Φ-Value analysis of a three-state protein folding pathway by NMR relaxation dispersion spectroscopy

Philipp Neudecker[†‡], Arash Zarrine-Afsar[†], Alan R. Davidson[†], and Lewis E. Kay[†‡§]

[†]Departments of Medical Genetics and Biochemistry, University of Toronto, Toronto, ON, Canada M5S 1A8; and [‡]Department of Chemistry, University of Toronto, Toronto, ON, Canada M5S 3H6

Edited by David Baker, University of Washington, Seattle, WA, and approved August 20, 2007 (received for review May 30, 2007)

Experimental studies of protein folding frequently are consistent with two-state folding kinetics. However, recent NMR relaxation dispersion studies of several fast-folding mutants of the Fyn Src homology 3 (SH3) domain have established that folding proceeds through a low-populated on-pathway intermediate, which could not be detected with stopped-flow experiments. The dispersion experiments provide precise kinetic and thermodynamic parameters that describe the folding pathway, along with a detailed site-specific structural characterization of both the intermediate and unfolded states from the NMR chemical shifts that are extracted. Here we describe NMR relaxation dispersion Φ-value analysis of the A39V/N53P/V55L Fyn SH3 domain, where the effects of suitable point mutations on the energy landscape are quantified, providing additional insight into the structure of the folding intermediate along with per-residue structural information of both rate-limiting transition states that was not available from previous studies. In addition to the advantage of delineating the full three-state folding pathway, the use of NMR relaxation dispersion as opposed to stopped-flow kinetics to quantify Φ values facilitates their interpretation because the obtained chemical shifts monitor any potential structural changes along the folding pathway that might be introduced by mutation, a significant concern in their analysis. Φ-Value analysis of several point mutations of A39V/N53P/V55L Fyn SH3 establishes that the β_3–β_4-hairpin already is formed in the first transition state, whereas strand β_1, which forms nonnative interactions in the intermediate, does not fully adopt its native conformation until after the final transition state. The results further support the notion that on-pathway intermediates can be stabilized by nonnative contacts.

chemical exchange | Carr–Purcell–Meiboom–Gill | Fyn SH3 | protein folding intermediate | transition state

An understanding of the process(es) by which a protein folds from its unfolded state ensemble, U, into its native conformation, F, requires knowledge of the structural properties of the various states that populate its folding pathway(s), along with the kinetic and thermodynamic parameters that describe the energy landscape that drives the folding reaction. Although a description of the structural and energetic properties of F often is readily obtained by using standard biophysical approaches, intermediate states, I, that participate in the folding reaction are in general more difficult to quantify because of the fact that they usually are low-populated and may be formed only transiently (1, 2). Information about such "elusive" states, including transition states along the folding pathway, can be obtained indirectly, however, by monitoring the effects of changes in temperature, pressure, denaturant concentration, solvent, and/or point mutations on the folding/unfolding kinetics (1).

Solution-based NMR spectroscopy is a particularly powerful technique for quantifying processes that involve conformational exchange (3), such as protein folding. We previously have shown that Carr–Purcell–Meiboom–Gill (CPMG) relaxation dispersion NMR spectroscopy can be used to study folding events that occur on the millisecond time scale, so long as the exchange process involves states that are populated to 0.5% or higher (4–7). The NMR experiments enjoy a number of advantages over other spectroscopic measures that frequently deliver information that is averaged over all probes in the molecule or restricted to only one or two sites where probes are centered. First, by monitoring the line widths of NMR signals that derive from residues undergoing conformational exchange as a function of the application of radio-frequency pulses, the complete kinetics of the exchange process and the chemical shifts of the exchanging states can be obtained simultaneously from fits of the so-called relaxation dispersion data. Moreover, the chemical shifts of the excited states are site-specific probes of the extent of structure formation along the folding reaction. The fact that information about the folding process is obtained at many sites in the protein facilitates a distinction between two-state and three-state folding events (4, 5). In the case where the NMR relaxation data are best fit to a three-state folding model, it is even possible to distinguish between folding schemes of the form U ↔ I ↔ F, U ↔ F ↔ I, or I ↔ U ↔ F (7). Finally, relaxation dispersion experiments can be recorded in the solvent of choice, i.e., without denaturants that are necessary for most spectroscopic measurements of folding/unfolding rates that can render the detection of I states difficult.

In a series of previous studies, our laboratory used ^{15}N CPMG relaxation dispersion NMR spectroscopy to study a number of fast folding mutants of the Fyn and Abp1p Src homology 3 (SH3) domains that fold via previously undetected low-populated on-pathway intermediates, and the structures of the excited states involved in the folding process have been characterized in detail (4–7). ^{15}N chemical shifts demonstrate that state U is indeed unfolded, with very little or no residual structure. By contrast, however, the central β-sheet comprising strands β_2, β_3, and β_4 (Fig. 1) is at least partially formed in I for both SH3 domains. Although qualitative information about the transition state ensembles (TS) connecting U, I and I, F could be obtained from the temperature dependencies of the extracted four rate constants that describe the U ↔ I ↔ F folding process, more detailed site-specific knowledge of the formation of structure in each TS was not available. In principle, such information can be obtained, however, through the protein-engineering method referred to as Φ-value analysis (1, 8–11). Here the effect of mutations on the folding/unfolding rates is monitored and recast in terms of changes in free energies that, so long as care is taken, can be interpreted in terms of the formation of structure in the TS at the site of mutation. Indeed, site-specific information

Author contributions: P.N., A.Z.-A., A.R.D., and L.E.K. designed research; P.N. and A.Z.-A. performed research; P.N. and A.Z.-A. analyzed data; and P.N., A.Z.-A., A.R.D., and L.E.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Abbreviations: SH3, Src homology 3; CPMG, Carr–Purcell–Meiboom–Gill; wt, wild type.

§To whom correspondence should be addressed at: Department of Medical Genetics and Microbiology, University of Toronto, 1 King's College Circle, Toronto, ON, Canada M5S 1A8. E-mail: kay@pound.med.utoronto.ca.

This article contains supporting information online at www.pnas.org/cgi/content/full/0705097104/DC1.

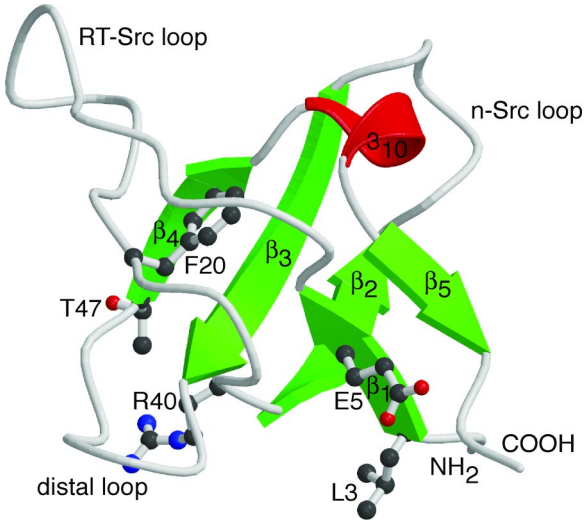

Fig. 1. Schematic representation of the secondary structure of a homology model of the A39V/N53P/V55L *G. gallus* Fyn SH3 domain, featuring the characteristic SH3 domain β-sandwich fold formed by the terminal (strands β_1 from Leu-3 to Ala-6 and β_5 from Leu-55 to Pro-57) and the approximately orthogonal central β-sheets (strands β_2 from Asp-25 to Asn-30, β_3 from Trp-36 to Ser-41, and β_4 from Thr-47 to Ile-50), along with a helical turn with 3_{10} geometry from Pro-51 to Tyr-54. The residues Leu-3, Glu-5, Phe-20, Arg-40, and Thr-47 mutated in this study are shown in ball-and-stick representation. The homology model was created as described previously (7).

about structure formation can be obtained for each state along the folding pathway that can include a pair of TS and I, in the case of a three-state folding process (see below).

In practice, Φ values for intermediate states, I, have been published for several proteins such as Im7 (12), but only very few examples exist in the literature where Φ values have been measured for both TS and I in studies of three-state folders (13–16), in part because of the difficulty of obtaining all four rate constants from stopped-flow-based spectroscopic approaches. By contrast, our experience with dispersion studies of SH3 domains suggests that NMR data can be analyzed to extract accurate rates (4–7, 17). NMR relaxation-dispersion-based Φ-value analysis thus appears to be a particularly powerful method for the analysis of more complex folding pathways than two-state. Below we present such an analysis of the folding of the A39V/N53P/V55L Fyn SH3 domain, which has been shown in a previous NMR study to fold by a three-state process, U $\leftrightarrow$ I $\leftrightarrow$ F (7). The obtained Φ values confirm the picture of structure formation along the folding pathway that was obtained from ^{15}N chemical shifts of the I state previously and extend it by providing information about formation of structure in the early TS, along with an indication of how structure formation "evolves" during the late (second) TS. The combined relaxation dispersion/Φ-value approach facilitates the determination of complete three-state protein folding pathways at atomic detail.

Methodology of Φ-Value Analysis

Before a discussion of the results of the relaxation dispersion study of A39V/N53P/V55L Fyn SH3, we briefly will review the essential features of Φ-value analysis (1, 8–11) and its extension to three-state folding. Most experimental studies of protein folding can be interpreted by using a simple two-state model,

$$\mathrm{U} \underset{k_{\mathrm{FU}}}{\overset{k_{\mathrm{UF}}}{\rightleftharpoons}} \mathrm{F},$$

where according to transition-state theory, the temperature dependence of the rate of transition from state X to state Y is given by the Eyring equation (18):

$$k_{\mathrm{XY}} = \frac{\kappa k_{\mathrm{B}} T}{h} e^{-\frac{\Delta G_{\mathrm{X}\rightarrow\mathrm{TS(XY)}}}{\mathrm{RT}}}, \qquad [1]$$

where κ is a transmission coefficient, k_{B} and h are the Boltzmann and Planck constants, respectively, and $\Delta G_{\mathrm{X}\rightarrow\mathrm{TS(XY)}} = G_{\mathrm{TS(XY)}} - G_{\mathrm{X}}$ denotes the free energy difference between the rate-limiting transition state TS(XY) and state X. Thus, the kinetic constants that are measured can be recast in terms of a simple one-dimensional energy landscape that describes the folding reaction. The introduction of a point mutation at a particular site in the protein modulates its stability so that the free energy of the mutant changes relative to the wild type (wt) by an amount $\Delta\Delta G_{\mathrm{U}\rightarrow\mathrm{F}} = \Delta G_{\mathrm{U}\rightarrow\mathrm{F}}(\mathrm{mutant}) - \Delta G_{\mathrm{U}\rightarrow\mathrm{F}}(\mathrm{wt}) = -\mathrm{RT}\ln(k_{\mathrm{UF}}(\mathrm{mutant})/k_{\mathrm{UF}}(\mathrm{wt})) + \mathrm{RT}\ln(k_{\mathrm{FU}}(\mathrm{mutant})/k_{\mathrm{FU}}(\mathrm{wt}))$. Similarly, the interactions in the rate-limiting transition state, TS(UF), can be modulated, so its stability changes by an amount $\Delta\Delta G_{\mathrm{U}\rightarrow\mathrm{TS(UF)}} = \Delta G_{\mathrm{U}\rightarrow\mathrm{TS(UF)}}(\mathrm{mutant}) - \Delta G_{\mathrm{U}\rightarrow\mathrm{TS(UF)}}(\mathrm{wt}) = -\mathrm{RT}\ln(k_{\mathrm{UF}}(\mathrm{mutant})/k_{\mathrm{UF}}(\mathrm{wt}))$. The Φ value is defined as

$$\Phi = \frac{\Delta\Delta G_{\mathrm{U}\rightarrow\mathrm{TS(UF)}}}{\Delta\Delta G_{\mathrm{U}\rightarrow\mathrm{F}}} = \frac{\Delta G_{\mathrm{TS(UF)}} - \Delta G_{\mathrm{U}}}{\Delta G_{\mathrm{F}} - \Delta G_{\mathrm{U}}}, \qquad [2]$$

where $\Delta G_{\mathrm{X}} = G_{\mathrm{X}}(\mathrm{mutant}) - G_{\mathrm{X}}(\mathrm{wt})$, X $\in$ {U, TS(UF), F}. Values of $\Phi \approx 0$ ($\Delta\Delta G_{\mathrm{U}\rightarrow\mathrm{TS(UF)}} \approx 0$) and $\Phi \approx 1$ ($\Delta\Delta G_{\mathrm{U}\rightarrow\mathrm{TS(UF)}} \approx \Delta\Delta G_{\mathrm{U}\rightarrow\mathrm{F}}$) indicate that the interactions at the position of the mutation in the TS are unfolded-like or native-like, respectively, independent of changes that the mutation might introduce to the U state. By contrast, Φ values that are not 0 or 1 must be interpreted with caution because they can derive from interactions that only are partially formed in the TS, multiple folding pathways, formation of nonnative interactions, or perturbations to any residual structure in the U state (1, 10, 11, 19). Fersht and Sato (11) describe these points in detail and suggest appropriate mutations such that Φ values can be analyzed rigorously.

Φ-Value analysis can be readily extended to more complex folding processes, such as those that are operative for the Fyn SH3 domain (described below). In the case of the pathway

$$\mathrm{U} \underset{k_{\mathrm{IU}}}{\overset{k_{\mathrm{UI}}}{\rightleftharpoons}} \mathrm{I} \underset{k_{\mathrm{FI}}}{\overset{k_{\mathrm{IF}}}{\rightleftharpoons}} \mathrm{F}$$

one can define the Φ value for state A $\in$ {I, TS(UI), TS(IF)}, where TS(UI) and TS(IF) denote the transition states connecting states U/I and I/F, respectively, as

$$\Phi_{\mathrm{A}} = \frac{\Delta\Delta G_{\mathrm{U}\rightarrow\mathrm{A}}}{\Delta\Delta G_{\mathrm{U}\rightarrow\mathrm{F}}} = \frac{\Delta G_{\mathrm{A}} - \Delta G_{\mathrm{U}}}{\Delta G_{\mathrm{F}} - \Delta G_{\mathrm{U}}},$$

with ΔG values obtained directly from the measured rates via Eq. **1**. Fig. 2 *Top* shows a schematic of a pair of one-dimensional energy landscapes corresponding to the folding pathways of mutant (blue) and wt (black), along with the differences in energies attributable to mutation for each state A $\in$ {U, TS(UI), I, TS(IF), F} from which the Φ_{A} values are calculated.

Results and Discussion

Choice of System and Point Mutants for Φ-Value Analysis. In a previous publication, we analyzed the folding pathway of the A39V/N53P/V55L Fyn SH3 domain by using ^{15}N relaxation dispersion NMR spectroscopy (7). Dispersion profiles were recorded as a function of temperature and could be fit to a two-state exchange process, I $\leftrightarrow$ F, with the intermediate populated to $\approx$2%, at temperatures below $\approx$25°C. At higher tem-

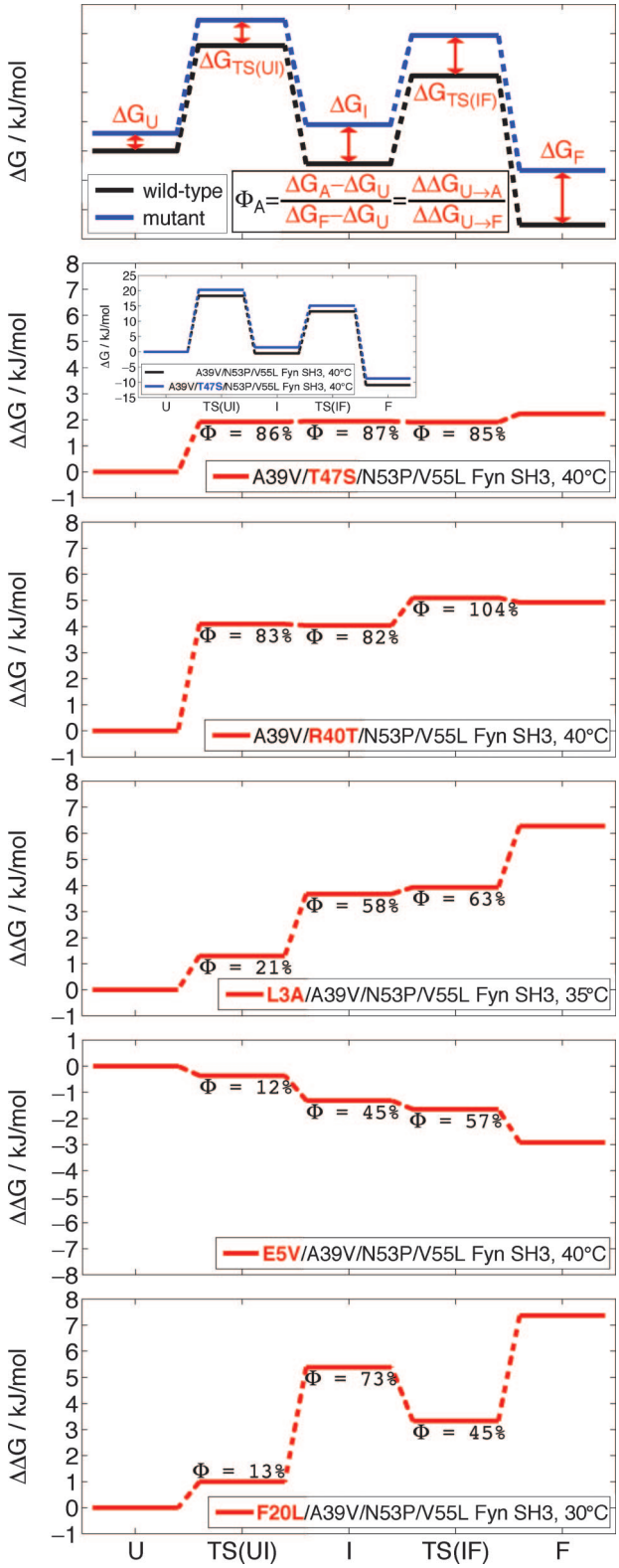

Fig. 2. Changes in ΔG used to calculate Φ_A (*Top*), along with changes in free energy upon point mutation, $\Delta\Delta G$, along the folding pathway for several mutants of the A39V/N53P/V55L Fyn SH3 domain based on the parameters of SI Table 3. Values of G are referenced with respect to the unfolded state U that is arbitrarily assigned a value of 0. TS(UI) and TS(IF) denote the rate-limiting transition states between states U, I and I, F respectively. *Inset* for the A39V/**T47S**/N53P/V55L mutant shows the pair of ΔG profiles (pseudo-wt and mutant) from which $\Delta\Delta G$ values are obtained, with the free energies of U states both assigned arbitrarily to 0, which is strictly true only if $\Delta G_U = 0$.

peratures, the population of U increases so that profiles are only well fit by a three-state model, U ↔ I ↔ F, and this model was shown to be superior to either of I ↔ U ↔ F or U ↔ F ↔ I. The fact that temperature can be used as a variable to "simplify" the analysis is an attractive feature, especially when the folding kinetics of several mutants are to be analyzed. We have chosen, therefore, to use A39V/N53P/V55L Fyn SH3 as the "pseudo-wt background" in which to place additional mutations for relaxation-dispersion-based Φ-value analyses. Interestingly, in the presence of denaturant, I and F equilibrate too fast to allow detection of the intermediate in stopped-flow fluorescence experiments. Using the rate constants for folding of the A39V/N53P/V55L Fyn SH3 domain obtained from NMR studies, we have shown that the stopped-flow measurements in fact probe the TS connecting U and I (7). This finding further emphasizes the unique role of relaxation dispersion NMR in the study of the complete folding pathway of this system.

In the present work, we have focused on mutations that are known from previous stopped-flow studies to modulate the stability of the wt Fyn SH3 domain by ≈2 kJ/mol ≤ $|\Delta\Delta G_{U\to F}|$ ≤ 6 kJ/mol (20, 21) [supporting information (SI) Table 1]. Values of $|\Delta\Delta G_{U\to F}|$ in this range ensure that the stability change is large enough for extraction of meaningful Φ values (11), while keeping structural perturbations to a minimum. Several SH3 domains have been characterized extensively by Φ-value analysis by using stopped-flow methods under the assumption of a two-state folding equilibrium, U ↔ F (20–23). These studies suggest that formation of the β_3–β_4-hairpin (Fig. 1) is the earliest event in folding, and it indeed appears to be fully formed in I based on ^{15}N relaxation dispersion measurements (7). Therefore, we examined the mutations **R40T** and **T47S** (for clarity, mutations used for Φ-value analysis are highlighted in boldface) of two surface-exposed side chains opposite each other on strands β_3 and β_4 (see Fig. 1) that have been used in a previous study to probe β-sheet propensity (21). The mutations, **L3A** and **E5V**, of two neighboring surface-exposed side chains in strand β_1 were analyzed to probe formation of the terminal β-sheet and nonnative interactions in this region, noted in the previous relaxation dispersion study of A39V/N53P/V55L Fyn SH3 (7). Finally, the mutation **F20L** was made to investigate hydrophobic core packing along the folding pathway of the domain (20). The positions of each of the sites of mutation are illustrated in Fig. 1.

Measurement and Interpretation of Φ Values for A39V/N53P/V55L Fyn SH3. Relaxation dispersion profiles were recorded and analyzed for each of the mutants mentioned above (SI Fig. 4). In addition to the four exchange rates obtained in the analysis of dispersion profiles according to the

$$\mathrm{U} \underset{k_{IU}}{\overset{k_{UI}}{\rightleftharpoons}} \mathrm{I} \underset{k_{FI}}{\overset{k_{IF}}{\rightleftharpoons}} \mathrm{F}$$

exchange model, a pair of chemical-shift differences, $\Delta\varpi_{F\to I} = \varpi_I - \varpi_F$ and $\Delta\varpi_{F\to U} = \varpi_U - \varpi_F$ also are obtained that, of course, are site-specific. Consistent with previous results for A39V/N53P/V55L Fyn SH3 (7), the low-temperature dispersion profiles (15–20°C) of each of the mutants examined in the present study could be fit to a two-state exchange process, I ↔ F, from which $\Delta\varpi_{F\to I}$ was determined, although for **L3A** and **F20L** it was clear that even at these temperatures there was a nonnegligible population of the U state (SI Table 2). The values of $\Delta\varpi_{F\to I}$ subsequently were used in fits of dispersion data at higher temperatures where the U ↔ I ↔ F model was required. The fits of the dispersion profiles for the mutants (SI Table 2 and SI Fig. 4) were of similar quality to those obtained in the analysis of A39V/N53P/V55L Fyn SH3 (7), except for the **F20L** mutant, in which the slightly poorer agreement between dispersions pre-

dicted by the model and the experimental data (SI Table 2) may reflect the lower quality of the dispersion spectra for this variant (SI Fig. 5).

As expected, the three-state folding kinetics obtained from fitting the CPMG data (SI Table 2) are sufficiently precise to allow calculation of well defined Φ values (SI Table 3), which are listed at the temperature where they could be extracted with the highest reliability and precision for each of the five mutants (Fig. 2). As described above and elsewhere (11), values of Φ that are close to either 0 or 1 are most easily interpreted; Φ values >80% were measured for TS(UI), I, and TS(IF) for the mutations **T47S** and **R40T** (SI Table 3 and Fig. 2), which modify a surface-exposed side chain in strand β_4 and β_3, respectively (Fig. 1). These values support our working model of the A39V/N53P/V55L Fyn SH3 folding pathway based on relaxation dispersion measurements where ^{15}N chemical shifts of residues comprising the β_3–β_4-hairpin in the I and F states are similar, suggesting that this structure already is formed in I (7) (SI Fig. 6). The results also are consistent with previous stopped-flow experiments on the Fyn SH3 domain that are interpreted in terms of a two-state folding model and probe the transition state corresponding to TS(UI) (ref. 21 and see above) where $\beta_{TS(UF)}$ = 67% and 86% were obtained for Thr-47 and Arg-40, respectively (mutants in the wt background). Unfortunately, Φ values could not be measured by stopped-flow fluorescence for any of the mutants used in the present study that are in the A39VN53P/V55L background because the dead time of our stopped-flow instrument precluded collection of a sufficient number of data points along the folding arm of the chevron plot. The high $\Phi_{TS(IF)}$ values for both A39V/**T47S**/N53P/V55L and A39V/**R40T**/N53P/V55L Fyn SH3 demonstrate that the β-hairpin region remains native-like in TS(IF). Thus, the results from the present study establish that the β_3–β_4 structure forms early on in the folding pathway and is maintained throughout.

To analyze the formation of structure during folding for residues comprising strand β_1 in the folded protein (Fig. 1), a pair of mutations, **L3A** and **E5V**, were made. Position 3 is surface-exposed in the Fyn SH3 domain folded structure (Fig. 1), and **L3A** is considered a nondisruptive deletion recommended for Φ value analysis (1, 11). Nevertheless, the reduction in stability of the native conformation that accompanies the **L3A** mutation is complex because $\Delta\Delta G_{U\rightarrow F}$ = 6.3 kJ/mol (SI Table 3 *Middle*) derives from a combination of the decrease in β-sheet propensity that is predicted to be destabilizing by ≈1.9 kJ/mol (table 17.3 in ref. 1), and the deletion of the equivalent of approximately two buried methylene groups that form favorable van der Waals' interactions in the F state (24). The situation is even more complex for the surface mutation **E5V**. Here the net stabilization by $\Delta\Delta G_{U\rightarrow F}$ = −2.9 kJ/mol (SI Table 3 *Top*) involves contributions from the enhancement in β-sheet propensity, predicted to be −2.9 kJ/mol (table 17.3 in ref. 1), the deletion of a negative charge from a domain with a net negative charge, and the removal of a salt bridge with Lys-25, noted in the crystal structure of a Fyn SH3 mutant (N. Neculai, A.Z.-A., P. Howell, and A.R.D., unpublished data). In addition, although the x-ray structure of E5V Fyn SH3 (*Homo sapiens* Fyn SH3) shows that the additional aliphatic surface that accompanies the mutation, equivalent to approximately one methylene group, is solvent-exposed in F (24), and thus solvation effects should not contribute to $\Delta\Delta G_{U\rightarrow F}$ (assuming that this position is fully solvated in U), the mutation may nevertheless stabilize potential nonnative hydrophobic contacts in the intermediate as observed previously (7). As expected for mutations modulating several different interactions, **L3A** and **E5V** yield fractional Φ values (SI Table 3 and Fig. 2); nevertheless, we have made a cautious attempt to interpret these values within the framework of our model of folding derived from previous studies (7). The low values of $\Phi_{TS(UI)}$ of 21% and 12% for **L3A**/A39V/N53P/V55L and **E5V**/A39V/N53P/V55L Fyn SH3, respectively, provide strong evidence that strand β_1 is not formed in the first transition state. This result is consistent with studies that have been performed on the single-point mutants **L3A** and **E5V** of the wt Fyn SH3 domain by using stopped-flow fluorescence where values of $\Phi_{TS(UF)}$ = 0% and 38% were obtained, respectively (SI Fig. 7 and SI Table 1). As described above, we have shown previously, based on the kinetic constants isolated from dispersion measurements of A39V/N53P/V55L Fyn SH3, that stopped-flow studies on this domain effectively probe structure in the first transition state, TS(UI) (7). The similarity between $\Phi_{TS(UI)}$ values obtained in the present study, where each of the states along the pathway is probed individually, and the $\Phi_{TS(UF)}$ values reported from stopped-flow fluorescence further supports our interpretation.

Our previous NMR study of the folding pathway of A39V/N53P/V55L Fyn SH3 provided strong evidence based on ^{15}N chemical shifts that residues in strands β_1 and β_5 were stabilized in the intermediate state by nonnative long-range interactions. Moreover, the temperature-dependent kinetic data established a heat-capacity difference between states I and F, ΔC_p(F→I), of ≈0 J/mol·K, implying a collapsed intermediate state (7). In this context, values of Φ_I = 58% and Φ_I = 45% for **L3A**/A39V/N53P/V55L and **E5V**/A39V/N53P/V55L Fyn SH3, respectively, are consistent with the fact that strand β_1 still is not formed in the intermediate, whereas extensive and partly nonnative hydrophobic interactions are. As has been described previously, a complication in the interpretation of fractional Φ values is that they can arise because of multiple folding pathways (1, 11). The similar Φ_I (and $\Phi_{TS(IF)}$, see below) values observed for the **L3A** and **E5V** mutants here, which are expected to destabilize (**L3A**) and stabilize (**E5V**) hydrophobic interactions, respectively, argue against the case of parallel folding pathways that generate states with (path A) or without (path B) significant hydrophobic interactions involving these sites. If this were the case, Φ_I (and $\Phi_{TS(IF)}$) for the **E5V** mutant would be expected to be larger than the corresponding value for **L3A** (25). This interpretation must be considered tentative given the small amount of data; nevertheless, the fractional Φ values most likely reflect contributions to stability that arise from several interactions that are perturbed differentially by mutation.

The transition state connecting I and F has a larger heat capacity than the states themselves, ΔC_p(F→TS(IF)) = 1.46 kJ/mol·K, which has been interpreted as reflecting the breaking of nonnative interactions along the folding pathway that leads to a transition state that is transiently hydrated before dehydration and repacking to the native structure (7). The values $\Phi_{TS(IF)}$ = 63% and 57% for **L3A**/A39V/N53P/V55L and **E5V**/A39V/N53P/V55L Fyn SH3, respectively (SI Table 3 and Fig. 2), are consistent with this model and may reflect formation of strand β_1 with only limited tertiary interactions in TS(IF), as predicted by computer simulations (26). This "picture" also is supported by the measured Φ values for the hydrophobic core mutant **F20L**/A39V/N53P/V55L (SI Table 3 *Bottom* and Fig. 2). The low $\Phi_{TS(UI)}$ value (13%), which is similar in magnitude to $\Phi_{TS(UF)}$ = −1% obtained from a stopped-flow study of **F20L** Fyn SH3 (20), suggests that this position is not well structured in the first transition state. The high value for Φ_I = 73% and the intermediate value for $\Phi_{TS(IF)}$ = 45% are in keeping with expectations based on a collapsed intermediate with nonnative interactions connecting the central sheet β_2–β_3–β_4 with the rest of the domain that must then be broken during the second transition state, TS(IF), before formation of the native conformation. Further evidence for nonnative interactions in the intermediate is provided by the ^{15}N chemical-shift changes in the I state that accompany the **F20L** mutation. The largest changes in shifts (other than at positions 20 and 21) all map to strand β_1 (Leu-3, Glu-5, Ala-6, and Leu-7) (SI Figs. 8 and 9), which, as mentioned

above, has been shown previously to form hydrophobic nonnative interactions in the intermediate state (7). Conversely, the largest chemical-shift changes that accompany the **E5V** mutation include those from Leu-18 and Ser-19. In the structure of the native state, position 20 does not make contacts with the first strand (24).

Probing the Effects of Mutation on Structure. A big advantage of NMR relaxation dispersion spectroscopy in the study of protein folding is the availability of chemical shifts to monitor the formation of structure as folding proceeds. In the context of Φ-value analysis, a comparison of chemical shifts can be used to ascertain potential changes to the structure(s) of F, I, and U states as a result of mutation. It is expected that there will be a chemical-shift change at the site of mutation that arises purely from the substituted side chain, and this effect can be minimized by carrying out comparisons on the basis of secondary shifts, i.e., the differences $\varpi_X - \varpi_{RC}$ of the measured chemical shifts for state X, ϖ_X, with respect to random-coil chemical shifts, ϖ_{RC} (27). Inspection of the NOE patterns used for sequence-specific resonance assignment provided no indications for major rearrangements of the native structure caused by any of the point mutations **L3A**, **E5V**, **F20L**, **R40T**, or **T47S**. Accordingly, the ^{15}N secondary shifts of the F states of the mutants are very close to those of the F state of the A39V/N53P/V55L Fyn SH3 domain (rmsd $= \frac{1}{N}\sqrt{\Sigma\{(\varpi_X - \varpi_{RC})_{mut} - (\varpi_X - \varpi_{RC})_{wt}\}^2} = 0.70$ ppm for **F20L**, which modulates hydrophobic core packing including ring-current effects (28), whereas rmsd < 0.25 ppm for the other mutations; in all cases, the mutated residue and its successor are excluded). There is one notable exception: the mutation **R40T** in strand β_3 causes large changes in the ^{15}N chemical shifts of Thr-47 and Gly-48 on the neighboring strand β_4 (Fig. 1) toward lower field (Fig. 3 *Top* and SI Fig. 10). Such shifts are highly indicative of a change in side-chain conformation of Thr-47 modulating the so-called χ_1 effect on the ^{15}N chemical shifts of the neighboring backbone amides (29). The NOE patterns involving Thr-47 Hβ and Hγ2* of A39V/N53P/V55L Fyn SH3 (data not shown) are inconsistent with the gauche$^-$ ($\chi_1 \approx -60°$) rotamer found in the crystal structure of *H. sapiens* Fyn SH3 (24); in solution, the side chain of Thr-47 preferentially populates the gauche$^+$ rotamer ($\chi_1 \approx +60°$). The point mutation **R40T**, however, shows a strongly altered NOE pattern now favoring the gauche$^-$ rotamer, which rotates the hydroxyl group of Thr-47 toward the newly introduced hydroxyl group of Thr-40, most likely to form an interstrand hydrogen bond. Interestingly, the large ^{15}N shifts for Thr-47 and Gly-48 to lower field also are observed for the I state (Fig. 3 *Middle*), providing very strong evidence that, in the intermediate, Thr-47 interacts with sequence position 40 as it does in the native conformation and, hence, that the β_3–β_4-hairpin already is formed.

The secondary ^{15}N chemical shifts for the unfolded states of all mutants also are extremely close to those of the unfolded A39V/N53P/V55L Fyn SH3 domain (rmsd values between 0.19 ppm and 0.44 ppm). That the effects of mutations introduce little perturbation to the U state structural ensemble can be appreciated by comparing the rmsd of secondary chemical shifts over the complete sequence (excluding the site of mutation and the successive residue) with the corresponding rmsd values obtained for a seven-residue window that surrounds, but does not include, the site of mutation. For **L3A**, **E5V**, **F20L**, **R40T**, and **T47S**, average rmsd values over the whole sequence are 0.27, 0.33, 0.44, 0.37, and 0.20 ppm, whereas averaged over the seven-residue window the rmsd values are 0.56, 0.59, 0.45, 0.46, and 0.26 ppm. Given the typical range of ^{15}N chemical shifts upon formation of secondary and tertiary structure (29, 30), the small differences observed in rmsd values suggest strongly that even if there is minor residual structure in the U state ensemble, it is perturbed very little by mutation. As for the folded and unfolded states, the small ^{15}N chemical-shift changes of the

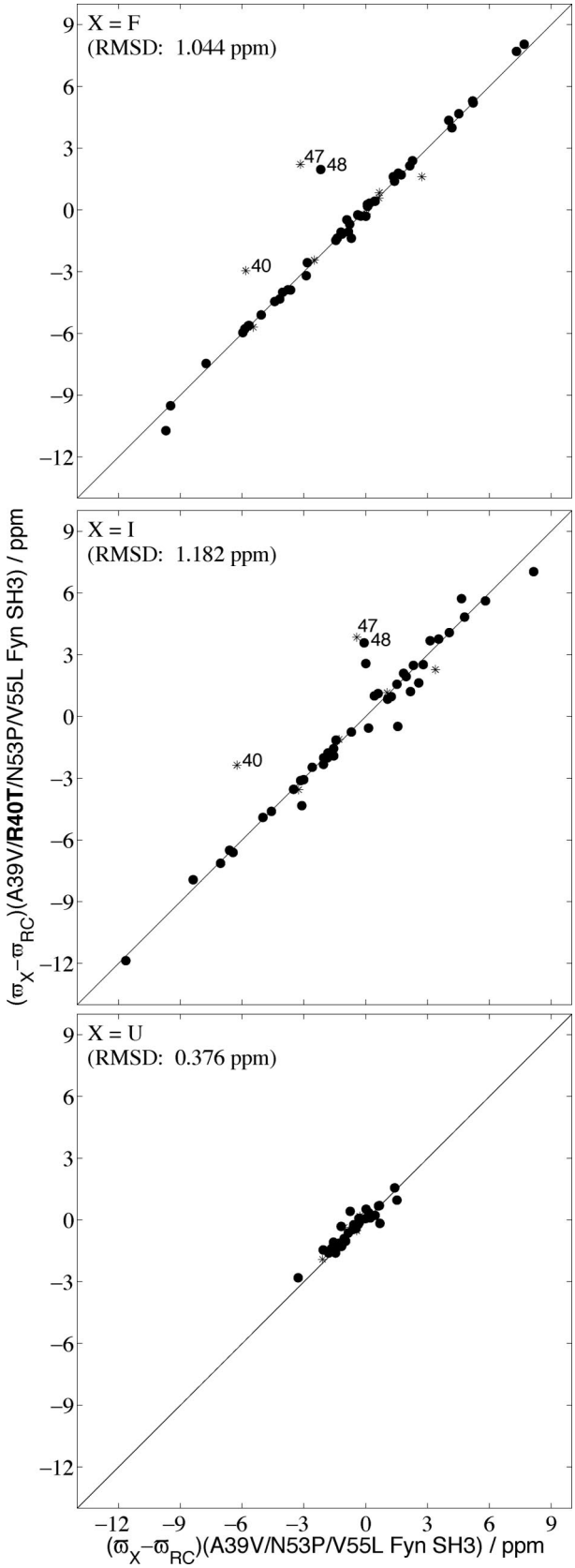

Fig. 3. Comparison of secondary ^{15}N chemical shifts ($\varpi_X - \varpi_{RC}$; ref. 27, see text) of A39V/**R40T**/N53P/V55L Fyn SH3 and A39V/N53P/V55L Fyn SH3 for the folded (*Top*), intermediate (*Middle*), and unfolded (*Bottom*) states. Residues in the immediate vicinity of the mutated side chain in the native structure (Fig. 1) are indicated by asterisks.

BIOPHYSICS

intermediate (rmsd values between 0.59 ppm and 0.79 ppm if the mutated residue and its successor are excluded, plus residues Thr-2 to Leu-7 for **F20L** or Thr-47 and Gly-48 for **R40T**; see above) verify that the structure of the I state is essentially unperturbed by the point mutations. In summary, the chemical shifts extracted from a relaxation dispersion analysis provide a powerful probe of structural changes in states U, I, and F that are introduced by the point mutations and that are an obvious concern in Φ-value analysis. In cases where shift changes are small, such as for the mutants considered in the present study, this concern is minimized. Clearly, the site-specific chemical shifts are superior to probes of structural integrity obtained from other spectroscopies such as CD or tryptophan fluorescence, where averaged structural measures are obtained, or measures from only one or two sites. Chemical shifts also can be used to provide qualitative information about whether the mutation changes the folding pathway. For example, the similar shifts for the I states of pseudo-wt and the mutants observed here (see Fig. 3 and SI Figs. 6, 8, 9, and 11) argue that the folding pathway is not significantly perturbed, which facilitates interpretation of the Φ values markedly.

In summary, we have introduced a relaxation-dispersion-based Φ-value analysis that allowed a residue-specific structural investigation of the complete three-state folding pathway of the A39V/N53P/V55L Fyn SH3 domain, including the pair of transition states that connect U with F. As with other relaxation dispersion analyses of three-state folding (4–7, 31) chemical-shift information can be used to obtain insight into the structure of the intermediate state. Φ values provide additional data on the I state and also probe structure formation in both transition states that previously only could be measured very qualitatively through temperature-dependent relaxation dispersion studies. The results of the present study support and extend the notion that nonnative interactions play a role in the stabilization of the on-pathway intermediate state in the folding of the Fyn SH3 domain (7). Finally, the work significantly extends previous stopped-flow fluorescence Φ-value analyses of Fyn SH3 that were based on the assumption of two-state folding, where the measured Φ values are those associated with the first transition state, TS(UI) (7). The combination of Φ-value analysis and relaxation dispersion NMR spectroscopy promises to greatly increase the level of detail that can be obtained in the study of exchange processes that include, but are not limited to, protein folding.

Materials and Methods

Details of sample preparation and CD melting are given in *SI Materials and Methods*. Sequence-specific ^{1}H and ^{15}N resonance assignments and the integrity of the native structure for all *Gallus gallus* Fyn SH3 domain mutants were confirmed by using [^{1}H,^{15}N]-NOESY-heteronuclear single-quantum coherence (HSQC) spectra (32) recorded at 20°C and processed as described previously (7). ^{15}N single-quantum CPMG relaxation dispersion experiments (33, 34) were recorded, processed, and quantified as described previously (7). Taking into account the different stabilities of the individual Fyn SH3 domain mutants, data sets were recorded at the following temperatures by using both 500- and 800-MHz NMR spectrometers: (*i*) A39V/**R40T**/N53P/V55L and A39V/**T47S**/N53P/V55L at 20°C, 35°C, and 40°C; (*ii*) **E5V**/A39V/N53P/V55L at 20°C, 35°C, 40°C, and 45°C; and (*iii*) **L3A**/A39V/N53P/V55L and **F20L**/A39V/N53P/V55L at 15°C, 20°C, 25°C, 30°C, and 35°C. The relaxation dispersion profiles at all temperatures were fit together to extract global exchange parameters ($k_{ex,FI} = k_{FI} + k_{IF}$, $k_{ex,IU} = k_{IU} + k_{UI}$, and the populations of states I, p_I and U, p_U) for each temperature along with residue-specific values ($\Delta\varpi_{F\rightarrow I}$ and $\Delta\varpi_{F\rightarrow U}$, both assumed to be temperature-invariant and intrinsic ^{15}N relaxation rates for all temperatures) by numerical nonlinear least-squares fitting using in-house written software as described previously (4, 5, 7). Errors of the fitted parameters were calculated from the covariance matrix (35). Absolute signs of $\Delta\varpi_{F\rightarrow I}$ and $\Delta\varpi_{F\rightarrow U}$ were inferred from the ^{15}N chemical shifts of A39V/N53P/V55L Fyn SH3 (7), which is justified by the excellent correlation between shifts of the pseudo-wt (A39V/N53P/V55L Fyn SH3) and mutants (Fig. 3 and SI Fig. 8). Absolute signs of $\Delta\varpi_{F\rightarrow I}$ in **L3A**/A39V/N53P/V55L Fyn SH3 were confirmed experimentally from differences in ^{15}N peak positions at 15°C (where the dispersion profiles are dominated by the $I \leftrightarrow F$ exchange) in [^{1}H,^{15}N]-heteronuclear multiple-quantum coherence (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 and 800 MHz (36), as described previously (7).

This work was supported by a grant from the Canadian Institutes of Health Research (to L.E.K.). P.N. acknowledges postdoctoral support from the Deutsche Forschungsgemeinschaft (NE 1197/1-2) and the Canadian Institutes of Health Research. A.Z.-A. was supported by a doctoral Canada Graduate Scholarship (CGS-D3) from the Natural Sciences and Engineering Research Council of Canada. L.E.K. holds a Canada Research Chair in Biochemistry.

1. Fersht A (1999) *Structure and Mechanism in Protein Science* (Freeman, New York).
2. Brockwell DJ, Radford SE (2007) *Curr Opin Struct Biol* 17:30–37.
3. Palmer AG (2004) *Chem Rev* 104:3623–3640.
4. Korzhnev DM, Salvatella X, Vendruscolo M, Di Nardo AA, Davidson AR, Dobson CM, Kay LE (2004) *Nature* 430:586–590.
5. Korzhnev DM, Neudecker P, Mittermaier A, Orekhov VY, Kay LE (2005) *J Am Chem Soc* 127:15602–15611.
6. Korzhnev DM, Neudecker P, Zarrine-Afsar A, Davidson AR, Kay LE (2006) *Biochemistry* 45:10175–10183.
7. Neudecker P, Zarrine-Afsar A, Choy W-Y, Muhandiram DR, Davidson AR, Kay LE (2006) *J Mol Biol* 363:958–976.
8. Matthews CR (1987) *Methods Enzymol* 154:498–511.
9. Goldenberg DP, Frieden RW, Haack JA, Morrison TB (1989) *Nature* 338:127–132.
10. Fersht AR, Matouschek A, Serrano L (1992) *J Mol Biol* 224:771–782.
11. Fersht A, Sato S (2004) *Proc Natl Acad Sci USA* 101:7976–7981.
12. Capaldi AP, Kleanthous C, Radford SE (2002) *Nat Struct Biol* 9:209–216.
13. Matouschek A, Serrano L, Fersht AR (1992) *J Mol Biol* 224:819–835.
14. Khan F, Chuang JI, Gianni S, Fersht AR (2003) *J Mol Biol* 333:169–186.
15. Nölting B, Golbik R, Fersht AR (1995) *Proc Natl Acad Sci USA* 92:10668–10672.
16. Nölting B, Golbik R, Neira JL, Soler-Gonzalez AS, Schreiber G, Fersht AR (1997) *Proc Natl Acad Sci USA* 94:826–830.
17. Neudecker P, Korzhnev DM, Kay LE (2006) *J Biomol NMR* 34:129–135.
18. Eyring H (1935) *Chem Rev* 17:65–77.
19. Cho J-H, Raleigh DP (2006) *J Am Chem Soc* 128:16492–16493.
20. Northey JGB, Di Nardo AA, Davidson AR (2002) *Nat Struct Biol* 9:126–130.
21. Zarrine-Afsar A, Dahesh S, Davidson AR (August 9, 2007) *J Mol Biol*, 10.1016/j.jmb.2007.07.059.
22. Martinez JC, Serrano L (1999) *Nat Struct Biol* 6:1010–1016.
23. Riddle DS, Grantcharova VP, Santiago JV, Alm E, Ruczinski I, Baker D (1999) *Nat Struct Biol* 6:1016–1024.
24. Noble MEM, Musacchio A, Saraste M, Courtneidge SA, Wierenga RK (1993) *EMBO J* 12:2617–2624.
25. Sosnick TR, Krantz BA, Dothager RS, Baxa M (2006) *Chem Rev* 106:1862–1876.
26. Guo W, Lampoudi S, Shea J-E (2003) *Biophys J* 85:61–69.
27. Wishart DS, Bigam CG, Holm A, Hodges RS, Sykes BD (1995) *J Biomol NMR* 5:67–81.
28. Wüthrich K (1986) *NMR of Proteins and Nucleic Acids* (Wiley, New York).
29. Wishart DS, Case DA (2001) *Methods Enzymol* 338:3–34.
30. Wishart DS, Sykes BD, Richards FM (1991) *J Mol Biol* 222:311–333.
31. Grey MJ, Tang Y, Alexov E, McKnight CJ, Raleigh DP, Palmer AG (2006) *J Mol Biol* 355:1078–1094.
32. Zhang O, Kay LE, Olivier JP, Forman-Kay JD (1994) *J Biomol NMR* 4:845–858.
33. Loria JP, Rance M, Palmer AG (1999) *J Am Chem Soc* 121:2331–2332.
34. Tollinger M, Skrynnikov NR, Mulder FA, Forman-Kay JD, Kay LE (2001) *J Am Chem Soc* 123:11341–11352.
35. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) *Numerical Recipes in C* (Cambridge Univ Press, Cambridge, UK), 2nd Ed.
36. Skrynnikov NR, Dahlquist FW, Kay LE (2002) *J Am Chem Soc* 124:12352–12360.