Supporting Information

Wani and Udgaonkar 10.1073/pnas.0908617106

SI Methods

Buffers and Reagents. All of the chemicals used were of ultrapure grade. D_2O (99.9%), DCl, NaOD, sodium phosphate, Tris and glycine were from Sigma. GdnHCl was from USB Corp. 10 or 20 mM sodium phosphate was used as the buffer at pH 7.2. For exchange at pH 8.2, 20 mM Tris was used as the buffer; 100 or 200 mM glycine was used as the buffer to quench the exchange at pH 2.6. All of the pH values reported were recorded on a Thermo Orion 420 pH meter. The values reported for D_2O buffers are uncorrected for any isotopic effect.

Deuteration of the SH3 Domain. The protein was deuterated by dissolving 4 mg of lyophilized protein in 600 μl of D_2O buffer (10 mM sodium phosphate pH 7.2). The solution was incubated for 24 h at 25 °C, and ESI-MS indicated an increase in mass by 158 ± 2 Da. The mass did not increase upon further incubation. The increase in mass was the same as when 1 mg of lyophilized protein was dissolved in 1 ml of D_2O-buffer (10 mM sodium phosphate pH 7.2), and heated at 70, 75, or 80 °C for 10 min.

Native-State Exchange at pH 8.2. 15 μl of deuterated protein (approximately 700 μM protein concentration) in D_2O buffer (10 mM sodium phosphate, pH 7.2) were diluted 15-fold into 210 μl of labeling buffer (20 mM Tris pH 8.25, in H_2O), and after different time intervals exchange was quenched at pH 2.6 by adding 225 μl of quench buffer (200 mM glycine pH 2.5 H_2O). The samples were then processed for mass spectrometry.

Processing of HX Samples for Mass Spectrometry. All of the samples were desalted immediately after quenching using a Sephadex G-25 HiTrap desalting column equilibrated with water at pH 2.6 (pH adjusted with formic acid) in conjunction with an Akta Basic HPLC system. The samples were analyzed by ESI-MS with no delay. The time taken for the processing of a sample was 2 ± 0.5 min.

In every experiment, 3 control samples were also analyzed (1). A protonated protein sample in the same solvent as that of the experimental samples (water at pH 2.6) (2). A deuterated protein sample (10 μl) in D_2O buffer was diluted 20-fold into 185 μl of D_2O and 5 μl of 2% formic acid just before injecting into mass spectrometer. This was done to confirm that the protein was completely deuterated before the start of the experiment (3). A sample corresponding to the native-state at pH 7.2, prepared as discussed above, was also analyzed to check that the native-state had 19 ± 1 deuteriums protected.

Pepsin Digestion. A column having a bed volume of 400 μl was packed manually with pepsin-agarose beads (from Sigma). The column was incubated on ice and equilibrated with ice-cold water at pH 2.6 (pH adjusted with formic acid). 400 μl of exchanged, quenched, and desalted sample were loaded onto the pepsin-agarose column, and the column was centrifuged at 2200 rpm for 18 s to introduce the sample into the pepsin-agarose matrix. After 30 s of incubation, 400 μl of ice-cold water at pH 2.6 were added on top of the matrix and the column was again centrifuged for 18 s at 2200 rpm, and 400 μl of the digested sample were collected in a 1.5 ml centrifuge tube. The digested samples were immediately analyzed by ESI-MS. An identical procedure was followed for samples labeled in presence of 1.8 M GdnHCl.

Protonated protein in water at pH 2.6 was digested by following same procedure, the peptides obtained were identified by exact mass measurement and collision induced dissociation (CID) MS/MS. The intact protein sample corresponding to the native-state, intermediate in the absence of denaturant, and intermediate in presence of 1.8 M GdnHCl were also analyzed at the same time. The extent of labeling observed was the same, as mentioned earlier. The intact samples were incubated in water at pH 2.6 after desalting for 90 ± 10 s, the time required for pepsin digestion.

Data Acquisition by ESI-MS. The samples were injected at a flow rate of 18 μl/min using a 250 μl Hamilton glass syringe in conjunction with a Harvard pump. For acquisition of MS data of the intact as well as of the digested samples, the capillary voltage was set to + 2 kV, the source temperature to 80 °C, and the desolvation temperature to 200 °C. No organic solvent was added to any of the samples. A typical 1 s scan contained about 50 ion-counts, and data from the first 75 scans were combined. For each sample, the syringe and flow-line were washed thoroughly, and a blank spectrum was acquired between 2 successive samples to check that there was no left-over from the previous sample.

Data Analysis. The first 75 scans of each mass spectrum were averaged, smoothened, and background subtracted using the MassLynx 4.0 software. The data were normalized by making the sum of all intensities of the + 9 charge state equal to 1, for each mass profile. The list spectrum option in MassLynx software was used to copy the data to the Igor-Pro statistical software.

Determination of the Rate Constant of Formation of Fully Protonated Molecules. For unfolding/HX in 1.8 M GdnHCl, the fraction of molecules fully protonated at any time point was determined as the area under the peak centered at m/z = 1041.68 ± 0.06 relative to the area under the full mass spectrum. To obtain areas, the mass profile was fitted to the sum of 2 Gaussian distributions. At each denaturant concentration, the plot of the fraction of protonated molecules (U) versus time was compared with the respective simulated curve and rate constants of formation of fully protonated molecules were those used in simulation (Fig. S1).

A similar analysis was carried out of the mass profiles describing unfolding in 0, 0.5, and 1 M GdnHCl (Fig. S6), although caution has to be exercised in the interpretation because, unlike in the case of unfolding in 1.8 M GdnHCl, the higher mass and lower mass peaks both change not only in relative intensity but also in position and width during exchange/unfolding in native-like conditions.

Peak Width Analysis of the Mass Profiles. For HX/unfolding in 0, 0.5, and 1 M GdnHCl, the overall rate constant at which all 19 amide hydrogen sites get protonated was obtained from an estimate of the half-life ($t_{1/2}$) of the observed exchange process. This was obtained as the time of exchange at which the width of the mass distribution is maximum (1). The width of the mass profile was determined using the MassLynx 4.0 software and plotted versus the time of labeling (Fig. 3). The rate constant of unfolding/HX is estimated as $0.693/t_{1/2}$.

Text

Hydrogen Exchange Theory. Structural fluctuations in proteins break and reform hydrogen bonds, leading to unfolding and folding events having rate constants of k_u and k_f, respectively.

When not hydrogen bonded, and when accessible to solvent, the amide deuterium exchanges with a solvent proton. The rate constant of intrinsic exchange, k_{int} is base catalyzed when the pH of the solvent is about >4.

$$\mathrm{ND(folded)} \underset{k_f}{\overset{k_u}{\rightleftharpoons}} \mathrm{ND(unfolded)} \xrightarrow{k_{int}} \mathrm{NH(exchanged)}$$

k_{int} depends upon the pH, the nature of the amino acid residue, its flanking residues, and temperature. If k_u and k_f are independent of pH, then the observed exchange rate constant, k_{HX} is given by:

$$k_{HX} = \frac{k_u \cdot k_{int}}{k_u + k_f + k_{int}}$$

For exchange in the EX1 limit, where $k_u < k_f \ll k_{int}$, every unfolding event leads to exchange of one or more deuteriums, and the observed exchange rate constant, $k_{HX} = k_u$. In the EX2 limit, where $k_f \gg k_{int}$, every unfolding event does not result in exchange, and an equilibrium get established between the folded and the exchange-competent unfolded state. The observed exchange rate constant, $k_{HX} = (k_u/k_f)$. k_{int}. The intrinsic exchange rate constant, k_{int} increases 10-fold for every increase in pH by 1 unit. Hence, the observed exchange rate constant increases with an increase in pH for exchange in EX2 limit, while it is independent of pH for exchange in the EX1 limit.

Mechanism of Unfolding

Unfolding/HX in 1.8 M GdnHCl. The mechanism for the unfolding of the SH3 domain in 1.8 M GdnHCl has to account for the following observations and assumptions: (*i*) 15% of N molecules form I_U within 10 ms of unfolding, whose fluorescence properties are identical to those of U (2). (*ii*) The remaining 85% of N molecules are transformed into the native-like I_N at the same time (Fig. 1). (*iii*) The fluorescence change during unfolding occurs in 2 kinetic phases. 15% of the change occurs in a 10 ms burst phase, corresponding to the 15% accumulation of I_U. Hence, the fluorescence properties of I_N must be the same as that of N. (*iv*) Unfolded protein is formed in a single slow kinetic phase, as shown by interrupted unfolding experiments. The rate constant for the formation of U is the same as the rate constant for the disappearance of I_U and the same as the observed rate constant for the slow phase of fluorescence change ($0.02\ s^{-1}$). (*v*) The average value of the intrinsic HX rate constant, k_{int} is calculated to be approximately $20\ s^{-1}$ in zero denaturant at pH 7.2 and 25 °C, with values for the individual residues falling in the range of $2.6\ s^{-1}$ to $84\ s^{-1}$ (3). (*vi*) Fifteen percent of N molecules get fully protonated in a fast step, and the remaining 85% get fully protonated in a slow step (Fig. 1). The observed rate constant for the slow step is the same (within a factor of approximately 2) as the observed rate constant for the formation of U. It appears that the fast phase of complete protonation corresponds to the protonation of the molecules that have unfolded to I_U, while the slow phase of complete protonation corresponds to the formation of U. (*vii*) The transformation of the 85% molecules that initially form partially protonated I_N into fully protonated U occurs in a slow, concerted, apparently 2-state manner, as indicated by the iso-*m*/*z* point in the bimodal mass spectra. Bimodal spectra are observed because of the slow unfolding of I_N to U, and because both I_N and U have distinctive exchange properties, giving rise to the 2 peaks (4). (*viii*) The observation of bimodal mass spectra during the transformation of the 85% molecules that initially form partially protonated I_N into fully protonated protein might suggest that HX into U occurs in the EX1 limit (5). This is possible only if the rate constant of folding of U is $\ll k_{int}$. Hence, the observation that the fluorescence-monitored rate constant of the folding of U is $0.003\ s^{-1}$ in 1.8 M GdnHCl (2), which is much slower than k_{int}, is consistent with HX into U being in the EX1 limit.

Kinetic simulations using the program KINSIM (6) were used to distinguish between the simplest possible mechanisms that involve 2 unfolding intermediates.

$N \leftrightarrow I_N \leftrightarrow I_U \leftrightarrow U$ Mechanism. In this mechanism, I_N precedes I_U in the direction of unfolding because it is less unfolded than I_U. Hence, I_N like I_U would form within 10 ms of unfolding. But because k_{int} is much faster (see *Unfolding/HX in 1.8 M GdnHCl*) than the rate constant of the $I_U \rightarrow U$ transition, which is $0.1 s^{-1}$ in 1.8 M GdnHCl (2), complete protonation of all protein molecules would occur in I_U itself in one fast step, instead of in a fast step (for 15% of the molecules) followed by a slow step (for 85% of the molecules), as is observed (Fig. 1). Hence, this mechanism can be ruled out.

$I_N \leftrightarrow N \leftrightarrow I_U \leftrightarrow U$ Mechanism. In this mechanism, rapid kinetic partitioning of N into 15% I_U and 85% I_N would need to occur at rate constants of approximately $800\ s^{-1}$ and $200\ s^{-1}$, respectively, because I_U is formed within 10 ms (2). The rate constants of the $U \rightarrow I_U$ and $I_U \rightarrow U$ transitions are 0.003 and $0.1\ s^{-1}$, respectively, in 1.8 M GdnHCl (2). The $I_N \rightarrow N$ transition would need to be slow (approximately $0.1\ s^{-1}$) (this is expected in unfolding conditions), because it would be rate limiting for the complete protonation of the 85% molecules that initially form I_N. This mechanism predicts, however, that the slow fluorescence change accompanying the unfolding of the 85% molecules that initially form I_N should begin in a lag phase, during which I_N molecules slowly refold to N. No lag phase is, however, observed (2). Hence, this mechanism is also very unlikely.

Branched Mechanism. In this mechanism, all N molecules rapidly transform into I_U, which very rapidly converts to I_N so that 15% I_U and 85% I_N are present from 10 ms to 5 s. Only I_U can transform into U. An additional direct pathway of N unfolding to U is not included because no U is seen to form initially during unfolding (2). This mechanism is different from the $N \leftrightarrow I_N \leftrightarrow I_U \leftrightarrow U$ mechanism only in the order in which I_N and I_U form very rapidly. Thereafter, the 2 mechanisms become equivalent in describing further unfolding and HX. Hence, the branched mechanism is rejected for the same reasons that the $N \leftrightarrow I_N \leftrightarrow I_U \leftrightarrow U$ mechanism is rejected.

Two Pathway Mechanism. In this mechanism, the $N \leftrightarrow I_N \leftrightarrow U$ pathway competes with the $N \leftrightarrow I_U \leftrightarrow U$ pathway. As in the $I_N \leftrightarrow N \leftrightarrow I_U \leftrightarrow U$ mechanism, rapid (10 ms) kinetic partitioning into 15% I_U and 85% I_N occurs, and U is formed from both I_N and I_U. Rapid preequilibration between N, I_N and I_U does not occur because the $I_N \rightarrow N$ and $I_U \rightarrow N$ transitions are much slower than the $N \rightarrow I_N$ and $N \rightarrow I_U$ transitions, which is expected in unfolding conditions.

Fig. S1 shows the 2-pathway mechanism with the rate constants used to simulate the data in Fig. 1. The 2-pathway mechanism accounts for all of the experimental observations listed at the beginning of this section. This mechanism can also account for the observation that I_U accumulates to larger extents at higher GdnHCl concentrations (2) if kinetic partitioning favors the $N \leftrightarrow I_U \leftrightarrow U$ pathway at higher GdnHCl concentrations, and the $N \leftrightarrow I_N \leftrightarrow U$ pathway at lower GdnHCl concentrations.

HX/Unfolding in Native-Like Conditions. The native protein unfolds very transiently to I_N and U in native-like conditions too. In native-like conditions, the rate constants for the $N \rightarrow I_N$ and $N \rightarrow I_U$ transitions will be very slow because they are unfolding transitions. The rate constant of the $I_U \rightarrow N$ transition is

expected to be very fast; in 1 M GdnHCl itself it is >300 s^{-1} (2), which is much faster than k_{int}. Any HX that might occur into I_U would therefore occur by the EX2 mechanism, and it will be very slow because the N ↔ I_U equilibrium greatly favors N. Hence, very little labeling of I_U is expected to occur in native-like conditions. The rate constant of the I_N → N transition is not known, but it appears to be much slower than k_{int} because HX into I_N occurs in the EX1 limit, as indicated by its pH independence (Fig. S3 and Fig. S4).

The kinetic data in Figs. 2 and S6 describing complete protonation as a consequent of the transient sampling of U were simulated to the 2-pathway model (see *Two Pathway Mechanism*) using KINSIM. The simulations showed that only the N ↔ I_N ↔ U pathway is operative under native-like conditions. It is not necessary to include the N ↔ I_U ↔ U pathway at all. In the simulations, the values of the rate constants of the N → I_N and I_N → N transitions were adjusted so that I_N is generated in the times indicated in Figs. 2 and S2. The simulations could not determine whether I_N is an off-pathway intermediate or an on-pathway unfolding intermediate between N and U. It was assumed that I_N is on-pathway as it appears to be so for unfolding in 1.8 M GdnHCl.

Two-State Analysis of the Mass Distributions During the First Stage of Exchange/Unfolding. The mass spectra of N and I are each seen to fit well to a Gaussian distribution (Fig. S5.), as described by Eq. **1**:

$$P_{(x)} = a.e^{-0.5\left(\frac{x-m}{\sigma}\right)^2} \qquad [1]$$

m is the mean and σ is the width of the Gaussian-distributed mass ensemble. a is a prefactor. For N, the values obtained for a_N, m_N, and σ_N, were 0.00907, 1043.73, and 0.575, respectively. For I_N, the values obtained fo a_I, m_I, and σ_I were 0.0101, 1043.20, and 0.508, respectively.

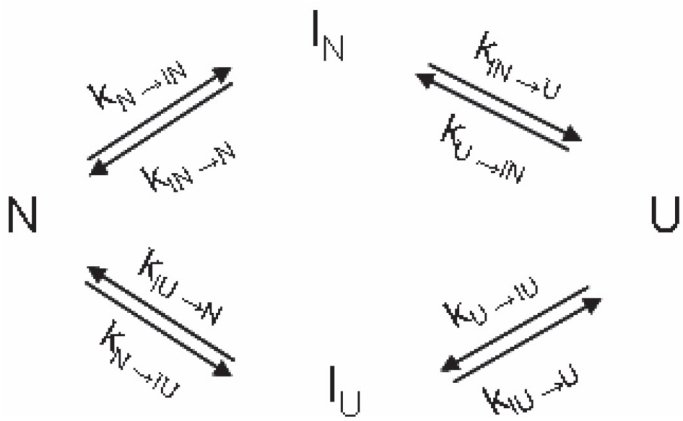
The gradual shifting of the unimodal mass spectra during the first stage of unfolding corresponding to the N to I_N transition was analyzed according to a 2-state N ↔ I mechanism. According to such a mechanism the fraction of molecules of I_N would increase at the cost of a decrease in the fraction of molecules of N, as the time of labeling increases. Consequently, the mass spectra at intermediate times of exchange would represent the weighted sum of the spectra of N and I_N. The mass profiles for labeling times between 5 s and 240 s were analyzed by a global fitting procedure using the Igor-Pro software, according to Eq. **2**:

$$y = a_N \cdot e^{-0.5\left(\frac{x-m_N}{\sigma_N}\right)^2} \cdot e^{-\frac{t}{\tau}} + a_I \cdot e^{-0.5\left(\frac{x-m_I}{\sigma_I}\right)^2} \cdot \left(1 - e^{-\frac{t}{\tau}}\right) \qquad [2]$$

t is the time of labeling, and τ is the time constant for the N ↔ I transition. Fig. S5 shows that a global fitting of the unimodal mass spectra describing the initial stage of exchange in 0 M GdnHCl, pH 7.2 is not satisfactory, indicating that more than one exchange-competent conformation is populated during the N → I_N transition. Fig. S5*f* shows that for a 2-state transition an iso-*m*/*z* point is expected, but such a point is not seen in the unimodal mass spectra (Fig. 2*A*). Similarly, the initial stages of exchange out in 0.5 and 1 M GdnHCl at pH 7.2 (Figs. S2 *A* and *C*), and in 0 M GdnHCl at pH 8.2 (Fig. S3*a*) are also not describable in terms of a 2-state transition between N and I.

1. Weis DD, Wales TE, Engen JR, Hotchko M, Ten Eyck LF (2006) Identification and characterization of EX1 kinetics in H/D exchange mass spectrometry by peak width analysis. *J Am Soc Mass Spectrom* 17:1498–1509.
2. Wani AH, Udgaonkar JB, (2009) Revealing a concealed intermediate that forms after the rate-limiting step of refolding of the SH3 domain of PI3 kinase. *J Mol Biol* 378:348–362.
3. Bai Y, Milne JS, Mayne L, Englander SW (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins* 17:75–86.
4. Ferraro DM, Lazo ND, Robertson AD (2004) EX1 hydrogen exchange and protein folding. *Biochemistry* 43:587–594.
5. Engen JR, Smithgall TE, Gmeiner WH, Smith DL (1997) Identification and localization of slow, natural, cooperative unfolding in the hematopoietic cell kinase SH3 domain by amide hydrogen exchange and mass spectrometry. *Biochemistry* 36:14384–14391.
6. Barshop BA, Wernn RF, Frieden C (1983) Analysis of numerical methods for computer simulation of kinetic processes: Development of KINSIM—a flexible, portable system. *Anal Biochem* 130:134–145.

a



b

[GdnHCl] (M)	$k_{N\rightarrow IN}$ (s^{-1})	$k_{IN\rightarrow N}$ (s^{-1})	$k_{IN\rightarrow U}$ (s^{-1})	$k_{U\rightarrow IN}$ (s^{-1})	$k_{N\rightarrow IU}$ (s^{-1})	$k_{IU\rightarrow N}$ (s^{-1})	$k_{IU\rightarrow U}$ (s^{-1})	$k_{U\rightarrow IU}$ (s^{-1})
0	0.017	2	0.00008	0.15	0	0	0	0
0.5	0.035	1.8	0.00026	0.1	0	0	0	0
1.0	0.13	1.5	0.001	0.06	0	0	0	0
1.8	1200	0.02	0.008	0.0002	160	0.12	0.1	0.003

Fig. S1. Mechanism of unfolding of the SH3 domain of PI3 kinase. (*a*) The N ↔ I_U ↔ U mechanism which accounts for previously obtained optical data (2) has been modified to accommodate the native-like intermediate, I_N, identified here by the HX studies. The data in Fig. 1, as well as previously obtained optical data, were simulated on the basis of this mechanism using the program KINSIM (6). (*b*) The table lists the values of the rate constants of the individual steps, which were used in the kinetic simulation to the mechanism shown in *a*.

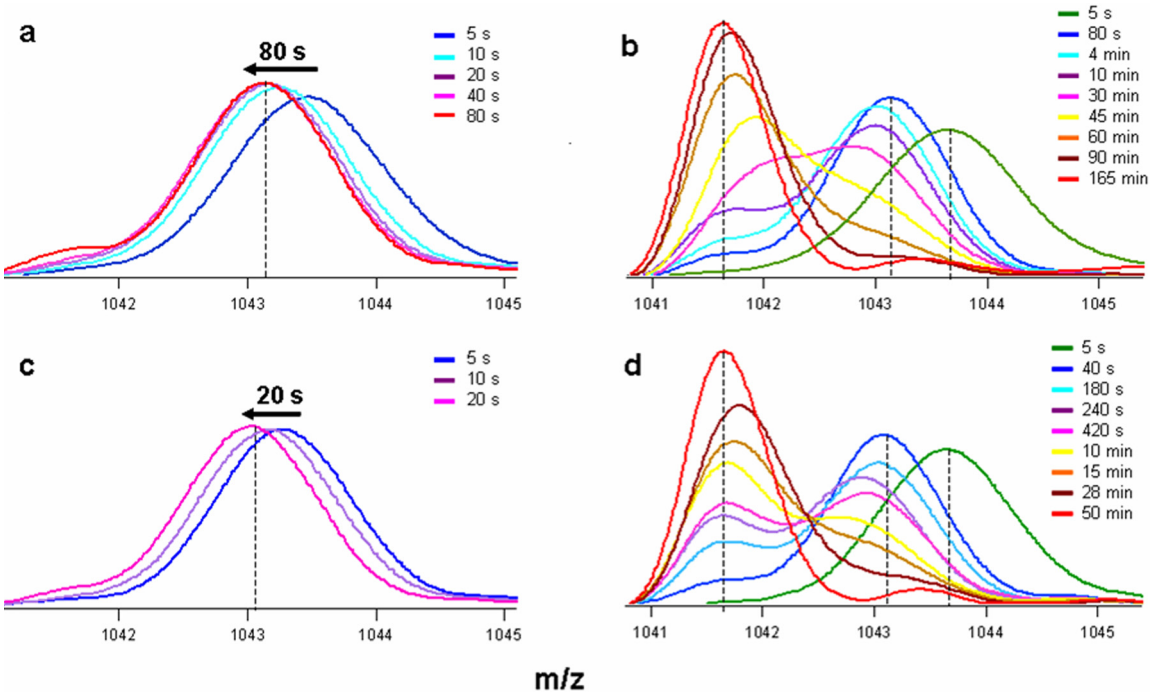


Fig. S2. Kinetics of exchange of the protected deuteriums from the native-state of SH3 domain at pH 7.2. Exchange was allowed to proceed for different times in 0.5 (*a* and *b*) and 1 M (*c* and *d*) GdnHCl. The mass spectra shown in green in (*b* and *d*) correspond to the native-state having 19 ± 1 deuteriums protected and was obtained by quenching the exchange after 5 s of labeling in the absence of GdnHCl. The mass spectra shown by the red lines in (*b* and *d*) correspond to the completely unfolded state, which retains 2 deuteriums due to the presence of residual (7%) D_2O during labeling. (*a* and *c*) show the unimodal mass spectra for exchange of the first 5 deuteriums, at different times of unfolding. The mass spectra shift gradually to the *m/z* value of I_N, having 14 deuteriums protected in the time indicated above the arrow. (*b* and *d*) show representative mass profiles for the exchange of the remaining 14 protected deuteriums. The mass spectra in all of the panels correspond to the + 9 charge state.

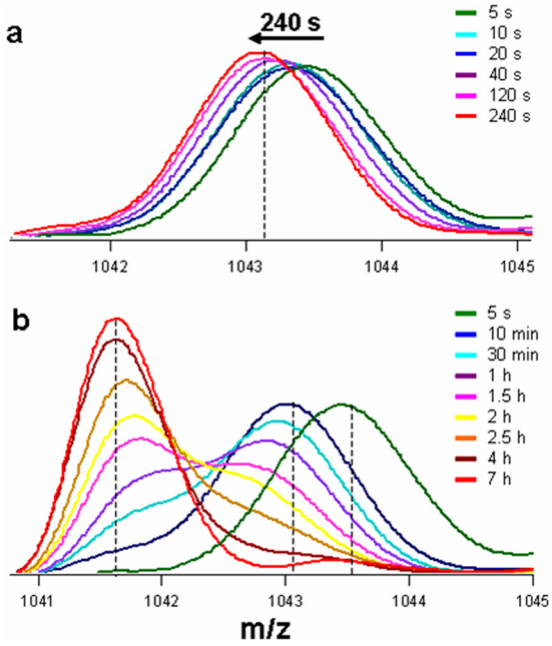


Fig. S3. Kinetics of exchange of protected deuteriums from the native-state of SH3 domain at pH 8.2. Exchange was allowed to proceed for different times in 0 M GdnHCl at pH 8.2. (*a*) shows representative unimodal mass spectra at early times of labeling. The time above the arrow indicates the time of labeling at which the mass spectrum attains the *m/z* value corresponding to the species having 14 deuteriums protected (I_N). (*b*) shows representative mass spectra at later times of labeling. The mass spectrum of the native-state shown in green was obtained after 5 s of labeling at pH 8.2. The mass spectrum shown by the red-line retains only 2 deuteriums indicating that all of the protein molecules have sampled the completely unfolded-state. All of the mass spectra correspond to the + 9 charge state.

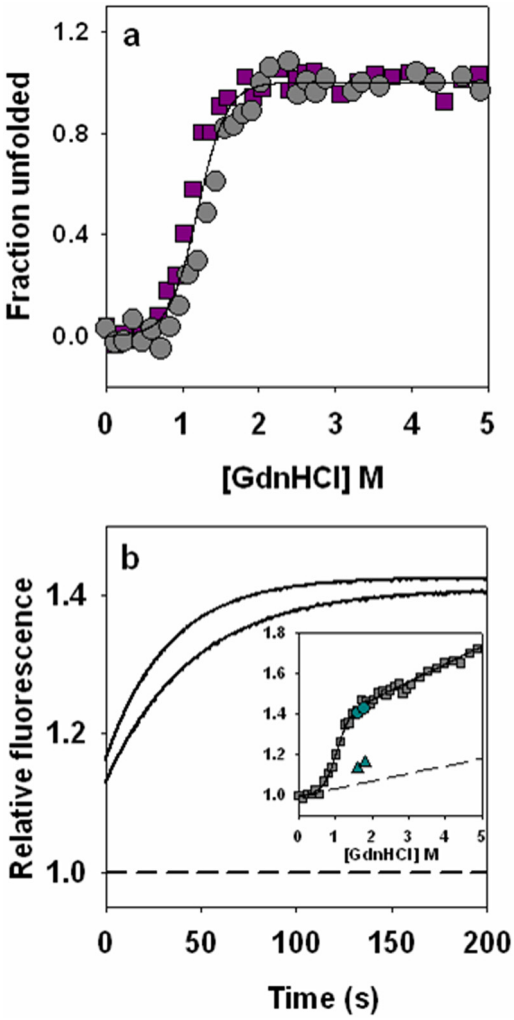


Fig. S4. Stability and kinetics of unfolding at pH 8.2. (*Upper*) GdnHCl-induced equilibrium unfolding transitions at pH 7.2 (●) and pH 8.2 (■) were determined from fluorescence measurements. The solid line through both unfolding curves has been drawn by inspection. (*Lower*) Fluorescence-monitored kinetic traces of unfolding in 1.6 and 1.8 M GdnHCl at pH 8.2. The dashed line indicates the fluorescence value of the native protein. The inset compares the kinetic and equilibrium amplitudes of unfolding at pH 8.2. The $t = 0$ (▲) and $t = \infty$ (●) fluorescence values of the kinetic traces are compared to the fluorescence values of equilibrium unfolded protein (■).

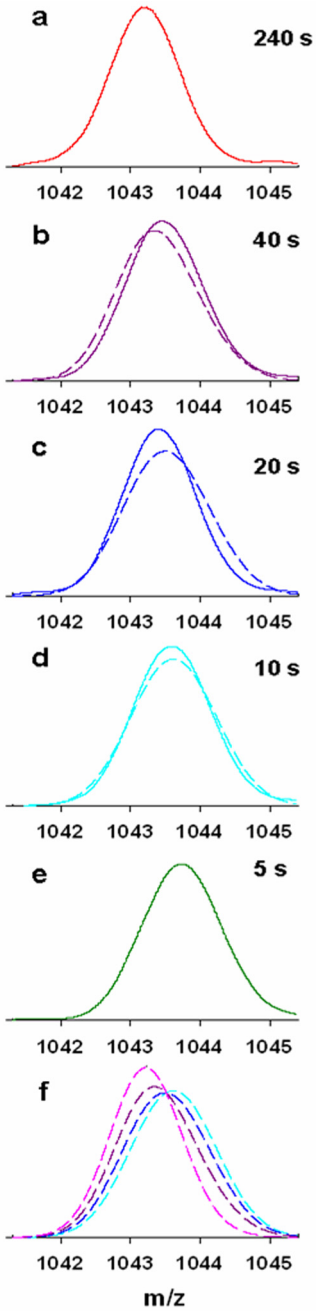


Fig. S5. Fitting of unimodal mass spectra to a 2-state N ↔ I mechanism. The mass spectra from Fig. 2*A* are shown here as continuous lines. The mass spectra shown in (*b*–*d*) were obtained after 40, 20, and 10 s of labeling, respectively. The mass spectra shown in red and green (*a* and *e*) correspond to the I_N and native-state, respectively. The dashed lines are the fits obtained by global analysis according to Eq. **2** (see *SI Data Analysis*). (*f*) shows an iso-*m*/*z* point, through which the fits from (*b*–*d*), and also the fit to the mass spectrum obtained after 2 min of labeling, cross each other.

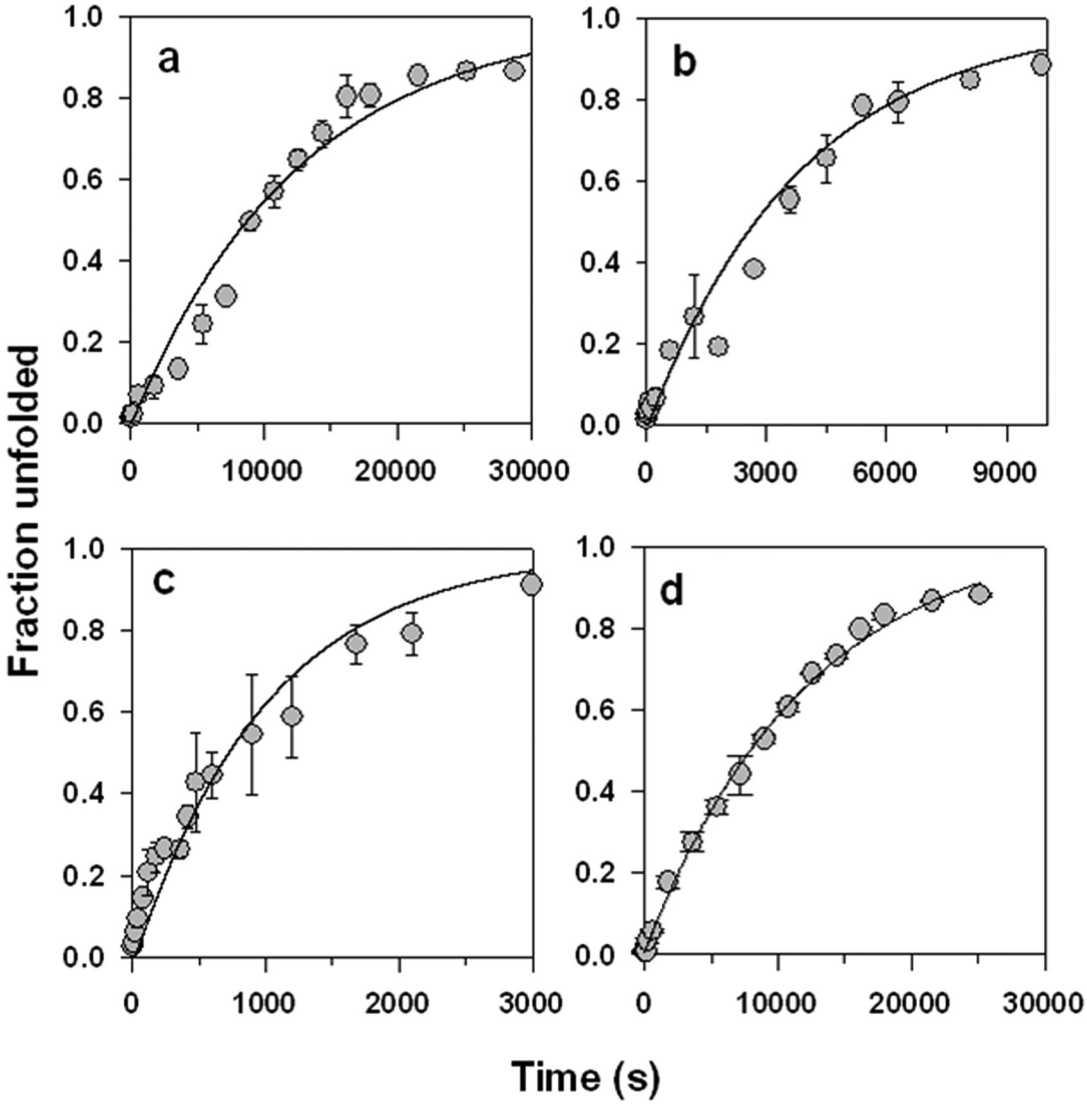


Fig. S6. Kinetics of formation of fully protonated molecules. The increase in the fraction of molecules that have become fully protonated with time of HX in 0 M GdnHCl, pH 7.2 (*a*); 0.5 M GdnHCl, pH 7.2 (*b*); 1 M GdnHCl, pH 7.2 (*c*), and 0 M GdnHCl, pH 8.2 (*d*) is shown. The fraction of molecules fully protonated at each time point was determined from the area under the lower mass peak arising from completely protonated molecules relative to the total area under the mass spectrum in Fig. 2 (for *a*), in Fig. S2 (for *b* and *c*), and in Fig. S3 (for *d*). In (*a*–*c*) the solid line through the data represents a kinetic simulation to the 2-pathway mechanism (Fig. S1), using values for the rate constants as shown. The error bars represent standard deviations determined from 3 separate experiments.

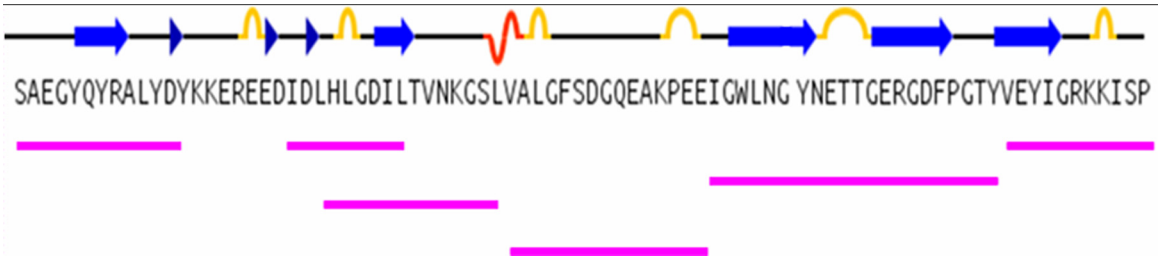


Fig. S7. Pepsin fragmentation map of the SH3 domain. The primary structure of the SH3 domain of PI3K is shown along with the location of the secondary structural elements. The pink lines represent the peptide fragmentation pattern of the SH3 domain, obtained upon pepsin-digestion at pH 2.6, as described in the text.