-Carlo analysis on a similar system[18] suggests typical error margins on the order of 1.0 ppm for $\Delta\varpi_{F\to I}$ and hence ϖ_I for N53P/V55L Fyn SH3 and 0.1 ppm for all other $\Delta\varpi_{F\to I}$, ϖ_I, $\Delta\varpi_{F\to U}$, and ϖ_U values with the exception of $\Delta\varpi_{F\to U}$ and ϖ_U of Ser52 of A39V/N53P/V55L Fyn SH3 (italics), where resonance overlap precluded extraction of relaxation dispersion profiles at 35 °C and above. Errors for the shift values of A39V/N53P/V55L Fyn SH3 are significantly smaller, since $\Delta\varpi_{F\to I}$ values could be extracted at 25 °C, where the exchange is two-state. The low RMSD values between ϖ_U and ϖ_{RC}, 1.20 ppm and 1.08 ppm for N53P/V55L and A39V/N53P/V55L Fyn SH3, respectively, establish that the U state is indeed random-coil like.

are larger when $p_I > p_U$ (25 °C) than when $p_I < p_U$ (45 °C). In the case of Asp16 $|\Delta\varpi_{F\rightarrow I}| \ll |\Delta\varpi_{F\rightarrow U}|$ so that a larger profile is observed when $p_I < p_U$.

Fitting three-site exchange: an application to the folding of A39V/N53P/V55L Fyn SH3

Extracting kinetic and thermodynamic parameters along with ^{15}N chemical shifts from relaxation dispersion data for a system exchanging between three states involves a highly complex non-linear global optimization and some effort has been devoted to developing a suitable minimization protocol.[7,8] Here, we wish to describe the approach used in the data analysis of A39V/N53P/V55L Fyn SH3 because this system is distinct from other SH3 modules that have been investigated previously[7–10] where the data (acquired over fairly limited temperature ranges in these cases) could only be fit using a three-site exchange model. By contrast, in the present case the steep temperature dependence of the populations of the excited states and the wide range of temperatures that were obtained can be used to help "dissect" the exchange process, in ways described below.

As indicated in Table 1, dispersion data for A39V/N53P/V55L Fyn SH3 could be well fit to the two-site exchange model, F↔I, at temperatures ≤25 °C. The values of $\Delta\varpi_{F\rightarrow I}$ that are extracted with high precision from such fits can be used to simplify data fitting to the three-site exchange model, F↔I↔U, that is necessary at higher temperatures (Tables 1 and 2). This ensures that robust exchange parameters can be extracted without the need to record additional types of dispersion data.[18] Using the 40 °C dispersion data exclusively a grid search was performed in the parameters $k_{ex,FI}$, $k_{ex,IU}$, p_I, and p_U as described[7,8] starting from values of $\Delta\varpi_{F\rightarrow U}$ and intrinsic relaxation rates $R_{2,0}$ extracted from a per-residue two-state fit of the data (40 °C) as well as the values of $\Delta\varpi_{F\rightarrow I}$ that are obtained from analysis of the low temperature data (≤25 °C). That is, for each grid point ($k_{ex,FI}$, $k_{ex,IU}$, p_I, p_U), values of $\Delta\varpi_{F\rightarrow U}$ and $R_{2,0}$ are minimized for every residue separately with $\Delta\varpi_{F\rightarrow I}$ fixed. In cases where dispersion curves become measurable only at temperatures of 30 °C or above $\Delta\varpi_{F\rightarrow I}$ is fixed to 0. Although the signs of $\Delta\varpi_{F\rightarrow U}$ and $\Delta\varpi_{F\rightarrow I}$ were determined experimentally, we have not used this information at this stage of analysis so that both relative sign combinations of $\Delta\varpi_{F\rightarrow I}$ and $\Delta\varpi_{F\rightarrow U}$ have been retained in starting values. Initially the 50 grid points with the lowest overall χ^2 (sum of all per-residue χ^2 values) served as starting points for 50 separate global (multiple-residue) χ^2 minimizations of the 40 °C data based on a Levenberg–Marquardt algorithm. In each of these separate fits the grid points $k_{ex,FI}$, $k_{ex,IU}$, p_I, p_U are now minimized along with $\Delta\varpi_{F\rightarrow U}$ and $R_{2,0}$, while still keeping $\Delta\varpi_{F\rightarrow I}$ fixed. Subsequently, the sign of $\Delta\varpi_{F\rightarrow I}$ was reversed one residue at a time, the minimization repeated and the new sign for $\Delta\varpi_{F\rightarrow I}$ retained in the few cases that a lower reduced χ^2 was obtained. Convergence of 11 of the 50 final solutions (that are all independently derived) to a model with $k_{ex,FI} = 5655(\pm164)/s$, $k_{ex,IU} = 1624(\pm46)/s$, $p_I = 1.84(\pm0.06)$ %, and $p_U = 1.42(\pm0.01)$ % (reduced χ^2 of 0.89) is obtained at this stage. All 50 solutions at 40 °C were subsequently used as starting points for global χ^2 minimizations that now included data recorded at temperatures of 35 °C and 45 °C. As before the sign of $\Delta\varpi_{F\rightarrow I}$ for each residue was inverted individually, followed by minimization with the solution retained in cases where χ^2 was lowered. All of the 50 solutions now converge to a single model either without (reduced χ^2 of 1.07) or with (reduced χ^2 of 1.08) the assumption of a rate *versus* temperature profile described by transition state theory. Data recorded at 50 °C was omitted from the analysis because of relatively poor fits (reduced χ^2 of 1.26 without using transition state theory), suggesting that additional states might become populated as the temperature begins to approach that of the thermal denaturation midpoint (66.2 °C; unpublished data). By contrast, dispersion profiles obtained at 30 °C were easily fit together with data recorded at 35 °C, 40 °C and 45 °C (reduced χ^2 of 1.03), even with the assumption of transition-state theory (reduced χ^2 of 1.07). In the final minimization series data from 10 °C to 25 °C were included; $k_{ex,IU}$ and p_U are ill-defined at these lower temperatures and were therefore fixed to 0, but $\Delta\varpi_{F\rightarrow I}$ values were now used as adjustable parameters, resulting in the final solution shown in Table 2 (reduced χ^2 of 1.02). The low reduced χ^2 value obtained supports the assumptions inherent in the model used, including the temperature invariance of $\Delta\varpi$ over the range considered. It is noteworthy that the values of shift differences change very little compared to differences from fits that were performed for the low (≤25 °C) and high (>25 °C) temperature profiles independently, with RMSD values of 0.29 ppm and 0.09 ppm for $\Delta\varpi_{F\rightarrow I}$ and $\Delta\varpi_{F\rightarrow U}$, respectively. Finally, we did not assume any rate *versus* temperature dependence in minimizations that included data from 10 °C–45 °C because of significant curvature in the Eyring plots (Figure 7) that reflect the non-negligible heat capacity differences between F or I and the intervening transition state (see below).

As in the case of N53P/V55L Fyn SH3, the folding model I↔F↔U does not reproduce experimental data as well as does F↔I↔U (for example, reduced χ^2 of 1.23 *versus* 0.89 at 40 °C), and in addition, the solutions for the F↔U↔I model have non-physical ^{15}N chemical shifts. Thus, there is strong evidence that folding of A39V/N53P/V55L Fyn SH3 also proceeds predominantly *via* an on-pathway intermediate under native conditions.

^{15}N chemical shifts and structural analysis of excited states

Since the dispersion data for A39V/N53P/V55L Fyn SH3 effectively derive from a two-state exchange process, F↔I, at temperatures below 25 °C it is possible to verify the robustness of the extracted chemical shift parameters prior to their interpretation in terms of structure. For example, while the relative signs of $\Delta\varpi_{F\rightarrow I}$ and $\Delta\varpi_{F\rightarrow U}$ can be obtained from three-state fits of the data,[7,8] the absolute signs can be measured directly from small exchange contributions to resonance positions in ^{1}H-^{15}N correlation maps[19] and the consistency of the two sets of data compared. Thus, the signs of $\Delta\varpi_{F\rightarrow I}=\varpi_I-\varpi_F$ and $\Delta\varpi_{F\rightarrow U}=\varpi_U-\varpi_F$ for the triple mutant were determined from spectra at 25 °C (where relaxation dispersions profiles are dominated by F↔I) and at 45 °C (where the F↔U process dominates), respectively. At 25 °C 24 resonances showed measurable chemical shift changes and the absolute signs of $\Delta\varpi_{F\rightarrow I}$ could be established in these cases. Magnitudes of $\Delta\varpi_{F\rightarrow U}$ are large for most resonances, so their signs could be readily extracted directly from experiment at 45 °C; these values correspond well to predictions based on random coil values of chemical shifts (Table 3). A match in the relative signs obtained from peak positions in correlation maps and from the three-site data fit was obtained for 18 of the 22 resonances for which the absolute signs of both $\Delta\varpi_{F\rightarrow I}$ and $\Delta\varpi_{F\rightarrow U}$ were measured and the four exceptions (Leu3, Leu29, Trp37, and Tyr49) all have relatively small values of $\Delta\varpi_{F\rightarrow I}$, $|\Delta\varpi_{F\rightarrow I}|<1.3$ ppm (Table 3). By contrast, there are only three cases where the relative signs of $\Delta\varpi_{F\rightarrow I}$ and $\Delta\varpi_{F\rightarrow U}$ values obtained from fitting the dispersion data recorded at 40 °C using the model I↔F↔U agree with those determined experimentally from peak positions in spectra, along with 15 conflicting cases, arguing against this exchange model still further.

The dependence of backbone ^{15}N chemical shifts on a variety of factors such as primary sequence effects, backbone angles of the preceding and observed residue, side-chain orientations of the preceding residue and hydrogen bonds[20] complicates their interpretation in terms of structure. Accordingly, changes in shifts of up to 2 ppm upon side-chain repacking are not uncommon even if the overall tertiary structure remains conserved (Figure 2). The RMSD between ϖ_U and ϖ_{RC} (Wishart *et al.*[16]; note that the nearest-neighbor sequence corrections of Table 7 and Table 8 in this publication were used here) of about 1.1 ppm for both mutants (Table 3) is similar to differences observed between backbone ^{15}N random-coil chemical shift values and those from urea-denatured apomyoglobin,[21] demonstrating that the U states of both mutant SH3 domains studied here are unfolded, with little or no population of any residual structure. In contrast, $|\Delta\varpi_{F\rightarrow I}|$ values are small for the majority of residues, so that for A39V/N53P/V55L Fyn SH3 for example, only the NH_2-terminal region extending from Phe4 to Tyr10 and the COOH-terminal region from Ser52 to Val58 have stretches with $|\Delta\varpi_{F\rightarrow I}|>1.5$ ppm. This suggests that the RT-Src loop followed by strands β_2 to β_4 including the distal and n-Src loops (Figure 1) essentially retain their native backbone conformation in the intermediate state and that the small changes in ^{15}N chemical shifts between F and I are likely the result of side-chain repacking and/or incomplete formation of stable hydrogen bonds, that do not occur until after the establishment of tertiary structure in the intermediate states of other proteins (reviewed by Capaldi & Radford).[13] Values of $|\Delta\varpi_{F\rightarrow I}|$ for A39V/N53P/V55L Fyn SH3 are plotted on a ribbon representation of the domain in Figure 6, showing regions of structure in the I state of this mutant, as described above.

Of particular interest are the ϖ_I values of Phe4 to Tyr10 (strand β_1 plus a short COOH-terminal extension) and Tyr54 to Ala56 (strand β_5) that deviate significantly from random coil values (see above and Table 3), exemplified by the strong downfield shift of Glu5 in the I states for both N53P/V55L and A39V/N53P/V55L Fyn SH3. Although ϖ_I values for the double mutant are not extracted with the same precision as for the triple mutant, they clearly show the same features, providing strong evidence that strands β_1 and β_5 adopt a non-native conformation in the intermediate folding state of both mutants. Similar studies involving G48M and G48V Fyn SH3 domains have suggested the presence of on-pathway folding intermediates as well, with the major differences in secondary structure between folded and intermediate states residing in the conformations of strands β_1 and β_5.[7–9] However, the pronounced downfield shift of Glu5 ϖ_I (Table 3) is clearly absent in the G48 mutants. Since the primary structure of all these mutants is identical in strand β_1, the unusual ϖ_I values observed for N53P/V55L and A39V/N53P/V55L Fyn SH3 must be caused by non-native long-range interactions that are not present in the G48M and G48V Fyn SH3 domain folding intermediates (or *vice versa*). Interestingly, stopped-flow folding kinetics experiments on a variety of Fyn SH3 variants with a mutation at sequence position 53 indicate a strong positive correlation between the hydrophobicity of the engineered side-chain and stability of the rate-limiting transition state (unpublished results), despite the fact that the side-chain of Asn53 is solvent-exposed in the native state (Figure 1). Molecular dynamics simulations based on a native-centric model supplemented by a hydrophobicity term showed that increasing the side-chain hydrophobicity at residue 53 leads to a higher probability of interaction with Leu3, Phe4, Glu5, Ala6, Phe26, Gln27, and Ile28 in the transition state ensemble, in addition to the contacts between strands β_1 and β_5 that are already present in the native state (S. Wallin &

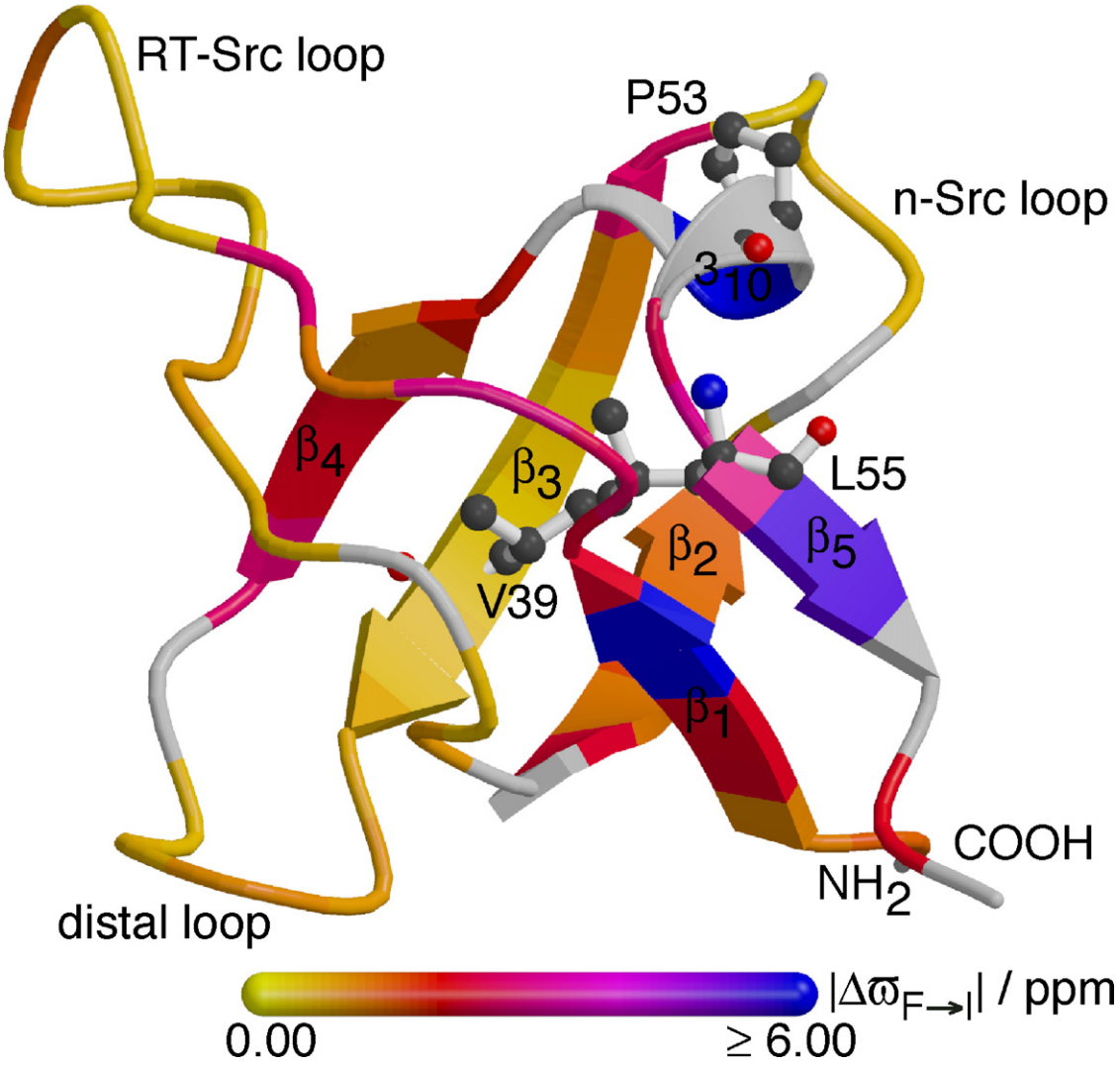

Figure 6. Schematic representation of the secondary structure of a homology model of the A39V/N53P/V55L *G. gallus* Fyn SH3 domain (same view as in Figure 1) colored according to the backbone amide ^{15}N chemical shift difference between intermediate and folded states, $|\Delta\varpi_{F\rightarrow I}|$ (Table 3), from 0.0 ppm (yellow) to 6.0 ppm (blue), or gray if $\Delta\varpi_{F\rightarrow I}$ could not be determined due to missing resonance assignments, resonance overlap, or if the exchange contributions to transverse relaxation were of insufficient magnitude. The residues Val39, Pro53, and Leu55 mutated in this study are shown in ball-and-stick representation. Values of $|\Delta\varpi_{F\rightarrow I}| > 1.5$ ppm (red to blue) are localized primarily to a region including the proximal NH_2 and COOH termini comprising strands β_1 and β_5, indicating that the rest of the domain essentially retains native-like backbone conformation in the intermediate state. The homology model was created on the basis of the major conformation (SH3-1) of the wild-type *H. sapiens* Fyn SH3 domain[12] (Figure 1) with the Swiss-PdbViewer 3.7[53] and the rotamers of the mutated side-chains were adjusted to match the observed NOE pattern if necessary using the NIH version 1.2.1[54] of X-PLOR 3.851.[55] The Figure was drawn with MolScript 2.1.2[51] and rendered with Raster3D 2.7.[52]

H. S. Chan, personal communication). These results, in total, suggest that the intermediate states of N53P/V55L and A39V/N53P/V55L Fyn SH3 are stabilized by non-native hydrophobic long-range contacts involving Leu3 to Phe10, Pro53 to Ala56, and other hydrophobic core side-chains.

Thermodynamic analysis

From the temperature dependence of the kinetic parameters that have been extracted from the CPMG dispersion experiments described above insight can be obtained into the energy landscape that governs the detected chemical exchange process. Briefly, states X, Y ∈ {F, I, U} are populated according to the Boltzmann distribution:

$$\frac{p_Y}{p_X} = e^{-\frac{\Delta G_{X\rightarrow Y}}{RT}} \quad (1)$$

where R is the gas constant, T the absolute temperature and the temperature dependence of

the free energy difference between states is given by:

$$\begin{aligned} \Delta G_{X\rightarrow Y}(T) &= G_Y(T) - G_X(T) \\ &= \Delta H_{X\rightarrow Y}(T_m) + \Delta C_p(X\rightarrow Y) \times (T - T_m) \\ &\quad - T\Delta S_{X\rightarrow Y}(T_m) - T\Delta C_p(X\rightarrow Y)\ln\frac{T}{T_m} \end{aligned} \quad (2)$$

where $\Delta H_{X\rightarrow Y}(T_m)$, $\Delta S_{X\rightarrow Y}(T_m)$, and $\Delta C_p(X\rightarrow Y)$ are the enthalpy, entropy, and heat capacity differences, respectively, between states Y and X at an arbitrary reference temperature T_m.[1] Assuming transition-state theory, the temperature dependence of the rate of transition from state X to state Y is given by the Eyring equation:[22]

$$k_{XY} = \kappa\frac{k_B T}{h}e^{-\frac{\Delta G_{X\rightarrow XY}}{RT}} \quad (3)$$

where κ is a transmission coefficient, k_B and h are the Boltzmann and Planck constants, respectively, and the temperature dependence of the free energy difference between the rate-limiting transition state XY and state X, $\Delta G_{X\rightarrow XY}(T) = G_{XY}(T) - G_X(T)$, again follows equation (2). A value of $\kappa = 1.6\times10^{-7}$ has been used in the present study, based on an empirical estimate for protein folding.[23] It is noteworthy that κ shifts the calculated entropy, S_{XY}, of all transition states by $-R\ln(\kappa) = +130$ J/mol/K, and hence the calculated free energy, G_{XY}, by $+RT\ln(\kappa) = -41.4$ kJ/mol at 45 °C, while all other extracted thermodynamic parameters, including differences between transition state entropies or free energies, are independent of κ (as long as a single value of κ is used for all rates).

Because only small dispersion profiles are observed for N53P/V55L Fyn SH3 in general, and only a narrow temperature range (40 °C to 50 °C) could be employed, the values of $\Delta H_{X\rightarrow Y}(T_m)$ and $\Delta S_{X\rightarrow Y}(T_m)$ that are extracted from a temperature dependent analysis of this double mutant are error prone. In addition, it is not possible to provide robust estimates for $\Delta C_p(X\rightarrow Y)$. Indeed, dispersion data could be well fit by a model which assumes $\Delta C_p(X\rightarrow Y) = 0$ from 40 °C –50 °C (i.e. a model which assumes a linear dependence of $\Delta G_{X\rightarrow Y}$ with temperature). Table 4 lists the thermodynamic parameters, along with their uncertainties determined from a covariance matrix analysis of the errors[24] but it is likely that the error estimates are lower bounds.

By contrast, relaxation dispersion data over a much larger temperature range could be collected and analyzed for A39V/N53P/V55L Fyn SH3, leading to more robust estimates of thermodynamic parameters for this system (Table 4). Because data obtained from 10 °C to 45 °C were analyzed together, the assumption that $\Delta C_p(X\rightarrow Y) = 0$ is no longer valid, and we have fit dispersion profiles without any assumptions regarding the temperature dependence of the kinetic parameters. Subsequently, the thermodynamic parameters were obtained in a second fitting stage from the temperature dependence of the extracted rates, k_{FI}, k_{IF}, k_{IU}, and k_{UI} (Figure 7) and the curvature exhibited by the Eyring plots for k_{FI} and k_{IF} (top row) indicates non-negligible heat capacity differences between states F, I and the transition state connecting them. The narrow temperature range over which the dispersion data report on the unfolded state (temperatures ≥30 °C; see Table 2), however, renders quantification of the heat capacity differences between I, U, and the connecting transition state very difficult (Figure 7, bottom row).

Figure 8 illustrates the one-dimensional energy level diagrams for the folding reactions of both Fyn SH3 domains studied here at 45 °C. Not surprisingly, $\Delta G_{X\rightarrow Y}$ values are determined much more robustly than either $\Delta H_{X\rightarrow Y}$ or $\Delta S_{X\rightarrow Y}$ (obtained from the temperature dependence of $\Delta G_{X\rightarrow Y}$). Thus, free energy differences can be compared with high confidence. Of note, the A39V mutation results in little change to the overall stability of N53P/V55L Fyn SH3 ($\Delta\Delta G_{U\rightarrow F} = \Delta G_{U\rightarrow F}$(A39V/N53P/V55L)$-\Delta G_{U\rightarrow F}$(N53P/V55L) = −0.3(±0.2) kJ/mol; Table 4), similar to $\Delta\Delta G_{U\rightarrow F} = -0.7(\pm0.2)$ kJ/mol between A39V and wild-type Fyn SH3 domains.[6] However, the intermediate is stabilized

Table 4. Thermodynamics of the folding of N53P/V55L and A39V/N53P/V55L Fyn SH3 at 45 °C from fits of a folding model of the form F↔I↔U to the ^{15}N relaxation dispersion data

Mutant	N53P/V55L			A39V/N53P/V55L		
State	ΔH (kJ/mol)	$-T\Delta S$ (kJ/mol)	ΔG (kJ/mol)	ΔH (kJ/mol)	$-T\Delta S$ (kJ/mol)	ΔG (kJ/mol)
F	−105±12	96±12	−8.85±0.22	−104±5	94±5	−9.14±0.01
FI	−74±13	94±13	20.57±0.24	−14±4	27±4	13.36±0.23
I	−94±12	99±12	4.50±0.22	−119±3	121±3	1.24±0.21
IU	64±12	−41±12	23.42±0.23	−2±2	21±2	18.88±0.10
U	0	0	0	0	0	0

Enthalpic ΔH, entropic $-T\Delta S$, and free energy ΔG contributions to the folding of N53P/V55L Fyn SH3 were extracted directly from fits of relaxation dispersion data assuming that the temperature dependence of the kinetics is described by transition state theory (equations (2) and (3)) with $\Delta C_P(X\rightarrow Y) = 0$. Kinetic parameters were obtained from fits of dispersion profiles for A39V/N53P/V55L Fyn SH3 with no assumption about their temperature dependence; values of ΔH, −TΔS, and ΔG were subsequently obtained from fits of the Eyring plots shown in Figure 7. Note that ΔG is calculated directly from exchange rates and therefore determined with much better precision than ΔH and $-T\Delta S$. Values of H, TS and G are referenced with respect to U, for which values of 0 are arbitrarily assigned. FI and IU denote the rate-limiting transition states connecting F, I and I, U, respectively. Values of $\Delta C_p(U\rightarrow I) = \Delta C_p(U\rightarrow F) - 0.15(\pm0.17)$ kJ/mol/K and $\Delta C_p(U\rightarrow FI) = \Delta C_p(U\rightarrow F) + 1.46(\pm0.11)$ kJ/mol/K are extracted from the temperature dependence of k_{FI} and k_{IF} for A39V/N53P/V55L Fyn SH3 (Figure 7).

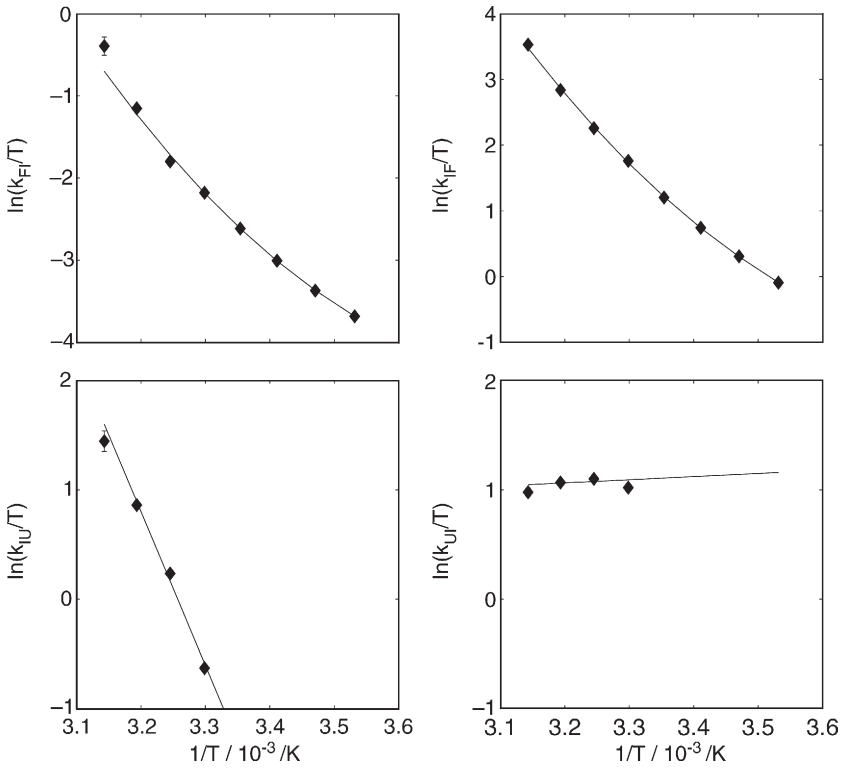

Figure 7. Eyring plots[22] showing the temperature dependence of each of the four rates extracted from fits of the three-site exchange model F↔I↔U to dispersion profiles of A39V/N53P/V55L Fyn SH3. Continuous lines are the best fits to the experimental data (diamonds) using equations (2) and (3) with a transmission coefficient of $\kappa = 1.6 \times 10^{-7}$, including non-zero ΔC_p terms for k_{FI} and k_{IF} to account for curvature.

by $\Delta\Delta G_{U\rightarrow I} = -3.3(\pm 0.3)$ kJ/mol, and the transition states even more ($\Delta\Delta G_{U\rightarrow FI} = -7.3(\pm 0.3)$ kJ/mol and $\Delta\Delta G_{U\rightarrow IU} = -4.5(\pm 0.3)$ kJ/mol; Table 4). All these values are similar in magnitude to those that have been quantified by stopped-flow tryptophan fluorescence assuming two-site exchange, where the (single) transition state was stabilized by $\Delta\Delta G_{U\rightarrow FU} = -5.1(\pm 0.2)$ kJ/mol for A39V Fyn SH3 and by $\Delta\Delta G_{U\rightarrow FU} = -4.5(\pm 0.3)$ kJ/mol for A39V/V55L Fyn SH3 relative to the WT Fyn domain (measured in the presence of 0.5 M guanidinium hydrochloride and 250 mM KCl at 25 °C).[6] These

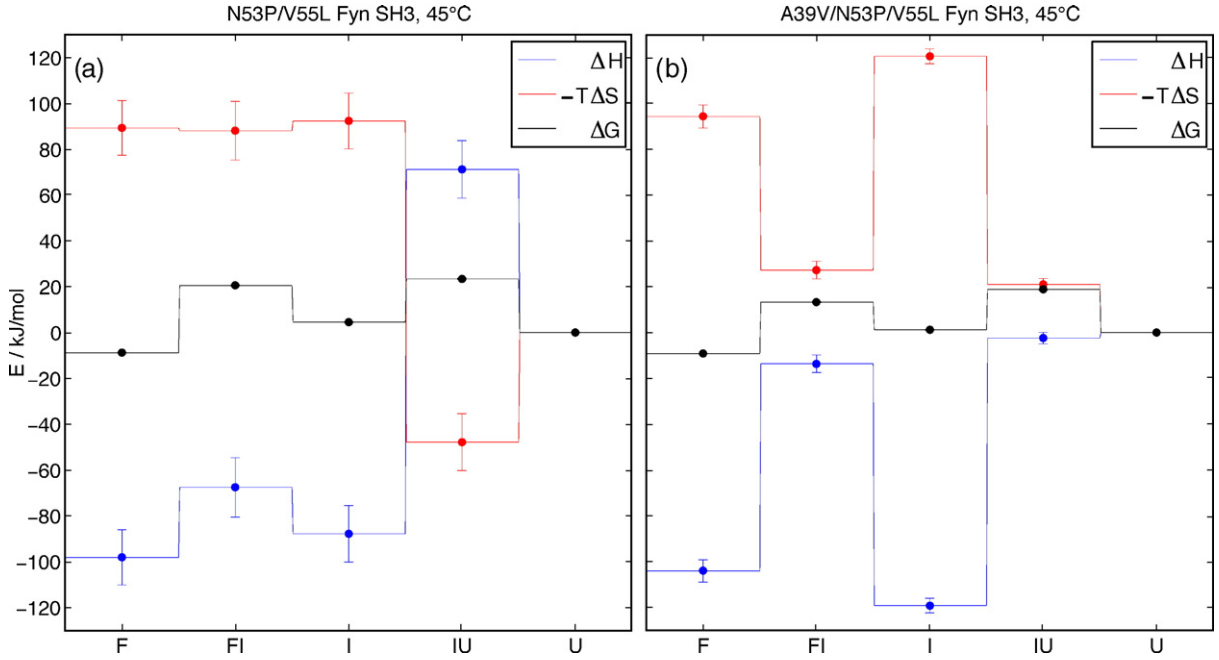

Figure 8. Profiles of free energy, ΔG (black), and the associated enthalpic ΔH (blue) and entropic $-T\Delta S$ (red) contributions to ΔG along the folding pathway for N53P/V55L Fyn SH3 (a) and A39V/N53P/V55L (b) Fyn SH3 at 45 °C, based on the parameters in Table 4. Values of G, H and TS are referenced with respect to the unfolded state U; G, H and TS for this state are arbitrarily assigned values of 0. FI and IU denote the rate-limiting transition states between states F, I and I, U, respectively. Values of $\Delta C_p(U\rightarrow I) = \Delta C_p(U\rightarrow F) - 0.15(\pm 0.17)$ kJ/mol/K and $\Delta C_p(U\rightarrow FI) = \Delta C_p(U\rightarrow F) + 1.46(\pm 0.11)$ kJ/mol/K are extracted from the temperature dependence of k_{FI} and k_{IF} for A39V/N53P/V55L Fyn SH3 (Figure 7).

values of $|\Delta\Delta G_{U\rightarrow FU}|$ are much larger than $|\Delta\Delta G_{U\rightarrow F}|$, suggesting energetically favorable non-native hydrophobic core packing of the Val39 side-chain in the rate-limiting transition state, thus increasing the folding rate by an order of magnitude ($e^{-\Delta\Delta G_{U\rightarrow FU}/RT}$). (Clearly, the results of this study are also consistent with non-native hydrophobic core packing of Val39 in the intermediate and in both transition states of A39V/N53P/V55L Fyn SH3. Given that a direct contact between the side-chains of Val39 and Pro53 or Leu55 is not expected based on the emerging structural model of the intermediate discussed above, it is not surprising that this feature of the WT Fyn SH3 domain is also observed in the N53P/V55L background.

Partitioning of free energy values into enthalpic and entropic contributions can be difficult, unless dispersion profiles are large and a wide range of temperatures are sampled. Nevertheless, a number of conclusions can be drawn with confidence. First, it is clear that there is significant compensation between entropy and enthalpy contributions to free energy in I and F states of both N53P/V55L and A39V/N53P/V55L Fyn SH3 (H, S and G for U are set to 0). Second, the slightly more favorable enthalpy and slightly less favorable entropy of the I state of the triple mutant relative to the F state are consistent with a collapsed intermediate state with favorable non-native interactions, as discussed above. Third, the heat capacities of the I and F states of the triple mutant are the same to within error, $\Delta C_p(U\rightarrow I) = \Delta C_p(U\rightarrow F) - 0.15(\pm 0.17)$ kJ/mol/K. Heat capacity values depend on a number of factors, but are thought to be most sensitive to hydration levels and exposure of hydrophobic residues to solvent,[25] so the similar C_p values for the I and F states indicate that the hydrophobic side-chains of the intermediate are buried to a similar extent as in the native state. By contrast, the transition state between I and F has a significantly higher heat capacity, $\Delta C_p(U\rightarrow FI) = \Delta C_p(U\rightarrow F) + 1.46(\pm 0.11)$ kJ/mol/K, presumably because non-native interactions in the I state must be broken along the folding pathway, leading to transient hydration of hydrophobic residues (and subsequent dehydration) prior to their repacking to form the native conformation. A $\Delta C_p(U\rightarrow F)$ value of −2.84 kJ/mol/K has been obtained for the wild-type Fyn SH3 domain in 250 mM KCl, 10 mM Tris–HCl (pH 8.0) at the thermal denaturation midpoint 80 °C.[3] Assuming that this value is unaltered for the A39V/N53P/V55L Fyn SH3 domain, it is possible to estimate the relative solvent-inaccessible surface area of the transition state between the I and F states of the protein as roughly half that of the folded protein ([−2.84+1.46]/−2.84 = 49%). By comparison, $\Delta C_p(U\rightarrow F) = -3.60$ kJ/mol/K and $\Delta C_p(U\rightarrow F) = -1.86$ kJ/mol/K $= \Delta C_p(U\rightarrow F) + 1.74$ kJ/mol/K were measured for the wild-type Fyn SH3 domain (20 mM Tris–HCl (pH 7.2), at 25 °C)[2] so that a similar value for the solvent-inaccessible surface area of the rate-limiting transition state for the folding of this protein, 52%, is obtained.[2] Another measure for the relative solvent-inaccessible surface area of the rate-limiting transition state is the Tanford β value $= m_f/(-m_u + m_f)$, where m_f and m_u are the derivatives of the natural logarithm of the folding and unfolding rates, respectively, with respect to denaturant concentration.[1] The stopped-flow kinetics experiments on A39V/N53P/V55L Fyn SH3 reported here (Figure 3) yield a β value of 74% that is only slightly higher than the range of values, 68% – 72%, that have been reported for other SH3 domains.[2,6]

Comparison of stopped-flow and CPMG folding kinetics

As described in detail above, the relaxation dispersion profiles recorded on the Fyn SH3 domains studied here in the absence of denaturants are consistent with a three-state folding mechanism, where the U and I states are populated at only low levels. Traditionally, folding kinetics obtained from stopped-flow experiments are reported in so called chevron plots,[1] ln (k_f+k_u) *versus* dentaturant concentration, and in what follows below we show that with a reasonable set of assumptions the chevron plot for A39V/N53P/V55L Fyn SH3 (Figure 3) can be well reproduced with the kinetic parameters from the NMR relaxation dispersion data.

Consider the set of coupled differential equations given by equation (4) that describes the recovery of equilibrium populations for each of the states along the folding reaction at a given denaturant concentration as monitored by the stopped-flow experiment:

$$\frac{d}{dt}\begin{pmatrix}\Delta p_F\\ \Delta p_I\\ \Delta p_U\end{pmatrix} = -\begin{pmatrix}k_{FI} & -k_{IF} & 0\\ -k_{FI} & k_{IF}+k_{IU} & -k_{UI}\\ 0 & -k_{IU} & k_{UI}\end{pmatrix}\begin{pmatrix}\Delta p_F\\ \Delta p_I\\ \Delta p_U\end{pmatrix} \quad (4)$$

Here Δp_X is the deviation of the population of state $X \in \{F, I, U\}$ from its equilibrium value at a particular denaturant concentration. The general solution is biexponential[26] with rate constants given by the two non-trivial eigenvalues of the kinetic rate matrix in equation (4):

$$\lambda_\pm = \frac{1}{2}\left(k_{ex,FI} + k_{ex,IU}\right) \times \left(1 \pm \sqrt{1 - 4\frac{k_{FI}k_{IU} + k_{FI}k_{UI} + k_{IF}k_{UI}}{\left(k_{ex,FI} + k_{ex,IU}\right)^2}}\right) \quad (5)$$

The kinetic rate constants for the folding of A39V/N53P/V55L Fyn SH3 at 23 °C can be readily obtained from equations (2) and (3) using the thermodynamic parameters reported in Table 4 to give $k_{FI} = 18.4(\pm 0.1)$/s, $k_{IF} = 822(\pm 5)$/s, $k_{IU} = 55.1(\pm 2.4)$/s and $k_{UI} = 901(\pm 37)$/s. In the absence of denaturant, $k_{ex,FI}$ and $k_{ex,IU}$ are similar in magnitude and there is not a direct mapping

between the pair of eigenvalues above and the two individual exchange processes F↔I and I↔U. Also, both these eigenvalues ($\lambda_+ = 1119/s$ and $\lambda_- = 678/s$) that describe the folding/unfolding kinetics in aqueous solution are too fast so that signal change would not be observed after the dead-time of the stopped-flow device of about 12 ms.

It is well known that denaturants such as guanidinium hydrochloride or urea stabilize the unfolded state of a protein much more than the folded state.[1] Since the I state of A39V/N53P/V55L Fyn SH3 is collapsed (i.e. folded-like, see Discussion in the previous sections), while the transition state connecting I and F is less compact, with increased exposure of hydrophobic residues to water (based on ΔC_p values; see above), addition of denaturant will preferentially stabilize the transition state relative to I and F, leading to an increase in $k_{ex,FI}$ with denaturant. By contrast, it is reasonable to expect that $k_{ex,IU}$ will increase less substantially with denaturant than $k_{ex,FI}$ (or even decrease) because the partially collapsed transition state between the intermediate and unfolded states is expected to be less stabilized by denaturant than the unfolded state. Under the conditions of the stopped-flow experiments, therefore, F↔I equilibrates very rapidly compared to I↔U ($k_{ex,FI} >> k_{ex,IU}$) and the square root in equation (5) can be expanded in a truncated Taylor series to good approximation:

$$\lambda_{\pm} \approx \frac{1}{2} k_{ex,FI} \left(1 \pm 1 \mp 2 \frac{k_{FI}k_{IU} + k_{FI}k_{UI} + k_{IF}k_{UI}}{k_{ex,FI}^2} \right) \quad (6)$$

The component of the signal which decays with a rate given by the eigenvalue $\lambda_+ \approx k_{ex,FI}$ is lost within the instrument dead-time, so that the observed time evolution of the fluorescence signal is effectively monoexponential with a rate constant given by the smaller eigenvalue:

$$\lambda_- \approx \frac{k_{FI}k_{IU} + k_{FI}k_{UI} + k_{IF}k_{UI}}{k_{ex,FI}} \quad (7)$$

At low denaturant concentrations the dominant term in equation (7) is $k_{IF}k_{UI}/k_{ex,FI} \approx k_{UI}$ (see rate values above) so that the folding arm of the chevron plot reports $k_f \approx k_{UI}$ and its slope gives the difference in *m*-values of the transition state connecting states U and I and the U state, $m_f \approx m_{U\rightarrow IU} = m_{IU} - m_U$. At high denaturant concentrations, however, the dominant term becomes $k_u \approx k_{FI}k_{IU}/k_{ex,FI} = k_{IU}p_I/(p_I+p_F)$, so that the unfolding arm of the chevron plot effectively reports on k_{IU} weighted by the population of the intermediate. The small heat capacity difference between I and F states (see above) indicates that the accessible surface areas of these two states are similar so that their population ratio remains fairly constant upon addition of denaturant. Thus, the slope of the unfolding arm of the chevron plot, m_u, provides a measure of the difference in *m*-values of the transition state between I and U and the I state, $m_u \approx m_{I\rightarrow IU}$. A more precise estimate of m_u can be obtained by noting that at zero denaturant concentration $k_{FI} = 18/s \ll k_{IF} = 822/s$, and assuming that the difference in *m*-values of the I and F states is small (that follows from $\Delta C_p(I\rightarrow F) \approx 0$) $k_{FI} \ll k_{IF}$ holds at all denaturant concentrations. Thus, we can write:

$$\ln k_u \approx \ln\left(\frac{k_{FI}}{k_{ex,FI}} k_{IU}\right) \approx \ln\left(\frac{k_{FI}}{k_{IF}} k_{IU}\right) = \ln k_{FI} - \ln k_{IF} + \ln k_{IU} \quad (8)$$

The slope of $\ln k_u$ with respect to denaturant concentration is therefore $m_u \approx m_{F\rightarrow FI} - m_{I\rightarrow FI} + m_{I\rightarrow IU} = m_{F\rightarrow I} + m_{I\rightarrow IU}$, and the calculated Tanford β value $= m_f/(-m_u + m_f) \approx m_{U\rightarrow IU}/(-m_{F\rightarrow I} - m_{I\rightarrow IU} + m_{U\rightarrow IU}) = m_{U\rightarrow IU}/m_{U\rightarrow F}$. Thus, the β value obtained from analysis of the chevron plot of Figure 3 using a two-state folding model reports on the solvent-inaccessible surface area of the transition state between states I and U relative to F. A value of 74% has been measured for A39V/N53P/V55L Fyn SH3, consistent with a relatively compact transition state. Interestingly, the ratio of the rate constants determined from the unfolding and folding arms that are extrapolated to zero denaturant concentration $(k_{FI}k_{IU})/(k_{IF}k_{UI}) = p_I/p_F \times p_U/p_I = p_U/p_F$ faithfully reports on the relative equilibrium population of the unfolded and folded states in the absence of denaturant.

Using the rate constants listed above for the three-state folding reaction of A39V/N53P/V55L Fyn SH3 at 23 °C, effective folding and unfolding rates (in the absence of denaturant) of $k_{UI} = 901(\pm 37)/s$ and $k_{FI}k_{IU}/k_{ex,FI} = 1.21(\pm 0.05)/s$, respectively, are calculated, in good agreement with the extrapolated values from the chevron plot of Figure 3, $k_f = 791(\pm 102)/s$ and $k_u = 1.46(\pm 0.19)/s$. Figure 3 also shows the calculated chevron plot, $\ln \lambda_-$ *versus* denaturant, that has been obtained from the predicted values for the effective folding and unfolding rates based on the model described above, the assumptions that the I and F states have similar *m*-values as do the two transition states (connecting U, I and I, F) and using values of $m_{U\rightarrow IU} = -1.85(\pm 0.08)/M$, $m_{I\rightarrow IU} = 0.639(\pm 0.024)/M$, obtained from the stopped-flow data.

Implications for the folding pathway

Over the past years the folding properties of SH3 domains have been studied extensively using a range of techniques, including mutagenesis, a variety of spectroscopic probes of kinetics and thermodynamics, calorimetry, computer simulations, and NMR spectroscopy. The folding pathways of all SH3 domains are believed to share essential common features, despite folding rates that vary over several orders of magnitude (reviewed by Capaldi & Radford).[13] Φ Value analysis based on stopped-flow experiments has established that a folding nucleus consisting of the central β-sheet, more precisely strand β_2 and the hairpin β_3/distal loop/β_4 (Figure 1), is already formed in the rate-

limiting transition state.[4–6] Although all these experiments are consistent with a two-state folding mechanism, computer simulations of folding of the Src-SH3 domain have found two separate barriers, indicating that an intermediate state exists on the folding pathway.[27] According to these simulations, the central β-sheet is formed in the first transition state, while the first and last β-strands and the RT–src loop remain disordered. The second transition state involves the association of hydrophobic surfaces, one comprising strands β_2–β_4 and the other β_1, β_5 and the RT–src loop, with a desolvation process where about ten water molecules are expelled prior to formation of the hydrophobic core of the native state. CPMG relaxation dispersion experiments have detected a low populated intermediate I on the folding pathway of the fast-folding mutants, G48M and G48V Fyn SH3.[7–9] In agreement with Φ value analyses, the NMR studies indicate a native-like backbone conformation for the central β-sheet in the intermediate state but not for the terminal strands β_1, β_5 and the RT loop. Variation in folding rates of G48M Fyn SH3 with protein deuteration further suggests that some hydrophobic contacts are already formed in the first transition state between U and I[9] and ^{13}C relaxation dispersion studies on this mutant showed that the hydrophobic core of the intermediate state in the folding reaction is largely collapsed but not tightly packed, with features similar to a "molten globule",[9] so that its partial molar volume is between that of U and F.[10] Consistent with hydrophobic core desolvation occurring during the transition between I and F, the I state is likely to be partially hydrated.[10]

The kinetic, thermodynamic, and structural information obtained from analyses of CPMG relaxation dispersion data recorded on both N53P/V55L and A39V/N53P/V55L Fyn SH3 domains strongly supports this emerging picture of the SH3 domain folding pathway. The large temperature range over which the triple mutant was studied and the fact that at temperatures lower than 25 °C the two-state exchange reaction F↔I predominates facilitates the extraction of robust exchange and thermodynamic parameters for this mutant, in particular, along with accurate chemical shift differences. The ^{15}N chemical shifts of the folding intermediates of both domains studied here are particularly interesting because although they indicate that the central β-sheet forms already at this stage, there are some noteworthy differences relative to the structures of the I states of both G48M and G48V Fyn SH3. In particular, the backbone conformation of the RT loop appears to be native-like in the N53P/V55L and A39V/N53P/V55L Fyn SH3 intermediates, and the hydrophobic mutations in the COOH-terminal region cause non-native hydrophobic long-range interactions between the side-chain of Pro53 and strands β_1 and β_2. The substantial stabilization of the intermediate and both transition states caused by the A39V mutation in the N53P/V55L background is of similar magnitude to that observed for the A39V substitution in wild-type Fyn determined by different methods. The stabilization underscores the importance of hydrophobic interactions in all stages of SH3 folding and is even larger than expected based solely on the increase in hydrophobicity from alanine to valine,[6] supporting the notion of favorable non-native packing interactions of Val39 in both intermediate and transition states. In this regard, the intermediate of the A39V/N53P/V55L Fyn SH3 domain is unusually stable enthalpically and destabilized entropically (Table 4 and Figure 8) compared to intermediates of other Fyn mutants.[7] These findings suggest that hydrophobic core packing of the folding intermediate of A39V/N53P/V55L Fyn SH3 might be tighter than for the G48M mutant,[9,10] although the experimental data obtained here cannot rule out a similar "molten globule" like hydrophobic core.

Computer simulations suggest that the intermediate state is a general property of the folding pathway of SH3 domains rather than just a consequence of the Gly48 mutation (that was present in previously studied Fyn SH3 domains[7–9]). The present study supports this notion because (i) a pair of Fyn SH3 domain mutants with glycine at position 48 are investigated and (ii) both of these mutants are shown to fold *via* on-pathway intermediates with features similar to those reported for G48M and G48V Fyn SH3.[7–9] NMR dispersion experiments can detect such intermediate states, while many other techniques cannot, because multiple probes are employed and the folding reaction can be studied in the absence of denaturants that increase the concentration of U relative to I. For example, for G48M, G48V, and N53P/V55L Fyn SH3 the I state is higher in energy than U even under native conditions and therefore never sufficiently populated to be observed in experiments based on circular dichroism or fluorescence. In the case of A39V/N53P/V55L Fyn SH3, equation (4) predicts a transient buildup of the I state in stopped-flow experiments to about 35% after 1 ms at 25 °C in low denaturant; the pronounced bi-exponential time evolution of the fluorescence leads to a slight "rollover" in the chevron plot (see predicted plots in Figure 3) that may well be difficult to detect, even using instrumentation with a dead-time well under 1 ms. In addition, the temperature dependence of the folding rates of this triple mutant argue that the F and I states have similar *m*-values ($\Delta C_P(F \rightarrow I) \approx 0$) so that it is most unlikely that comparison of folding and unfolding *m*-values with a value determined from equilibrium unfolding experiments[1] would allow discrimination between a two and three-state folding process. Thus, despite unequivocal evidence from CPMG dispersion data sets for a more complex process, the stopped-flow data are consistent with a two-state folding mechanism, where the rate-limiting transition state between U and I is the one characterized by this technique, as discussed above.

Our observation of non-native interactions in the folding pathway of Fyn SH3 domain mutants adds to the growing list of theoretical[28–32] and experimental[33–36] studies reporting the formation of non-native contacts during folding. These studies

collectively suggest that non-native interactions may be more widespread than previously recognized. Moreover, contrary to the common perception that non-native interactions "stall" the folding process, a subset of the above studies[28–33,35] suggest that the formation of such interactions could result in enhancement of protein folding rates. In light of observations that non-native interactions can assist in folding, the widely accepted notion that folding pathways are "exclusively" driven by partial or complete formation of native contacts needs to be revisited and the results of folding studies based exclusively on Φ value analysis and/or native-centric (Gō-type) folding models should be interpreted with caution. Of course, assumptions about formation of native-like contacts are unnecessary in the interpretation of NMR relaxation dispersion data sets where structural information about the states along the folding pathway is derived quite naturally from their chemical shifts.

In summary, the study presented here highlights the advantages of relaxation dispersion NMR spectroscopy in providing site-specific information about excited states along protein folding pathways. The methodology is of considerable utility in elucidating general features of folding, while at the same time providing insights into the differences in folding intermediates resulting from individual mutations. Of particular note is the detection and characterization of low populated intermediates that cannot be observed using more traditional approaches for studying protein folding.

Materials and Methods

Sample preparation

The site-directed variants of the *Gallus gallus* Fyn SH3 domain used here were constructed *via* PCR-mediated mutagenesis, expressed and purified using nickel affinity chromatography as described.[3] NMR samples contained 0.9 mM [U-^{15}N] N53P/V55L or A39V/N53P/V55L Fyn SH3, 0.2 mM EDTA, 0.05% (w/v) NaN_3, 50 mM sodium phosphate (pH 7.0) in $H_2O/^2H_2O$ (9:1). Note that the only difference between wild-type versions of *Gallus gallus* and *Homo sapiens* Fyn SH3 is at position 5 where valine is replaced by glutamate (*Gallus gallus*).

Stopped-flow folding kinetics

Folding kinetics experiments were carried out on a SFM-4 stopped-flow device (BioLogic, Claix, France) equipped with a photomultiplier tube monitoring the recovery of tryptophan fluorescence upon guanidinium hydrochloride denaturation of the protein. Folding and unfolding rates were measured at room temperature (23 °C) in 0.2 mM EDTA, 50 mM sodium phosphate (pH 7.0). At each denaturant concentration, c_{denat}, at least five separate fluorescence recovery curves were averaged. Appropriate monoexponential functions were fit using the software BioKine™ to obtain the relaxation rate $k_{ex}(c_{denat})$. The chevron plot, ln $k_{ex}(c_{denat})$ *versus* c_{denat}, was fit in Kaleidagraph (Synergy Software) to extract the folding and unfolding rates extrapolated to zero denaturant concentration, k_f and k_u, respectively, and the corresponding m-values m_f and m_u:

$$\ln k_{ex}(c_{denat}) = \ln\left(e^{\ln k_f + m_f c_{denat}} + e^{\ln k_u + m_u c_{denat}}\right) \quad (9)$$

Sequence-specific NMR resonance assignments

Sequence-specific ^{1}H and ^{15}N resonance assignments for the *G. gallus* Fyn SH3 mutants N53P/V55L and A39V/N53P/V55L were obtained using conventional ^{1}H-^{15}N heteronuclear NMR spectroscopy. Briefly, [^{1}H, ^{15}N]-TOCSY-HSQC (N53P/V55L Fyn SH3 only) with a DIPSI-2rc mixing scheme[37] (78 ms mixing time) and [^{1}H, ^{15}N]-NOESY-HSQC (150 ms mixing time) spectra[38] were recorded at 40 °C (N53P/V55L Fyn SH3) and 25 °C (A39V/N53P/V55L Fyn SH3) on a Varian Unity INOVA 500 MHz NMR spectrometer equipped with a room-temperature probe with z-axis pulsed field gradient capabilities. The H_2O resonance was suppressed by gradient coherence selection with quadrature detection in the indirect ^{1}H and ^{15}N dimensions achieved by States-TPPI[39] and the echo-antiecho method,[40,41] respectively. All NMR spectra were processed with NMRPipe[42] software and analyzed with NMRView 5.2.2.[43] ^{1}H chemical shifts were referenced with respect to external DSS in ^2H_2O and ^{15}N chemical shifts were referenced indirectly.[44]

Experiments probing conformational exchange

^{15}N single quantum CPMG relaxation dispersion experiments[45,46] were used to study millisecond exchange processes in both N53P/V55L and A39V/N53P/V55L Fyn SH3. Relaxation dispersion data sets were recorded for N53P/V55L and A39V/N53P/V55L Fyn SH3 in steps of 5 °C from 35 °C to 55 °C and from 5 °C to 50 °C, respectively, using both 500 and 800 MHz NMR spectrometers equipped with room-temperature probes. In each experiment 14 different CPMG frequencies $\nu_{CPMG} = 1/(2\delta)$, where δ is the time between consecutive refocusing pulses, ranging from 41.7 Hz to 1000.0 Hz were sampled during a constant-time relaxation interval of $T_{CPMG} = 48$ ms. Amide resonance intensities $I(\nu_{CPMG})$ were quantified by three-way decomposition using MUNIN[47,48] and converted into effective transverse relaxation rates, $R_{eff}(\nu_{CPMG}) = -\ln(I(\nu_{CPMG})/I_0)/T_{CPMG}$, where I_0 is the corresponding resonance intensity in a reference spectrum recorded without the constant-time relaxation interval. Error estimates ΔR_{-eff} for R_{-eff} were obtained from duplicate measurements at two or three different ν_{CPMG} values as described,[7] assuming a minimum relative error $\Delta R_{eff}/R_{eff}$ of 2.0% and 1.5% for the N53P/V55L and A39V/N53P/V55L Fyn SH3, respectively, to account for offset effects and other systematic experimental imperfections.[49] The relaxation dispersion profiles were fit to extract global exchange parameters, including exchange rates and populations of exchanging states, along with residue specific values such as ^{15}N chemical shift differences between exchanging states and intrinsic ^{15}N relaxation rates, $R_{2,0}$. This was accomplished by a non-linear least-squares fitting procedure whereby experimental dispersion profiles, $R_{eff}(\nu_{CPMG}) = R_{2,0} + R_{ex}(\nu_{CPMG})$, were fit to those calculated by deriving the evolution of magnetization during the CPMG interval by solving the Bloch–McConnell equations[50] numerically for the appropriate

two or three-site exchange model using in-house written software and a procedure described.[7,8] Errors of the fitted parameters were calculated from the covariance matrix,[24] an approach that has been shown to provide reasonable error estimates in a study similar to that presented here.[18] Absolute signs of ^{15}N chemical shift differences between exchanging states in A39V/N53P/V55L Fyn SH3 were obtained from differences in ^{15}N peak positions in [^{1}H, ^{15}N]-HMQC and [^{1}H, ^{15}N]-HSQC spectra recorded at 500 MHz and/or between two [^{1}H, ^{15}N]-HSQC spectra obtained at 500 MHz and 800 MHz.[19] To this end, spectra were recorded with high signal/noise ratios and digital resolution in the ^{15}N dimension (acquisition time 0.15 s and 0.13 s at 500 MHz and 800 MHz, respectively); resonances not subject to chemical exchange were used to correct for any residual differences in ^{15}N referencing.

Acknowledgements

The authors thank Dr D. M. Korzhnev for providing software for data analysis and for many helpful discussions. This work was supported by a grant from the Canadian Institutes of Health Research (CIHR) (to L.E.K.). P.N. acknowledges postdoctoral support from the Deutsche Forschungsgemeinschaft (NE 1197/1-1 and 1-2). A.Z.-A. is supported by a doctoral Canada Graduate Scholarship (CGS-D3) from the Natural Sciences and Engineering Research Council of Canada (NSERC). L.E.K. holds a Canada Research Chair in Biochemistry.

References

1. Fersht, A. (1999). *Structure and Mechanism in Protein Science.* Freeman, New York.
2. Plaxco, K. W., Guijarro, J. I., Morton, C. J., Pitkeathly, M., Campbell, I. D. & Dobson, C. M. (1998). The folding kinetics and thermodynamics of the Fyn-SH3 domain. *Biochemistry*, **37**, 2529–2537.
3. Maxwell, K. L. & Davidson, A. R. (1998). Mutagenesis of a buried polar interaction in an SH3 domain: Sequence conservation provides the best prediction of stability effects. *Biochemistry*, **37**, 16172–16182.
4. Martinez, J. C. & Serrano, L. (1999). The folding transition state between SH3 domains is conformationally restricted and evolutionarily conserved. *Nature Struct. Biol.* **6**, 1010–1016.
5. Riddle, D. S., Grantcharova, V. P., Santiago, J. V., Alm, E., Ruczinski, I. & Baker, D. (1999). Experiment and theory highlight role of native state topology in SH3 domain folding. *Nature Struct. Biol.* **6**, 1016–1024.
6. Northey, J. G. B., Di Nardo, A. A. & Davidson, A. R. (2002). Hydrophobic core packing in the SH3 domain folding transition state. *Nature Struct. Biol.* **9**, 126–130.
7. Korzhnev, D. M., Salvatella, X., Vendruscolo, M., Di Nardo, A. A., Davidson, A. R., Dobson, C. M. & Kay, L. E. (2004). Low-populated folding intermediates of Fyn SH3 characterised by relaxation dispersion NMR. *Nature*, **430**, 586–590.
8. Korzhnev, D. M., Neudecker, P., Mittermaier, A., Orekhov, V. Y. & Kay, L. E. (2005). Multiple-site exchange in proteins studied with a suite of six NMR relaxation dispersion experiments: An application to the folding of a Fyn SH3 domain mutant. *J. Am. Chem. Soc.* **127**, 15602–15611.
9. Mittermaier, A., Korzhnev, D. M. & Kay, L. E. (2005). Side-chain interactions in the folding pathway of a Fyn SH3 domain mutant studied by NMR relaxation dispersion NMR spectroscopy. *Biochemistry*, **44**, 15430–15436.
10. Bezsonova, I., Korzhnev, D. M., Prosser, R. S., Forman-Kay, J. D. & Kay, L. E. (2006). Hydration and packing along the folding pathway of SH3 domains by pressure-dependent NMR. *Biochemistry*, **45**, 4711–4719.
11. Korzhnev, D. M., Neudecker, P., Zarrine-Afsar, A., Davidson, A. R. & Kay, L. E. (2006). Abp1p and Fyn SH3 Domains Fold through Similar Low-Populated Intermediate States. *Biochemistry*, **45**, 10175–10183.
12. Noble, M. E. M., Musacchio, A., Saraste, M., Courtneidge, S. A. & Wierenga, R. K. (1993). Crystal structure of the SH3 domain in human Fyn; comparison of the three-dimensional structures of SH3 domains in tyrosine kinases and spectrin. *EMBO J.* **12**, 2617–2624.
13. Capaldi, A. P. & Radford, S. E. (1998). Kinetic studies of β-sheet protein folding. *Curr. Opin. Struct. Biol.* **8**, 86–92.
14. Morton, C. J., Pugh, D. J. R., Brown, E. L. J., Kahmann, J. D., Renzoni, D. A. C. & Campbell, I. D. (1996). Solution structure and peptide binding of the SH3 domain from human Fyn. *Structure*, **4**, 705–714.
15. Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids.* Wiley, New York.
16. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S. & Sykes, B. D. (1995). ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR*, **5**, 67–81.
17. Palmer, A. G., III, Kroenke, C. D. & Loria, J. P. (2001). Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods Enzymol.* **339**, 204–238.
18. Neudecker, P., Korzhnev, D. M. & Kay, L. E. (2006). Assessment of the effects of increased relaxation dispersion data on the extraction of 3-site exchange parameters characterizing the unfolding of an SH3 domain. *J. Biomol. NMR*, **34**, 129–135.
19. Skrynnikov, N. R., Dahlquist, F. W. & Kay, L. E. (2002). Reconstructing NMR spectra of "Invisible" excited protein states using HSQC and HMQC experiments. *J. Am. Chem. Soc.* **124**, 12352–12360.
20. Wishart, D. S. & Case, D. A. (2001). Use of chemical shifts in macromolecular structure determination. *Methods Enzymol.* **338**, 3–34.
21. Schwarzinger, S., Kroon, G. J. A., Foss, T. R., Chung, J., Wright, P. E. & Dyson, H. J. (2001). Sequence-dependent correction of random coil NMR chemical shifts. *J. Am. Chem. Soc.* **123**, 2970–2978.
22. Eyring, H. (1935). The activated complex and the absolute rate of chemical reactions. *Chem. Rev.* **17**, 65–77.
23. Hagen, S. J., Hofrichter, J., Szabo, A. & Eaton, W. A. (1996). Diffusion-limited contact formation in unfolded cytochrome c: Estimating the maximum rate of protein folding. *Proc. Natl. Acad. Sci. USA*, **93**, 11615–11617.
24. Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. (1992). *Numerical Recipes in C.* 2nd edit. Cambridge University Press, Cambridge.
25. Myers, J. K., Pace, C. N. & Scholtz, J. M. (1995).

Denaturant m values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **4**, 2138–2148.
26. Nölting, B. (2005). *Protein Folding Kinetics*. 2nd edit. Springer, Berlin.
27. Guo, W., Lampoudi, S. & Shea, J.-E. (2003). Posttransition state desolvation of the hydrophobic core of the src-SH3 protein domain. *Biophys. J.* **85**, 61–69.
28. Plotkin, S. S. (2001). Speeding protein folding beyond the Gō model: How a little frustration sometimes helps. *Proteins: Struct. Funct. Genet.* **45**, 337–345.
29. Treptow, W. L., Barbosa, M. A. A., Garcia, L. G. & Pereira de Araújo, A. F. (2002). Non-native interactions, effective contact order, and protein folding: A mutational investigation with the energetically frustrated hydrophobic model. *Proteins: Struct. Funct. Genet.* **49**, 167–180.
30. Li, L., Mirny, L. A. & Shakhnovich, E. I. (2000). Kinetics, thermodynamics and evolution of nonnative interactions in a protein folding nucleus. *Nature Struct. Biol.* **7**, 336–342.
31. Clementi, C. & Plotkin, S. S. (2004). The effects of nonnative interactions on protein folding rates: Theory and simulation. *Protein Sci.* **13**, 1750–1766.
32. Fan, K., Wang, J. & Wang, W. (2002). Folding of lattice protein chains with modified Gō potential. *Eur. Phys. J.* **B30**, 381–391.
33. Viguera, A. R., Vega, C. & Serrano, L. (2002). Unspecific hydrophobic stabilization of folding transition states. *Proc. Natl. Acad. Sci. USA*, **99**, 5349–5354.
34. Canet, D., Lyon, C. E., Scheek, R. M., Robillard, G. T., Dobson, C. M., Hore, P. J. & van Nuland, N. A. J. (2003). Rapid formation of non-native contacts during the folding of HPr revealed by real-time photo-CIDNP NMR and stopped-flow fluorescence experiments. *J. Mol. Biol.* **330**, 397–407.
35. Di Nardo, A. A., Korzhnev, D. M., Stogios, P. J., Zarrine-Afsar, A., Kay, L. E. & Davidson, A. R. (2004). Dramatic acceleration of protein folding by stabilization of a nonnative backbone conformation. *Proc. Natl. Acad. Sci. USA*, **101**, 7954–7959.
36. Friel, C. T., Beddard, G. S. & Radford, S. E. (2004). Switching Two-state to Three-state Kinetics in the Helical Protein Im9 *via* the optimisation of stabilising non-native interactions by design. *J. Mol. Biol.* **342**, 261–273.
37. Cavanagh, J. & Rance, M. (1992). Suppression of cross-relaxation effects in TOCSY spectra *via* a modified DIPSI-2 mixing sequence. *J. Magn. Reson.* **96**, 670–678.
38. Zhang, O., Kay, L. E., Olivier, J. P. & Forman-Kay, J. D. (1994). Backbone ^{1}H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. *J. Biomol. NMR*, **4**, 845–858.
39. Marion, D., Ikura, M., Tschudin, R. & Bax, A. (1989). Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *J. Magn. Reson.* **85**, 393–399.
40. Kay, L. E., Keifer, P. & Saarinen, T. (1992). Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**, 10663–10665.
41. Schleucher, J., Sattler, M. & Griesinger, C. (1993). Coherence selection by gradients without signal attenuation: application to the three-dimensional HNCO experiment. *Angew. Chem. Int. Ed. Engl.* **32**, 1489–1491.
42. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
43. Johnson, B. A. & Blevins, R. A. (1994). NMRView: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR*, **4**, 603–614.
44. Markley, J. L., Bax, A., Arata, Y., Hilbers, C. W., Kaptein, R., Sykes, B. D. *et al.* (1998). Recommendations for the presentation of NMR structures of proteins and nucleic acids (Recommendations 1998). *Pure Appl. Chem.* **70**, 117–142.
45. Loria, J. P., Rance, M. & Palmer, A. G., III (1999). A Relaxation-Compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. *J. Am. Chem. Soc.* **121**, 2331–2332.
46. Tollinger, M., Skrynnikov, N. R., Mulder, F. A., Forman-Kay, J. D. & Kay, L. E. (2001). Slow Dynamics in Folded and Unfolded States of an SH3 Domain. *J. Am. Chem. Soc.* **123**, 11341–11352.
47. Orekhov, V. Y., Ibraghimov, I. V. & Billeter, M. (2001). MUNIN: a new approach to multi-dimensional NMR spectra interpretation. *J. Biomol. NMR*, **20**, 49–60.
48. Korzhnev, D. M., Ibraghimov, I. V., Billeter, M. & Orekhov, V. Y. (2001). MUNIN: Application of three-way decomposition to the analysis of heteronuclear NMR relaxation data. *J. Biomol. NMR*, **21**, 263–268.
49. Ishima, R. & Torchia, D. A. (2005). Error estimation and global fitting in transverse-relaxation dispersion experiments to determine chemical-exchange parameters. *J. Biomol. NMR*, **32**, 41–54.
50. McConnell, H. M. (1958). Reaction rates by nuclear magnetic resonance. *J. Chem. Phys.* **28**, 430–431.
51. Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.
52. Merrit, E. A. & Murphy, M. E. P. (1994). Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Crystallog. sect. D*, **50**, 869–873.
53. Guex, M. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.
54. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. (2003). The Xplor-NIH NMR molecular structure determination package. *J. Mag. Reson.* **160**, 65–73.
55. Brünger, A. T. (1992). *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR.* Yale University Press, New Haven.

Edited by P. Wright

(Received 26 July 2006; received in revised form 16 August 2006; accepted 17 August 2006)
Available online 22 August 2006