Non-Native Interactions in the FF Domain Folding Pathway From an Atomic

Resolution Structure of a Sparsely Populated Intermediate: An NMR Relaxation

Dispersion Study

Dmitry M. Korzhnev, Robert M. Vernon, Tomasz L. Religa, Alexandar L. Hansen, David Baker, Alan R. Fersht and Lewis E. Kay

Full citation for reference 18:

Shen, Y.; Lange, O.; Delaglio, F.; Rossi, P.; Aramini, J. M.; Liu, G.; Eletsky, A.; Wu, Y.; Singarapu, K. K.; Lemak, A.; Ignatchenko, A.; Arrowsmith, C. H.; Szyperski, T.; Montelione, G. T.; Baker, D.; Bax, A. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 4685-90.
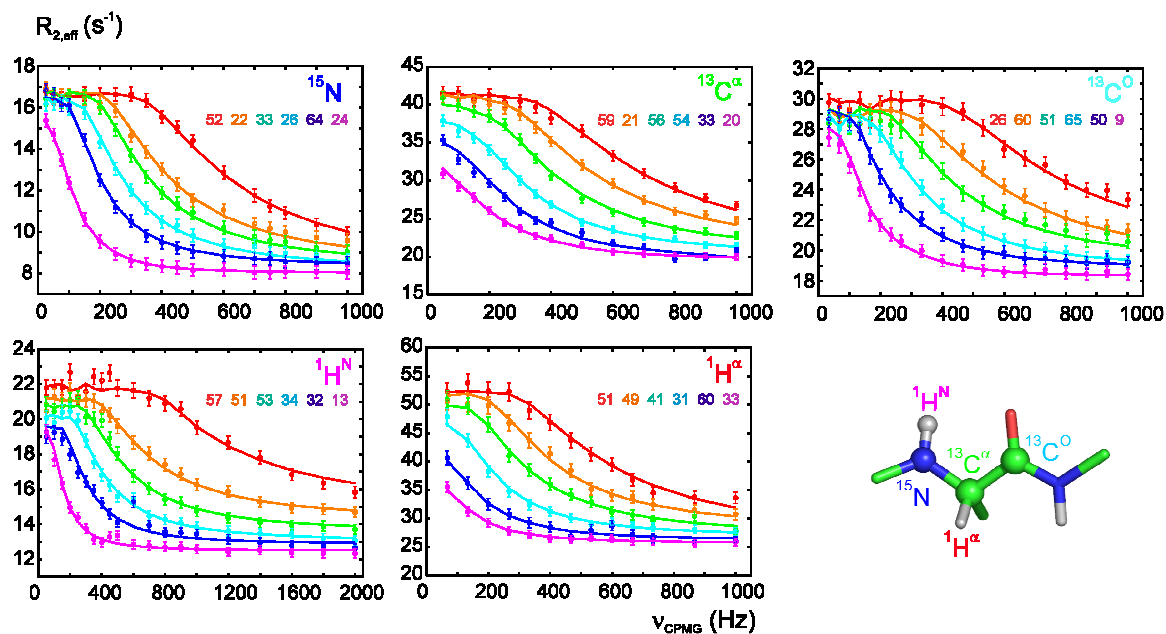
Figure S1. Experimental relaxation dispersion profiles for select backbone ^{15}N [1,2], ^{1}H^{N} [3], ^{13}C$^{\alpha}$ [4], ^{1}H$^{\alpha}$ [5] and ^{13}C^{O} [4,6] nuclei of the L24A FF domain measured at a static magnetic field of 18.8 T (open circles), along with best fits to a 2-site exchange model, I↔N, performed as described elsewhere [7,8] (solid lines).
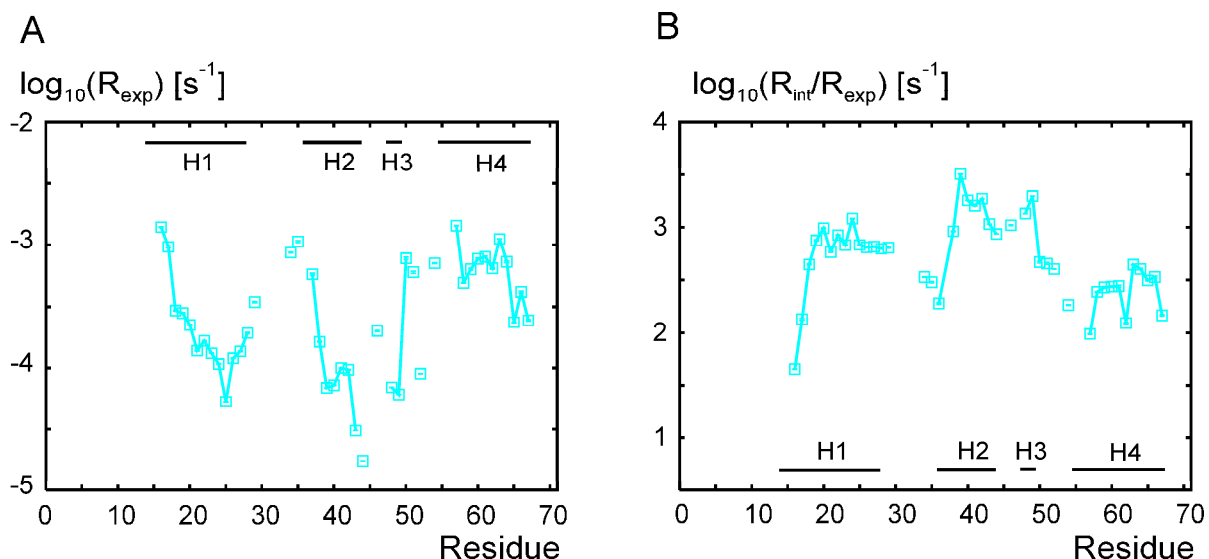
Figure S2. (**A**) Logarithm of experimental ^{1}H/^{2}H-exchange rates for the backbone amide groups of the L24A FF domain, $\log_{10}(R_{exp})$, obtained by dissolving lyophilized protein (^{1}H,^{15}N,selective-^{13}C$^{\alpha}$ labeled) in 100% D_2O buffer, 20°C, pD=5.3 (uncorrected). ^{1}H/^{2}H-exchange rates that are faster than 10^{-2}-10^{-3}/s cannot be measured because of the delay (approximately 10 minutes) between dissolving the sample and recording the first spectrum. (**B**) Logarithm of protection factors for the backbone amide groups of the L24A FF domain, $\log_{10}(R_{int}/R_{exp})$, where R_{int} are intrinsic random-coil ^{1}H/^{2}H-exchange rates for the amide groups calculated using the FBMME_HD.xls Microsoft Excel spreadsheet downloaded from http://hx2.med.upenn.edu [9,10]. Note that the secondary structure elements as they pertain to the N structure and *not* the I state are listed in the figure. As first outlined by Linderstrom-Lang [11] the ^{1}H/^{2}H-exchange reaction for a given amide group can be described by, $close \underset{R_C}{\overset{R_O}{\rightleftharpoons}} open \xrightarrow{R_{int}} exchange$, where '*close*' and '*open*' denote states that are inaccessible and accessible to exchange, respectively, and 'exchange' corresponds to that state where ^{1}H has been exchanged for ^{2}H. In the EX2 limit, $R_{int} << R_C = k_f$; this limit certainly applies here since $R_{int} \sim 10^{-1}$ - 10^{-2} s^{-1} and $R_C \sim 300$ s^{-1}. In this case the free energy of folding can be calculated from the relation $\Delta G = -RT\ln(R_{exp}/R_{int})$, using the largest protection factors measured in the protein [12]. Assuming that the maximum protection factor ($\log_{10}(R_{int}/R_{exp})\sim 3.5$) reflects the relative populations of states U and F while the protection factors for residues 55-65 (that form H4 in N but not I) are sensitive to the combined population of U and I, it can be calculated that $p_U \sim 1/7\, p_I$.
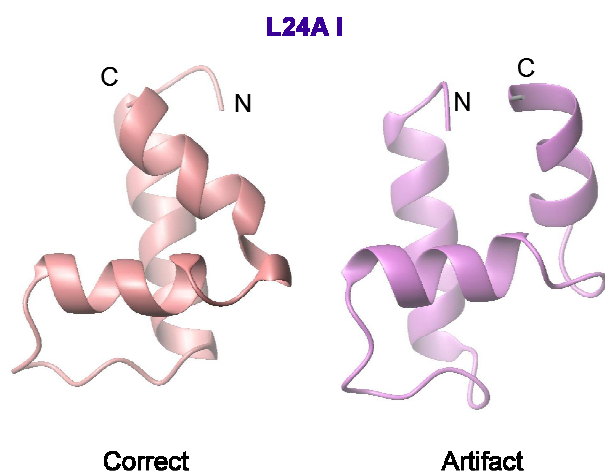
Figure S3. Comparison of the two lowest energy structures of L24A FF domain folding intermediate (residues 11-59) generated by CS-Rosetta. In the lowest energy model, shown on the right side of the figure, the N-terminus points in a direction that would cause steric clashes with helix H2 involving the first 10 residues that were not included in calculations. Therefore, the second lowest energy model, left hand side, that is consistent with all remaining lowest energy structures (from an ensemble of 10) has been chosen as representative and used to generate the energy funnel of Figure 1B.

A

L24A N

B

Energy

40

20

0

-20

-40

-60

-80

0   4   8   12

RMSD [Å]

Figure S4. CS-Rosetta modeling of the three-dimensional structure of the N state of the L24A FF domain (residues 11-70) based on exp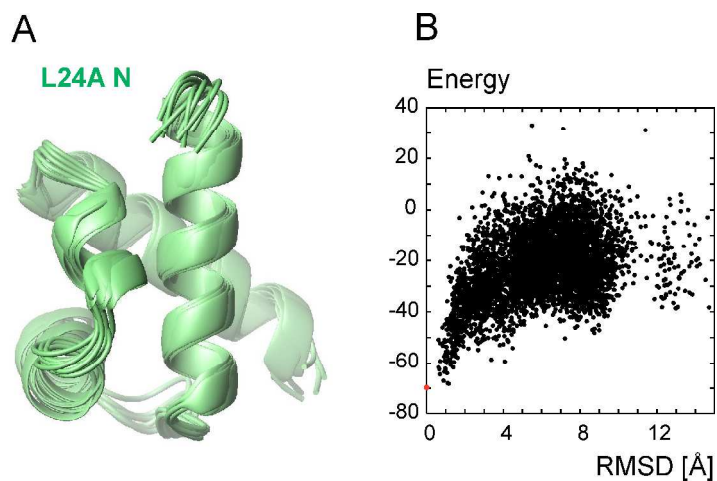erimental ^{15}N, ^{1}H^{N}, ^{13}C$^{\alpha}$, ^{1}H$^{\alpha}$ and ^{13}C^{O} chemical shifts. (**A**) Superposition of the 10 lowest energy structures. (**B**) Energy of the generated models (total 5,000) *vs.* rmsd to the lowest energy structure (red dot).

Table S1. ^{15}N and $^{1}H^{N}$ chemical shifts for the native, N, and the intermediate, I, states of the L24A FF domain measured on an $^{15}N/^{2}H$ sample (50mM NaAc, 100mM NaCl 10%D_2O/90%H_2O, pH=5.7; 20 ^{o}C).

rn	ϖ_N [ppm] ^{15}N	ϖ_N[ppm] $^{1}H^{N}$	ϖ_{RC}[ppm] ^{15}N	ϖ_{RC}[ppm] $^{1}H^{N}$	$\Delta\varpi_{IN}$ [ppm] ^{15}N	$\Delta\varpi_{IN}$ [ppm] $^{1}H^{N}$
3	122.98	8.45	122.5	8.41	-0.09±0.10	0.000±0.000
5	124.60	8.36	124.2	8.44	-0.06±0.15	-0.001±0.072
6	121.15	8.21	119.6	8.33	-0.08±0.11	-0.004±0.027
7	123.58	8.27	122.0	8.41	-0.20±0.05	0.027±0.004
8	116.37	8.04	115.2	8.27	**-0.49±0.03**	**-0.030±0.004**
9	122.51	8.23	122.7	8.25	0.12±0.08	-0.049±0.003
10	114.29	7.79	116.4	8.20	**1.47±0.04**	**0.072±0.005**
11	120.61	7.67	123.7	8.38	**1.71±0.04**	**0.279±0.011**
12	119.27	9.19	121.5	8.36	**0.71±0.02**	**-0.783±0.023**
13	108.45	7.69	114.3	8.23	**3.87±0.08**	**0.179±0.005**
14	122.61	8.94	122.8	8.42	0.16±0.06	-0.551±0.025
15	118.66	8.40	121.8	8.54	**2.19±0.05**	**-0.100±0.004**
16	119.58	7.90	121.4	8.56	**1.46±0.04**	**0.268±0.010**
17	124.73	8.05	125.0	8.38	**-1.04±0.03**	**-0.125±0.005**
18	116.97	8.01	119.6	8.33	**2.15±0.05**	**-0.047±0.004**
19	117.91	7.81	121.4	8.44	**1.93±0.05**	**0.193±0.006**
20	122.48	7.99	125.1	8.38	**0.91±0.03**	**0.015±0.010**
21	118.71	7.76	119.5	8.34	**0.56±0.02**	**0.284±0.008**
22	115.62	7.80	122.8	8.37	**5.04±0.11**	**0.114±0.004**
23	120.72	8.39	121.8	8.54	**-0.28±0.04**	**-0.316±0.009**
24	122.50	7.55	125.0	8.38	**1.28±0.03**	**0.504±0.013**
25	116.18	6.77	121.0	8.20	**3.14±0.07**	**1.118±0.027**
26	117.48	7.75	121.4	8.35	**3.30±0.08**	**0.055±0.004**
27	121.63	8.84	121.8	8.54	**-1.45±0.04**	**-0.820±0.021**
28	116.10	7.44	121.6	8.43	**3.75±0.08**	**0.660±0.017**
29	115.71	7.89	122.1	8.35	4.99±0.10	0.278±0.009
30	120.43	7.63	121.9	8.16	**0.77±0.02**	**0.332±0.010**
32	115.55	8.75	116.1	8.51	**0.51±0.03**	**-0.350±0.010**
33	114.84	7.74	120.6	8.52	**4.28±0.09**	**0.494±0.013**
34	122.76	7.49	124.5	8.32	**0.78±0.03**	**0.453±0.011**
35	115.61	8.27	114.9	8.35	-0.10±0.08	-0.013±0.008
36	122.38	9.19	123.2	8.37	**0.97±0.03**	**-1.106±0.025**
37	116.25	8.57	123.0	8.38	**2.76±0.06**	**-0.496±0.033**
38	118.09	7.40	121.0	8.46	**0.79±0.03**	**0.417±0.014**
39	121.90	8.01	125.1	8.38	**0.83±0.03**	**0.107±0.005**
40	115.22	8.10	118.8	8.32	**1.33±0.04**	**-0.099±0.005**
41	115.13	6.83	121.5	8.39	4.00±0.09	x
42	114.37	7.28	121.2	8.40	**3.53±0.08**	**0.450±0.013**
43	107.93	6.99	121.0	8.10	11.66±0.19	0.776±0.023
44	122.46	7.10	124.1	8.17	**-2.18±0.05**	**0.798±0.032**
45	115.60	7.74	122.9	8.57	**3.53±0.08**	**0.165±0.006**
46	125.82	8.11	122.1	8.39	**-3.09±0.07**	**-0.261±0.008**
48	115.83	9.49	120.9	8.43	**2.55±0.06**	**-1.148±0.028**
49	120.38	8.32	121.7	8.26	**-2.50±0.06**	**-0.618±0.018**
50	107.56	7.23	118.5	8.36	8.17±0.15	0.639±0.017
51	123.11	7.39	125.7	8.36	**1.54±0.04**	**0.846±0.022**
52	111.32	5.73	121.0	8.20	7.53±0.14	2.128±0.052
53	125.98	8.46	124.8	8.30	**-3.34±0.08**	**-0.651±0.017**
54	115.69	7.95	119.6	8.33	**2.67±0.06**	**-0.123±0.004**
55	127.57	8.89	123.4	8.28	**-5.95±0.13**	**-1.148±0.047**
56	111.65	8.58	116.7	8.37	**3.66±0.08**	**-0.613±0.018**
57	121.43	6.56	122.1	8.54	**0.98±0.03**	**1.503±0.041**
58	120.81	7.66	121.6	8.43	**1.02±0.03**	**0.440±0.014**
59	116.60	7.89	122.0	8.41	**5.45±0.13**	**0.318±0.011**
60	118.84	7.57	121.4	8.44	**2.27±0.05**	**0.684±0.020**
61	122.68	7.81	125.1	8.38	**2.12±0.05**	**0.421±0.017**
62	119.18	8.48	119.5	8.34	0.24±0.04	-0.490±0.019
63	118.61	8.49	121.1	8.48	**1.96±0.05**	**-0.381±0.011**
64	121.62	7.80	124.5	8.32	**2.39±0.05**	**0.184±0.006**
65	120.64	8.08	119.5	8.16	**-2.07±0.05**	**-0.198±0.006**
66	118.93	7.97	123.2	8.34	**4.07±0.09**	**-0.129±0.004**
67	114.44	7.20	120.8	8.15	7.04±0.14	0.830±0.020
68	120.24	7.58	123.7	8.50	**4.00±0.08**	**0.823±0.018**
69	113.86	7.75	114.9	8.29	**1.81±0.05**	**0.369±0.009**
70	123.12	7.98	122.6	8.55	**0.58±0.02**	**0.362±0.008**
71	127.02	7.77	121.6	8.43	-0.32±0.03	-0.169±0.004

^{15}N and $^{1}H^{N}$ chemical shifts in state N, ϖ_N, (**columns 2 and 3**) were measured directly from an $^{15}N/^{1}H^{N}$ HSQC data set recorded at 20°C, 800 MHz spectrometer (referenced indirectly).

Random coil (RC) ^{15}N and $^1H^N$ chemical shifts for the L24A FF domain (^{15}N: **column 4**; ^1H: **column 5**) predicted using the CSI module of the NMRView program ('Wishart(Peptides)' CSI option) [13,14].

^{15}N and $^1H^N$ chemical shift differences between states I and N, $\Delta\varpi_{IN} = \varpi_I - \varpi_N$ (**columns 6 and 7**) were obtained from the analysis of ^{15}N single-quantum (SQ) [1,2,15] or $^1H^N$ SQ [3] CPMG relaxation dispersion data measured at 500 and 800 MHz (minimal uncertainties of 2% or 0.2/s, whichever is greater, were assumed for $^1H^N$ SQ, ^{15}N SQ data). ^{15}N ($^1H^N$) SQ relaxation dispersion data collected at 20 ^oC were fit for all residues together using a model of 2-state exchange between states I and

N as described previously [7,8], $I \underset{k_{NI}}{\overset{k_{IN}}{\rightleftarrows}} N$. ^{15}N and $^1H^N$ dispersion data were fit separately, with results from $^1H^N$ data

shown in parenthesis in what follows. Values of $k_{ex,NI}=k_{IN}+k_{NI}$ = 223±5 s^{-1}, p_I=3.68±0.07% ($k_{ex,NI}$=252±9 s^{-1}, p_I=3.08±0.10%) and a reduced global χ^2 of 0.38(1.0) were obtained. The data for a total of 66(65) residues were included in global fits with 40(40) ^{15}N SQ (^1H SQ) data points per residue. ^1H SQ data for residue 41 were excluded from the analysis due to spectral overlap with side-chain resonances (indicated in the table by 'x').

Signs of ^{15}N $\Delta\varpi_{IN}$ values (**column 6**) were selected based on differences in peak positions in HSQC spectra recorded at 500 and 800 MHz, $\delta_{exp}=\varpi_{500}-\varpi_{800}$ [16]. The experimental differences δ_{exp} were compared to δ_{clc} predicted using $\Delta\varpi_{IN}$, $k_{ex,NI}$ and p_I extracted from CPMG dispersion data. In cases where both δ_{exp} and δ_{clc} > 0.001ppm the signs of $\Delta\varpi_{IN}$ were assigned according to δ_{exp} (chemical shift differences in columns 6 and 7 are marked by bold in these cases).

Signs of $^1H^N$ chemical shift differences ($\Delta\varpi_{IN}$, **column 7**) were obtained as described in detail previously [17] from (i) signs of ^{15}N $\Delta\varpi_{IN}$ values and (ii) from the relative signs of $^1H^N$, ^{15}N $\Delta\varpi_{IN}$ values that were determined experimentally on the basis of a comparison of χ^2 values obtained in per-residue fits of ^{15}N SQ, $^1H^N$ SQ, $^{15}N/^1H^N$ double-quantum (DQ) and zero-quantum (ZQ) CPMG data sets with $\Delta\varpi_{IN}(^1H^N)\times\Delta\varpi_{IN}(^{15}N)$ >0 and $\Delta\varpi_{IN}(^1H^N)\times\Delta\varpi_{IN}(^{15}N)$<0 (using fixed values of $k_{ex,NI}$ and p_I from the global fit of the data).

In cases where signs could not be determined experimentally and where values of $\Delta\varpi_{IN}$ fall below the uncertainty for chemical shift predictions of a given nucleus type (*i.e.*, uncertainty values of 2.53 ppm for ^{15}N, 0.51 ppm for $^1H^N$ that correspond to the root-mean-squared deviations between predicted *vs.* experimental chemical shifts in a protein database used by the SPARTA program [18]) the sign of $\Delta\varpi_{IN}$ was assumed to be the same as the sign of $\varpi_{RC}-\varpi_N$. This assumption follows from the empirical observation that for the great majority of residues for which the signs could be determined experimentally (88% for ^{15}N, 82% for $^1H^N$) the signs of $\Delta\varpi_{IN}$ and $\varpi_{RC}-\varpi_N$ are the same. We have also assumed the opposite signs for these residues in structure calculations, with little change in the structural properties of the ensemble, as was observed in a previous study of the WT FF domain folding intermediate [7]. For the several residues where $\Delta\varpi_{IN}$ values are above the SPARTA cutoff and for which signs could not be obtained experimentally we have checked to see if one of the two possibilities, $\varpi_I=\varpi_N\pm\Delta\varpi_{IN}$, lies outside the range of chemical shifts that are typically observed for these residue types (outside twice the standard deviation from the mean) and can thus be discarded [7].

The chemical shifts of the intermediate state, $\varpi_N+\Delta\varpi_{IN}$ (**columns 2, 3, 6, 7**) were used in CS-Rosetta calculations of I as described in Materials and Methods.

Table S2. $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ chemical shifts in the N and I states of the L24A FF domain (ϖ_N and $\varpi_I = \varpi_N + \Delta\varpi_{IN}$). $^{13}C^{\alpha}$ relaxation dispersion measurements [4] were performed on a U-^{15}N, selectively-^{13}C,^{1}H labeled L24A FF domain generated in $^{1}H_2O$ M9 medium using 2-^{13}C-pyruvate as the sole carbon source (50mM NaAc, 100mM NaCl, 100% D_2O, pD=5.3, 30°C; **sample 1**) [19]. $^{1}H^{\alpha}$ relaxation dispersion experiments [5] were measured on a U-^{15}N,^{13}C and partially-^{2}H labeled L24A FF domain sample produced in 50%H_2O/50%D_2O M9 medium using ^{2}H/^{13}C-glucose as the sole carbon source (50mM NaAc, 100mM NaCl 100% D_2O, pD=5.3, 30 °C; **sample 2**) [5]

rn	ϖ_N [ppm] $^{13}C^{\alpha}$	ϖ_N [ppm] $^{1}H^{\alpha}$	ϖ_{RC} [ppm] $^{13}C^{\alpha}$	ϖ_{RC} [ppm] $^{1}H^{\alpha}$	$\Delta\varpi_{IN}$ [ppm] $^{13}C^{\alpha}$	$\Delta\varpi_{IN}$ [ppm] $^{1}H^{\alpha}$
3	53.71	4.68	53.7	4.65	-0.23±0.02	-0.011±0.026
4	63.06	4.41	63.3	4.42	0.33±0.02	0.009±0.032
5	52.27	4.28	52.5	4.32	0.10±0.04	0.001±0.229
6	56.07	4.28	56.2	4.32	0.15±0.03	0.007±0.045
8	61.44	4.24	61.8	4.35	0.14±0.03	0.000±0.000
9	57.26	4.58	57.9	4.55	0.32±0.02	-0.037±0.010
10	60.63	3.98	61.8	4.35	**0.94±0.02**	**0.243±0.008**
11	56.46	4.85	57.5	4.66	**0.84±0.02**	**-0.195±0.007**
12	54.58	4.86	53.1	4.74	**-0.72±0.02**	**-0.165±0.006**
13	59.29	4.79	61.8	4.35	**1.82±0.05**	**-0.354±0.013**
14	58.49	4.17	56.2	4.32	-0.38±0.03	0.050±0.010
15	59.83	4.04	56.6	4.35	**-1.34±0.03**	**0.084±0.008**
16	59.06	4.12	56.6	4.35	**-1.00±0.03**	0.076±0.009
17	55.02	3.46	52.5	4.32	**-1.01±0.02**	**0.401±0.017**
18	60.17	3.76	56.2	4.32	**-1.46±0.04**	**0.305±0.012**
19	58.69	4.02	55.7	4.34	**-1.32±0.03**	x
20	54.96	4.12	52.5	4.32	**-0.81±0.02**	**0.068±0.009**
21	62.06	3.62	57.7	4.62	**-2.39±0.07**	**0.705±0.029**
22	60.24	3.90	56.2	4.32	**-1.33±0.03**	0.090±0.008
23	59.18	4.02	56.6	4.35	**-1.08±0.03**	**0.112±0.008**
24	55.52	3.96	52.5	4.32	**-1.27±0.03**	**0.138±0.008**
26	59.12	4.26	56.2	4.32	**-1.10±0.03**	**-0.117±0.008**
27	59.43	4.02	56.6	4.35	**-2.09±0.06**	**0.190±0.010**
28	54.83	4.24	56.2	4.32	**1.08±0.03**	0.020±0.024
29	55.86	3.88	56.0	4.34	**0.47±0.02**	**0.376±0.014**
30	61.77	3.93	59.8	4.44	x	x
31	63.21	4.67	63.3	4.42	**-0.36±0.02**	**-0.259±0.009**
32	60.90	4.02	58.3	4.47	**-1.56±0.04**	**0.304±0.013**
33	51.64	4.83	53.1	4.74	**1.15±0.03**	**-0.127±0.006**
34	52.49	4.44	52.5	4.32	0.36±0.02	**-0.149±0.007**
35	56.66	4.85	58.3	4.47	**1.32±0.03**	**-0.312±0.011**
36	59.86	3.80	57.5	4.66	-0.50±0.03	**0.517±0.026**
37	60.67	3.26	56.6	4.35	**-1.34±0.04**	**0.444±0.018**
38	58.12	3.84	55.7	4.34	x	**0.259±0.009**
39	54.93	3.67	52.5	4.32	**-0.46±0.02**	**0.390±0.015**
40	58.16	3.36	55.4	4.48	**-1.31±0.09**	**0.755±0.049**
41	58.84	3.67	56.2	4.32	**-0.38±0.02**	**0.366±0.014**
42	58.39	4.11	55.4	4.48	**-1.29±0.04**	**0.200±0.008**
43	61.39	4.49	61.1	4.17	x	**-0.437±0.017**
44	63.05	2.85	61.1	4.17	x	x
45	52.63	4.80	53.1	4.74	**0.69±0.02**	-0.084±0.007
46	51.65	4.92	52.2	4.90	0.21±0.03	-0.076±0.009
47	65.31	4.51	63.3	4.42	**-1.10±0.06**	-0.090±0.007
48	57.65	3.97	56.0	4.34	-0.42±0.03	**0.089±0.008**
49	61.18	4.14	57.9	4.55	**-3.01±0.09**	**0.467±0.021**
50	58.23	4.22	58.3	4.47	**2.43±0.07**	0.051±0.010
51	54.73	3.56	52.5	4.32	**-0.83±0.02**	**0.676±0.027**
53	54.45	4.04	52.5	4.32	**-0.76±0.02**	**0.151±0.006**
54	55.47	4.41	56.2	4.32	**1.44±0.03**	**-0.183±0.007**
56	61.00	3.85	58.3	4.47	**-1.92±0.05**	**0.384±0.024**
57	58.53	4.01	56.6	4.35	**-1.85±0.06**	x
58	59.57	3.01	56.2	4.32	**-2.80±0.08**	x
59	60.21	3.57	56.2	4.32	**-3.22±0.08**	**0.743±0.029**
60	59.12	4.08	55.7	4.34	**-2.33±0.08**	**0.170±0.008**
61	54.98	4.29	52.5	4.32	**-2.15±0.06**	0.010±0.046
62	60.44	4.62	57.7	4.62	**-2.28±0.07**	**-0.090±0.008**
63	56.37	4.12	53.1	4.74	**-2.99±0.13**	**0.445±0.018**
64	54.76	4.17	52.5	4.32	**-1.82±0.05**	0.022±0.023
65	60.54	4.21	57.9	4.55	**-2.53±0.07**	**0.270±0.010**
66	59.92	3.49	56.2	4.32	**-3.59±0.09**	**0.801±0.028**
67	63.54	3.98	62.2	4.12	x	0.068±0.007
68	56.65	4.26	55.7	4.34	**-0.82±0.02**	x
69	62.50	4.19	61.8	4.35	**-0.80±0.01**	**0.132±0.005**
70	56.71	4.28	56.6	4.35	-0.38±0.01	x

| 71 | 57.64 | 4.14 | 56.2 | 4.32 | -0.04±0.09 | 0.002±0.123 |

$^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ chemical shifts in the native state N, ϖ_N, (**columns 2 and 3**) obtained from a $^{13}C^{\alpha}$,$^{1}H^{\alpha}$ HSQC spectrum recorded on sample 1 at 30 ^{o}C, 800 MHz spectrometer (referenced indirectly).

Random coil (RC) $^{13}C^{\alpha}$ (**column 4**) and $^{1}H^{\alpha}$ (**column 5**) chemical shifts for the L24A FF domain were predicted using the CSI module of the NMRView program ('Wishart(Peptides)' CSI option) [13,14].

$^{13}C^{\alpha}$ chemical shift differences between I and N states, $\Delta\varpi_{IN}$ (**column 6**) were obtained from the analysis of $^{13}C^{\alpha}$ SQ CPMG relaxation dispersion data measured on sample 1 at 30 ^{o}C, 600 and 800 MHz spectrometer frequencies (^{1}H). Minimal uncertainties of 2% or 0.2/s, whichever is greater, were assumed for individual relaxation rates of the dispersion profiles. From a global fit of $^{13}C^{\alpha}$ SQ CPMG data for 60 residues (26 data points per residue) values of $k_{ex,NI}$=784±14s^{-1}, p_I=2.68±0.03% and a reduced global χ^2 of 1.22 were obtained. The data for the following residues were excluded from the analysis, denoted by 'x' in the table: 38 (overlapped with folded peak); 30, 43, 44, 67 (these are Ile, Val residues; the labeling scheme used results in $^{13}C^{\alpha}$-$^{13}C^{\beta}$ pairs for these residues that renders the $^{13}C^{\alpha}$ dispersion experiment difficult to analyze in these cases [19]. Leu 25, 52, 55 could not be quantified because Leu residues are not labeled at the $^{13}C^{\alpha}$ position using the 2-^{13}C-pyruvate labeling scheme [19].

$^{1}H^{\alpha}$ chemical shift differences between the I and N states, $\Delta\varpi_{IN}$, (**column 7**) were obtained from the analysis of $^{1}H^{\alpha}$ SQ CPMG dispersion data measured for sample 2 at 30 ^{o}C, 500 and 800 MHz spectrometers (minimal uncertainties of 3% of 0.5/s, whichever is greater, were assumed for each dispersion point). Parameters from a global fit of $^{1}H^{\alpha}$ SQ CPMG data for 58 residues (26 data points per residue) resulted in $k_{ex,NI}$=625±20s^{-1}, p_I=4.02±0.08% and χ^2 =0.79. The data for the following residues were excluded from the analysis due to resonance overlap: 19/57, 68/70, and 44, 58 ('x' in table). Additional resonance overlap was noted in data sets recorded at 500 MHz for residues 16/27, 28/61, 64 (with a folded peak); the data recorded at this field were therefore excluded for these amino acids. Dispersion profiles for T8, T10, T13, S32, S50, S56 and S69 were included in the analysis despite the potential for modulation due to $^{1}H^{\alpha}$-$^{1}H^{\beta}$ couplings that are not refocused, since the high level of deuteration at the β position (~66% for Thr, ~88% for Ser) ensures that the artifacts are minimal [5].

Signs of $^{13}C^{\alpha}$ chemical shift differences between I and N states, $\Delta\varpi_{IN}$, (**column 6**) were selected based on differences in peak positions along the ^{13}C dimension, δ_{exp}, in (1) $^{1}H^{\alpha}$-$^{13}C^{\alpha}$ HSQC spectra recorded at 500 and 800 MHz and in (2) $^{1}H^{\alpha}$-$^{13}C^{\alpha}$ HMQC and HSQC spectra recorded at 500 MHz and 800 MHz (all spectra were recorded at 30 ^{o}C). The differences were considered significant if both experimental and predicted differences, δ_{exp} and δ_{clc}, were greater then 0.002ppm. Signs of $\Delta\varpi_{IN}$ were taken as reliable and assigned according to signs of experimental peak shifts, δ_{exp}, if all of the data were self-consistent. If only 1 value of δ_{exp} > 0.002ppm was obtained for a given residue the sign was assumed correct only if it was the same as ϖ_{RC}-ϖ_N, where ϖ_{RC} is the $^{13}C^{\alpha}$ random-coil chemical shift predicted using the CSI module of the NMRView program [13,14]. In addition, signs were also obtained by off-resonance $R_{1\rho}$ experiments as described previously [20,21]. There was complete agreement in the signs from the two methods. All $\Delta\varpi_{IN}$ values for which experimental signs could be obtained are marked by bold in the Table.

Signs of $^{1}H^{\alpha}$ chemical shift differences between the I and N states, $\Delta\varpi_{IN}$, (**column 7**) were obtained by off-resonance $R_{1\rho}$ experiments as described previously previously [20,21].

In cases where experimentally determined signs could not be obtained and where $\Delta\varpi_{IN}$ values were less than the uncertainty of chemical shift predictions by the SPARTA program [18] (0.97 ppm and 0.27 ppm for $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$, respectively) the signs of $\Delta\varpi_{IN}$ were assumed to be identical to those of ϖ_{RC}-ϖ_N. It is worth noting that all of the signs of $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ $\Delta\varpi_{IN}$ values that were determined experimentally were in agreement with the signs of ϖ_{RC}-ϖ_N. CS-Rosetta calculations were performed with these signs inverted as well and very similar structures were obtained.

The chemical shifts of the intermediate, ϖ_N+$\Delta\varpi_{IN}$ (**columns 2, 3, 6, 7**) were used in CS-Rosetta calculations of the I state (see Materials and Methods).

Table S3. $^{13}C^O$ chemical shifts for the N and I states of the L24A FF domain (ϖ_N and $\varpi_N + \Delta\varpi_{IN}$) measured on an $^{15}N/^{13}C/^2H$ labeled sample (50mM NaAc, 100mM NaCl 10%D_2O/90%H_2O, pH=5.7; 20 ^{o}C).

rn	ϖ_N [ppm] $^{13}C^O$	ϖ_{RC}[ppm] $^{13}C^O$	$\Delta\varpi_{IN}$ [ppm] $^{13}C^O$
2	174.20	174.6	0.08±0.02
4	176.56	177.2	0.08±0.02
5	177.76	177.9	0.06±0.03
6	176.42	176.6	0.09±0.02
7	176.05	176.4	**0.27±0.01**
8	173.50	174.5	**0.28±0.01**
9	174.73	175.7	**0.75±0.02**
10	172.73	174.1	x
11	175.12	175.9	**0.77±0.02**
12	175.26	175.7	x
13	175.21	174.5	**-0.24±0.01**
14	178.89	176.4	-1.12±0.03
15	179.65	176.4	**-1.66±0.04**
16	179.53	176.6	**-1.72±0.05**
17	178.61	177.8	**-0.12±0.02**
18	178.93	176.6	-0.90±0.02
19	178.00	175.8	**-1.05±0.03**
20	178.69	177.7	-0.24±0.01
21	177.16	175.8	**-0.97±0.02**
22	178.97	176.1	-0.92±0.02
23	178.93	176.4	**-1.56±0.04**
24	179.20	177.8	**-0.49±0.01**
25	177.70	177.6	**0.90±0.02**
26	181.55	176.5	**-3.89±0.08**
27	178.75	176.4	**-2.12±0.06**
28	174.64	176.6	**1.38±0.03**
29	176.61	176.1	**-0.58±0.01**
31	177.52	177.2	**-0.35±0.01**
32	174.02	174.7	**0.60±0.02**
33	175.85	175.0	**-0.63±0.02**
34	177.57	177.6	**0.67±0.02**
35	175.24	174.6	-0.11±0.02
36	176.16	175.9	**0.92±0.02**
37	179.17	177.1	**-0.85±0.02**
38	178.18	176.0	**-0.58±0.01**
39	178.89	177.7	**0.23±0.01**
40	177.58	176.3	**-0.15±0.01**
41	178.58	176.4	**-1.04±0.03**
42	178.36	176.1	**-1.19±0.03**
43	176.67	176.2	-0.11±0.02
44	174.52	176.2	**1.54±0.04**
45	175.00	175.0	x
47	178.58	177.3	**-0.53±0.01**
48	177.86	176.4	**-0.37±0.01**
49	176.84	175.7	**-0.32±0.01**
50	174.27	174.0	**1.10±0.03**
51	176.74	177.6	**2.20±0.06**
52	174.84	177.6	**3.19±0.07**
53	178.91	177.7	**-0.36±0.01**
54	177.83	176.6	**-0.36±0.01**
55	178.50	177.4	**-0.56±0.02**
56	177.41	174.5	x
57	179.22	176.4	x
58	178.30	176.6	**-1.48±0.04**
59	179.15	176.4	**-2.12±0.06**
60	178.98	175.8	-2.97±0.07
61	179.95	177.7	**-2.23±0.06**
62	176.37	175.8	**-0.81±0.02**
63	177.63	174.7	x
64	179.85	177.6	**-2.29±0.06**
65	177.37	175.9	**-1.65±0.04**
66	177.93	176.0	**-1.84±0.05**
67	177.24	176.1	**-0.97±0.02**
68	176.58	175.9	**-0.56±0.01**
69	174.64	174.6	**-0.20±0.01**
70	175.51	176.4	0.10±0.02

$^{13}C^O$ chemical shifts in the native state N, ϖ_N, (**column 2**) were obtained from a $^{13}C^O,^1H^N$ HSQC spectrum recorded at 20 ^{o}C, 800 MHz spectrometer (referenced indirectly).

Random coil (RC) $^{13}C^O$ chemical shifts for the L24A FF domain (**column 3**) were predicted using the CSI module of the NMRView program (Wishart (Peptides)' CSI option) [13,14].

$^{13}C^O$ chemical shift differences between the I and N states, $\Delta\varpi_{IN}$, (**column 4**) were obtained from the analysis of $^{13}C^O$ SQ relaxation dispersion profiles measured at 500 and 800 MHz. Minimal uncertainties of 2% or 0.2/s (whichever is greater) were assumed for individual data points (relaxation rates) of the dispersion curves. From a global fit of $^{13}C^O$ dispersion data values of $k_{ex,NI}$=193$\pm$5 s^{-1}, p_I=4.90$\pm$0.11% and a reduced global χ^2 of 0.73 were obtained. The data from a total of 60 residues were included in the fit with 36 $^{13}C^O$ SQ data points per residue. Residues 10, 12 (Asn), 45 (Asn), 56, 57 and 63 (Asn) were excluded from the analysis ('x' in table) - peaks from residues 10 and 57 are overlapped, while dispersion profiles for all Asn residues (except for Asn 33) are artifactual due to the 3-bond coupling to the side-chain $^{13}C^{O\,6,22}$, leading to large contributions to the global χ^2. Poor quality dispersions were also obtained for residue Ser 56 and these data were excluded from analysis as well.

Signs of $^{13}C^O$ chemical shift differences between I and N states, $\Delta\varpi_{IN}$, (**column 4**) were selected based on differences in peak positions along the $^{13}C^O$ dimension in $^{13}C^O,^1H^N$ HSQC spectra recorded at 500 and 800 MHz, δ_{exp}, at (1) 20 ^{o}C and (2) at 30 ^{o}C, and based on differences in F_1 peak positions in $^{13}C^O,^1H^N$ HSQC and HMQC (^{15}N-$^{13}C^O$ MQ coherences) data sets measured at 500 MHz at 20 ^{o}C (3) and at 30 ^{o}C (4). The differences were considered significant if both experimental and predicted differences, δ_{exp} and δ_{clc}, were greater then 0.003ppm at 30 ^{o}C or 0.002ppm at 20 ^{o}C. δ_{clc} values at 20 and 30 ^{o}C were predicted using exchange parameters obtained from a global fit of ^{15}N SQ CPMG data recorded at 5 temperatures ranging form 10 to 25 ^{o}C (10, 15, 20, 22.5, 25 ^{o}C) assuming an Arrhenius temperature dependence of populations and rate constants, $k_{ex,NI}$=242/s, p_I=3.53% at 20 ^{o}C and $k_{ex,NI}$=407/s, p_I=9.27% at 30 ^{o}C. Signs of $\Delta\varpi_{IN}$ were considered reliable and obtained from experimental peak shifts, δ_{exp}, if all of the data were self-consistent. In cases where only 1 value of $\delta_{exp} >$ 0.002ppm (20 ^{o}C) or 0.003 ppm (30 ^{o}C) the sign was assumed correct only if it was the same as ϖ_{RC}-ϖ_N, where ϖ_{RC} is $^{13}C^\alpha$ random-coil chemical shift predicted using the CSI module of the NMRView program [13,14]. All $\Delta\varpi_{IN}$ values for which experimental signs could be obtained are marked by bold in the table. In cases where experimentally determined signs could not be obtained and where $\Delta\varpi_{IN}$ values were less than the uncertainty of chemical shift predictions by the SPARTA program [18] (1.08 ppm) the signs of $\Delta\varpi_{IN}$ were assumed to be identical to those of ϖ_{RC}-ϖ_N. Notably 95% of the signs of $^{13}C^O$ $\Delta\varpi_{IN}$ values that were determined experimentally were in agreement with the signs of ϖ_{RC}-ϖ_N. For the two residues for which $\Delta\varpi_{IN}$ values are greater than the SPARTA uncertainty and for which sign information was not forthcoming from experiment (14, 60), the signs were obtained from those measured experimentally previously for the WT FF domain folding intermediate [7].

$^{13}C^O$ chemical shifts ϖ_N+$\Delta\varpi_{IN}$ (**columns 2, 4**) were used in CS-Rosetta calculations of the structure of the I state, as described in Methods. A second set of structure calculations with the signs of $\Delta\varpi_{IN}$ inverted in cases where experimental information is not obtained and where the sign has been assumed to be the same as ϖ_{RC}-ϖ_N has been performed, with the structures essentially unchanged.

Table S4. ^{15}N-$^1H^N$ residual dipolar couplings in the N and I states measured on a sample of the $^{15}N/^{13}C/^2H$ L24A FF domain weakly aligned in a PEG $C_{12}E_5$/hexanol mixture, 20 ^oC (~14.1/4.8µl PEG $C_{12}E_5$/hexanol, D_2O splitting 16.6Hz; 50mM NaAc, 100mM NaCl 10%D_2O/90%H_2O, pH=5.7).

rn	J_N(n-al)	J_N(al)	$\Delta\varpi_{IN}$ [Hz]	ΔD_{IN} [Hz]	Comment
3	-93.29	-96.72	-0.08±0.05	-3.48±7.79	S^2<0.6
5	-93.32	-97.05	0.08±0.04	1.21±6.63	S^2<0.6
6	-93.05	-95.75	-0.11±0.03	-0.86±5.13	S^2<0.6
7	-93.03	-98.66	-0.20±0.02	-0.82±2.91	S^2<0.6
8	-92.78	-97.12	**-0.49±0.01**	**-1.80±1.48**	S^2<0.6
9	-92.50	-96.20	0.12±0.03	0.26±4.74	S^2<0.6
10	-92.70	-95.81	**1.22±0.01**	**0.37±1.78**	S^2<0.6
11	-92.83	-95.37	x	x	S^2<0.6
12	-91.13	-93.29	**0.60±0.01**	**3.25±1.48**	S^2<0.6
13	-92.26	-98.25	**3.17±0.03**	**8.00±5.09**	S^2<0.6
14	-93.21	-100.70	0.17±0.03	-10.05±4.37	
15	-93.21	-99.01	**1.91±0.02**	**7.24±2.93**	
16	-93.53	-96.02	**1.28±0.01**	**5.73±2.04**	
17	-93.67	-101.16	**-1.15±0.01**	**6.68±1.87**	
18	-93.52	-101.42	**1.95±0.02**	**5.99±3.03**	
19	-93.39	-97.60	**1.68±0.02**	**4.20±2.63**	
20	-94.05	-98.38	**0.74±0.01**	**5.10±1.57**	
21	-93.62	-101.71	**0.67±0.01**	**8.80±1.56**	
22	-93.19	-101.82	**4.45±0.05**	**7.10±7.53**	
23	-93.82	-96.77	**-0.52±0.01**	**6.92±1.63**	
24	-94.35	-100.13	**1.09±0.01**	**9.30±1.83**	
25	-92.87	-101.11	**2.89±0.04**	**8.06±5.91**	
26	-93.56	-99.37	**3.15±0.03**	**6.84±5.14**	
27	-93.43	-97.54	**-1.83±0.02**	**13.42±3.10**	
28	-93.26	-99.95	**2.98±0.03**	**15.90±4.89**	S^2<0.6
29	-93.83	-99.37	3.97±0.04	-3.56±6.15	S^2<0.6
30	-94.22	-94.00	**0.29±0.01**	**9.48±2.40**	S^2<0.6
32	-92.43	-95.22	**0.51±0.01**	**12.92±1.59**	S^2<0.6
33	-92.87	-97.32	**4.02±0.04**	**10.01±6.41**	S^2<0.6
34	-93.88	-95.72	**0.77±0.01**	**23.37±1.67**	S^2<0.6
35	-93.02	-99.41	-0.14±0.03	-25.89±4.60	
36	-93.10	-87.72	**1.22±0.02**	**41.83±2.15**	
37	-92.98	-84.23	**2.42±0.03**	**29.76±4.00**	
38	-93.66	-85.27	**0.85±0.01**	**44.81±2.09**	
39	-94.40	-89.02	**0.92±0.01**	**46.09±2.13**	
40	-93.62	-86.63	**1.10±0.01**	**38.67±2.13**	
41	-93.60	-82.64	3.83±0.05	46.64±6.68	
42	-93.85	-87.36	**3.27±0.04**	**56.56±4.91**	
43	-92.37	-87.63	11.46±0.11	56.67±16.52	
44	-91.97	-86.86	**-3.64±0.04**	**68.79±6.16**	
45	-92.85	-90.55	**2.73±0.03**	**39.84±4.27**	
46	-93.95	-86.05	**-2.75±0.03**	**-20.62±4.37**	
48	-91.70	-97.08	**2.50±0.03**	**30.13±3.82**	
49	-93.06	-95.01	**-2.68±0.03**	**-10.00±4.37**	
50	-92.04	-90.11	8.08±0.09	9.73±14.69	
51	-93.44	-97.55	**1.26±0.01**	**5.92±2.04**	
52	-90.46	-99.10	7.64±0.07	12.55±11.33	
53	-93.26	-100.05	**-3.60±0.04**	**31.13±5.43**	
54	-93.65	-99.24	**2.39±0.02**	**11.98±3.52**	
55	-93.46	-94.25	**-6.13±0.07**	1.60±12.20	
56	-93.16	-93.64	**3.30±0.04**	**19.06±5.48**	S^2<0.6
57	-92.09	-87.68	**0.84±0.01**	**3.49±1.62**	S^2<0.6
58	-93.19	-94.31	x	x	S^2<0.6
59	-95.76	-94.92	**5.34±0.07**	2.33±11.27	S^2<0.6
60	-93.43	-90.50	**2.16±0.02**	**9.50±3.31**	S^2<0.6
61	-93.29	-91.43	**2.14±0.02**	**3.43±3.45**	S^2<0.6
62	-93.60	-94.45	0.12±0.03	-0.64±5.13	S^2<0.6
63	-93.71	-94.63	**1.87±0.02**	**12.04±3.19**	S^2<0.6
64	-93.28	-88.97	**2.34±0.02**	**-4.84±3.62**	S^2<0.6
65	-93.39	-93.44	**-2.09±0.02**	**2.57±3.24**	S^2<0.6
66	-93.18	-95.23	**4.03±0.04**	**10.52±6.41**	S^2<0.6
67	-92.59	-91.55	7.09±0.07	-0.54±12.26	S^2<0.6
68	-93.24	-91.32	**3.99±0.04**	**3.71±6.04**	S^2<0.6
69	-92.67	-95.20	**1.86±0.02**	**2.94±2.67**	S^2<0.6
70	-92.76	-94.86	**0.58±0.01**	**2.11±1.42**	S^2<0.6
71	-92.27	-92.97	-0.33±0.01	1.49±1.90	S^2<0.6

^{15}N-$^1H^N$ couplings in state N measured without and with alignment, J_{NH}(n-al) and J_{HN}(al), respectively, (**columns 2 and 3**) were obtained from ^{15}N-$^1H^N$ IPAP spectra recorded at 800 MHz.

^{15}N chemical shift and ^{15}N-$^1H^N$ RDC differences, $\Delta\varpi_N$ and ΔD_{NH}, between I and N states of the L24A FF domain (**columns 4 and 5**) were obtained from a global fit of ^{15}N single-quantum (SQ), TROSY (TR) and anti-TROSY (AT) CPMG relaxation dispersion data measured at 500 and 800 MHz (minimal uncertainties of 2% or 0.2/s, whichever is greater, were assumed for ^{15}N SQ, TR and AT data). The data for 64 residues (99 points per residue) were fit together to a global 2-state exchange model, with values of $k_{ex,NI}$=217±2 s^{-1}, p_I=5.89±0.04% and χ^2 =0.61 (residues 11 and 58 were excluded due to spectral overlap, 'x').

Relative signs of $\Delta\varpi_{IN}$ and ΔD_{IN} were determined by fitting ^{15}N SQ/TR/AT CPMG dispersion data as described previously [2]. Signs of $\Delta\varpi_{IN}$ were assumed to be the same as those obtained from measurements without alignment media (see Table S1; $\Delta\varpi_{IN}$ and ΔD_{IN} where signs determined from experiment are marked by bold).

Only RDC values for residues with S^2>0.6 were used in CS-Rosetta calculations of the structure of the folding intermediate; residues with low order parameters are marked by 'S^2<0.6' in **column 6**). ^{15}N-$^1H^N$ dipolar couplings, D_N, in the N state were calculated as J_N(al)-J_N(n-al) (**columns 2, 3**), and the corresponding values in the I state were obtained from the relation D_N+ΔD_{IN} (**columns 2, 3, 5**).

References

(1)   Hansen, D.F.; Vallurupalli, P.; Kay, L.E. *J. Phys. Chem. B* **2008**, *112*, 5898.

(2)   Vallurupalli, P.; Hansen, D.F.; Stollar, E.; Meirovitch, E.; Kay, L.E. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 18473.

(3)   Ishima, R.; Torchia, D.A. *J. Biomol. NMR* **2003**, *25*, 243.

(4)   Hansen, D.F.; Vallurupalli, P.; Lundstrom, P.; Neudecker, P.; Kay, L.E. *J. Am. Chem. Soc.* **2008**, *130*, 2667.

(5)   Lundstrom, P.; Hansen, D.F.; Vallurupalli, P.; Kay, L.E. *J. Am. Chem. Soc.* **2009**, *131*, 1915.

(6)   Lundstrom, P.; Hansen, D.F.; Kay, L.E. *J. Biomol. NMR* **2008**, *42*, 35.

(7)   Korzhnev, D.M.; Religa, T.L.; Banachewicz, W.; Fersht, A.R.; Kay, L.E. *Science* **2010**, *329*, 1312.

(8)   Korzhnev, D.M.; Religa, T.L.; Lundström, P.; Fersht, A.R.; Kay, L.E. *J. Mol. Biol.* **2007**, *37*, 497.

(9)   Bai, Y.W.; Milne, J.S.; Mayne, L.; Englander, S.W. *Proteins* **1993**, *17*, 75.

(10)   Connelly, G.P.; Bai, Y.W.; Jeng, M.F.; Englander, S.W. *Proteins* **1993**, *17*, 87.

(11)   Englander, S.W.; Mayne, L.; Bai, Y.; Sosnick, T.R. *Protein Sci.* **1997**, *6*, 1101.

(12)   Englander, S.W. *Annu. Rev. Biophys. Biomolec. Struct.* **2000**, *29*, 213.

(13)   Wishart, D.S.; Bigam, C.G.; Holm, A.; Hodges, R.S.; Sykes, B.D. *J. Biomol. NMR* **1995**, *5*, 67.

(14)   Johnson, B.A.; Blevins, R.A. *J. Biomol. NMR* **1994**, *4*, 603.

(15)   Loria, J.P.; Rance, M.; Palmer, A.G. *J. Am. Chem. Soc.* **1999**, *121*, 2331.

(16)   Skrynnikov, N.R.; Dahlquist, F.W.; Kay, L.E. *J. Am. Chem. Soc.* **2002**, *124*, 12352.

(17)   Korzhnev, D.M.; Neudecker, P.; Mittermaier, A.; Orekhov, V.Y.; Kay, L.E. *J. Am. Chem. Soc.* **2005**, *127*, 15602.

(18)   Shen, Y.; Bax, A. *J. Biomol. NMR* **2007**, *38*, 289.

(19)   Lundstrom, P.; Teilum, K.; Carstensen, T.; Bezsonova, I.; Wiesner, S.; Hansen, D.F.; Religa, T.L.; Akke, M.; Kay, L.E. *J. Biomol. NMR* **2007**, *38*, 199.

(20)   Korzhnev, D.M.; Orekhov, V.Y.; Kay, L.E. *J. Am. Chem. Soc.* **2005**, *127*, 713.

(21)   Auer, R.; Neudecker, P.; Muhandiram, D.R.; Lundstrom, P.; Hansen, D.F.; Konrat, R.; Kay, L.E. *J. Am. Chem. Soc.* **2009**, *131*, 10832.

(22)   Ishima, R.; Baber, J.; Louis, J.M.; Torchia, D.A. *J. Biomol. NMR* **2004**, *29*, 187.