Originally published 20 April 2012; corrected 23 April 2012

www.sciencemag.org/cgi/content/full/336/6079/362/DC1

Supplementary Materials for

Structure of an Intermediate State in Protein Folding and Aggregation

Philipp Neudecker, Paul Robustelli, Andrea Cavalli, Patrick Walsh, Patrik Lundström, Arash Zarrine-Afsar, Simon Sharpe, Michele Vendruscolo, Lewis E. Kay*

*To whom correspondence should be addressed. E-mail: kay@pound.med.utoronto.ca

Published 20 April 2012, *Science* **336**, 362 (2012)
DOI: 10.1126/science.1214203

This PDF file includes:

Correction: Some equations did not display properly in the original file. They have been corrected here.

Supporting Online Material
Materials, Methods, and Results
Sample preparation

The A39V/N53P/V55L *G. gallus* Fyn SH3 domain (*19*) was expressed from a modified pET vector carrying an NH_2-terminal hexahistidine tag followed by a TEV protease cleavage site. [U-^{2}H,^{15}N] and [U-^{2}H,^{13}C,^{15}N] enriched A39V/N53P/V55L Fyn SH3 domain was obtained by IPTG-induced overexpression in *E. coli* BL21(DE3) in 1 l M9 minimal medium with 1.5 g $^{15}NH_4Cl$ and 3.0 g D-glucose-$^{12}C_6$-$^{2}H_7$ or D-glucose-$^{13}C_6$-$^{2}H_7$, respectively, in 99.9% D_2O at 37°C for 3 to 4 h. In addition, a sample of [U-50% ^{2}H;U-^{13}C,^{15}N] A39V/N53P/V55L Fyn SH3 was expressed in 1 l M9 with 1.5 g $^{15}NH_4Cl$ and 3.0 g D-glucose-$^{13}C_6$-$^{2}H_7$ in 50.0% D_2O (*33*). Additional samples were expressed in 1 l H_2O-based M9 using 1.5 g $^{15}NH_4Cl$ as the nitrogen source and either 3.0 g D-glucose-1-^{13}C or 3.0 g D-glucose-2-^{13}C as the carbon source (*34*); these samples are referred to as [1-^{13}C-glc;U-^{15}N] or [2-^{13}C-glc;U-^{15}N] A39V/N53P/V55L Fyn SH3 in what follows. The expression product was purified using nickel affinity chromatography and refolded by dialysis as described previously (*35*), followed by cleavage with histidine-tagged TEV protease at room temperature overnight, removal of cleaved hexahistidine-tag and TEV protease by nickel affinity chromatography, and final purification by size-exclusion chromatography. The TEV-cleaved construct comprises an unstructured seven-residue NH_2-terminal cloning artifact, the A39V/N53P/V55L mutated SH3 domain from Thr2 (Thr84 of full-length *G. gallus* Fyn) to Asp59, and the COOH-terminal residue Arg60. Site-directed variants of the Fyn SH3 domain were constructed using the QuikChange Site-Directed Mutagenesis Kit from Stratagene (Agilent, Mississauga, ON, Canada) according to the manufacturer's instructions or were custom-made by GenScript (Piscataway, NJ, USA). Samples of [U-^{13}C,^{15}N] A39V/N53P/V55L/Δ(57-60) Fyn SH3, [U-^{13}C,^{15}N] Δ(56-60) Fyn SH3 and [U-^{13}C,^{15}N] Δ(57-60) Fyn SH3 were expressed in 1 l M9 minimal medium with 1.5 g $^{15}NH_4Cl$ and 3.0 g D-glucose-$^{13}C_6$ in H_2O and purified as described above; the yield for these aggregation-prone Fyn SH3 truncation mutants could be improved by performing the overexpression at 25°C for about 10 h, as well as the refolding and subsequent purification steps at 4°C. NMR samples contained between 0.5 mM and 1.2 mM isotope-enriched Fyn SH3 mutant, 0.2 mM EDTA, 0.05% NaN_3, 50 mM sodium phosphate (pH 7.0), in either H_2O/D_2O (9:1) or D_2O.

Sequence-specific NMR resonance assignments for the native state

The sequence-specific ^{1}H and ^{15}N resonance assignments for the native A39V/N53P/V55L Fyn SH3 domain at 25°C, generated from ^{15}N-edited experiments reported previously (*19*), were verified and transferred to 20°C with a [^{1}H, ^{15}N]-NOESY-HSQC spectrum (*36*; 150 ms mixing time) recorded on a sample of 0.7 mM [U-^{15}N] A39V/N53P/V55L Fyn SH3, as described in more detail previously (*19*). Sequence-specific backbone ^{13}C resonance assignments were obtained from HNCO and HNCA spectra (*37*) recorded on a sample of 1.0 mM [U-50% ^{2}H;U-^{13}C,^{15}N] A39V/N53P/V55L Fyn SH3 at 20°C on a Varian Unity INOVA 500 MHz NMR spectrometer. [^{1}Hα, ^{13}Cα] correlations were verified with a [^{1}H, ^{13}C]-CTHSQC spectrum (*38*) recorded on a sample of 1.0 mM [U-50% ^{2}H;U-^{13}C,^{15}N] A39V/N53P/V55L Fyn SH3 in D_2O at 20°C on a Varian Unity INOVA 800 MHz NMR spectrometer. The NMR spectrometers were equipped with room-temperature probes with z-axis and triple-axis pulsed field gradient capabilities, respectively. Methyl group ^{1}H and ^{13}C assignments were obtained as reported previously

(*39*, *40*). The H_2O resonance was suppressed by gradient coherence selection with quadrature detection in the indirect ^{13}C and ^{15}N dimensions achieved by States-TPPI (*41*) and the echo-antiecho method (*42*, *43*), respectively. All NMR spectra were processed with NMRPipe (*44*) software and analyzed with NMRViewJ (*45*). ^{1}H chemical shifts were referenced with respect to external DSS in D_2O, ^{13}C and ^{15}N chemical shifts were referenced indirectly (*46*). ^{15}N chemical shifts were corrected for ^{2}H isotope shifts as appropriate (*47*). These experiments resulted in unambiguous assignment of all ^{1}H, ^{15}N, ^{13}CO, and $^{13}C\alpha$ backbone resonances from $^{1}H\alpha$,$^{13}C\alpha$ of Ala-4 to $^{1}H\alpha$,$^{13}C\alpha$ of Arg60 (Figure S3A), with the exception of Ile50 ^{13}CO, Pro51 ^{15}N, Ser52 ^{13}CO, Pro53 ^{15}N, Ala56 ^{13}CO, Pro57 ^{15}N, which are not observable in these experiments. In addition, the $^{1}H\varepsilon1$-$^{15}N\varepsilon1$ indole resonances of the two Trp residues in the protein (Trp36 and Trp37), the side-chain NH_2 groups of Gln-1, Gln27, and Asn30, the $^{1}H\varepsilon$-$^{15}N\varepsilon$ guanidinium resonances of the three Arg residues (Arg13, Arg40, and Arg60), most of the side-chain ^{1}H and several side-chain ^{13}C resonances, as well as the ^{1}H and ^{13}C resonances of all 34 methyl groups (*39*, *40*) could be assigned unambiguously. ^{1}H, ^{13}C, and ^{15}N chemical shifts of the native A39V/N53P/V55L Fyn SH3, along with further experimental details, have been deposited with the BioMagResBank (access code: 17149).

Sequence-specific ^{1}H, ^{13}C, and ^{15}N resonance assignments for the native monomeric state of the A39V/N53P/V55L/Δ(57-60) Fyn SH3 domain were obtained from [^{1}H, ^{15}N]-HSQC (*36*), [^{1}H, ^{15}N]-TOCSY-HSQC (*36*) with a DIPSI-2rc mixing scheme (*48*; 60 ms mixing time), [^{1}H, ^{15}N]-NOESY-HSQC (*36*; 150 ms mixing time), HNCO (*37*), and HNCA (*37*) spectra recorded on a sample of 0.6 mM [U-^{13}C,^{15}N] A39V/N53P/V55L/Δ(57-60) Fyn SH3 at 10°C (500 MHz). In addition, a [^{1}H, ^{13}C]-CTHSQC spectrum (*38*) was recorded on the same sample at 10°C on a Varian Unity INOVA 600 MHz NMR spectrometer equipped with a cryogenic probe with z-axis pulsed field gradient capabilities. Despite severe line-broadening due to chemical exchange, these experiments allowed unambiguous assignment of all detectable backbone amide resonances; the backbone amide resonances of Leu3, Glu5, Leu7, Trp37, Ser52, and Ala56 were too line-broadened to be observed. In addition, the ^{13}CO resonances from Gly-5 to Tyr49 (except Thr2 to Leu7, Arg13, Trp36), the $^{13}C\alpha$ resonances from Gly-5 to Ile50 (except Leu3 to Leu7), the $^{1}H\alpha$ resonances from Gly-5 to Leu55 (except Phe4 to Ala6, Pro51 to Pro53), the $^{1}H\varepsilon1$-$^{15}N\varepsilon1$ indole resonances of the two Trp residues in the protein (Trp36 and Trp37), the side-chain NH_2 groups of Gln-1, Gln27, and Asn30, the $^{1}H\varepsilon$-$^{15}N\varepsilon$ guanidinium resonances of the two Arg residues (Arg13 and Arg40), most of the side-chain ^{1}H and several side-chain ^{13}C resonances - including 13 of the 32 methyl groups - could also be assigned unambiguously.

Experiments probing conformational exchange

Carr-Purcell-Meiboom-Gill relaxation dispersion experiments (*49*-*52*) were used to quantify chemical exchange in the A39V/N53P/V55L Fyn SH3 domain. ^{15}N SQ CW CPMG (*53*), ^{15}N TROSY/AntiTROSY CPMG (*54*), ^{1}HN SQ CPMG (*55*), and ^{1}HN/^{15}N ZQ/DQ CPMG (*56*) RD experiments (Fig. S1, Fig. S4A,B) were recorded on a sample of 1.0 mM [U-^{2}H,^{15}N] A39V/N53P/V55L Fyn SH3 in H_2O/D_2O (9:1) at 20°C to study millisecond exchange processes affecting the backbone amide groups; data sets were obtained at both 500 and 800 MHz. ^{13}CO SQ CPMG (*57*) RD experiments (Fig. S4E) were recorded on a sample of 1.0 mM [U-50% ^{2}H;U-^{13}C,^{15}N] A39V/N53P/V55L Fyn SH3 at 20°C (500 and 800 MHz). $^{13}C\alpha$ SQ CPMG (*58*) RD experiments (Fig. S4C) for Ile/Leu

residues and for the remaining residues with the exception of Val were recorded on samples of 1.2 mM [1-^{13}C-glc;U-^{15}N] A39V/N53P/V55L Fyn SH3 and 1.0 mM [2-^{13}C-glc;U-^{15}N] A39V/N53P/V55L Fyn SH3, respectively, in D_2O at 20°C (600 and 800 MHz). ^{1}Hα SQ CPMG (*33*) RD experiments (Fig. S4D) were recorded on a sample of 1.0 mM [U-50% ^{2}H;U-^{13}C,^{15}N] A39V/N53P/V55L Fyn SH3 in D_2O at 20°C (600 and 800 MHz). For enhanced accuracy the ^{1}Hα SQ CPMG experiments were recorded in duplicate using two different elements that refocus ^{1}Hα-^{1}H scalar couplings (B and C in Fig. 2 of ref. *33*) and fit together (Fig. S4D). Methyl group ^{13}C SQ CPMG experiments were recorded on a sample of 1.0 mM [1-^{13}C-glc;U-^{15}N] A39V/N53P/V55L Fyn SH3 in H_2O/D_2O (9:1) and have been reported previously (*39*, *40*). In each CPMG experiment 14 to 17 different CPMG frequencies $\nu_{CPMG} = 1/(2\delta)$, where δ is the time between consecutive refocusing pulses, ranging up to 1000.0 Hz (2000.0 Hz for ^{1}HN SQ and ^{1}Hα SQ) were sampled during a constant-time relaxation interval of T_{CPMG} = 40 ms (30 ms for ^{1}HN SQ, ^{1}HN/^{15}N ZQ/DQ, and ^{13}Cα SQ; 24 ms for ^{1}Hα SQ).

Sequence-specific NMR resonance assignments for the intermediate state

Resonance intensities from CPMG RD experiments, $I(\nu_{CPMG})$, were quantified by three-way decomposition using MUNIN (*59*, *60*) 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 (*61*). Error estimates ΔR_{eff} for R_{eff} were obtained from duplicate measurements at 3 different ν_{CPMG} values as described previously (*62*), assuming a minimum relative error $\Delta R_{eff}/R_{eff}$ of 1.0% for ^{15}N SQ, ^{15}N TROSY/AntiTROSY, and ^{1}HN SQ, 2.0% for ^{13}CO SQ, ^{13}Cα SQ, and ^{1}Hα SQ, or 4.0% for ^{1}HN/^{15}N ZQ/DQ, and a minimum absolute error ΔR_{eff} of 0.2/s for ^{15}N SQ and ^{15}N TROSY/AntiTROSY, 0.3/s for ^{1}HN SQ, 0.4/s for ^{13}CO SQ and ^{13}Cα SQ, or 0.5/s for ^{1}Hα SQ to account for systematic experimental imperfections (*63*). All RD profiles for any given sample (note that the extent of ^{2}H enrichment as well as the ^{2}H content of the solvent have a minor influence on the folding equilibrium) were fit together to extract global exchange parameters along with residue specific values such as ^{1}HN, ^{15}N, ^{13}CO, ^{13}Cα, and ^{1}Hα chemical shift differences between exchanging states and intrinsic ^{1}HN, ^{15}N, ^{13}CO, ^{13}Cα, and ^{1}Hα relaxation rates, $R_{2,0}$. From the global fit an exchange rate of 836/s ± 4/s and an intermediate state population of 1.98% ± 0.01% was obtained for [U-^{2}H,^{15}N] A39V/N53P/V55L Fyn SH3 in H_2O/D_2O (9:1) at 20°C; corresponding values of 800/s (average) ± 98/s (standard deviation) and 1.94% ± 0.28% were obtained for the 25 amide groups with large ^{15}N relaxation dispersions ($R_{eff}(\nu_{CPMG})$ - R_{eff}(1000 Hz) > 4/s) fit on a per-residue basis. For enhanced accuracy, the ^{13}CO SQ RD profiles were fit together with ^{15}N SQ CW CPMG dispersion data (*53*) recorded on the same sample at 20ºC (500 MHz) as described above. In prior work we have shown that the population of the unfolded state is negligibly small below about 30°C so that RD data are well approximated by a 2-state model describing the exchange between native and intermediate states (*19*). Parameter extraction was therefore accomplished by a non-linear least-squares fitting procedure whereby experimental dispersion profiles, $R_{eff}(\nu_{CPMG}) = R_{2,0} + R_{ex}(\nu_{CPMG})$, at 20°C were fit to those calculated by deriving the evolution of magnetization during the CPMG interval by solving the Bloch-McConnell equations (*64*) numerically for a 2-site exchange model using in-house written software and procedures

described previously (*58*, *62*, *65*). Errors of the fitted parameters were calculated from the covariance matrix (*66*), an approach that has been shown to provide reasonable error estimates in a study similar to that presented here (*67*). Absolute signs of ^{1}HN, ^{15}N, ^{13}CO, ^{13}Cα, and ^{1}Hα chemical shift differences between the intermediate and the native state were determined from a comparison of resonance positions (derived from single quantum and multiple quantum coherences) recorded at different static magnetic field strengths (*68*), ^{1}HN/^{15}N ZQ/DQ CPMG RD profiles (*56*), and 1D off-resonance $R_{1\rho}$ experiments (*69*, *70*) as described in detail previously (*19*, *69-71*). The chemical shift for any given nucleus X of the intermediate state I was then calculated according to $\varpi_{X,I} = \varpi_{X,F} + \Delta\varpi_{X,F\rightarrow I}$ with an uncertainty of $s(\varpi_{X,I}) = s(\Delta\varpi_{X,F\rightarrow I})$, where $\varpi_{X,F}$ is the chemical shift of nucleus X in the native state and $\Delta\varpi_{X,F\rightarrow I} = \varpi_{X,I} - \varpi_{X,F}$ is the chemical shift difference extracted from the RD profiles. When the chemical shift difference was too small to allow reliable determination of its sign, $\varpi_{X,I}$ was assumed to be identical to $\varpi_{X,F}$ within a margin of error of $s(\varpi_{X,I}) = ((\Delta\varpi_{X,F\rightarrow I})^2 + (s(\Delta\varpi_{X,F\rightarrow I}))^2)^{1/2}$. A minimum error of 0.10 ppm (0.01 ppm for ^{1}HN) was assumed for the intermediate state chemical shifts. ^{1}H, ^{13}C, and ^{15}N chemical shifts reconstructed for the intermediate state of the A39V/N53P/V55L Fyn SH3 domain (Fig. S3B), along with further experimental details, have been deposited with the BioMagResBank (access code: 17149).

Residual anisotropic interactions

Backbone amide residual dipolar couplings (RDCs) for the native state, $D_{NH,F}$, were obtained by weak alignment of [U-^{2}H,^{15}N] A39V/N53P/V55L Fyn SH3 via the addition of 24 mg/ml Pf1 filamentous phage (*72*; ASLA Biotech, Riga, Latvia) or a liquid crystalline C12E5-PEG/hexanol mixture (*73*). $D_{NH,F}$ values were calculated from the observed doublet splitting in ^{15}N-IPAP spectra (*74*) by subtracting the $^{1}J_{NH}$ scalar coupling contribution as observed in reference spectra of the unaligned sample. The alignment was verified to be independent of spectrometer field strength in both cases (*75*), resulting in residual solvent ^{2}H quadrupole couplings of 27 Hz and 14 Hz, respectively. The measured values for $D_{NH,F}$ range from -26.5 Hz to +22.2 Hz in Pf1 (-30.3 Hz to +23.4 Hz in PEG/hexanol), corresponding to an alignment tensor (*76*) with a magnitude of D_a = -13.3 Hz and a rhombicity of R = 0.45 (D_a = -15.2 Hz, R = 0.36 in PEG/hexanol).

Changes in backbone amide RDCs between native and intermediate states, $\Delta D_{NH,F\rightarrow I} = D_{NH,I} - D_{NH,F}$, were extracted (*54*) from fits of ^{15}N SQ CW CPMG (*53*) and ^{15}N TROSY/AntiTROSY CPMG (*54*) RD experiments (Fig. S4F) recorded on samples of [U-^{2}H,^{15}N] A39V/N53P/V55L Fyn SH3 weakly aligned by addition of Pf1 or PEG/hexanol. Data sets were recorded, processed, and analyzed in the same fashion as those obtained for the unaligned sample described above. There is no indication of any specific interactions of the two alignment media with either of the exchanging states and the extracted exchange parameters and chemical shift differences are in excellent agreement with the corresponding values for the unaligned sample. As expected for well-defined fits, $\Delta D_{NH,F\rightarrow I}$ values are not correlated with the corresponding chemical shift changes $\Delta\varpi_{N,F\rightarrow I}$, whereas their experimental uncertainties, $s(\Delta D_{NH,F\rightarrow I})$, as estimated from the covariance matrix (*66*) are distributed as observed previously (*54*), with an optimal precision of $s(\Delta D_{NH,F\rightarrow I}) \approx$ 1.9 Hz for $|\Delta\varpi_{N,F\rightarrow I}| \approx$ 1.5 ppm. The backbone amide RDCs for the intermediate were then calculated according to $D_{NH,I} = D_{NH,F} + \Delta D_{NH,F\rightarrow I}$ with an uncertainty of $s(\Delta D_{NH,I}) = ((s(D_{NH,F}))^2 + (s(\Delta D_{NH,F\rightarrow I}))^2)^{1/2} \approx s(\Delta D_{NH,F\rightarrow I})$. The measured values for $D_{NH,I}$ correlate

poorly with $D_{NH,F}$, indicating a difference in molecular alignment between native and intermediate states, and range from -32.1 Hz to +24.7 Hz in Pf1 (-28.1 Hz to +26.6 Hz in PEG/hexanol), corresponding to D_a = -16.1 Hz and R = 0.35 (D_a = -14.0 Hz, R = 0.60 in PEG/hexanol).

Backbone carbonyl ^{13}C residual chemical shift anisotropies (RCSAs) for the native state, $\delta\varpi_{CO,F}$, were obtained from differences in ^{13}CO peak positions in H(N)CO spectra recorded at 800 MHz on [U-^{2}H,^{13}C,^{15}N] A39V/N53P/V55L Fyn SH3 samples either unaligned or weakly aligned by 36 mg/ml Pf1 filamentous phage (*72*; ASLA Biotech, Riga, Latvia). The alignment was verified to be independent of spectrometer field strength (*75*), resulting in a residual solvent ^{2}H quadrupole coupling of 38 Hz. The measured values for $\delta\varpi_{CO,F}$ range from -134 ppb to +149 ppb. In structure calculations a ^{13}CO chemical shift anisotropy (CSA) tensor with σ_{11} = -74.7 ppm, σ_{22} = -11.8 ppm, σ_{33} = 86.5 ppm (*77*) was used along with an effective N-H bond length of 1.04 Å (*77*) and a molecular alignment tensor identical to the one determined from $D_{NH,F}$ values measured in 24 mg/ml Pf1 but scaled to take into account the higher Pf1 phage concentration in the RCSA measurements (43% higher magnitude D_a based on a comparison of $D_{NH,F}$ values recorded on samples aligned with 24 mg/ml and 36 mg/ml phage).

Backbone carbonyl ^{13}C chemical shift changes $\Delta\varpi^{iso}_{CO,F\to I}$ and $\Delta\varpi^{Pf1}_{CO,F\to I}$ were extracted from ^{13}CO SQ CPMG (*57*) RD experiments recorded on the same samples used for measuring $\delta\varpi_{CO,F}$ (see above) at 20°C. For enhanced accuracy, the ^{13}CO SQ RD profiles were fit together with additional ^{15}N SQ CW CPMG (*53*) experiments recorded on the same samples at 500 MHz. Backbone carbonyl ^{13}C RCSAs for the intermediate were then calculated (*78*) according to

$$\delta\varpi_{CO,I} = \delta\varpi_{CO,F} + \Delta\varpi^{Pf1}_{CO,F\to I} - \Delta\varpi^{iso}_{CO,F\to I}$$

with an uncertainty of

$$s(\delta\varpi_{CO,I}) = \left((s(\delta\varpi_{CO,F}))^2 + (s(\Delta\varpi^{Pf1}_{CO,F\to I}))^2 + (s(\Delta\varpi^{iso}_{CO,F\to I}))^2 \right)^{1/2}$$

The measured values for $\delta\varpi_{CO,I}$ range from -240 ppb ± 53 ppb to +232 ppb ± 32 ppb. $D_{NH,I}$ RDCs and $\delta\varpi_{CO,I}$ RCSAs reconstructed for the intermediate state of the A39V/N53P/V55L Fyn SH3 have been deposited with the PDB (access code: 2L2P).

Unlike RDCs and RCSAs for the native state, which can be measured with very high precision, RDCs and RCSAs for the intermediate carry substantial experimental uncertainties. Quality factors (*79*) quantifying the agreement of RDCs back-calculated from structures calculated for the intermediate, $D^{calc}_{NH,I}$, with the experimental RDCs were therefore adjusted by the estimated uncertainties according to

$$Q_{err} = \sqrt{\frac{\sum_i (\Delta_{NH,I}(i))^2}{\sum_i (D_{NH,I}(i))^2}} \quad \text{with}$$

$$\Delta_{NH,I}(i) = \begin{cases} 0 & \text{if } |D_{NH,I}(i) - D^{calc}_{NH,I}(i)| \le s(D_{NH,I}(i)) \\ |D_{NH,I}(i) - D^{calc}_{NH,I}(i)| - s(D_{NH,I}(i)) & \text{otherwise} \end{cases}$$

where the summation extends over all measured RDCs; an analogous equation was used for RCSAs.

Structure calculation of the native state

An initial structural model for the native state of the A39V/N53P/V55L Fyn SH3 domain was created on the basis of the crystal structure of N53I/V55L Fyn SH3 (PDB 3CQT; *20*). That the N53I/V55L Fyn SH3 crystal structure is a valid model for the A39V/N53P/V55L Fyn SH3 native state is established by the fact that experimental $D_{NH,F}$ values agree with those predicted from this model to a quality factor (*79*) of Q = 30.3% in Pf1 (Q = 31.6% in PEG/hexanol), with a Q factor of 39.9% for measured $\delta\varpi_{CO,F}$ values. Thus, the native backbone conformation is virtually identical in solution and unaffected by the two side-chain mutations. To elucidate any structural differences in detail we calculated the solution structure of the native A39V/N53P/V55L Fyn SH3 domain from NMR-spectroscopic experimental restraints. To this end, ^{1}H-^{1}H distance restraints were derived from the peak list of a [^{1}H, ^{15}N]-NOESY-HSQC spectrum (*36*; 150 ms mixing time) recorded at 20°C on a sample of 0.7 mM [U-^{15}N] A39V/N53P/V55L Fyn SH3 using the CHESHIRE-YAPP protocol (*80*), which is guided by the structural information inherent in the experimental ^{1}H, ^{13}C and ^{15}N backbone chemical shifts.

CHESHIRE consists of a three-phase computational procedure (*81*). In the first phase, the chemical shifts and the intrinsic secondary structure propensities of amino acid triplets are used to predict the secondary structure of the protein. In the second phase, the secondary structure predictions and the chemical shifts are used to predict backbone torsion angle restraints for the protein. These angles are screened against a database to create a library of trial conformations of three and nine residue fragments spanning the sequence of the protein. In the third phase, a molecular fragment replacement strategy is used to assemble these fragments into low-resolution structural models guided by the information provided by chemical shifts. The resulting structures are refined with a hybrid molecular dynamics and Monte Carlo conformational search using a scoring function defined by: (1) the agreement between experimental and calculated chemical shifts, and (2) the energy of a molecular mechanics force field. This scoring function ensures that a structure is associated with a low CHESHIRE score only if it has a low value of the molecular mechanics energy and is highly consistent with experimental chemical shifts. 50,000 structures were generated using this method.

The CHESHIRE-YAPP protocol uses the best scoring 500-1000 high-resolution structures generated by CHESHIRE to select compatible NOEs from the unassigned NOESY peak lists in an iterative fashion. In the first step, atoms are assigned to each spectral dimension using a chemical shift tolerance of 0.03 ppm for ^{1}H and 0.3 ppm for ^{15}N. Then, chemical shift-based assignments that are violated by more than 2.0 Å in 50 or more of the best 500 CHESHIRE structures are removed. The remaining restraints are used to refine the best scoring 100 CHESHIRE structures. The last two steps are repeated four times with decreasing violation thresholds of 1.5, 1.0, 0.5, and 0.2 Å. In order to check the consistency of the remaining NOE distance restraints, they were used in combination with backbone torsion dihedral angle restraints derived from chemical shifts using 3PRED (*81*) and TALOS+ (*82*) to obtain a low-resolution structure with a backbone RMSD of approximately 2 Å from the initial model based on the crystal structure of the N53I/V55L Fyn SH3 domain (see above). These calculations were performed by torsion angle molecular dynamics in *almost* 1.0.4 (A. C.; http://www-almost.ch.cam.ac.uk/site) with the ParallHDG force field (*83*) and harmonic NOE distance restraints. The approach described

above generates a set of consistent restraints that are subsequently used in a refinement protocol to generate high-resolution structural models, as described below.

The resulting NOE distance restraints, supplemented with several additional NOE distance restraints assigned manually from the [^{1}H, ^{15}N]-NOESY-HSQC spectrum, were used together with the backbone torsion dihedral angle restraints and the experimental $D_{NH,F}$ RDC restraints to refine the initial model based on the crystal structure of N53I/V55L Fyn SH3 (see above). The initial model was subjected to an initial 1000 steps of steepest-descent minimization of the total target function, consisting of the AMBER03 force-field (*84*) with a generalized Born implicit solvation model (*85*) and backbone dihedral angle restraints from chemical shifts, supplemented with harmonic potentials for the NOE distance restraints and the RDC restraints (*86*, *87*). This energy-minimized model was used as a starting structure for generating 1000 structures by 10,000 integration steps of torsion angle molecular dynamics simulated annealing from 350 K to 300 K in *almost* 1.0.4 (A. C.), followed by 2000 steps of steepest-decent minimization. The 10 best structures out of the 1000 structures generated were selected for further analysis. As expected, the agreement with the $D_{NH,F}$ improves further to Q = 18.9% (average) ± 0.7% (standard deviation) in Pf1 (Q = 19.8% ± 0.7% in PEG/hexanol) and the agreement with the measured $\delta\varpi_{CO,F}$ (not used in the structure calculation) is only very marginally affected (Q = 41.9% ± 1.7%). This final ensemble of 10 structures describing the native A39V/N53P/V55L Fyn SH3 domain (Fig. S2, Tab. S1) is well-defined and virtually identical to the initial model based on the crystal structure of N53I/V55L Fyn SH3 (see above) with RMSDs of 0.78 Å (average) ± 0.07 Å (standard deviation) for the backbone and 1.16 Å ± 0.10 Å for all heavy atoms (for residues 2 to 59). The most significant difference is the solvent-exposed side-chain of Thr47, which preferentially populates the gauche^{+} rotamer in solution rather than the gauche^{-} rotamer found in the crystal, as we have noted previously (*22*).

Structure calculation of the intermediate state

The chemical shift differences, $\Delta\varpi_{X,F\rightarrow I}$, for the RT-src loop and the central β-sheet (strands β_2, β_3, β_4) are very small (Fig. S5A), indicating the absence of any major structural rearrangements between the native and intermediate states in these regions. Thus, the conformational sampling required for structure calculation of I is limited to topologies close to the native state. We have demonstrated previously that in such cases correct atomic-resolution structures of satisfactory quality can be obtained from backbone chemical shifts as the only experimental restraints using chemical shift restrained molecular dynamics (CS-MD) simulations starting from partially unfolded structures (*24*). In this approach chemical shifts are implemented as restraints in a molecular dynamics conformational search via energy functions that penalize differences between experimental and calculated chemical shifts, in a manner analogous to conventional NMR structure calculations. Additional NMR restraints such as RDCs and RCSAs are readily incorporated using standard algorithms. Accordingly, the high-resolution structure of the folding intermediate was calculated essentially as described previously (*24*) by replica-exchange molecular dynamics simulations (Fig. S6; *88*) in *almost* 1.0.4 (A. C.), using the AMBER03 force-field (*84*) with bond lengths and angles kept fixed by the SHAKE algorithm (*89*), a generalized Born implicit solvation model (*85*), and CamShift 1.2 (Fig. S7; *90*) chemical shift restraints supplemented with RDC and RCSA restraints. Only chemical shifts defined to better than 0.25 ppm (for ^{15}N, ^{13}CO, ^{13}Cα) or 0.11 ppm (for ^{1}HN, ^{1}Hα), D_{NH} RDCs defined to better than 7.1 Hz, and ^{13}CO RCSAs defined to better than 80 ppb were used as experimental

restraints. RDCs and RCSAs close to zero as a result of conformational averaging in flexible parts of the structure were also excluded from the structure calculation. RDC and RCSA restraints were imposed via standard flat-bottom square well potentials (*86*, *87*) with a width equal to the experimental uncertainties and force constants of 0.020 kcal/mol/Hz^2 (optimized with respect to the quality factor of RDCs set aside for cross-validation in a similar calculation; *24*) and 0.001 kcal/mol/ppb^2 (adjusted to generate similar energetic penalties for RDCs and RCSAs for similar quality factors), respectively. Magnitude, D_a, and rhombicity, R, of the two molecular alignment tensors in Pf1 and PEG/hexanol were fixed to the values estimated as described above while the variable orientations of these tensors were represented by two artificial tetra-atomic molecular coordinate systems (*86*). As described above, a ^{13}CO CSA tensor with σ_{11} = -74.7 ppm, σ_{22} = -11.8 ppm, σ_{33} = 86.5 ppm (*77*), an effective N-H bond length of 1.04 Å (*77*), and an identical molecular alignment tensor for the ^{13}CO RCSAs as for the D_{NH} RDCs in Pf1 but with 43% higher magnitude D_a were assumed.

The initial model of the native state of A39V/N53P/V55L Fyn SH3 created on the basis of the crystal structure of the N53I/V55L Fyn SH3 (see above) was partially unfolded by 100 ps torsion angle molecular dynamics at 10,000 K with a time step of 20 fs (*24*) in *almost* 1.0.4 (A. C.) using the AMBER03 force-field (*84*). A structure with a backbone RMSD of 5.37 Å (for residues 2 to 59) to the native state was randomly selected from the denaturing trajectory as a starting structure for the CS-MD simulations (Figs. S6, S7). Several additional structures randomly selected from the trajectory were used for control CS-MD simulations to establish that the results are robust with respect to the choice of initial conditions. Starting structures were subjected to 4 ns replica-exchange CS-MD simulations (initially without RDC and RCSA restraints) with swap attempts between the 21 replicas spanning temperatures from 270 K to 515 K every 6 ps (*24*). Snapshots of the lowest-temperature replica were saved every 0.6 ps. Each structure from this ensemble of snapshots was re-scored using a modified target function $E_{sel} = E_{FF} + 5\times E_{CS}$ with five times higher weight on the CamShift potential E_{CS} (Fig. S7) relative to the force-field E_{FF} than during the CS-MD simulation itself, which we have found to significantly improve selection of structures that agree with the chemical shift restraints (*24*). The 21 structures with lowest E_{sel} were then used as starting structures for refinement by further 4 ns of replica-exchange CS-MD simulations (Fig. S6) with RDC restraints E_{RDC} and RCSA restraints E_{RCSA}, followed by 500 steps of steepest-descent minimization with the modified target function $E_{sel} = E_{FF} + 5\times E_{CS} + E_{RDC} + E_{RCSA}$ for selection of the 10 best structures from the lowest-temperature replica for further analysis, and 5000 steps of steepest-descent minimization without chemical shift restraints to relax any minor local distortions that can occur around aromatic rings when the CamShift potential is weighted very strongly. Structural statistics were evaluated using NIH version 2.25 (*91*) of X-PLOR 3.851 (*92*) and PROCHECK-NMR 3.5.4 (*93*).

As a positive control we tested this structure calculation protocol on the native state first. Even before refinement with RDC and RCSA restraints the CS-MD simulations with the native state chemical shifts produced a cluster of 21 structures with backbone RMSDs of only 1.30 Å (average) ± 0.12 Å (standard deviation) (for residues 2 to 59) from the initial model and 1.30 Å ± 0.12 Å from the average solution structure of the native state of A39V/N53P/V55L Fyn SH3 (Fig. S8A), but with only poor agreement with the $D_{NH,F}$ (Q = 50.6% ± 5.6% in Pf1 and Q = 53.9% ± 6.4% in PEG/hexanol) and $\delta\varpi_{CO,F}$ (Q = 67.0% ±

6.3%). Refinement with the native state RDC and RCSA restraints followed by energy minimizations as described above improved the agreement of the resulting ensemble of 10 structures with the $D_{NH,F}$ to Q = 22.2% ± 2.1% in Pf1 (Q = 22.1% ± 1.6% in PEG/hexanol) and with the $\delta\varpi_{CO,F}$ to Q = 27.4% ± 2.7%. These quality factors are typical for the target function used in our CS-MD protocol (*24*). Inclusion of the RDC and RCSA restraints also further improved the accuracy of the structures, to backbone RMSDs of 0.76 Å ± 0.11 Å (for residues 2 to 59) from the initial model and 0.75 Å ± 0.10 Å from the average solution structure of the native state.

By contrast, initial CS-MD simulations with the intermediate state chemical shifts extracted from CPMG RD profiles, while still sampling native-like structures (backbone RMSD around 1.3 Å from the native state) extensively, preferentially sample a substantially different backbone conformation approximately 3.5 Å away from the native state (Fig. S8A). Notably, these conformations are favored by the chemical shift penalty function E_{CS} - especially for the amide proton chemical shifts (Fig. S8B), which are particularly sensitive to tertiary interactions such as hydrogen bonds and ring-current shifts - whereas the force-field E_{FF} favors native-like conformations. These initial structures already exhibit all the essential structural features of the intermediate discussed below. In particular, the COOH-terminal strand β_5 is not yet formed in the folding intermediate. In fact, the chemical shifts indicate virtually complete flexibility of Pro57 to Asp59 with little or no detectable structure in the intermediate (Fig. S5B). Exclusion of any artificial tertiary contacts of the flexible COOH-terminal residues with the structured remainder of the domain became an – albeit minor - concern for a structure calculation protocol with a force-field that includes attractive long-range interactions such as AMBER03. To this end all further structure calculations for the folding intermediate were carried out on a truncated construct comprising residues 1 to 56 only; in addition, any artificial electrostatic interactions *in silico* involving the terminal NH_3^+ and COO^- groups were excluded by addition of acetyl (PDB hetero-compound ACE) and N-methylamide (NME) protection groups, respectively.

For this truncated construct, CS-MD simulations with the intermediate state chemical shifts generate an energy landscape funneled to a cluster of 10 structures with backbone RMSDs of 1.37 Å ± 0.17 Å (for residues 2 to 55) from the initial model and 1.34 Å ± 0.21 Å from the average solution structure of the native state of the A39V/N53P/V55L Fyn SH3 (Fig. S8C); note that the backbone RMSDs between intermediate and native states are smaller for the truncated construct as most of the backbone variability in the full-length construct derives from the COOH-terminal region. As in the case of the control calculations for the native state (see above) the agreement with the $D_{NH,I}$ (Q_{err} = 50.6% ± 5.9% in Pf1 and Q_{err} = 80.6% ± 5.3% in PEG/hexanol) and $\delta\varpi_{CO,I}$ (Q_{err} = 60.0% ± 6.4%) is still poor. The 21 structures with lowest E_{sel} were further refined by replica-exchange CS-MD with RDC and RCSA restraints as described above, resulting in an even narrower energy landscape (Fig. S8D). After energy minimizations as described above we obtained a final set of 10 well-defined structures (Fig. 1A, Tab. S2) with backbone RMSDs of 1.17 Å ± 0.07 Å (for residues 2 to 55) from the initial model and 1.15 Å ± 0.06 Å from the average solution structure of the native state of the A39V/N53P/V55L Fyn SH3 and excellent agreement with the $D_{NH,I}$ (Q_{err} = 16.3% ± 1.4% in Pf1 and Q_{err} = 35.1% ± 3.1% in PEG/hexanol) and $\delta\varpi_{CO,I}$ (Q_{err} = 14.7% ± 0.4%). Together with the experimental RDC and RCSA restraints the atomic coordinates of this set

of 10 structures of the folding intermediate of the A39V/N53P/V55L Fyn SH3 have been deposited with the PDB (access code: 2L2P).

H/D exchange NMR experiments

The structures calculated for the folding intermediate are consistent with all the experimental data but nevertheless we decided to seek independent verification. In light of the sparse number of experimental restraints, setting aside the $D_{NH,I}$ values measured in PEG/hexanol for cross-validation was not considered to be an attractive option because even highly precise D_{NH} RDC restraints in one alignment medium often do not carry enough information for insightful cross-validation (*94*, *95*) and the $D_{NH,I}$ RDCs extracted from CPMG RD profiles have elevated experimental uncertainties. Measurement of additional RDCs, such as D_{NCO} or $D_{C\alpha H\alpha}$, using suitable CPMG experiments was hampered by the relatively low equilibrium population of the intermediate ($p_I \approx 2\%$). However, knowledge of the kinetics and thermodynamics of the three-state folding equilibrium $F \leftrightarrow I \leftrightarrow U$ (*19*) and of the high-resolution structures and hence the hydrogen bond networks of both F and I (calculated from CPMG RD profiles, see above) allows quantitative prediction of the H/D exchange kinetics (*23*, *96-97*), which is readily testable under suitable experimental conditions.

To this end, a sample of 500 µl 1.1 mM [1-^{13}C-glc;U-^{15}N] A39V/N53P/V55L Fyn SH3 was buffer-exchanged into 0.2 mM EDTA, 50 mM sodium phosphate (pH 6.0) in H_2O/D_2O (9:1). This pH value, one unit lower than in the other experiments, was chosen to enable quantification of exchange rates that would otherwise be too rapid. An ^{15}N SQ CW CPMG (*53*) RD experiment was recorded at 20°C (500 MHz) to verify that the pH change has virtually no influence on the kinetics and thermodynamics of the $F \leftrightarrow I$ equilibrium nor on the chemical shifts of either of the two exchanging states. Thus the structures of F and I are invariant to pH over this range. The sample was freeze-dried, dissolved in D_2O and a series of [^{1}H, ^{15}N]-HSQCs was recorded over 13 h at 15°C (500 MHz). Resonance intensities were quantified by three-way decomposition using MUNIN (*59*, *60*) and the net H/D exchange rates k_{ex} determined by exponential fitting (*97*). Intrinsic (unprotected) H/D exchange rates k_{int} for these experimental conditions were predicted from the amino acid sequence (*98*) using the SPHERE server (http://www.fccc.edu/research/labs/roder/sphere) and are of the order of magnitude of 1/min to 10/min. On this time-scale the folding/unfolding reaction $F \leftrightarrow I \leftrightarrow U$ is always in full equilibrium because it rapidly equilibrates within milliseconds (*19*) so the observed net H/D exchange rate in this limit (EX2) is the population-weighted average over the states involved (*23*):

$$k_{ex} = p_F k_F + p_I k_I + p_U k_{int}$$

where k_F and k_I are the direct H/D exchange rates from F and I, respectively, U is assumed to be completely unstructured (*19*) and hence $k_U = k_{int}$, and neglecting any back-exchange due to residual solvent protonation. The calculated net protection factor PF therefore follows

$$1/PF = k_{ex}/k_{int} = p_F/PF(F) + p_I/PF(I) + p_U$$

where $PF(F) = k_{int}/k_F$ and $PF(I) = k_{int}/k_I$ are the structural protection factors (*99*) that report on the hydrogen bond networks of F and I, respectively. From the CPMG RD experiments recorded in D_2O (see above) we obtain the relative populations of F and I under the given experimental conditions to be $p_F \approx 98.6\%$ and $p_I \approx 1.4\%$, respectively, and the relative

population of the unfolded state can be extrapolated from the known thermodynamics of the folding equilibrium (*19*) to $p_U \approx 0.04\%$.

Because all these quantities are positive the exchange rates and protection factors have to satisfy several inequalities. First, $k_{ex} \geq p_F\ k_F = p_F/PF(F) \times k_{int} = 0.986/PF(F) \times k_{int}$ so amide protons without significant protection in the native state ($PF(F) \approx 1$) exchange on the time-scale of k_{int} and only protected amide protons are observable beyond the experimental dead time of a few minutes; the results of the H/D exchange experiment can therefore be color-coded onto the hydrogen bond network of the native state (Fig. S9). Second, $PF \leq 1/p_U \approx 2500$ is an estimate for the upper bound on the net protection factors. Third, any given net protection factor PF implies a lower bound for the protection in the intermediate according to $PF(I) \geq p_I \times PF \geq 0.014 \times PF$. In practice, structural protection factors are typically highly polarized, with $PF(X) \geq 10^6$ for amide groups protected by stable hydrogen bonds (note that local fluctuations causing some "fraying out" of the apparent hydrogen bond stabilities – i. e., decreased protection factors - at the termini of regular secondary structure elements are not uncommon; *100*) in state X or $PF(X) \leq 5$ for residues not involved in hydrogen bonds (*97*). Therefore, the net protection factors are predicted to be either $PF \approx 1/p_U \approx 2500$ for amide groups already involved in stable hydrogen bonds in the folding intermediate ($PF(F) = PF(I) = \infty$) or $PF \approx 1/p_I \approx 70$ for amide groups whose hydrogen bonds form after the folding intermediate ($PF(F) = \infty$, $PF(I) \approx 1$). This is indeed the pattern that is observed experimentally (Fig. S9), thus corroborating the atomic-resolution structure of the folding intermediate calculated from CPMG RD profiles.

CD spectroscopy

For CD spectroscopy, samples were diluted to about 30 μM. A39V/N53P/V55L/Δ(57-60) Fyn SH3 aggregates were separated from the monomeric SH3 domain (supernatant) by ultracentrifugation. CD spectra from 185 nm to 290 nm were recorded in 0.04 mM EDTA, 10 mM sodium phosphate (pH 7.0) on a Jasco J-810 Spectropolarimeter using a step width of 0.2 nm, a band width of 1 nm, and a scanning speed of 50 nm/min in cuvettes with a light path of 1 mm (Hellma, Müllheim, Germany) at room temperature. To increase the signal to noise ratio 10 scans were accumulated. The residue ellipticity Θ_{MRW} was calculated from the measured ellipticity Θ according to $\Theta_{MRW} = \Theta/c/d/N$, where c denotes the protein concentration, d the light path, and N the number of peptide bonds. A thermal unfolding profile (Fig. S10) was obtained by monitoring changes in ellipticity at 220 nm of a sample consisting of 32.7 μM A39V/N53P/V55L/Δ(57-60) Fyn SH3 domain, 0.2 mM EDTA, 50 mM sodium phosphate (pH 7.0) in H_2O on an Aviv 202 CD spectrometer (Aviv Biomedical, Lakewood, NJ, USA) and analyzed as described previously (*22*, *35*).

Electron microscopy

Samples to be analyzed by transmission electron microscopy (TEM) were deposited on fresh continuous carbon films prepared from copper rhodium grids (Electron Microscopy Sciences (EMS), Hatfield, PA, USA). Grids were charged using a glow discharger (EMS) for 15 s at 30 mA negative discharge before the addition of samples. Protein solutions were diluted 1000-fold and were allowed to adsorb to the grids for 2 minutes. The solution was then blotted away from the grid using Whatman No. 2 filter paper. Samples were washed in a 10 μl drop of water, then stained twice with 10 μl drops

of freshly filtered 2% uranyl acetate, with blotting between and after staining. Negative stain TEM images were obtained using a JEOL JEM-1011 microscope operating at 80 kV.

Congo red assay

Congo red binding assays were carried out using a modified protocol described by Klunk *et al.* (*101*). A fresh stock of 35 mM Congo red (Sigma-Aldrich) solution was prepared in 0.04 mM EDTA, 10 mM sodium phosphate buffer (pH 7.0). This stock was diluted to 20 μM in 250 μl aliquots containing this buffer alone or SH3 domain mutants in this buffer, incubated for 30 minutes at room temperature, transferred into a clean quartz cuvette with a 1 cm light path, and tested for the spectrophotometric hyperchromicity and red shift of the major absorbance band of Congo red around 500 nm typical for amyloid fibril binding (*101*, *102*). This was achieved by comparing the absorption spectra, recorded from 240 nm to 600 nm on a Biochrom Ultrospec 2100 pro spectrophotometer at room temperature, in the absence and presence of SH3 domain mutants.

Supporting Tables

Table S1. Summary of the structure calculation of the A39V/N53P/V55L Fyn SH3 domain native state. Except for the experimental restraints and Ramachandran plot statistics all values are averages over the 10 accepted structures in the form average value ± standard deviation.

Experimental restraints used for the structure calculation	
Intraresidual NOEs	116
Interresidual NOEs	
Sequential	176
Medium-range	42
Long-range	82
Dihedral angle restraints	73
Residual dipolar coupling (RDC) restraints	
$D_{NH,F}$ in Pf1	54
$D_{NH,F}$ in PEG/hexanol	53
Molecular dynamics simulation statistics	
Energies / kcal/mol	
Total	-2572 ± 13
Bond lengths	24.4 ± 0.2
Bond angles	95.6 ± 1.2
Dihedral angles and dihedral angle restraints	471.5 ± 1.4
Improper angles	2.57 ± 0.14
Electrostatic interactions	-1344 ± 33
Van der Waals interactions	-283.6 ± 2.3
Solvation	-1558 ± 15
Distance restraints	0.26 ± 0.02
Residual dipolar coupling (RDC) restraints	20.1 ± 1.6
RMSDs from RDC restraints[a] / Hz	
$D_{NH,F}$ in Pf1	2.41 ± 0.09 (2.25 ± 0.09)
$D_{NH,F}$ in PEG/hexanol	2.84 ± 0.10 (2.70 ± 0.10)
Ramachandran plot statistics according to PROCHECK-NMR 3.5.4 (*93*)	
Residues in most favored regions	98.0%
Residues in additional allowed regions	2.0%
Residues in generously allowed regions	0.0%
Residues in disallowed regions	0.0%

	Atomic RMSDs from the average structure	
	Backbone / Å	**Heavy atoms / Å**
Well-ordered part of structure[b]	0.28 ± 0.07	0.47 ± 0.10
	Comparison with the initial model of the native state	
	Backbone / Å	**Heavy atoms / Å**
Well-ordered part of structure[b]	0.78 ± 0.07	1.16 ± 0.10

[a] Values in parentheses were calculated from $\Delta_{NH,F}$ values adjusted by the experimental uncertainties as described above and in analogous manner for RCSA restraints.
[b] Residues 2-59.

Table S2. Summary of the structure calculation of the A39V/N53P/V55L Fyn SH3 domain folding intermediate. Except for the experimental restraints and Ramachandran plot statistics all values are averages over the 10 accepted structures in the form average value ± standard deviation.

Experimental restraints used for the structure calculation	
Chemical shift restraints	
^{15}N	48
^{1}HN	48
^{13}CO	45
$^{13}C\alpha$	38
$^{1}H\alpha$	32
Residual dipolar coupling (RDC) restraints	
$D_{NH,I}$ in Pf1	44
$D_{NH,I}$ in PEG/hexanol	42
Residual chemical shift anisotropy (RCSA) restraints	
$\delta\varpi_{CO,I}$ in Pf1	31
Chemical shift restrained molecular dynamics simulation statistics	
Energies / kcal/mol	
Total	-1543 ± 4
Bond lengths	177 ± 7
Dihedral angles	495 ± 6
Improper angles	7.4 ± 0.7
Electrostatic interactions	-1170 ± 120
Van der Waals interactions	-261 ± 9
Solvation	-830 ± 120
CamShift restraints	10.8 ± 0.9
Residual dipolar coupling (RDC) restraints	19.1 ± 1.6
Residual chemical shift anisotropy (RCSA) restraints	7.4 ± 0.4
RMSDs from RDC restraints[a] / Hz	
$D_{NH,I}$ in Pf1	4.48 ± 0.27 (2.53 ± 0.22)
$D_{NH,I}$ in PEG/hexanol	6.49 ± 0.36 (4.00 ± 0.35)
RMSDs from RCSA restraints[a] / ppb	
$\delta\varpi_{CO,I}$ in Pf1	49.5 ± 9.6 (15.5 ± 0.4)
Ramachandran plot statistics according to PROCHECK-NMR 3.5.4 (*93*)	
Residues in most favored regions	89.4%
Residues in additional allowed regions	9.0%
Residues in generously allowed regions	1.7%
Residues in disallowed regions	0.0%

	Atomic RMSDs from the average structure	
	Backbone / Å	**Heavy atoms / Å**
Well-ordered part of structure[b]	0.59 ± 0.17	1.12 ± 0.23
	Comparison with the average structure of the native state	
	Backbone / Å	**Heavy atoms / Å**
Well-ordered part of structure[b]	1.15 ± 0.06	2.02 ± 0.21

[a] Values in parentheses were calculated from $\Delta_{NH,I}$ values adjusted by the experimental uncertainties as described above and in analogous manner for RCSA restraints.
[b] Residues 2-55.

Supporting Figures

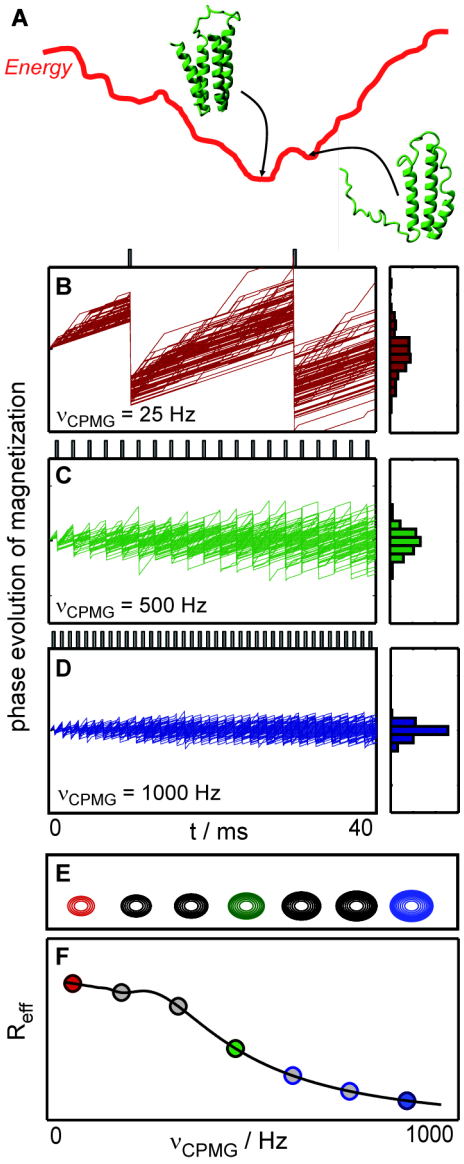

Fig. S1. Basic principles of CPMG RD NMR spectroscopy. (**A**) One-dimensional (highly schematized) free energy landscape of a "generic" protein showing a ground state conformation along with a low-lying excited state conformation that is thermally accessible via a low free energy barrier. If the states are populated to ≥ 0.5% in equilibrium and interconvert on a time-scale of approximately 100 to 2000/s the exchanging system is readily amenable to study by CPMG RD NMR spectroscopy (*16*). In the case of the A39V/N53P/V55L Fyn SH3 domain studied in this work the exchange "reaction" is well approximated by a two-state interconversion between the native ground state N and an excited folding intermediate I

$$N \underset{k_{IF}}{\overset{k_{FI}}{\rightleftharpoons}} I$$

with an exchange rate, $k_{ex,FI} = k_{FI} + k_{IF} = 836/s \pm 4/s$, and a fractional population of I, $p_I = 1.98\% \pm 0.01\%$, at 20°C (see above). The A39V/N53P/V55L mutant was chosen because $k_{ex,FI}$ and p_I are well within the window of the CPMG technique, in contrast to the wild-type (wt) protein (*20*); the N53P mutation increases the populations of the minor conformers (unfolded and intermediate states) relative to the wt (*20*), while A39V/V55L speeds up folding by an order of magnitude (*103*). (**B**) In CPMG RD NMR spectroscopy, a series of experiments are performed in which variable numbers of refocusing pulses are applied to magnetization as it evolves under the influence of chemical shift that varies stochastically due to the exchange process. The effect of the pulses, applied at the times denoted by the vertical bars above each plot, is illustrated schematically in (**B**)-(**D**), as a function of $\nu_{CPMG} = 1/(2\tau)$, where τ is the time between successive refocusing pulses. In these figures signal phase is plotted as a function of time so that the chemical shift is the slope. Each of the refocusing pulses inverts the signal phase. Note that in the absence of exchange, chemical shift is completely refocused by the action of each of the pulses, restoring the phase of the magnetization to its initial value. By contrast, the sets of colored traces correspond to signals from nuclei undergoing exchange between two different environments associated with two different chemical shifts, as would be the case for the exchanging system illustrated in (**A**). Due to the stochastic fluctuations in chemical shift (slope), each nucleus follows a slightly different trajectory, leading to dephasing of the magnetization, (**B**), peaks with low intensity in NMR spectra, and correspondingly large effective transverse relaxation rates, R_{eff}. As the number of refocusing pulses is increased, the trajectories are prevented from

deviating greatly from the initial position (**C**, **D**) so that dephasing is suppressed and peak intensities increase (**E**), corresponding to smaller values of R_{eff}. In this work CPMG RD profiles are presented, R_{eff} versus ν_{CPMG} (**F**). The shapes of the resultant RD profiles (**F**) depend on the populations of the interconverting states, rates of exchange, and the difference in resonance positions (chemical shifts) for nuclei in each of the exchanging conformations. Experiments can be performed either in isotropic media ("normal" buffer) such that the probability of each molecular orientation is equal or in media that impart a small degree of solute alignment (typically 0.1%; *104*) such that certain orientations are preferred (*54*). In the former case differences in resonance positions ($\Delta\nu$ in Hz, corresponding to $\Delta\varpi$ in ppm) are related directly to the chemical shift differences between ground and excited states, from which chemical shifts of the excited state are calculated directly (*68*). When experiments are performed under condition of fractional alignment the differences in resonance positions correspond to $\Delta\nu \pm 0.5\times\Delta D$, where ΔD is the difference in residual dipolar couplings in the interconverting pair of states. As described in detail previously it thus becomes possible to extract residual dipolar couplings of spin pairs (such as amide ^{15}N-^{1}HN) in the excited state (*54*). In summary, fits of CPMG RD data provide thermodynamic and kinetic information for the exchange reaction, as well as structural information for the exchanging states, even in cases where the excited states are populated so marginally that only the dominant conformer is directly observable in NMR spectra, as in the study presented here (Fig. S3). Additional details can be found in the literature (*15*, *16*, *105*, *106*). Figure adapted from references (*106*, *107*).

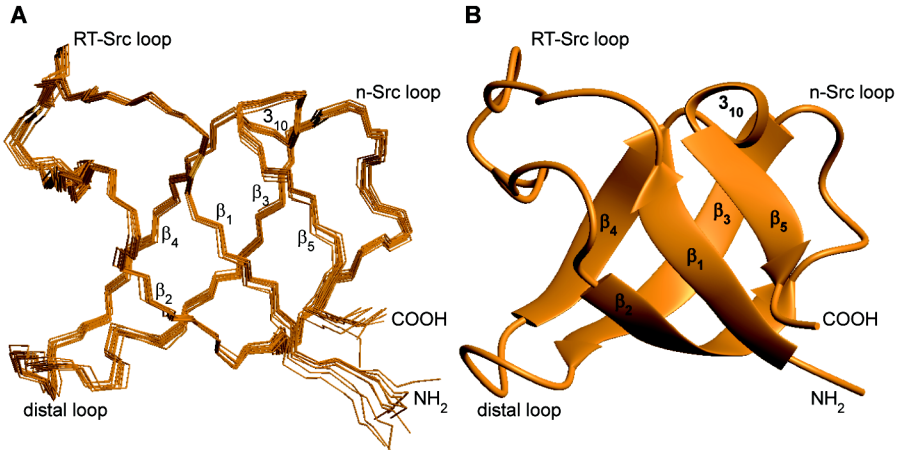

Fig. S2. Structure of the native state. (**A**) Backbone overlay of the 10 lowest-energy structures of the native A39V/N53P/V55L Fyn SH3 domain from Ser1 to Asp59 (Tab. S1). (**B**) Schematic representation of the lowest-energy solution structure of the native state (orange). Same view as in Fig. 1A, B.

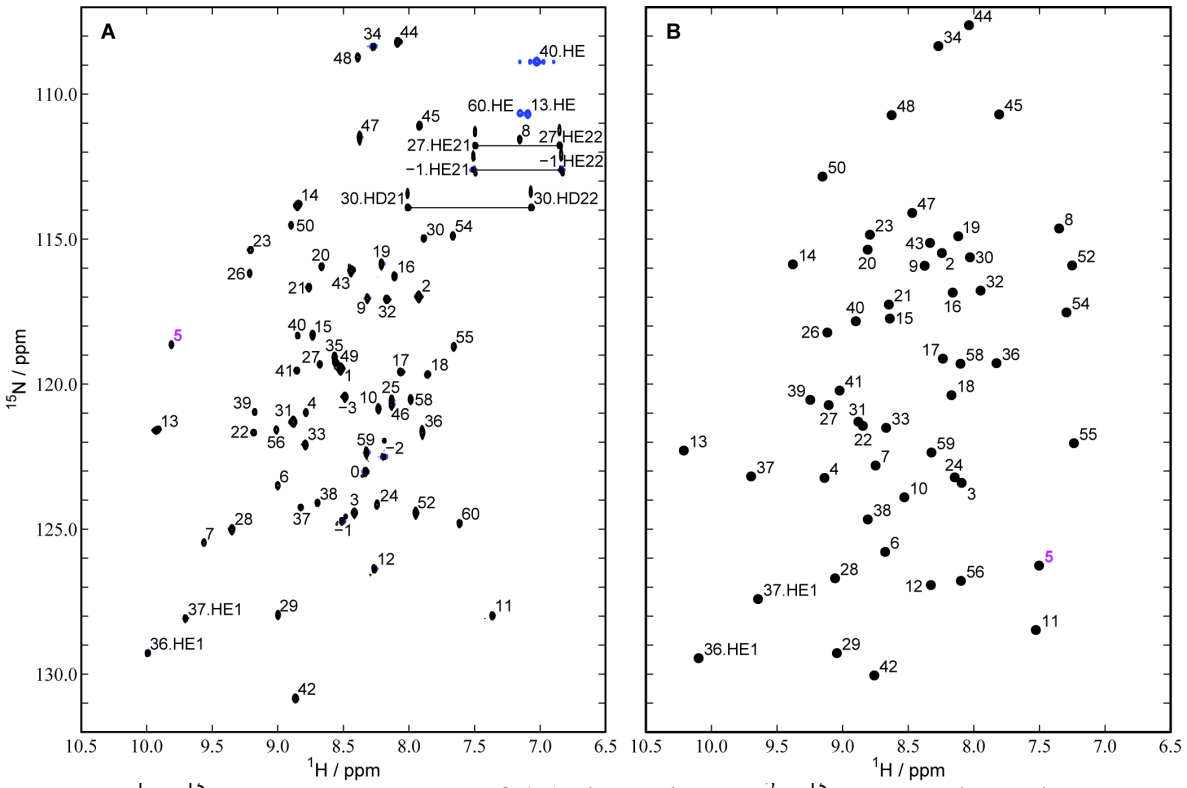

Fig. S3. [^{1}H, ^{15}N]-HSQC spectrum of (**A**) the native [U-^{2}H,^{15}N] A39V/N53P/V55L Fyn SH3 domain recorded at 800 MHz and of (**B**) the folding intermediate of [U-^{2}H,^{15}N] A39V/N53P/V55L Fyn SH3 reconstructed from $\Delta\varpi_{N,F\rightarrow I}$ and $\Delta\varpi_{HN,F\rightarrow I}$ values extracted from the backbone amide CPMG RD data (Fig. S4) at 20°C. Amide proton resonances are labeled according to their residue numbers. The spectrum of the intermediate shows good chemical shift dispersion and several backbone amide resonances, e. g. Glu5 (magenta), shift dramatically compared to the native state, consistent with the fact that the intermediate adopts a non-native but overall well-defined conformation.

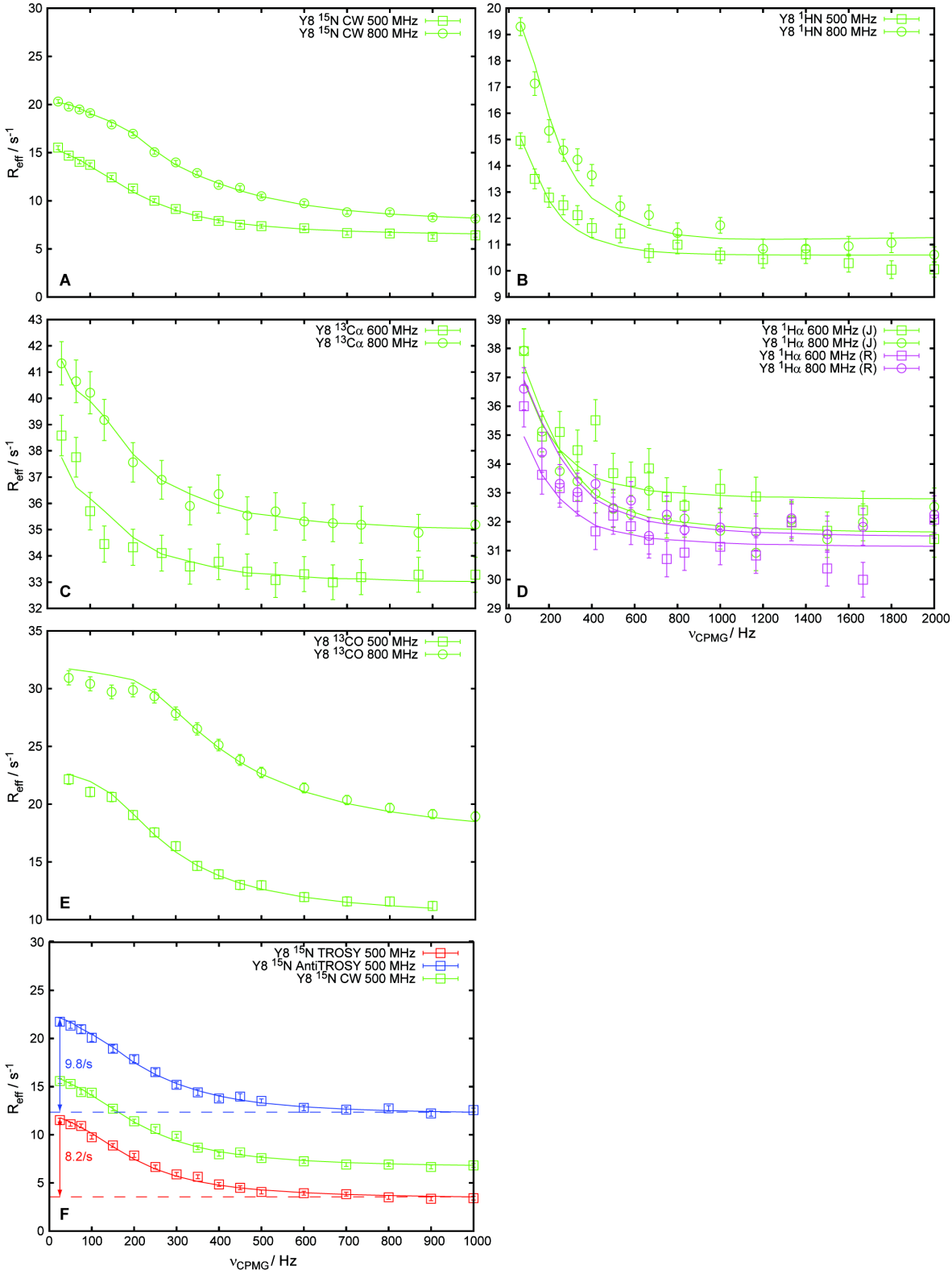

Fig. S4. (**A**) ^{15}N CW SQ, (**B**) ^{1}HN SQ, (**C**) ^{13}Cα SQ, (**D**) ^{1}Hα SQ, and (**E**) ^{13}CO SQ relaxation rates R_{eff} as a function of CPMG frequency ν_{CPMG} for Tyr8 of suitably isotope-

enriched A39V/N53P/V55L Fyn SH3 measured at 800 MHz (circles) and 500 or 600 MHz (squares) at 20°C. Continuous lines are the RD profiles predicted using parameters generated from fitting all RD data available for any given sample together as described in the Materials and Methods section. For enhanced accuracy the $^{1}H\alpha$ SQ RD profiles recorded with refocusing elements B (green) and C (magenta) in Fig. 2 of ref. (*33*) were also fit together. (**F**) As highlighted by the arrows, weak alignment by addition of 24 mg/ml Pf1 causes a difference in exchange contributions to ^{15}N TROSY (red) and AntiTROSY (blue) relaxation rates $R_{eff}(\nu_{CPMG})$ (*54*) of Tyr8 due to the change in backbone amide RDCs between native and intermediate states, $\Delta D_{NH,F\rightarrow I} = +23.1$ Hz ± 3.2 Hz. Dashed horizontal lines indicate the value predicted for $R_{eff}(1000\ Hz)$ as an approximation for the intrinsic relaxation rate. The exchange contribution to the ^{15}N CW SQ CPMG RD profile (green) is between the contributions to the ^{15}N TROSY and AntiTROSY curves.

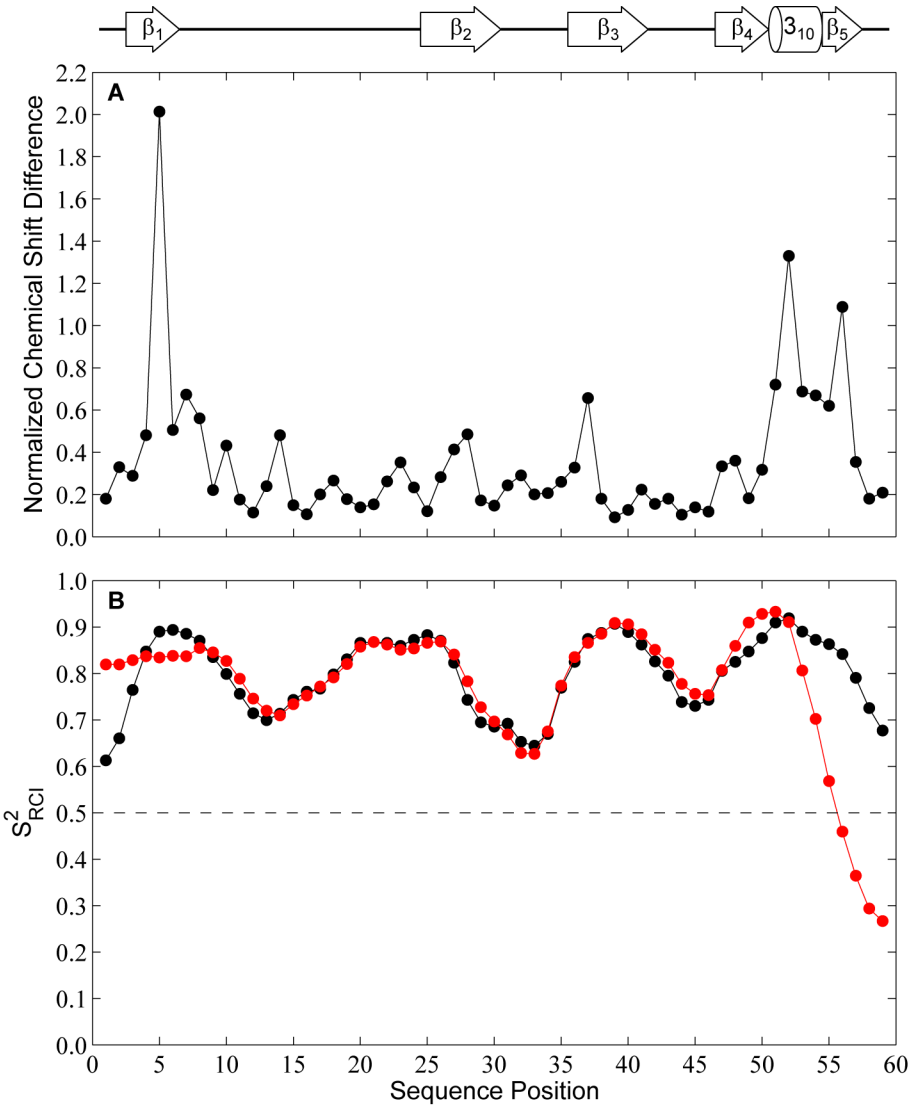

Fig. S5. (**A**) Backbone chemical shift differences between native and intermediate states normalized with respect to nucleus-specific variations in chemical shifts for each residue type as given in the BioMagResBank, $\Delta\varpi_i^{std}$, i ∈ {N, HN, CO, Cα, Hα}, according to

$$\Delta\varpi_{RMS,F\to I} = \sqrt{\frac{1}{N}\sum_i\left(\frac{\Delta\varpi_{i,F\to I}}{\Delta\varpi_i^{std}}\right)^2}$$

where the summation extends over all experimentally determined backbone chemical shift differences (N ≤ 5) for any given sequence position (*17*). Large differences are localized primarily to a region including the proximal NH_2- and COOH-termini comprising strands β_1 and β_5, consistent with the rest of the domain essentially retaining a native-like backbone conformation in the intermediate state. (**B**) Backbone order parameters predicted from the Random Coil Index (RCI), S^2_{RCI}, for the native (black) and intermediate (red) states of A39V/N53P/V55L Fyn SH3, calculated using the RCI Webserver with the default parameters (*108*). The dashed horizontal line indicates the cutoff of 0.5 used by TALOS+

(*82*) to identify highly dynamic protein segments. Unlike in the native state, the backbone chemical shifts of Pro57 to Asp59 in the intermediate are very close to random coil values (*109*), leading to very low S^2_{RCI} values. The regular secondary structure elements of the native A39V/N53P/V55L Fyn SH3 domain are indicated above the figure.

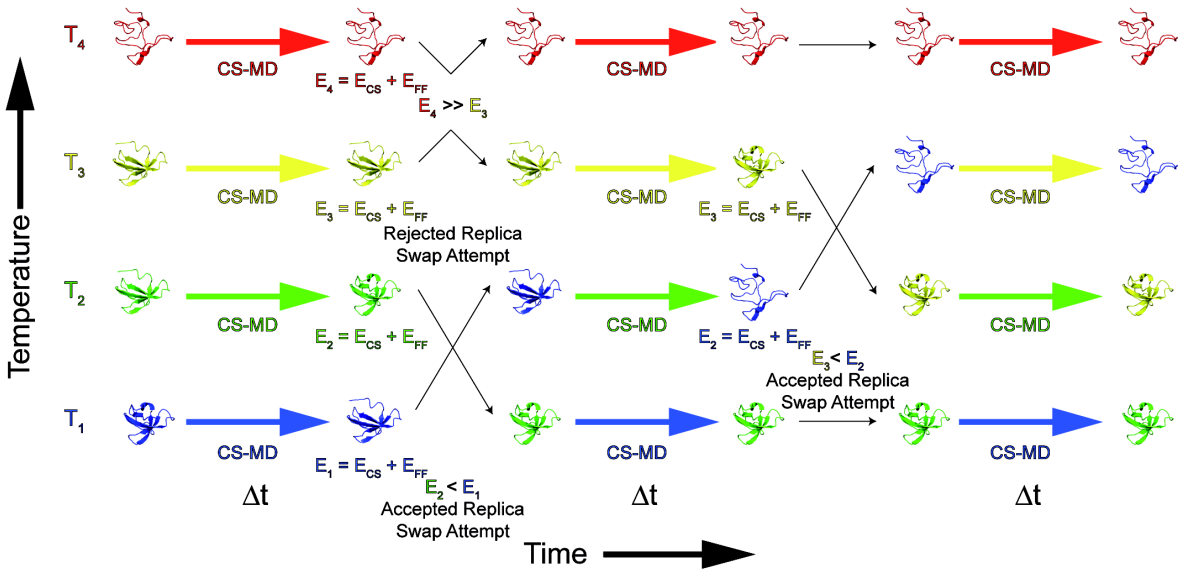

Fig. S6. Schematic illustration of the replica-exchange protocol that has been used to perform molecular dynamics simulations with chemical shift restraints (CamShift-MD). In this protocol chemical shifts are implemented as restraints to bias the molecular dynamics trajectory towards structures that are maximally compatible with the experimentally measured chemical shifts, in a manner similar to the use of NOE restraints in conventional NMR structure calculations (Fig. S7; *24*). Chemical shift restrained molecular dynamics (CS-MD) conformational searches are implemented using the replica-exchange algorithm (*88*) to enable a more thorough sampling of accessible conformational space. In this approach, multiple independent CS-MD simulations are run at different temperatures (21 temperatures, ranging from 270 K to 515 K, in this investigation). At a designated time interval Δt (6 ps in this investigation) replica swaps are attempted between the structures currently sampled by simulations running at adjacent temperatures. Before a replica swap attempt, the energies of the structures being considered for the swap are evaluated. The energy E_n of a structure sampled at a temperature T_n consists of a molecular mechanics force-field energy term (E_{FF}) and an additional chemical shift penalty term (E_{CS}), which penalizes deviations from the experimentally measured chemical shifts (Fig. S7). The differences between the energies of structures, normalized by the temperatures of their respective CS-MD simulations, dictate the probability that a swap attempt is successful. For structures sampled at temperatures T_i and T_k and with energies E_i and E_k, respectively, the swap is accepted if $(E_i - E_k)\times(1/(k_BT_k) - 1/(k_BT_i)) \leq 0$, where k_B is the Boltzmann constant. If $(E_i - E_k)\times(1/(k_BT_k) - 1/(k_BT_i)) > 0$ the swap is accepted with a probability equal to $\exp(-(E_i - E_k)\times(1/(k_BT_k) - 1/(k_BT_i)))$. If a swap attempt is accepted the structures are switched such that the structure previously sampled at temperature T_i will now be sampled at T_k and vice versa. If the swap attempt is rejected, the structures remain at their initial temperatures. As the sampling progresses, structures with the most favorable force-field energies and smallest chemical shift restraint penalties will preferentially be sampled at the lowest temperature T_{low} and structures with unfavorable force-field energies and large chemical shift penalties are sampled at higher temperatures, where they are more likely to undergo the large conformational adjustments necessary to improve their energies. Structures sampled at T_{low} are saved at regular intervals (0.6 ps) throughout the simulations

and after a designated sampling time the lowest-energy structures sampled at T_{low} are taken as the final structures.

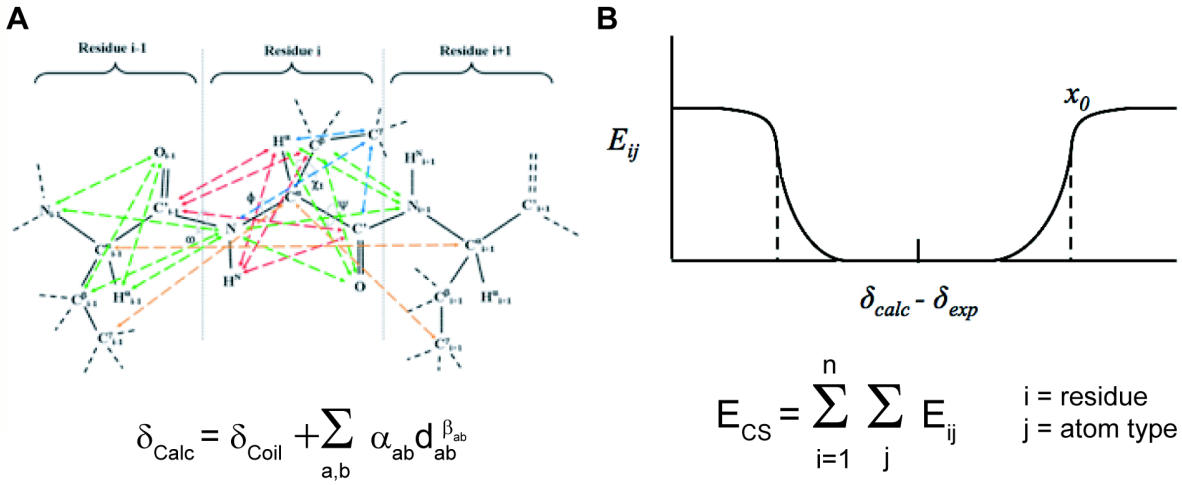

Fig. S7. Principle of the CamShift-MD protocol to perform molecular dynamics simulations with chemical shift restraints. (**A**) Chemical shifts are calculated from a given structure through the CamShift method based on differentiable functions that depend on interatomic distances (*90*). To calculate a chemical shift δ_{calc} for a given backbone query atom on residue i in a protein CamShift considers distances from the query atom to the backbone atoms of the neighboring residues i-1 and i+1, distances from the query atom to the side-chain atoms of residue i, distances to all non-bonded atoms within a 5 Å radius sphere originating from the query atom, a set of additional distances between atom pairs in residues i-1, i, and i+1 not involving the query atom (as depicted in **A**) and additional semi-empirical terms to capture the effects of aromatic ring currents and hydrogen bonds. For each pair of atoms a and b considered, the value of the interatomic distance d_{ab} makes a contribution to δ_{calc} based on the parameters α_{ab} and β_{ab}, which have been optimally fit from a database of structures and shifts for each pair of atoms considered in the calculations. The sum of these contributions is added to the random coil chemical shift δ_{coil} and combined with additional contributions from ring current effects and hydrogen bonds. (**B**) Illustration of the function used to convert CamShift backbone chemical shift calculations to an energetic penalty used to restrain molecular dynamics conformational searches (*24*). At each step in a molecular dynamics conformational search an energetic penalty contribution E_{ij} is calculated for each backbone atom for which an experimental chemical shift is available. The sum of these energetic penalties, E_{CS}, represents how well the current structure of the molecular dynamics conformational search agrees with the experimental chemical shifts. Differences between CamShift calculated chemical shifts, δ_{calc}, and experimental chemical shifts, δ_{exp}, are converted into forces which drive the molecular dynamics conformational search towards structures in better agreement with experimental chemical shifts. Force vectors are generated between all atom pairs a and b used as inputs in the CamShift calculation of δ_{calc} by evaluating the gradient of the chemical shift penalty function E_{CS} with respect to the Cartesian coordinates of a and b. This has the effect of pushing atom pairs apart or pulling atom pairs together in such a way that drives the values of δ_{calc} closer to δ_{exp} and lowers the energetic penalties E_{CS}. Figures adopted from (*24*, *90*).

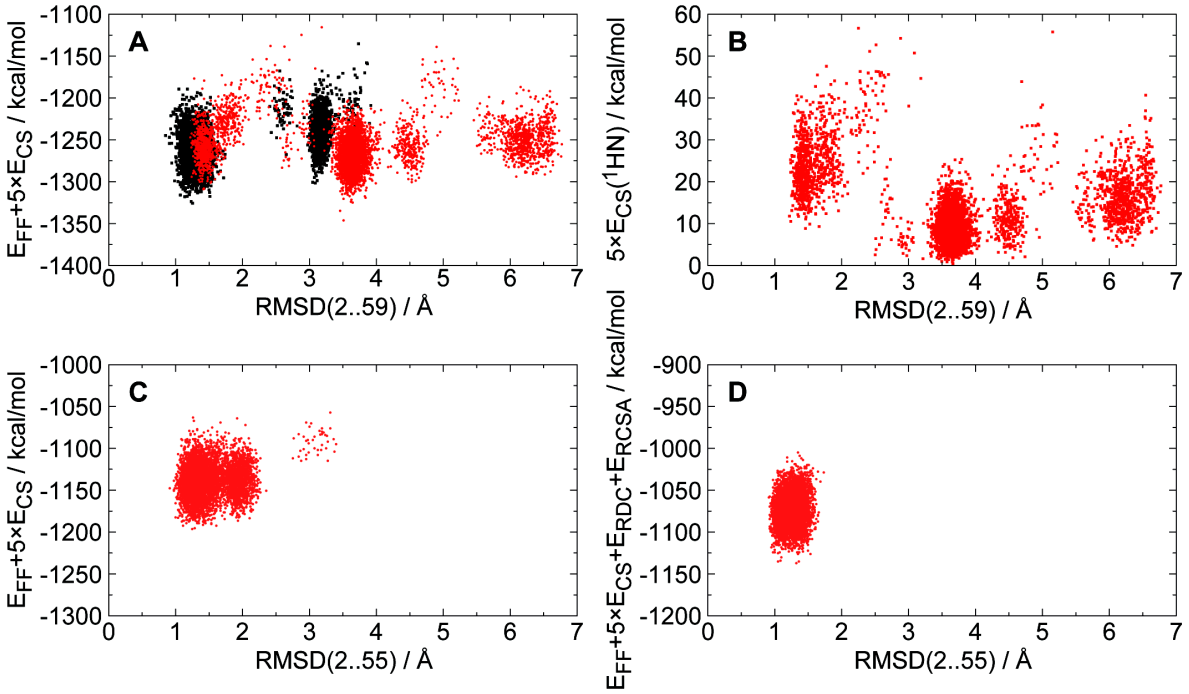

Fig. S8. (**A**) Comparison of the energy landscapes of the structures generated in the lowest-temperature replica of CS-MD simulations with native (black) and intermediate (red) state chemical shifts as a function of the backbone RMSD for residues 2 to 59 from the initial X-ray crystallography model of the native state (see above), which was used as a reference during the structure calculations (backbone RMSDs from the average solution structure of the native state are almost identical, see text). A cluster close to the native state is sampled extensively in both CS-MD simulations but favored by the target function $E_{sel} = E_{FF} + 5\times E_{CS}$ only in the positive control based on the native state chemical shifts. In contrast, the CS-MD simulations with the intermediate state chemical shifts favor a substantially different backbone conformation approximately 3.5 Å away. (**B**) The energy landscape for the same structures generated with intermediate state chemical shifts re-scored using only the CamShift penalty function for amide protons demonstrates that only the latter conformation is fully consistent with the amide proton chemical shifts of the intermediate state. (**C**) Energy landscape of the structures generated in the lowest-temperature replica of CS-MD simulations using the intermediate state chemical shifts as input as a function of the backbone RMSD for residues 2 to 55 from the X-ray crystallography model of the native state. The flexible COOH-terminal residues were excluded from the simulations. Note that the backbone RMSDs from the native state calculated for this shorter construct are smaller (≈1.5 Å vs. ≈3.5 Å, compare with **A**) because most of the backbone variability between intermediate and native state is located in the COOH-terminal region. (**D**) Refinement CS-MD simulations with intermediate state chemical shifts, D_{NH} RDCs, and ^{13}CO RCSAs narrow this energy landscape further (compare **D** with **C**), thereby yielding a set of well-defined structures representing the intermediate state at atomic resolution (Fig. 1A).

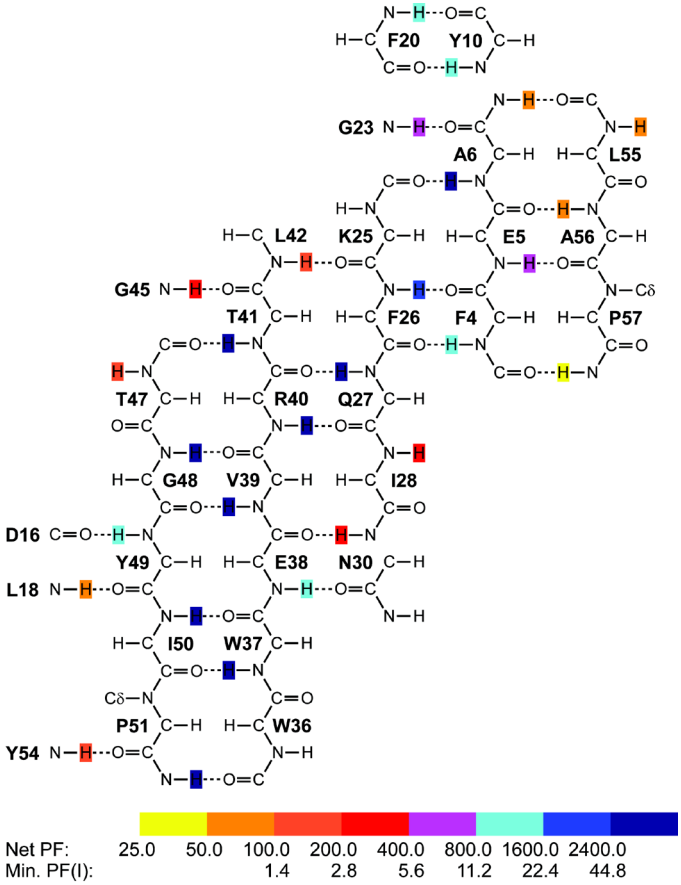

Fig. S9. Backbone hydrogen bond network of the native A39V/N53P/V55L Fyn SH3 domain with amide protons color-coded according to the net protection factors PF measured in H/D exchange NMR experiments. As described in detail under the section "H/D exchange NMR experiments" in Materials and Methods, the observed net H/D exchange rate for the experimental conditions here, k_{ex}, is given by the population-weighted average

$$k_{ex} = p_F\, k_F + p_I\, k_I + p_U\, k_{int}$$

where k_F and k_I are the H/D exchange rates for F and I, respectively, and we have assumed that $k_U = k_{int}$ (i. e., U is unstructured). The net protection factor PF can hence be written as

$$1/PF = k_{ex}/k_{int} = p_F/PF(F) + p_I/PF(I) + p_U$$

with $p_F \approx 98.6\%$, $p_I \approx 1.4\%$ and $p_U \approx 0.04\%$. From the above expression, net protection factors can be calculated and subsequently compared with experiment. For example, for an amide group in a region of structure that is well protected from exchange in both the native and intermediate states ($PF(F) = PF(I) = \infty$, $PF(U) = 1$) the net PF is given by $PF = 1/p_U \approx 2500$. By contrast, if there is no protection in the I state ($PF(F) = \infty$, $PF(I) = PF(U) = 1$) then the net $PF = 1/(p_I + p_U) \approx 70$. The predicted range of net PF values, ≈70-2500, is in excellent agreement with the experimentally determined range, indicated in the first line of the color-coding bar graph below the hydrogen bonding network. Therefore, the measured net PF values can be used to calculate the extent of protection in the intermediate state (assuming that $PF(F) = \infty$), which can then be compared with the structure of the intermediate, as a means of cross-validation. As described in the text, the folding

intermediate structure establishes that the topology of strands β_1, β_2, β_3, and β_4 is native-like, so that the net H/D exchange protection factors in this region, based on the arguments given above, are predicted to be on the order of 2500. Notably, the experimentally measured net PFs for amide groups in hydrogen bonds connecting strands β_1-β_2, β_2-β_3 and β_3-β_4 are ≈2500 (blue color-coding). By contrast, experimental net PF values for amide groups in hydrogen bonds between strands β_1 and β_5 are on the order of 70 (yellow/orange colors), indicating an absence of protection in the intermediate state. This is fully consistent with the structure of the folding intermediate, where strands β_1 and β_5 do not interact. Notably, the amide of Glu5 shows a considerably higher level of protection from exchange than what might naively be expected given that it is part of strand β_1. However, although strands β_1 and β_5 are not hydrogen bonded in the intermediate, the side-chain of Phe4 changes conformation to occupy the void created by the absence of strand β_5 and in so doing can protect the Glu5 amide hydrogen from solvent exchange (Fig. 1D). Again, this result is consistent with expectations based on the structure of the intermediate. The H/D exchange results thus provide strong cross-validation of the proposed structure for the A39V/N53P/V55L Fyn SH3 domain folding intermediate.

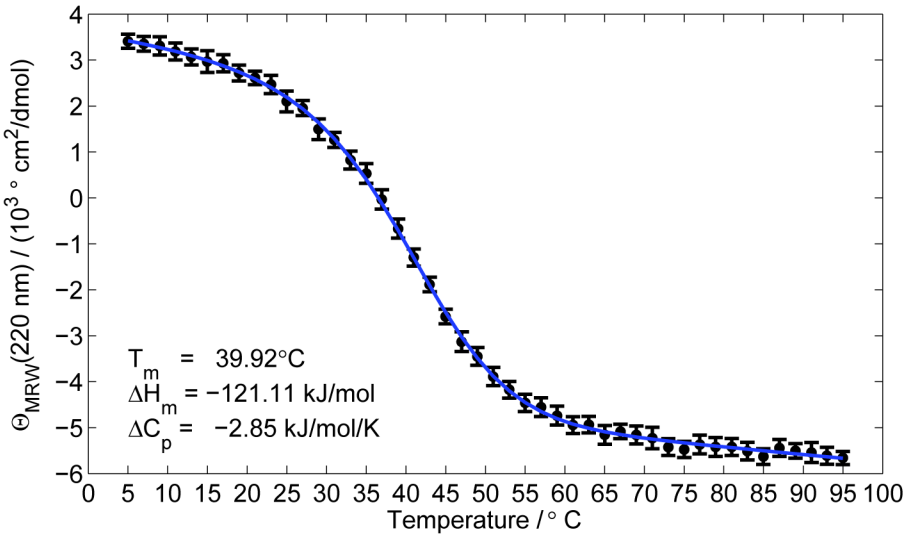

Fig. S10. Thermal unfolding profile of the A39V/N53P/V55L/Δ(57-60) Fyn SH3 domain monitored by CD ellipticity at 220 nm. The continuous line represents the two-state folding model with the specified temperature midpoint, T_m, van't Hoff enthalpy of folding, ΔH_m, specific heat of folding, $\Delta C_p(U \rightarrow F)$, and linear baselines for the two states that was fitted to the measured profile.

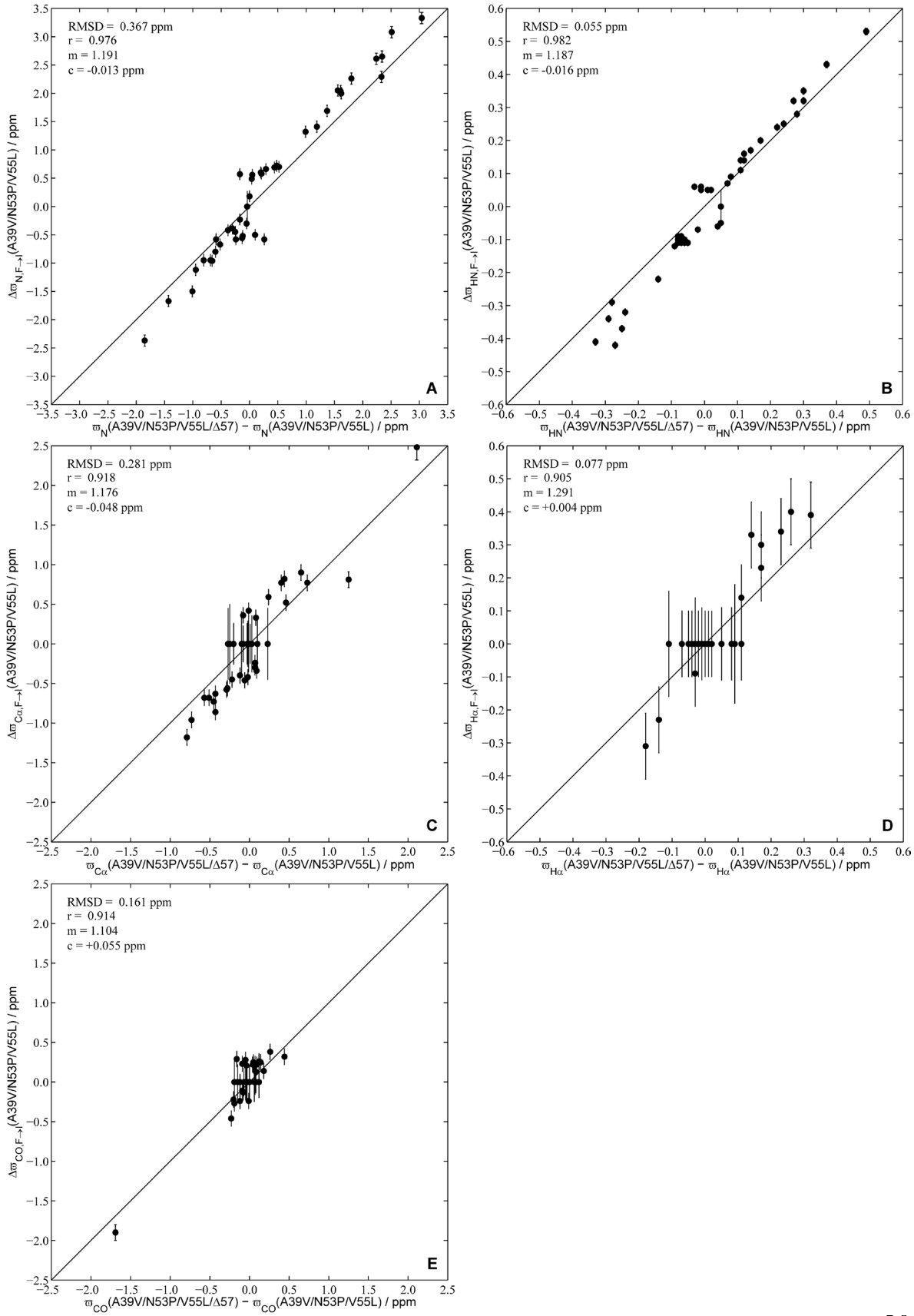
RMSD = 0.367 ppm
r = 0.976
m = 1.191
c = -0.013 ppm
A
$\Delta\varpi_{N,F\rightarrow}$(A39V/N53P/V55L) / ppm
ϖ_N(A39V/N53P/V55L/Δ57) − ϖ_N(A39V/N53P/V55L) / ppm
RMSD = 0.055 ppm
r = 0.982
m = 1.187
c = -0.016 ppm
B
$\Delta\varpi_{HN,F\rightarrow}$(A39V/N53P/V55L) / ppm
ϖ_{HN}(A39V/N53P/V55L/Δ57) − ϖ_{HN}(A39V/N53P/V55L) / ppm
RMSD = 0.281 ppm
r = 0.918
m = 1.176
c = -0.048 ppm
C
$\Delta\varpi_{C\alpha,F\rightarrow}$(A39V/N53P/V55L) / ppm
$\varpi_{C\alpha}$(A39V/N53P/V55L/Δ57) − $\varpi_{C\alpha}$(A39V/N53P/V55L) / ppm
RMSD = 0.077 ppm
r = 0.905
m = 1.291
c = +0.004 ppm
D
$\Delta\varpi_{H\alpha,F\rightarrow}$(A39V/N53P/V55L) / ppm
$\varpi_{H\alpha}$(A39V/N53P/V55L/Δ57) − $\varpi_{H\alpha}$(A39V/N53P/V55L) / ppm
RMSD = 0.161 ppm
r = 0.914
m = 1.104
c = +0.055 ppm
E
$\Delta\varpi_{CO,F\rightarrow}$(A39V/N53P/V55L) / ppm
ϖ_{CO}(A39V/N53P/V55L/Δ57) − ϖ_{CO}(A39V/N53P/V55L) / ppm

Fig. S11. Comparison of the chemical shift differences between the directly visible ground state resonances of the A39V/N53P/V55L/Δ(57-60) and A39V/N53P/V55L Fyn SH3 domains (abscissa), $\Delta\varpi_{X,\Delta 57} = \varpi_X$(A39V/N53P/V55L/Δ(57-60)) – ϖ_X(A39V/N53P/V55L), with the differences between native and intermediate states, $\Delta\varpi_{X,F\rightarrow I}$, of the A39V/N53P/V55L Fyn SH3 extracted from CPMG RD experiments (ordinate) for the following backbone nuclei: X = ^{15}N (**A**), ^{1}HN (**B**), ^{13}Cα (**C**), ^{1}Hα (**D**), ^{13}CO (**E**). The excellent correlation demonstrates that the A39V/N53P/V55L/Δ(57-60) Fyn SH3 domain is a faithful structural mimic of the folding intermediate of the A39V/N53P/V55L Fyn SH3. The diagonals are indicated by continuous lines. Best-fit slopes, m, and intercepts, c, were calculated by linear regression. Note that a slope slightly larger than unity is expected for a system undergoing chemical exchange because exchange contributions move the apparent peak positions in the directly visible spectra slightly towards the excited state (*68*) so that the difference between the peak positions for the two mutants, $\Delta\varpi_{X,\Delta 57}$, slightly underestimates the true difference in chemical shifts between their respective ground states.

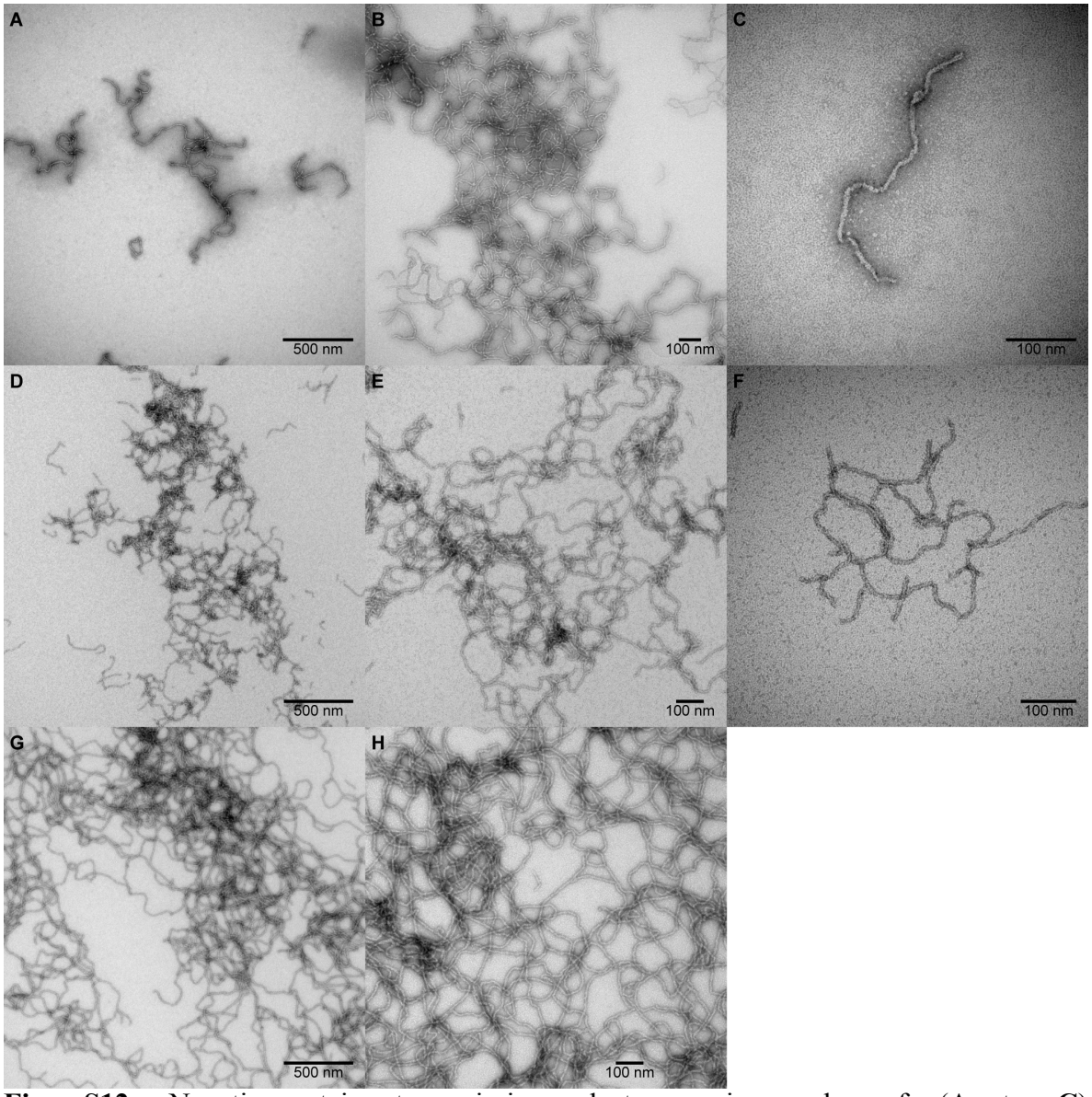

Fig. S12. Negative stain transmission electron micrographs of (**A** to **C**) A39V/N53P/V55L/Δ(57-60), (**D** to **F**) Δ(56-60), and (**G** and **H**) Δ(57-60) Fyn SH3 aggregates at different magnifications as indicated by the scale bars. Note that in **D** to **H** the Fyn SH3 constructs are wild-type up to the truncation at residue 56 or 57.

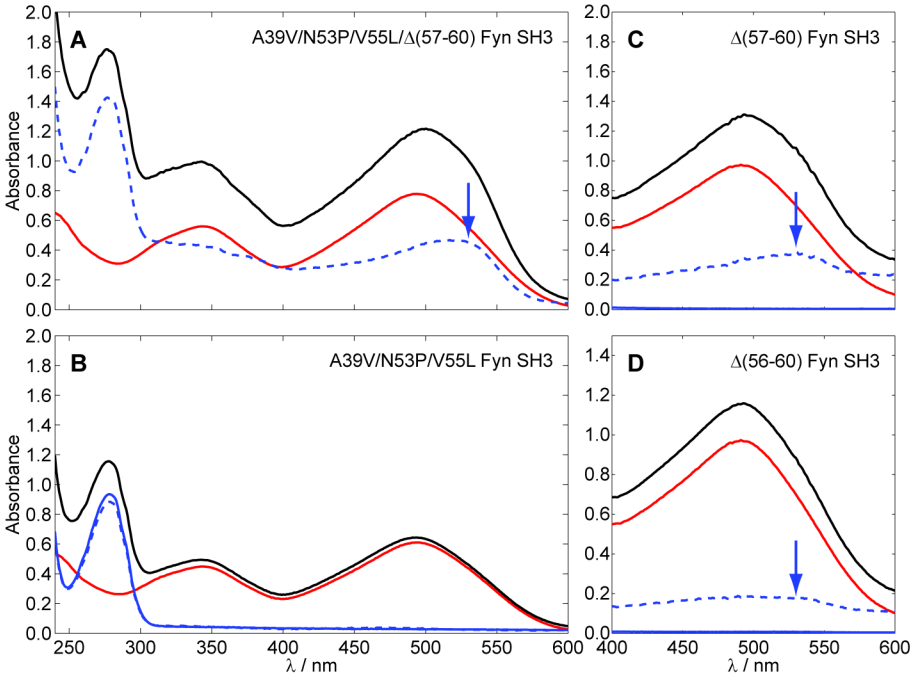

Fig. S13. Overlay of the absorption spectra of 20 μM Congo red in the absence (red) and presence (black) of (**A**) A39V/N53P/V55L/Δ(57-60) Fyn SH3 aggregates (same graph as Fig. 3E), (**B**) A39V/N53P/V55L Fyn SH3 as a negative control, (**C**) Δ(57-60) Fyn SH3, and (**D**) Δ(56-60) Fyn SH3. The difference spectra (dashed blue) show a clear hyperchromicity and red shift of the absorbance band around 500 nm (arrow), except in the case of the monomeric negative control (**B**), in which the difference spectrum (dashed blue) is identical to the absorption spectrum of the SH3 domain alone (continuous blue). Note that (**C**) and (**D**) show only the region of interest.

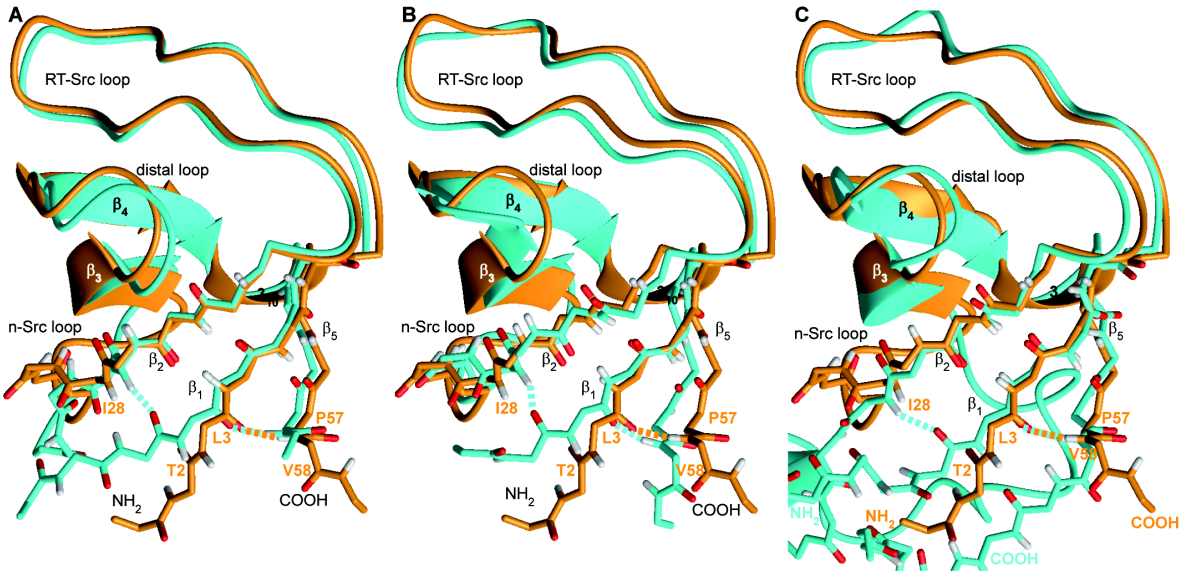

Fig. S14. Backbone overlay in schematic representation of the lowest-energy solution structure of the native state of the A39V/N53P/V55L Fyn SH3 domain (orange) with the (**A**) Yes (PDB 2HDA; *32*), (**B**) N47A α-spectrin (PDB 1QKX; *110*), and (**C**) PI3K (PDB 2PNI; *111*) SH3 domains (cyan). Same view as in Fig. 1C. Comparison with Fig. 1C reveals that all three SH3 domains form the hydrogen bond (dashed lines) corresponding to the non-native interaction between Ile28 HN and Thr2 CO stabilizing the folding intermediate of the A39V/N53P/V55L Fyn SH3 domain.

References

1. C. M. Dobson, Protein folding and misfolding. *Nature* **426**, 884 (2003). doi:10.1038/nature02261 Medline

2. D. J. Selkoe, Folding proteins in fatal ways. *Nature* **426**, 900 (2003). doi:10.1038/nature02264 Medline

3. F. Chiti, C. M. Dobson, Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333 (2006). doi:10.1146/annurev.biochem.75.101304.123901 Medline

4. F. Chiti, C. M. Dobson, Amyloid formation by globular proteins under native conditions. *Nat. Chem. Biol.* **5**, 15 (2009). doi:10.1038/nchembio.131 Medline

5. M. Bucciantini *et al*., Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* **416**, 507 (2002). doi:10.1038/416507a Medline

6. R. Tycko, Solid-state NMR studies of amyloid fibril structure. *Annu. Rev. Phys. Chem.* **62**, 279 (2011). doi:10.1146/annurev-physchem-032210-103539 Medline

7. V. N. Uversky, A. L. Fink, Conformational constraints for amyloid fibrillation: The importance of being unfolded. *Biochim. Biophys. Acta* **1698**, 131 (2004). Medline

8. R. Kayed *et al*., Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* **300**, 486 (2003). doi:10.1126/science.1079469 Medline

9. C. Haass, D. J. Selkoe, Soluble protein oligomers in neurodegeneration: Lessons from the Alzheimer's amyloid β-peptide. *Nat. Rev. Mol. Cell Biol.* **8**, 101 (2007). doi:10.1038/nrm2101 Medline

10. R. Roychaudhuri, M. Yang, M. M. Hoshi, D. B. Teplow, Amyloid β-protein assembly and Alzheimer disease. *J. Biol. Chem.* **284**, 4749 (2009). doi:10.1074/jbc.R800036200 Medline

11. T. R. Jahn, M. J. Parker, S. W. Homans, S. E. Radford, Amyloid formation under physiological conditions proceeds via a native-like folding intermediate. *Nat. Struct. Mol. Biol.* **13**, 195 (2006). doi:10.1038/nsmb1058 Medline

12. C. M. Eakin, A. J. Berman, A. D. Miranker, A native to amyloidogenic transition regulated by a backbone trigger. *Nat. Struct. Mol. Biol.* **13**, 202 (2006). doi:10.1038/nsmb1068 Medline

13. A. Fersht, *Structure and Mechanism in Protein Science* (Freeman, New York, 1999).

14. A. P. Capaldi, S. E. Radford, Kinetic studies of β-sheet protein folding. *Curr. Opin. Struct. Biol.* **8**, 86 (1998). doi:10.1016/S0959-440X(98)80014-6 Medline

15. A. G. Palmer 3rd, NMR characterization of the dynamics of biomacromolecules. *Chem. Rev.* **104**, 3623 (2004). doi:10.1021/cr030413t Medline

16. P. Neudecker, P. Lundström, L. E. Kay, Relaxation dispersion NMR spectroscopy as a tool for detailed studies of protein folding. *Biophys. J.* **96**, 2045 (2009). doi:10.1016/j.bpj.2008.12.3907 Medline

17. P. Vallurupalli, D. F. Hansen, L. E. Kay, Structures of invisible, excited protein states by relaxation dispersion NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 11766 (2008). doi:10.1073/pnas.0804221105 Medline

18. D. M. Korzhnev, T. L. Religa, W. Banachewicz, A. R. Fersht, L. E. Kay, A transient and low-populated protein-folding intermediate at atomic resolution. *Science* **329**, 1312 (2010). doi:10.1126/science.1191723 Medline

19. P. Neudecker *et al*., 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. *J. Mol. Biol.* **363**, 958 (2006). doi:10.1016/j.jmb.2006.08.047 Medline

20. A. Zarrine-Afsar *et al*., Theoretical and experimental demonstration of the importance of specific nonnative interactions in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 9999 (2008). doi:10.1073/pnas.0801874105 Medline

21. D. S. Riddle *et al*., Experiment and theory highlight role of native state topology in SH3 folding. *Nat. Struct. Biol.* **6**, 1016 (1999). doi:10.1038/14901 Medline

22. P. Neudecker, A. Zarrine-Afsar, A. R. Davidson, L. E. Kay, Φ-value analysis of a three-state protein folding pathway by NMR relaxation dispersion spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15717 (2007). doi:10.1073/pnas.0705097104 Medline

23. Y. Bai, Protein folding pathways studied by pulsed- and native-state hydrogen exchange. *Chem. Rev.* **106**, 1757 (2006). doi:10.1021/cr040432i Medline

24. P. Robustelli, K. Kohlhoff, A. Cavalli, M. Vendruscolo, Using NMR chemical shifts as structural restraints in molecular dynamics simulations of proteins. *Structure* **18**, 923 (2010). doi:10.1016/j.str.2010.04.016 Medline

25. G. G. Tartaglia *et al*., Prediction of aggregation-prone regions in structured proteins. *J. Mol. Biol.* **380**, 425 (2008). doi:10.1016/j.jmb.2008.05.013 Medline

26. J. S. Richardson, D. C. Richardson, Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 2754 (2002). doi:10.1073/pnas.052706099 Medline

27. N. Carulla *et al*., Experimental characterization of disordered and ordered aggregates populated during the process of amyloid fibril formation. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 7828 (2009). doi:10.1073/pnas.0812227106 Medline

28. M. J. Bayro *et al*., High-resolution MAS NMR analysis of PI3-SH3 amyloid fibrils: Backbone conformation and implications for protofilament assembly and structure. *Biochemistry* **49**, 7474 (2010). doi:10.1021/bi100864t Medline

29. L. Varela, B. Morel, A. I. Azuaga, F. Conejero-Lara, A single mutation in an SH3 domain increases amyloid aggregation by accelerating nucleation, but not by destabilizing thermodynamically the native state. *FEBS Lett.* **583**, 801 (2009). doi:10.1016/j.febslet.2009.01.033 Medline

30. B. Morel, L. Varela, A. I. Azuaga, F. Conejero-Lara, Environmental conditions affect the kinetics of nucleation of amyloid fibrils and determine their morphology. *Biophys. J.* **99**, 3801 (2010). doi:10.1016/j.bpj.2010.10.039 Medline

31. J. M. Martín-García, I. Luque, P. L. Mateo, J. Ruiz-Sanz, A. Cámara-Artigas, Crystallographic structure of the SH3 domain of the human c-Yes tyrosine kinase: Loop flexibility and amyloid aggregation. *FEBS Lett.* **581**, 1701 (2007). doi:10.1016/j.febslet.2007.03.059 Medline

32. S. Pechmann, E. D. Levy, G. G. Tartaglia, M. Vendruscolo, Physicochemical principles that regulate the competition between functional and dysfunctional association of proteins. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 10159 (2009). doi:10.1073/pnas.0812414106 Medline

33. P. Lundström, D. F. Hansen, P. Vallurupalli, L. E. Kay, Accurate measurement of alpha proton chemical shifts of excited protein states by relaxation dispersion NMR spectroscopy. *J. Am. Chem. Soc.* **131**, 1915 (2009). doi:10.1021/ja807796a Medline

34. P. Lundström *et al*., Fractional ^{13}C enrichment of isolated carbons using [1-^{13}C]- or [2-^{13}C]-glucose facilitates the accurate measurement of dynamics at backbone Cα and side-chain methyl positions in proteins. *J. Biomol. NMR* **38**, 199 (2007). doi:10.1007/s10858-007-9158-6 Medline

35. K. L. Maxwell, A. R. Davidson, Mutagenesis of a buried polar interaction in an SH3 domain: Sequence conservation provides the best prediction of stability effects. *Biochemistry* **37**, 16172 (1998). doi:10.1021/bi981788p Medline

36. O. Zhang, L. E. Kay, J. P. Olivier, J. D. Forman-Kay, 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 (1994). doi:10.1007/BF00398413 Medline

37. S. Grzesiek, A. Bax, Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein. *J. Magn. Reson.* **96**, 432 (1992).

38. G. W. Vuister, A. Bax, Resolution enhancement and spectral editing of uniformly ^{13}C-enriched proteins by homonuclear broadband ^{13}C decoupling. *J. Magn. Reson.* **98**, 428 (1992).

39. D. F. Hansen, P. Neudecker, P. Vallurupalli, F. A. A. Mulder, L. E. Kay, Determination of Leu side-chain conformations in excited protein states by NMR relaxation dispersion. *J. Am. Chem. Soc.* **132**, 42 (2010). doi:10.1021/ja909294n Medline

40. D. F. Hansen, P. Neudecker, L. E. Kay, Determination of isoleucine side-chain conformations in ground and excited states of proteins from chemical shifts. *J. Am. Chem. Soc.* **132**, 7589 (2010). doi:10.1021/ja102090z Medline

41. D. Marion, M. Ikura, R. Tschudin, A. Bax, Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins. *J. Magn. Reson.* **85**, 393 (1989).

42. L. E. Kay, P. Keifer, T. Saarinen, Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc.* **114**, 10663 (1992). doi:10.1021/ja00052a088

43. J. Schleucher, M. Sattler, C. Griesinger, Coherence selection by gradients without signal attenuation: Application to the three-dimensional HNCO experiment. *Angew. Chem. Int. Ed. Engl.* **32**, 1489 (1993). doi:10.1002/anie.199314891

44. F. Delaglio *et al*., NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277 (1995). doi:10.1007/BF00197809 Medline

45. B. A. Johnson, Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.* **278**, 313 (2004). Medline

46. J. L. Markley *et al*., Recommendations for the presentation of NMR structures of proteins and nucleic acids (IUPAC recommendations 1998). *Pure Appl. Chem.* **70**, 117 (1998). doi:10.1351/pac199870010117

47. K. H. Gardner, M. K. Rosen, L. E. Kay, Global folds of highly deuterated, methyl-protonated proteins by multidimensional NMR. *Biochemistry* **36**, 1389 (1997). doi:10.1021/bi9624806 Medline

48. J. Cavanagh, M. Rance, Suppression of cross-relaxation effects in TOCSY spectra via a modified DIPSI-2 mixing sequence. *J. Magn. Reson.* **96**, 670 (1992).

49. H. Y. Carr, E. M. Purcell, Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev.* **94**, 630 (1954). doi:10.1103/PhysRev.94.630

50. S. Meiboom, D. Gill, Modified spin-echo method for measuring nuclear relaxation times. *Rev. Sci. Instrum.* **29**, 688 (1958). doi:10.1063/1.1716296

51. A. Allerhand, H. S. Gutowsky, J. Jonas, R. A. Meinzer, Nuclear magnetic resonance methods for determining chemical-exchange rates. *J. Am. Chem. Soc.* **88**, 3185 (1966). doi:10.1021/ja00966a001 Medline

52. C. Deverell, R. E. Morgan, J. H. Strange, Studies of chemical exchange by nuclear magnetic relaxation in the rotating frame. *Mol. Phys.* **18**, 553 (1970). doi:10.1080/00268977000100611

53. D. F. Hansen, P. Vallurupalli, L. E. Kay, An improved ^{15}N relaxation dispersion experiment for the measurement of millisecond time-scale dynamics in proteins. *J. Phys. Chem. B* **112**, 5898 (2008). doi:10.1021/jp074793o Medline

54. P. Vallurupalli, D. F. Hansen, E. Stollar, E. Meirovitch, L. E. Kay, Measurement of bond vector orientations in invisible excited states of proteins. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 18473 (2007). doi:10.1073/pnas.0708296104 Medline

55. R. Ishima, D. A. Torchia, Extending the range of amide proton relaxation dispersion experiments in proteins using a constant-time relaxation-compensated CPMG approach. *J. Biomol. NMR* **25**, 243 (2003). doi:10.1023/A:1022851228405 Medline

56. V. Y. Orekhov, D. M. Korzhnev, L. E. Kay, Double- and zero-quantum NMR relaxation dispersion experiments sampling millisecond time scale dynamics in proteins. *J. Am. Chem. Soc.* **126**, 1886 (2004). doi:10.1021/ja038620y Medline

57. P. Lundström, D. F. Hansen, L. E. Kay, Measurement of carbonyl chemical shifts of excited protein states by relaxation dispersion NMR spectroscopy: Comparison between uniformly and selectively ^{13}C labeled samples. *J. Biomol. NMR* **42**, 35 (2008). doi:10.1007/s10858-008-9260-4 Medline

58. D. F. Hansen, P. Vallurupalli, P. Lundström, P. Neudecker, L. E. Kay, Probing chemical shifts of invisible states of proteins with relaxation dispersion NMR spectroscopy: how well can we do? *J. Am. Chem. Soc.* **130**, 2667 (2008). doi:10.1021/ja078337p Medline

59. V. Y. Orekhov, I. V. Ibraghimov, M. Billeter, MUNIN: A new approach to multi-dimensional NMR spectra interpretation. *J. Biomol. NMR* **20**, 49 (2001). doi:10.1023/A:1011234126930 Medline

60. D. M. Korzhnev, I. V. Ibraghimov, M. Billeter, V. Y. Orekhov, MUNIN: Application of three-way decomposition to the analysis of heteronuclear NMR relaxation data. *J. Biomol. NMR* **21**, 263 (2001). doi:10.1023/A:1012982830367 Medline

61. F. A. A. Mulder, N. R. Skrynnikov, B. Hon, F. W. Dahlquist, L. E. Kay, Measurement of slow (μs-ms) time scale dynamics in protein side chains by ^{15}N relaxation dispersion NMR spectroscopy: Application to Asn and Gln residues in a cavity mutant of T4 lysozyme. *J. Am. Chem. Soc.* **123**, 967 (2001). doi:10.1021/ja003447g Medline

62. D. M. Korzhnev *et al*., Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* **430**, 586 (2004). doi:10.1038/nature02655 Medline

63. R. Ishima, D. A. Torchia, Error estimation and global fitting in transverse-relaxation dispersion experiments to determine chemical-exchange parameters. *J. Biomol. NMR* **32**, 41 (2005). doi:10.1007/s10858-005-3593-z Medline

64. H. M. McConnell, Reaction rates by nuclear magnetic resonance. *J. Chem. Phys.* **28**, 430 (1958). doi:10.1063/1.1744152

65. D. M. Korzhnev, P. Neudecker, A. Mittermaier, V. Y. Orekhov, L. E. Kay, 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 (2005). doi:10.1021/ja054550e Medline

66. W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes* (Cambride Univ. Press, Cambride, ed. 3, 2007).

67. P. Neudecker, D. M. Korzhnev, L. E. Kay, 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 (2006). doi:10.1007/s10858-006-0001-2 Medline

68. N. R. Skrynnikov, F. W. Dahlquist, L. E. Kay, Reconstructing NMR spectra of "invisible" excited protein states using HSQC and HMQC experiments. *J. Am. Chem. Soc.* **124**, 12352 (2002). doi:10.1021/ja0207089 Medline

69. R. Auer *et al*., Measuring the signs of $^{1}H^{\alpha}$ chemical shift differences between ground and excited protein states by off-resonance spin-lock $R_{1\rho}$ NMR spectroscopy. *J. Am. Chem. Soc.* **131**, 10832 (2009). doi:10.1021/ja904315m Medline

70. R. Auer *et al*., Measurement of signs of chemical shift differences between ground and excited protein states: A comparison between H(S/M)QC and $R_{1\rho}$ methods. *J. Biomol. NMR* **46**, 205 (2010). doi:10.1007/s10858-009-9394-z Medline

71. G. Bouvignies *et al*., A simple method for measuring signs of $^{1}H^{N}$ chemical shift differences between ground and excited protein states. *J. Biomol. NMR* **47**, 135 (2010). doi:10.1007/s10858-010-9418-8 Medline

72. M. R. Hansen, L. Mueller, A. Pardi, Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nat. Struct. Biol.* **5**, 1065 (1998). doi:10.1038/4176 Medline

73. M. Rückert, G. Otting, Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *J. Am. Chem. Soc.* **122**, 7793 (2000). doi:10.1021/ja001068h

74. M. Ottiger, F. Delaglio, A. Bax, Measurement of *J* and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.* **131**, 373 (1998). doi:10.1006/jmre.1998.1361 Medline

75. M. Zweckstetter, A. Bax, Characterization of molecular alignment in aqueous suspensions of Pf1 bacteriophage. *J. Biomol. NMR* **20**, 365 (2001). doi:10.1023/A:1011263920003 Medline

76. G. M. Clore, A. M. Gronenborn, A. Bax, A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information. *J. Magn. Reson.* **133**, 216 (1998). doi:10.1006/jmre.1998.1419 Medline

77. G. Cornilescu, A. Bax, Measurement of proton, nitrogen, and carbonyl chemical shielding anisotropies in a protein dissolved in a dilute liquid crystalline phase. *J. Am. Chem. Soc.* **122**, 10143 (2000). doi:10.1021/ja0016194

78. P. Vallurupalli, D. F. Hansen, L. E. Kay, Probing structure in invisible protein states with anisotropic NMR chemical shifts. *J. Am. Chem. Soc.* **130**, 2734 (2008). doi:10.1021/ja710817g Medline

79. G. Cornilescu, J. L. Marquardt, M. Ottiger, A. Bax, Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. *J. Am. Chem. Soc.* **120**, 6836 (1998). doi:10.1021/ja9812610

80. A. Rosato *et al*., CASD-NMR: Critical assessment of automated structure determination by NMR. *Nat. Methods* **6**, 625 (2009). doi:10.1038/nmeth0909-625 Medline

81. A. Cavalli, X. Salvatella, C. M. Dobson, M. Vendruscolo, Protein structure determination from NMR chemical shifts. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 9615 (2007). doi:10.1073/pnas.0610313104 Medline

82. Y. Shen, F. Delaglio, G. Cornilescu, A. Bax, TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213 (2009). doi:10.1007/s10858-009-9333-z Medline

83. M. Nilges, G. M. Clore, A. M. Gronenborn, Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Lett.* **229**, 317 (1988). doi:10.1016/0014-5793(88)81148-7 Medline

84. Y. Duan *et al*., A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **24**, 1999 (2003). doi:10.1002/jcc.10349 Medline

85. D. Bashford, D. A. Case, Generalized born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem.* **51**, 129 (2000). doi:10.1146/annurev.physchem.51.1.129 Medline

86. G. M. Clore, A. M. Gronenborn, N. Tjandra, Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *J. Magn. Reson.* **131**, 159 (1998). doi:10.1006/jmre.1997.1345 Medline

87. R. S. Lipsitz, N. Tjandra, Carbonyl CSA restraints from solution NMR for protein structure refinement. *J. Am. Chem. Soc.* **123**, 11065 (2001). doi:10.1021/ja016854g Medline

88. Y. Sugita, Y. Okamoto, Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **314**, 141 (1999). doi:10.1016/S0009-2614(99)01123-9

89. J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **23**, 327 (1977). doi:10.1016/0021-9991(77)90098-5

90. K. J. Kohlhoff, P. Robustelli, A. Cavalli, X. Salvatella, M. Vendruscolo, Fast and accurate predictions of protein NMR chemical shifts from interatomic distances. *J. Am. Chem. Soc.* **131**, 13894 (2009). doi:10.1021/ja903772t Medline

91. C. D. Schwieters, J. J. Kuszewski, G. M. Clore, Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **48**, 47 (2006). doi:10.1016/j.pnmrs.2005.10.001

92. A. T. Brünger, *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, 1992).

93. R. A. Laskowski, J. A. C. Rullmannn, M. W. MacArthur, R. Kaptein, J. M. Thornton, AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477 (1996). doi:10.1007/BF00228148 Medline

94. A. Bax, A. Grishaev, Weak alignment NMR: A hawk-eyed view of biomolecular structure. *Curr. Opin. Struct. Biol.* **15**, 563 (2005). doi:10.1016/j.sbi.2005.08.006 Medline

95. J.-C. Hus *et al*., 16-fold degeneracy of peptide plane orientations from residual dipolar couplings: Analytical treatment and implications for protein structure determination. *J. Am. Chem. Soc.* **130**, 15927 (2008). doi:10.1021/ja804274s Medline

96. M. M. G. Krishna, L. Hoang, Y. Lin, S. W. Englander, Hydrogen exchange methods to study protein folding. *Methods* **34**, 51 (2004). doi:10.1016/j.ymeth.2004.03.005 Medline

97. K. Teilum, B. B. Kragelund, F. M. Poulsen, in *Protein Folding Handbook*, J. Buchner, T. Kiefhaber, Eds. (Wiley-VCH, Weinheim, Germany, 2008), vol. I, chap. 18.

98. Y. Bai, J. S. Milne, L. Mayne, S. W. Englander, Primary structure effects on peptide group hydrogen exchange. *Proteins* **17**, 75 (1993). doi:10.1002/prot.340170110 Medline

99. W. A. Houry, J. M. Sauder, H. Roder, H. A. Scheraga, Definition of amide protection factors for early kinetic intermediates in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4299 (1998). doi:10.1073/pnas.95.8.4299 Medline

100. S. W. Englander, L. Mayne, M. M. G. Krishna, Protein folding and misfolding: Mechanism and principles. *Q. Rev. Biophys.* **40**, 287 (2007). doi:10.1017/S0033583508004654 Medline

101. W. E. Klunk, J. W. Pettegrew, D. J. Abraham, Two simple methods for quantifying low-affinity dye-substrate binding. *J. Histochem. Cytochem.* **37**, 1293 (1989). doi:10.1177/37.8.2666512 Medline

102. W. E. Klunk, R. F. Jacob, R. P. Mason, Quantifying amyloid β-peptide (Aβ) aggregation using the Congo red-Aβ (CR-Aβ) spectrophotometric assay. *Anal. Biochem.* **266**, 66 (1999). doi:10.1006/abio.1998.2933 Medline

103. J. G. B. Northey, A. A. Di Nardo, A. R. Davidson, Hydrophobic core packing in the SH3 domain folding transition state. *Nat. Struct. Biol.* **9**, 126 (2002). doi:10.1038/nsb748 Medline

104. A. Bax, G. Kontaxis, N. Tjandra, Dipolar couplings in macromolecular structure determination. *Methods Enzymol.* **339**, 127 (2001). doi:10.1016/S0076-6879(01)39313-8 Medline

105. D. M. Korzhnev, L. E. Kay, Probing invisible, low-populated states of protein molecules by relaxation dispersion NMR spectroscopy: An application to protein folding. *Acc. Chem. Res.* **41**, 442 (2008). doi:10.1021/ar700189y Medline

106. D. F. Hansen, P. Vallurupalli, L. E. Kay, Using relaxation dispersion NMR spectroscopy to determine structures of excited, invisible protein states. *J. Biomol. NMR* **41**, 113 (2008). doi:10.1007/s10858-008-9251-5 Medline

107. A. K. Mittermaier, L. E. Kay, Observing biological dynamics at atomic resolution using NMR. *Trends Biochem. Sci.* **34**, 601 (2009). doi:10.1016/j.tibs.2009.07.004 Medline

108. M. V. Berjanskii, D. S. Wishart, A simple method to predict protein flexibility using secondary chemical shifts. *J. Am. Chem. Soc.* **127**, 14970 (2005). doi:10.1021/ja054842f Medline

109. D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, B. D. Sykes, ^{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 (1995). doi:10.1007/BF00227471 Medline

110. M. C. Vega, L. Serrano, J. C. Martínez, Thermodynamic and structural characterization of Asn and Ala residues in the disallowed II' region of the Ramachandran plot. *Protein Sci.* **9**, 2322 (2000). doi:10.1110/ps.9.12.2322 Medline

111. G. W. Booker *et al*., Solution structure and ligand-binding site of the SH3 domain of the p85α subunit of phosphatidylinositol 3-kinase. *Cell* **73**, 813 (1993). doi:10.1016/0092-8674(93)90259-S Medline