JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Revealing a Concealed Intermediate that Forms after the Rate-limiting Step of Refolding of the SH3 Domain of PI3 Kinase

Ajazul Hamid Wani and Jayant B. Udgaonkar*

National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bangalore 560065, India

Kinetic and equilibrium studies of the folding and unfolding of the SH3 domain of the PI3 kinase, have been used to identify a folding intermediate that forms after the rate-limiting step on the folding pathway. Folding and unfolding, in urea as well as in guanidine hydrochloride (GdnHCl), were studied by monitoring changes in the intrinsic fluorescence or in the far-UV circular dichroism (CD) of the protein. The two probes yield non-coincident equilibrium transitions for unfolding in urea, indicating that an intermediate, I, exists in equilibrium with native (N) and unfolded (U) protein, during unfolding. Hence, the equilibrium unfolding data were analyzed according to a three-state N ↔ I ↔ U mechanism. An intermediate is observed also in kinetic unfolding studies, and its presence leads to the unfolding reaction in urea as well as in GdnHCl, occurring in two steps. The fast step is complete within the initial 11 ms of unfolding and manifests itself in a burst phase change in fluorescence. At high concentrations of GdnHCl, the entire change in fluorescence during unfolding occurs during the 11 ms burst phase. CD measurements indicate, however, that I retains N-like secondary structure. An analysis of the kinetic and thermodynamic data, according to a minimal three-state N ↔ I ↔ U mechanism, positions I after the rate-limiting transition state, TS1, of folding, on the reaction coordinate of folding in GdnHCl. Hence, I is not revealed when folding is commenced from U, regardless of the nature of the probe used to follow the folding reaction. Interrupted unfolding experiments, in which the protein is unfolded transiently in GdnHCl for various lengths of time before being refolded, showed that I refolds to N much faster than does U, confirms the analysis of the direct folding and unfolding experiments, that I is formed after the rate-limiting step of refolding in GdnHCl.

© 2009 Elsevier Ltd. All rights reserved.

Keywords: SH3 domain; intermediates; transition state; burst phase; interrupted unfolding

Edited by K. Kuwajima

Introduction

There has been considerable controversy about the role of folding intermediates in protein folding, because of the discovery that many proteins fold in an apparently two-state U ↔ N manner, from the unfolded state (U) to the native state (N).[1] The controversy has abated now with the realization that: (1) the kinetic criteria used to define a protein as a two-state folder are oblivious to high-energy intermediates that may be populated transiently and marginally either before or after the rate-limiting step;[2,3] (2) such scarcely populated intermediates can now be detected by the application of high-resolution probes like NMR to the study of the folding of apparent two-state folders;[4] (3) for proteins that appear to fold or unfold by a two-state mechanism in one set of conditions, intermediates can be induced to accumulate in another set of conditions,[5–9] or upon mutation;[10] and (4) deviations from linearity in the dependence on denaturant concentration of the activation free energy of folding or unfolding or

*Corresponding author. E-mail address:
jayant@ncbs.res.in.

Abbreviations used: GdnHCl, guanidine hydrochloride; HX, hydrogen exchange; MS, mass spectrometry; ESI, electrospray ionization; PI3, phosphatidylionositol-3; SH3, Src homology-3.

both have been observed for a number of apparent two-state folders, and ascribed to sequential folding mechanisms with one or more intermediates between U and N.[11–15] It seems clear now that for many of the apparent two-state folders, it is merely a question of using the right folding conditions, or of using the right experimental probe, before a folding intermediate is revealed. Revealing an otherwise concealed intermediate is important because folding intermediates serve as signposts on folding pathways. In the absence of such a signpost, it is difficult to define a folding pathway, and to distinguish one pathway from an alternative pathway that may be used in different folding conditions.[16] The structural characterization of a late folding intermediate will allow the most direct determination of the sequence of events that occur late on the folding pathway.

To identify and characterize folding intermediates that are concealed after the rate-limiting step has been a difficult challenge. The possibility of such intermediates was first suggested by native state hydrogen exchange (HX) studies that indicated that the native state exists in equilibrium with many different partially unfolded forms.[17] In the case of ribonuclease A, it has been shown that one such partially unfolded form is an unfolding intermediate that accumulates on the direct unfolding pathway.[13] Such intermediates, which precede the rate-limiting step in unfolding, might be the same as the intermediates that are concealed after the rate-limiting step of folding.[18] Thus, one approach to detect intermediates that populate after the rate-limiting step of folding is to detect partially unfolded intermediates in kinetic unfolding studies. Partially unfolded intermediates have been detected in kinetic studies of the unfolding of several proteins.[9,12,1–21] It is important to characterize such intermediates because partially unfolded forms very often aggregate to form amyloid protofibrils and fibrils.[9,22]

A major group of two-state folding proteins is that of the SH3 domains,[23] which are found in signal transduction and cytoskeletal proteins,[24–27] where they provide the interaction surfaces for the assembly of multimeric protein–protein complexes.[28] SH3 domains are composed of 60–85 amino acid residues, and they all have a similar 3D fold, called the β-barrel fold. SH3 domains from many different proteins have been used as model systems to address various important aspects of protein folding, such as the nature of protein folding transition states,[29–31] and the relationship between protein topology and the folding pathway.[32] The folding of SH3 domains has been studied by protein engineering methods,[29–31] NMR,[4,33] as well as by computational studies.[34–36] Recently, the characterization of very scarcely populated folding intermediates preceding the rate-limiting step during the folding of the Fyn[4] and Abpl[33] SH3 domains, became possible with the advent of new NMR methods.[37] No unfolding intermediate has been reported for any of the SH3 domains, but partially unfolded conformations of many different SH3 domains have been observed under native conditions, when their dynamics were

studied by hydrogen/deuterium exchange in conjunction with mass spectrometry.[38]

The SH3 domain of PI3 kinase consists of 83 residues, threaded through five β-strands and two helix-like turns. The five β-strands are arranged in two β-sheets that are orthogonal to each other.[39] The folding of the SH3 domain of PI3 kinase was studied earlier using various optical probes and real time 1D NMR.[40] It appeared that the SH3 domain of PI3 kinase refolds and unfolds by a two-state mechanism, based on a comparison of the values of the free energy of unfolding (ΔG) and its dependence on denaturant concentration, obtained from kinetic and equilibrium studies.

Here, we have studied the folding and unfolding of the SH3 domain of the PI3 kinase, in urea as well as in GdnHCl, by monitoring the change in intrinsic fluorescence. An equilibrium unfolding intermediate, I, is shown to be populated during urea-induced unfolding. A burst phase change in fluorescence, whose amplitude increases with an increase in denaturant concentration, is observed during the unfolding of the protein in urea as well in GdnHCl. The product of the burst phase is shown to be a partially unfolded intermediate, which exposes less surface area to solvent than does U, and which refolds to N much faster than does U. All the thermodynamic and kinetic data can be described adequately on the basis of a simple and minimal three-state N $\leftrightarrow$ I $\leftrightarrow$ U mechanism, suggesting that the same intermediate I is seen in the equilibrium and kinetic studies. The analysis indicates that I forms after the rate-limiting step during the folding of this SH3 domain in GdnHCl.

Results

Equilibrium unfolding

Figure 1 shows the equilibrium unfolding transitions of the SH3 domain, with both urea (Fig. 1a–c) and GdnHCl (Fig. 1d–f) as denaturants. In the case of urea-induced unfolding, it is seen that the fluorescence-monitored and far-UV CD-monitored unfolding transitions are not coincident, indicating that an equilibrium unfolding intermediate, I, is populated during the unfolding of N to U. In the case of GdnHCl-induced unfolding, the far-UV CD-monitored unfolding transition has an anomalous shape, but it also appears not to be coincident with the fluorescence-monitored transition. Hence, both the urea-induced unfolding transitions and the GdnHCl-induced unfolding transitions need to be analyzed on the basis of a three-state N $\leftrightarrow$ I $\leftrightarrow$ U mechanism, according to Eq. (4). Since the loss of fluorescence is seen to precede the loss in far-UV CD during unfolding, the simplest assumption to make in the analysis is that the fluorescence properties of I are the same as that of U, and that the far-UV CD properties of I are the same as that of N. These assumptions amount to treating the fluorescence-
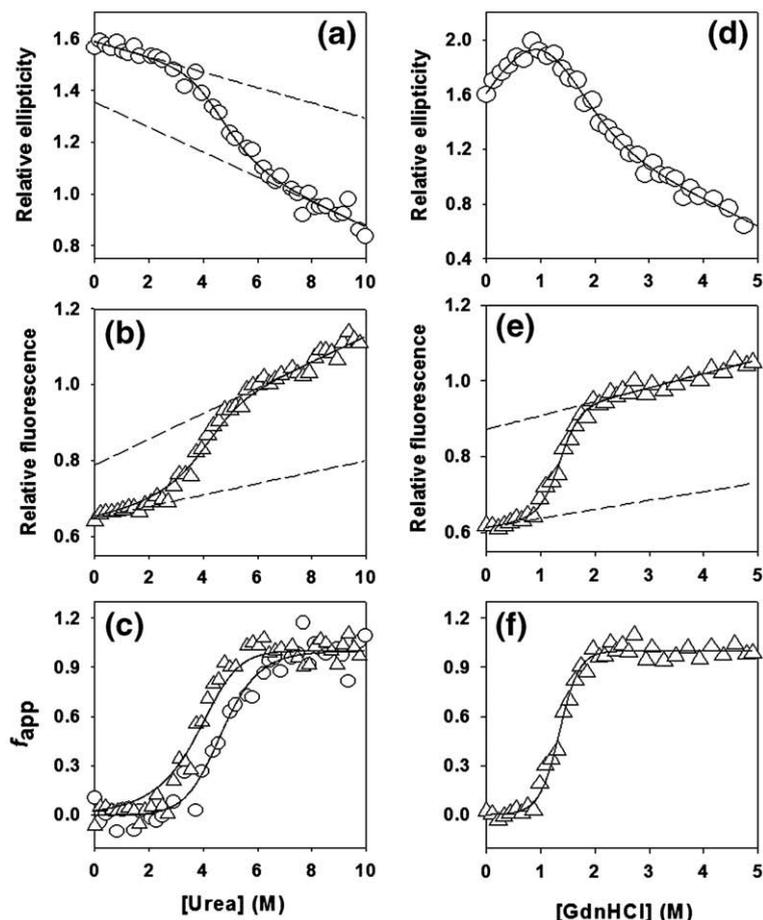
Fig. 1. Equilibrium unfolding of the SH3 domain of PI3 kinase at pH 7.2. Unfolding by urea (a, b and c) as well as GdnHCl (d, e and f) were monitored by changes in far-UV CD at 222 nm (a and d) and fluorescence at 300 nm (b and e). The broken lines in a, b, and e are the linearly extrapolated native protein and unfolded protein baselines. The continuous lines through the far-UV CD and fluorescence data in a and b are fits of the data to Eq. (1), with the value of Z_I set to 1 for the fluorescence data and to 0 for the CD data, and with the values of ΔG_{NI} and m^E_{NI} set to 2.1 kcal mol^{-1} and -0.48 kcal mol^{-1} M^{-1}, respectively (see the text). The fit yielded values for ΔG_{NU} and m^E_{NU} of 4.8 kcal mol^{-1} and −1.12 kcal mol^{-1} M^{-1}, respectively (see Table 1). The data in a and b were converted to f_{app} values using Eq. (3), and plotted against the concentration of urea in c. The continuous lines through the far-UV CD (o) and fluorescence (Δ) data in c are drawn according to Eq. (4), with the value of Z_I set to 1 for the fluorescence data and to 0 for the CD data, and with values for ΔG_{NI}, ΔG_{NU}, m^E_{NI} and m^E_{NU} of 2.1 kcal mol^{-1}, 4.8 kcal mol^{-1}, -0.48 kcal mol^{-1} M^{-1} and -1.12 kcal mol^{-1} M^{-1}, respectively (see Table 1). The solid line through the CD data in panel d was drawn by inspection only. The continuous line through the fluorescence data in e is a least-squares fit of the data to Eq. (1), with the value of Z_I set to 1, and the values of ΔG_{NI} and m^E_{NI} set to 3.7 kcal mol^{-1} and -1.53 kcal mol^{-1} M^{-1}, respectively (see the text). The fit yielded values for ΔG_{NU} and m^E_{NU} of 4.5 kcal mol^{-1} and –3.34 kcal mol^{-1} M^{-1}, respectively. The data in e were converted to f_{app} values using Eq. (3) and plotted against the concentration of GdnHCl in f. The continuous line through the fluorescence data in f is drawn according to Eq. (4), with the value of Z_I set to 1, and the values of ΔG_{NI}, ΔG_{NU}, m^E_{NI} and m^E_{NU} set to 3.7 kcal mol^{-1}, 4.5 kcal mol^{-1}, –1.53 kcal mol^{-1} M^{-1}, and –3.34 kcal mol^{-1} M^{-1}, respectively (see above).

monitored equilibrium unfolding curve as one monitoring the loss in tertiary structure, and the far-UV CD-monitored equilibrium unfolding curve as one monitoring loss of secondary structure. Nevertheless, even with this assumption it is difficult to extract reliably the values of all four thermodynamic parameters ΔG_{NI} (H$_2$O), ΔG_{IU} (H$_2$O), m^E_{NI}, and m^E_{IU} (see Data Analysis) from the three-state analysis of the urea-induced unfolding transitions, and especially for the GdnHCl-induced unfolding transitions.

Refolding and unfolding kinetics

Unfolded protein in 4 M GdnHCl or in 6 M urea was refolded by diluting the denaturant to different concentrations below the mid-point of the equilibrium unfolding transition, and the kinetics were monitored by the change in fluorescence at 300 nm. Figure 2a and b show representative kinetic refolding traces in different concentrations of urea and GdnHCl, respectively. A major refolding phase and a slower phase corresponding to prolyl peptide bond

isomerization was observed when refolding was carried out in final urea concentrations lower than 1.5 M. The kinetics in urea concentrations greater than 1.5 M are single-exponential, because the refolding rates appear to become comparable to the rate of prolyl peptide bond isomerization. In the case of GdnHCl also, a prolyl peptide bond isomerization phase was observed in addition to the major refolding phase, when refolding was carried out in final concentrations of GdnHCl below 0.9 M, and a single refolding phase was observed in final concentrations of GdnHCl above 0.9 M.

Figure 2c and d show representative fluorescence-monitored unfolding traces in different concentrations of urea and GdnHCl, respectively. In both denaturants, the kinetic traces do not extrapolate to the native protein signal at $t = 0$, but to values higher than the native protein signal, depending upon the denaturant concentration. Hence, there appears to be a burst phase change in fluorescence during the dead time of the stopped-flow mixer. Since the subsequent observable phase is relatively slow,
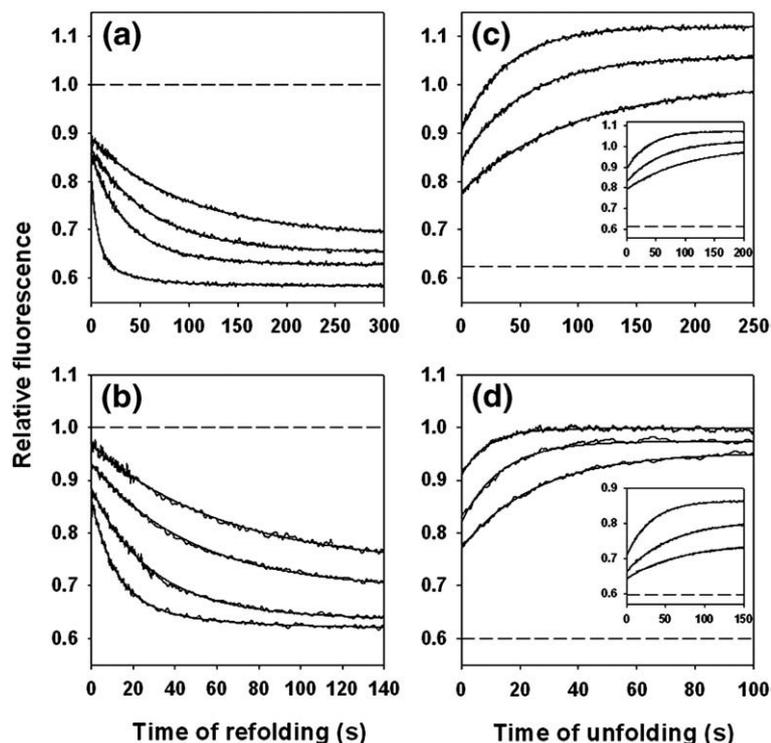
Fig. 2. Kinetics of refolding and unfolding of the SH3 domain of PI3 kinase. Refolding and unfolding at pH 7.2 were monitored by the change in fluorescence at 300 nm. (a) Representative kinetic refolding traces when unfolded protein in 6 M urea was refolded by diluting the urea to 2.96 M, 2.5 M, 2.0 M and 0.8 M (top to bottom). The refolding trace for 0.8 M urea was fit to a double-exponential equation, while the refolding traces for 2.0 M, 2.5 M and 2.96 M urea were fit to a single-exponential equation. The signal of the unfolded protein in 6 M urea, to which the kinetic curves were normalized is shown as a broken line. (b) Representative kinetic traces of refolding in 1.1 M, 1.0 M, 0.8 M and 0.53 M GdnHCl (top to bottom). The kinetic traces were normalized to the value of 1 for the unfolded protein signal in 4 M GdnHCl, shown as the broken line. The continuous lines show the fits to a single-exponential equation for the 1.1 M and 1.0 M GdnHCl data, and to a double-exponential equation for the 0.8 M and 0.53 M GdnHCl data. (c) Representative traces of unfolding in 8 M, 7 M and 6 M urea (top to bottom), from stopped-flow mixing experiments. All the traces were fit to a single-exponential equation, and normalized with respect to the $t = \infty$ point of the unfolding trace in 6 M urea. The native protein signal is shown as a broken line. The inset in c shows representative traces of unfolding in 8 M, 7 M and 6 M urea (top to bottom), from manual mixing experiments. (d) Representative traces of unfolding in 3.0 M, 2.5 M and 2.2 M GdnHCl (top to bottom). The continuous lines through the data represent fits to a single-exponential equation, and the broken line shows the native protein signal. Here also, the fluorescence signal of the protein in 4 M GdnHCl was taken as 1. The inset in d shows representative traces of unfolding in 2.0 M, 1.6 M and 1.4 M GdnHCl (top to bottom), from manual mixing experiments.

occurring in the 100 s time domain, the initial burst phase cannot be an experimental artifact of the sort that could conceivably arise in instances where the subsequent observable kinetic phase is very fast. The observed phase is well described as a single exponential. No unfolding trace was observed at concentrations of GdnHCl greater than 3.5 M, indicating that the entire fluorescence change during unfolding takes place during the 11 ms burst phase. It should be noted that the observation of a denaturant-dependent burst phase change in fluorescence during stopped-flow measurements of unfolding, seen here for the SH3 domain, is unusual,[12,41–45] and has not, for example, been seen in similar recent studies of the denaturant-induced unfolding of several other proteins.[9,13,46,47]

To confirm that the observed burst phase change in fluorescence observed in stopped-flow measurements was not an artifact, manual mixing experiments were done for all concentrations of GdnHCl and urea in which the observable unfolding rates could be measured precisely. When the observed unfolding rates were slower than 0.027 s^{-1}, and hence amenable to measurement by manual mixing experiments with a dead time of 10 s, the single-exponential fits to the kinetic unfolding traces could

be extrapolated reliably to $t = 0$, for determination of the burst phase changes in fluorescence. The insets in Fig. 2c and d show that the burst phase change in fluorescence is seen in stopped-flow mixing experiments, and in the manual mixing experiments. The relative amplitude of the burst phase fluorescence change, as well as the apparent rate constant of the observed kinetic phase, was the same for the manual mixing and stopped-flow mixing experiments at the same concentration of denaturant.

Figure 3a and b compare the kinetic and equilibrium amplitudes of fluorescence-monitored folding and unfolding, in urea, as well as in GdnHCl. The $t = \infty$ and $t = 0$ points of the kinetic refolding traces fall on the equilibrium unfolding curve, and the linearly extrapolated unfolded protein baseline, respectively. This indicates that the entire amplitude of the refolding reaction from U to N is captured in the kinetic experiments. For both denaturants, the $t = \infty$ points of the kinetic unfolding traces fall on the equilibrium unfolding transition, but the $t = 0$ points do not fall on the linearly extrapolated native protein baseline, as a result of the burst phase increase in fluorescence. For unfolding by GdnHCl, the $t = 0$ points show a typical sigmoidal dependence on the concentration of denaturant. For unfolding by urea,

the sigmoidal dependence could not be established clearly, because it was not possible to carry out unfolding experiments in sufficiently high concentrations of urea.

The sigmoidal dependence on denaturant concentration of the burst phase amplitude seen for unfolding in GdnHCl suggests that the burst phase increase in fluorescence accompanies the initial establishment of a pre-equilibrium between N and an unfolding intermediate, I, before the subsequent slower unfolding of I to U. Hence, the analysis was carried according to a two-state N ↔ I mechanism, in order to determine the values of ΔG_{NI} (H_2O) and m^E_{NI}. The native protein baseline used in this

analysis was that obtained for the equilibrium unfolding transition. The fluorescence properties of I appear to be similar to those of U at high concentrations of denaturant: in high concentrations of GdnHCl, the entire increase in fluorescence during unfolding occurs in the burst phase. Hence, it was assumed that the dependence on denaturant concentration of the fluorescence of I is the same as that of U. The values obtained for ΔG_{NI} (H_2O) and m^E_{NI} are 2.1 kcal mol^{-1} and -0.48 kcal mol^{-1} M^{-1}, respectively, when urea is used as the denaturant. In the case of GdnHCl, the value of ΔG_{NI} (H_2O) is 3.7 kcal mol^{-1}, and that of m^E_{NI} is –1.53 kcal $mol^{-1}M^{-1}$ (Table 1). For both urea-and GdnHCl-induced unfolding, the observation that the value of m^E_{NI} is less than the value of m^E_{NU}, suggests that the unfolding of N to I exposes less surface area than does the unfolding of N to U. In other words, I is not unfolded to the same extent as U, even though its fluorescence properties appear to be the same as that of U. Fig. 3c compares the kinetic and equilibrium amplitudes of far-UV CD-monitored unfolding, in urea. The $t = \infty$ and $t = 0$ points of the far-UV CD-monitored kinetic unfolding traces fall on the equilibrium unfolding curve, and the linearly extrapolated native protein baseline, respectively, indicating that the entire amplitude of the unfolding reaction from N to U is captured in the kinetic experiments (Fig. 3c).
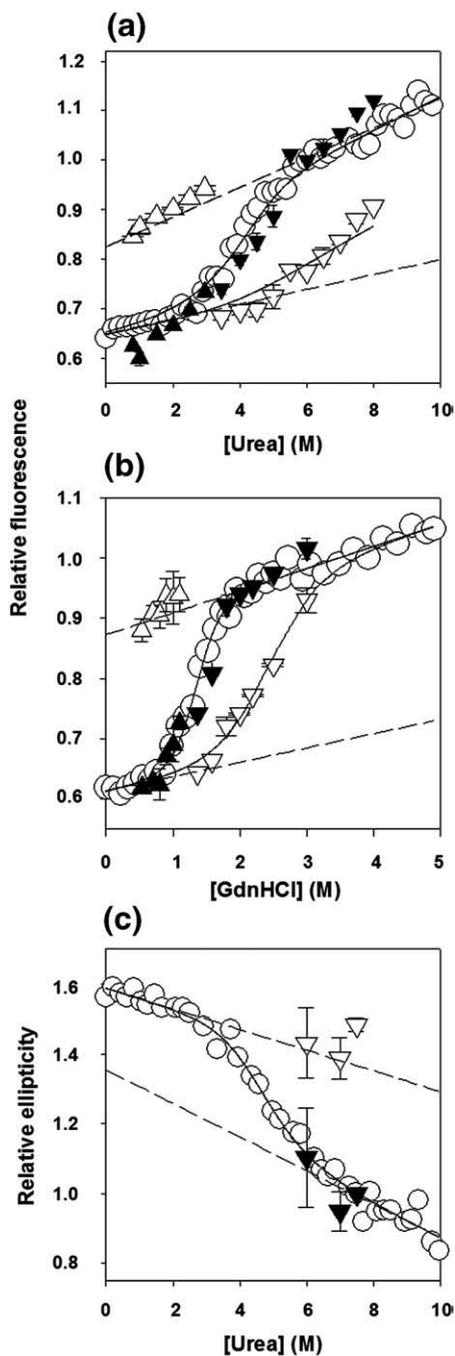
Fig. 3. Comparison of equilibrium and kinetic amplitudes. Equilibrium and kinetic amplitudes of refolding and unfolding monitored by the change in fluorescence at 300 nm (a and b) as well as by the change in ellipticity at 222 nm (c), are compared when urea (a and c) and GdnHCl (b) were used as the denaturant. (a–c) (O) equilibrium unfolding; (▲) $t=\infty$ and (△) $t=0$ points of the kinetic refolding traces; (▼) and (▽) $t=\infty$ and $t=0$ points of the kinetic unfolding traces. Kinetic data were normalized to the equilibrium data by using either the signal of the unfolded protein in 6 M urea (a) or in 4 M GdnHCl (b) or the signal of the unfolded protein in 7.5 M urea (c). The continuous lines through the urea-induced equilibrium unfolding data in a and c are drawn according to Eq. (1), using the values for ΔG_{NI}, ΔG_{NU}, m^E_{NI} and m^E_{NU} given in the legend to Fig. 1, and with $Z_I=1$ for the data in a, and with $Z_I=0$ for the data in c. (a and b) The continuous lines through the $t=0$ points of the kinetic unfolding traces are fits to Eq. (8) and yielded values for ΔG_{NI} and m^E_{NI} of 2.1 kcal mol^{-1} and –0.48 kcal mol^{-1} M^{-1}, respectively, for the urea data in a, and 3.7 kcal mol^{-1} and –1.53 kcal mol^{-1} M^{-1}, respectively, for the GdnHCl data in b. The continuous line through the GdnHCl-induced equilibrium unfolding data (b) is a non-linear, least-squares fit to Eq. (1), with the values of ΔG_{NI} and m^E_{NI} fixed to the values obtained above. The values obtained for ΔG_{NU} and m^E_{NU} were 4.5 kcal mol^{-1} and –3.34 kcal mol^{-1} M^{-1}, respectively (Table 1). In all panels, the broken lines represent the extrapolated native and unfolded protein baselines, and the error bars on the kinetic data points represent the spread in the data for two different sets of experiments.

Table 1. Comparison of thermodynamic and kinetic parameters governing the folding and unfolding of the SH3 domain of PI3 kinase, in GdnHCl and in urea

|	GdnHCl	Urea
ΔG_{NI} (H_2O) (kcal mol^{-1})	3.7 ($\pm$ 0.5)	2.1 ($\pm$ 3.7)
m^E_{NI} (kcal mol^{-1} M^{-1})	-1.53 ($\pm$ 0.16)	-0.48 ($\pm$ 0.6)
ΔG_{IU} (H_2O) (kcal mol^{-1})	0.8 ($\pm$ 0.64)	2.7($\pm$ 4.4)
m^E_{IU} (kcal mol^{-1} M^{-1})	-1.81 ($\pm$ 0.34)	-0.64 ($\pm$ 0.78)
ΔG_{NU} (H_2O) (kcal mol^{-1})	4.5 ($\pm$ 0.4)	4.8 ($\pm$ 2.4)
m^E_{NU} (kcal mol^{-1}M^{-1})	-3.34 ($\pm$ 0.3)	-1.12 ($\pm$ 0.5)
$\alpha_I = m^E_{UI}/m^E_{UN} = -m^E_{IU}/-m^E_{NU}$	0.54 ($\pm$ 0.1)	0.57 ($\pm$ 0.74)
λ_f (H_2O) (s^{-1})	0.34	0.3
m^k_{UI} (kcal mol^{-1} M^{-1})	1.52 ($\pm$ 0.2)	0.64 ($\pm$ 0.04)
λ_U (H_2O) (s^{-1})	0.00017	0.00009
m^k_{IU} (kcal mol^{-1} M^{-1})	-0.08 ($\pm$ 0.05)	-0.16 ($\pm$ 0.01)
$(m^k_{IU} - m^k_{UI})$ (kcal mol^{-1} M^{-1})	-1.6 ($\pm$ 0.2)	-0.80 ($\pm$ 0.04)
$-RT \ln (\lambda_U/\lambda_f)$ (kcal mol^{-1})	4.5	4.8
$\beta_T = m^k_{UI}/m^E_{UN}$	0.45 ($\pm$ 0.07)	0.57 ($\pm$ 0.25)

The errors shown represent errors from fitting data averaged from two sets of experiments with each denaturant. The relatively large errors in the parameters describing urea unfolding have probably arisen because the $t=0$ points of the kinetic unfolding traces, could not be determined over the entire range of urea concentrations over which the N $\leftrightarrow$ I transition occurs (Fig. 3). The values of ΔG_{NI}, ΔG_{IU}, m_{NI} and m_{IU} were determined from a three-state N $\leftrightarrow$ I $\leftrightarrow$ U analysis of the equilibrium unfolding experiments and from a two-state N $\leftrightarrow$ I analysis of the $t=0$ points of the kinetic unfolding traces (Fig. 3). It should be noted that these values of ΔG need to be corrected for the free energy change associated with the equilibrium that is established between unfolded molecules having the prolyl peptide bond in a *cis* configuration and unfolded molecules having the prolyl peptide bond in a *trans* configuration, but the correction needed is small (0.2 kcal mol^{-1}).[40] α_I is a measure of the position of I along the reaction coordinate relative to that of N. $-RT$ ln (λ_U/λ_f) yields the value of ΔG_{NU}, in native-like conditions where I is not populated. β_T is a measure of the position of the transition state separating U and I, along the reaction coordinate relative to that of N $m^E_{UN} = -(m^E_{NI+} m^E_{IU})$.

Three-state analysis of unfolding

Figures 1–3 show that an unfolding intermediate can be detected under equilibrium conditions, and in kinetic studies an unfolding intermediate can be seen to form very rapidly during unfolding. In both cases, the intermediate appears to retain N-like far-UV CD but to have lost N-like fluorescence. While it is possible that two different intermediates are detected in the equilibrium and kinetic studies, the simplest assumption is that the intermediates are one and the same, in either denaturant. This assumption is consistent with using the simplest and minimal model (N $\leftrightarrow$ I $\leftrightarrow$ U) that can adequately describe all the kinetic and thermodynamic data (see Discussion). Hence, the three-state analysis (Eq. (4)) of the fluorescence and far-UV CD-monitored equilibrium unfolding transitions, which both describe the equilibrium between N and I and U, was carried out using the values of ΔG_{NI} (H_2O) and m^E_{NI} obtained from an analysis of the denaturant dependence of the burst phase change in fluorescence seen in kinetic unfolding experiments (see above). When this was done for both urea-and GdnHCl-induced unfolding: (1) similar values were obtained for the stability of N relative to U, in either denaturant; and (2) similar values were obtained in both denaturants for α_I, a measure of the position of I on the reaction

coordinate relative to U and N, based on the degree of solvent-accessible surface area (Table 1). Nevertheless, the true test for the assumption that the same minimal N $\leftrightarrow$ I $\leftrightarrow$ U mechanism is adequate to describe both the kinetic and equilibrium data, is whether the thermodynamic parameters, which were determined from the analysis of the denaturant dependence of the equilibrium amplitudes as well as the kinetic burst phase amplitudes, are predicted by and are consistent with the rate constant data.

The dependence of the observed refolding and unfolding rate constants on denaturant concentration, shown in Fig. 4, were therefore analyzed on the basis of the N $\leftrightarrow$ I $\leftrightarrow$ U mechanism. The observation that no burst phase change in fluorescence occurs during folding in any concentration of denaturant suggested that I does not accumulate, and that the U to I transition is the rate-limiting step of folding in the range of denaturant concentrations studied. This inference is supported by the observation that the folding arms of the chevron are linear in both denaturants (Fig. 4). The observation that a burst phase change in fluorescence occurs during unfolding in either denaturant, suggested that the rate constants for the transition from N to I (k_{NI}) and for the transition from I to N (k_{IN}), are much faster than the rate constant for the transition from I to U (k_{IU}), at all concentrations of denaturant (see Data Analysis). The dependence on denaturant concentration of the observed folding and unfolding rate constants were fit to the sum of Eqs. (10) and (12). In using Eqs. (10) and (12) to obtain the values of $\Delta G^{\ddagger}_{IU}$, m^k_{IU}, $\Delta G^{\ddagger}_{UI}$ and m^k_{UI}, the values fixed for ΔG_{NI} and m^E_{NI} were
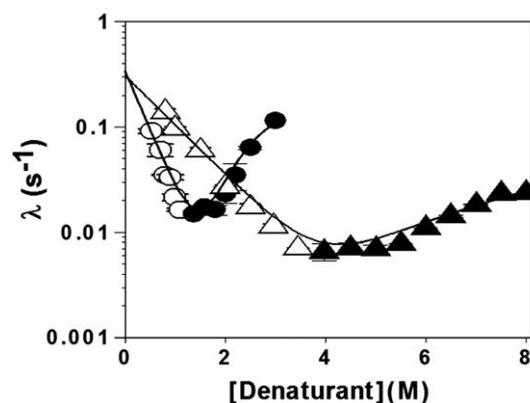
Fig. 4. Dependence of the observed unfolding and refolding rate constants on the concentration of denaturant. The observed refolding ($\bigcirc$, $\triangle$) and unfolding ($\bullet$, $\blacktriangle$) rate constants were measured in different concentrations of GdnHCl (circles) and urea (triangles) by monitoring the change in fluorescence at 300 nm. The continuous lines through the rate constants are fits to the sum of Eqs. (10) and (12) (see Data analysis). For fitting, the value of the pre-exponential factor A in Eqs. (10) and (12) was fixed at 10^6 s^{-1}, and the values for ΔG_{NI} and m^E_{NI} were fixed to the values obtained from the two-state analysis of the $t=0$ points of the kinetic unfolding traces (see Table 1). The values obtained for the various parameters are given in Table 1. The error bars represent the spread in the data for two different sets of experiments.

those obtained from the two-state N ↔ I analysis of the denaturant dependence of the $t=0$ points of the kinetic unfolding traces, as described above. The values obtained for the various parameters are given in Table 1.

Refolding and unfolding of the SH3 domain in urea *versus* GdnHCl

Table 1 compares the values of the various thermodynamic and kinetic parameters determined when refolding and unfolding was carried out in GdnHCl, to those obtained in urea. The values of ΔG_{NU}, obtained from the equilibrium as well as the kinetic studies, were similar for both the denatur-

ants. A denaturant-dependent burst phase during unfolding was observed for both denaturants. The value obtained for α_I, which is a measure of the extent of folding in I compared to that in N, is similar (~ 0.55) for folding in both the denaturants. Hence, it appears that the same intermediate, I is populated during folding or unfolding in urea and in GdnHCl. For folding in either denaturant, the value of the Tanford β value (β_T) for the rate-limiting transition state, TS1, is about the same (~ 0.5), implying that about 50% of the change in solvent-accessible surface area during folding has occurred in TS1.

The product of the burst phase of unfolding is a partially unfolded species

To determine whether the product of the burst phase during unfolding is a partially unfolded form, or whether a fraction of the native molecules has unfolded completely, interrupted unfolding experiments were carried out. Figure 5a shows representative kinetic refolding traces of protein that had been unfolded in 3 M GdnHCl for different lengths of time from 275 ms to 50 s, and then refolded by diluting the GdnHCl to 1 M. The $t=0$ points of the kinetic refolding traces do not fall on the direct unfolding trace obtained in 3 M GdnHCl, indicating that there is a burst phase change in fluorescence in addition to the observed phase of fluorescence change during refolding (Fig. 5b). The amplitude of the burst phase decreases with an increase in the time of unfolding in 3 M GdnHCl. The observed phase is described well by a single exponential, with an observed rate constant of 0.028 ± 0.003 s^{-1}. This is similar to the observed rate constant of refolding
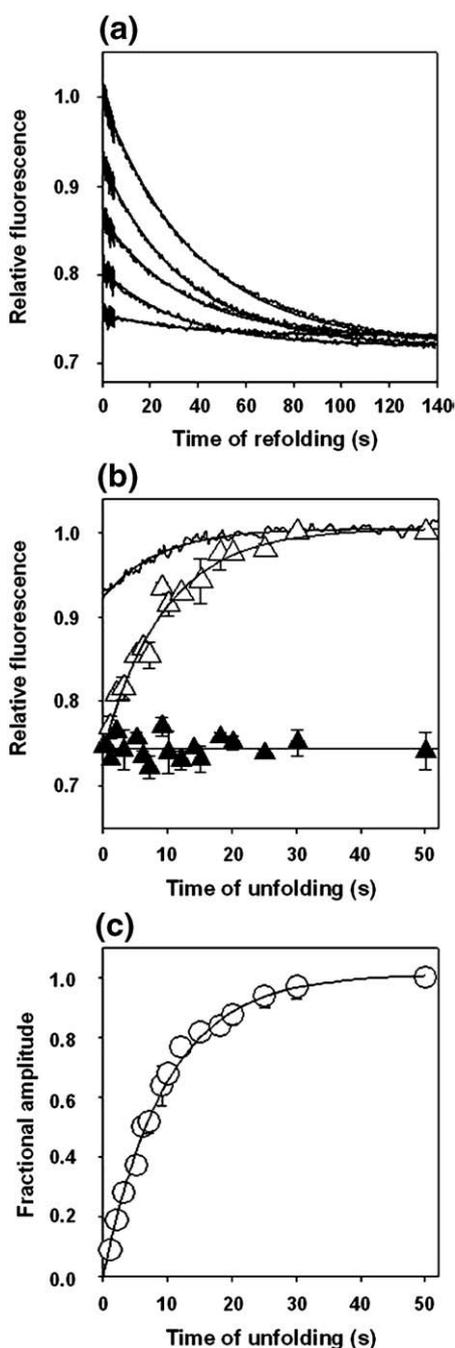
Fig. 5. Kinetics of formation of the U form from the N state. Native protein was unfolded in 3 M GdnHCl at pH 7.2 for variable lengths of time from 276 ms to 50 s, before refolding was initiated by diluting the concentration of denaturant to 1 M. (a) Representative refolding traces when the protein was refolded after unfolding for 50 s, 12 s, 6.2 s, 3.2 s and 1.2 s (top to bottom). Each refolding trace was fit to a single-exponential equation, which yielded an observed rate constant for refolding, of 0.028 ± 0.003 s^{-1}. All the kinetic traces were normalized to the $t=0$ point of the refolding trace obtained after unfolding for 50 s. (b) A comparison of the signal at the $t=0$ points of the kinetic refolding traces obtained after unfolding in 3 M GdnHCl for variable lengths of time, with the single-jump unfolding trace obtained when protein was unfolded in 3 M GdnHCl. ($\triangle$) The $t=0$ points; and (▲) the $t=\infty$ points of the kinetic refolding traces. The continuous line through the open triangles is a fit to a single-exponential equation, which yields an observed rate constant of 0.1 s^{-1}. The single-jump unfolding curve was also fit to a single-exponential equation, which yields an observed rate constant for unfolding, of 0.1 s^{-1}. (c) The increase in the relative amplitude of the observed phase with an increase in the time of unfolding before refolding was initiated. The amplitudes of all refolding traces were normalized to the amplitude of the refolding curve obtained when the protein was unfolded for 50 s before refolding. The fit to a single-exponential equation, shown as a continuous line, yielded an observed rate constant of 0.1 s^{-1}.

in 1 M GdnHCl of completely unfolded protein. Figure 5c shows that the relative amplitude of the observed refolding phase increases with an increase in the time of prior unfolding in 3 M GdnHCl. The amplitude of the observable phase increases with an observed rate constant of $0.1 \, s^{-1}$, which is similar to the directly observed rate constant of unfolding in 3 M GdnHCl. All molecules are found to be completely unfolded when unfolding is interrupted at 50 s: the resulting refolding trace does not display a burst phase, and the observable phase has an apparent folding rate constant similar to that seen in single-jump refolding in 1 M GdnHCl. This indicates that the observable phase of refolding after interrupted unfolding, corresponds to the refolding of completely unfolded protein, U, whose population increases with an increase in the time of unfolding. The burst phase of refolding after interrupted unfolding, represents the refolding of a partially unfolded form, whose population decreases to zero with increased time of unfolding.

Discussion

The SH3 domain of PI3 kinase is not a two-state folder

In an earlier study[40] of the folding of the SH3 domain of PI3 kinase, the folding reaction appeared to be two-state even when studied by multiple probes, including NMR. That study appeared therefore to preclude any significant accumulation of intermediates before the rate-limiting step of folding, but it did not preclude the formation of folding intermediates after the rate-limiting step. In the same study, no unfolding intermediate was detected, but GdnHCl was the only denaturant used in the equilibrium and kinetic studies, and it was not used at concentrations greater than 3 M. Moreover, multiple probes were not used to study unfolding, as they were used for studying refolding. The rate constant data appeared to fulfill the kinetic criteria for two-state folding: (1) the extrapolated values of the observed rate constants of folding and unfolding (λ_f and λ_u) in water predicted the stability (ΔG_{NU}) of the protein; (2) the dependence of λ_f and λ_u on GdnHCl concentration predicted the dependence on GdnHCl concentration of the free energy of unfolding (ΔG_{NU}); and (3) the concentration of GdnHCl at which λ_f and λ_u had identical values, was also the concentration corresponding to the midpoint of the equilibrium unfolding transition.

In this study, the observed rate constants of folding and unfolding in GdnHCl were identical with those reported earlier,[40] and hence, the values for λ_f and λ_u in GdnHCl, as well as in urea reported here also predict apparent two-state folding. But several observations indicate that the folding and unfolding reactions of the SH3 domain are not two-state but multi-state: (1) the equilibrium urea-induced unfolding transitions, monitored by fluorescence and far-UV CD, are not coincident. This indicates the population of an unfolding intermediate that exists in equilibrium with native and unfolded protein at intermediate urea concentrations; (2) in kinetic studies, a burst phase of fluorescence change is observed during unfolding in either urea or GdnHCl, at high concentrations of denaturant; and (3) a burst phase change of fluorescence is observed during the refolding of transiently unfolded protein. Hence, it is necessary to use, minimally, a three-state N ↔ I ↔ U mechanism to analyze the equilibrium and kinetic data.

Identity of the equilibrium and kinetic unfolding intermediates

In this study, an unfolding intermediate has been identified under equilibrium unfolding conditions, as well as in kinetic unfolding studies. It is shown that a three-state N ↔ I ↔ U model is the simplest and minimal model that can completely describe all aspects of both the kinetic and equilibrium data in an internal consistent manner, and that it is not necessary to include any additional intermediate for analysis of all the data. Thermodynamic parameters, such as ΔG and m, describing the model have been determined using the amplitudes of signal change seen during equilibrium unfolding in conjunction with the amplitudes of signal change seen in kinetic experiments, and analyzing the dependence of these amplitudes on denaturant concentration according to the three-state model. The same thermodynamic parameters are predicted, within experimental error, by application of the same three-state model to analyze the rate constant data from kinetic experiments, in either denaturant (Table 1). Hence, the assumption that the kinetic intermediate is identical with that observed under equilibrium conditions, which has also been made in the case of the folding of other proteins, appears to be valid in this case.[48,49]

The SH3 domain of PI3 kinase has seven tyrosine residues dispersed throughout its sequence (see Materials and Methods). Hence, the observation that the fluorescence of I is the same as that of U, indicates that I is solvated throughout its structure, and it is not the case that only a local segment of structure has become solvent-accessible in I. The latter possibility is in fact discounted by the analysis, which indicates, from the value obtained for the parameter α_I, that only about 55% of the solvent-accessible surface area has become buried in I (Table 1). It should be noted here that, although the SH3 domain does have a single tryptophan residue in its sequence, this residue is on the surface of the protein. If the intrinsic tryptophan fluorescence is used to probe for unfolding, a very small apparently non-cooperative transition is observed.

Nature of the burst phase product

Burst phase changes during protein unfolding have been reported for several proteins. The burst phase change can arise due to the accumulation of partially unfolded conformations.[12,41] In a few cases,

however, like that of canine milk lysozyme, the burst phase change arises due to a fraction of the native protein molecules unfolding very fast, when there is heterogeneity in the native state.[42] The burst phase change in fluorescence observed during both the urea-and GdnHCl-induced unfolding of the SH3 domain of PI3 kinase, has been interpreted to signify the very fast formation of an intermediate I and not the formation of U. This is because a two-state analysis of the burst phase amplitude indicates that the burst phase represents the unfolding of N to a form whose solvent-accessible surface area is only about 45% of that of U (see Results and Table 1). It should be noted that because I is formed at least 100 times faster than U over the entire range of denaturant concentrations studied, it was possible to consider a pre-equilibrium between N and I, for analysis of the kinetic data (see Data Analysis).

Heterogeneity in the native state has been reported for the α-spectrin SH3 domain.[50] Partial unfolding manifested as EX1 exchange has been observed for many SH3 domains when studied by hydrogen/deuterium exchange coupled to mass spectrometry.[38] In the case of the SH3 domain of PI3 kinase, it seems very unlikely that the burst phase change in fluorescence arises due to conversion of the native state to a native-like state, because there is a significant change in solvent-accessible surface area in going from the native state to the burst phase product (Table 1). The intermediate I is therefore not native-like in structure, even though its secondary structure, as indicated by far-UV CD at 222 nm appears to be the same as that of N (Fig. 3).

In interrupted unfolding experiments, different species populated during unfolding can be distinguished by their rates of refolding. At 276 ms of unfolding when mostly the burst phase product is populated, and no significant fraction of molecules is completely unfolded (the time constant of unfolding is 10 s), a refolding trace with the maximum burst phase amplitude and no significant amplitude for the observable phase is observed. This implies that the species populated at the end of the burst phase refolds much faster (within the 11 ms dead time) than does the completely unfolded protein. This is direct evidence that the product of the burst phase during the unfolding of the SH3 domain is a partially unfolded conformation.

The initial accumulation of the intermediate I during unfolding, which results in a burst phase change in fluorescence, is also seen to lead to a downward curvature in the unfolding arm of the chevron plot, for either denaturant (Fig. 4), as predicted by Eq. (9). Such curvatures in the unfolding branches of chevron plots have been seen also in unfolding studies of other proteins, where they have been attributed to the presence of high-energy intermediates.[11,13–15]

Mechanism of unfolding

The observation of three states N, I and U and two steps (burst phase and observable phase) during the unfolding of the SH3 domain can be explained by the minimal mechanism N ↔ I → U (Fig. 6). This mechanism accounts for the equilibrium unfolding data (Fig. 1) and for the kinetic unfolding data. During unfolding, a pre-equilibrium is established between N and I in the burst phase, and this equilibrium shifts towards I with an increase in denaturant concentration. From single-jump unfolding experiments, it appears that the fluorescence of I is similar to that of U. Thus, the change in fluorescence takes place when N unfolds to I, while the I to U transition is silent to fluorescence. The observable change in fluorescence during unfolding arises only because the slow I → U transition perturbs the N ↔ I pre-equilibrium such that more N molecules are pulled towards I. In 3.5 M GdnHCl, the equilibrium constant (K_{NI}) for the N ↔ I pre-equilibrium is high, and all the N molecules unfold to I during the burst phase itself. Hence, no observable phase is observed. In interrupted unfolding experiments, the amplitude of the observable phase of the refolding trace appears to increase at the expense of the burst phase amplitude, which suggests that I unfolds directly to U.

The alternative three-state mechanism is I ↔ N → U, in which I is an off-pathway intermediate. According to this mechanism, a pre-equilibrium is established between N and I in the burst phase, and
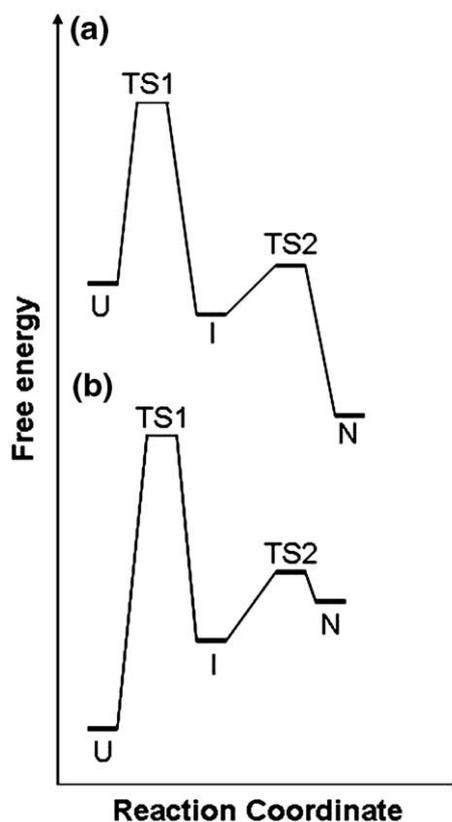
Fig. 6. Formation of an intermediate during the refolding and unfolding of the SH3 domain of PI3 kinase. (a) the free energy profile for folding in the absence of GdnHCl. A partially folded species is formed after the rate-limiting transition state during refolding. (b) The free energy profile for unfolding in 4 M GdnHCl. The partially unfolded intermediate accumulates during unfolding.

the protein molecules in the native state unfold directly to U in the observable phase. In the three-state analysis of the urea-induced equilibrium unfolding transitions (Fig. 1) according to an A ↔ B → C mechanism, the *m* values indicate that B has a solvent-accessible surface area intermediate between those of A and C. Hence, it is unlikely that B corresponds to N. Nevertheless, it is impossible to conclude only from equilibrium unfolding experiments, whether an intermediate is on-pathway or off-pathway.

The SH3 domain of PI3 kinase contains three proline residues that are in the *trans* conformation in the native protein.[39] It has been suggested that about 30% of unfolded protein molecules have Pro69 in the *cis* conformation. It should be noted, however, that the observed unfolding of the SH3 domain in two kinetic phases cannot be explained by a $N \leftrightarrow U_T \leftrightarrow U_C$ mechanism, in which U_T and U_C are two completely unfolded forms with the Phe68-Pro69 peptide bond in the native *trans* conformation and non-native *cis* conformation, respectively.[40] In such a mechanism, N would unfold to U_T very rapidly during the burst phase, and the $U_T \rightarrow U_C$ transition is a prolyl peptide isomerization reaction that would give rise to the observable phase. Such a mechanism is unlikely because: (1) the observed unfolding rate constants have a steep dependence on the denaturant concentration, which is not expected for a prolyl peptide bond isomerization reaction; (2) the observed rate constants are also faster than expected for a prolyl peptide bond isomerization reaction; and (3) all the change in solvent-accessible surface area would take place from N to U_T, but only about 45% of the surface area change occurs from N to the burst phase product.

Characterization and position of the rate-limiting transition state of folding

The Tanford β value (β_T) is a measure of the change in solvent-accessible surface area from U to the rate-limiting transition state, relative to the change in solvent-accessible surface area when U folds completely to N. β_T values have been used as a measure of the position of the transition state along the reaction coordinate. The β_T values of the SH3 domain of PI3 kinase have been determined to be 0.45 and 0.57 (Table 1) for folding in GdnHCl and urea, respectively. This indicates that the transition states in urea and GdnHCl possess somewhat similar solvent-accessible surface areas, and hence, very similar structures. It would appear therefore that unlike in the case of barstar,[15] the folding pathway of the SH3 domain in urea is the same as that in GdnHCl. It is possible, however, that the protein uses different pathways to fold in GdnHCl and urea, and that the solvent-accessible surface areas of the transition states on these pathways happen to be the same. In this context, it should be noted that although the solvent-exposure of residues in I is the same in both denaturants, the stabilities determined for I appear to be dissimilar (Table 1).

It is possible to reveal the range of folding and unfolding pathways available to a particular protein by studying different homologous members of the family,[32] by making different variants of the same protein,[30] or by changing solvent conditions using denaturants.[51] The transition states of parallel unfolding pathways have been distinguished by their β_T values.[51] In the case of SH3 domains, molecular dynamic simulations have indicated parallel pathways,[34] and the different transition states observed in different homologues or circular permutants may represent alternative folding pathways available to a protein.[52] The spread in β_T values seen for different SH3 domain proteins might, in fact, be due to different folding pathways preferred by different members of the family.[23] In future work, it will be important to determine whether the folding and unfolding pathways in the two denaturants are indeed the same.

Validating the population of an intermediate after the rate-limiting step of folding in GdnHCl

The values of β_T of the rate-limiting transition state TS1 and of the corresponding property, α_I, of the intermediate I represent the relative positions of TS1 and I along the reaction coordinate connecting U and N. When GdnHCl is the denaturant, the analysis of the thermodynamic and kinetic data according to the three-state $N \leftrightarrow I \leftrightarrow U$ model, clearly positions I after the rate-limiting transition state, TS1, of folding (Table 1). When urea is the denaturant, the values of β_T and α_I appear to be the same, but there are large experimental errors in their determination. It is therefore not possible at present, in the case of folding in urea, to determine whether I is a high-energy intermediate immediately preceding TS1, or whether it is populated after TS1.

In the case of folding in GdnHCl, it was important to confirm, in an independent manner, the result of the three-state analysis showing that I is populated after the rate-limiting transition state TS1. This was achieved by carrying out the interrupted unfolding experiments, which showed that I can fold directly to N (Fig. 6), and that the rate of refolding of I to N is much faster than that of U to N. This result validates the placement of the rate-limiting transition state TS1 between U and I, for folding in GdnHCl. If instead I were a high-energy intermediate preceding TS1 during refolding, then the refolding kinetics observed in the interrupted unfolding experiment would have been the same as in the direct refolding experiment because I would have equilibrated rapidly with U in refolding conditions.

Although the highest barrier exists between U and I for folding in GdnHCl, the actual change in fluorescence takes place only when I refolds to N. Since I is populated after the rate-limiting transition state, it remains undetected in single-jump refolding studies, even when multiple probes are used.[40] It is interesting to note that the formation of such an intermediate after the rate-limiting transition state during the folding of a SH3 domain was suggested

earlier by computational studies.[35,53] Those studies indicated that desolvation of the hydrophobic core occurs only as the intermediate refolds to the native state, in agreement with the observations made in this study. The computational studies indicated that secondary structure is formed upon crossing the main free energy barrier, and that tertiary contacts are formed only upon subsequent crossing of the smaller energy barrier. These results are in good agreement with the observation made in this study that the intermediate formed after the rate-limiting step possesses native-like secondary structure but has not attained native-like tertiary structure.

Intermediates populated after the rate-limiting transition state have been identified for several proteins by the use of native-state hydrogen exchange measurements.[3] The partially unfolded conformations detected by native-state HX usually remain undetected when folding is monitored by optical probes. An exception is the case of the bacteriophage T4 lysozyme, where the hidden intermediate manifests itself as an unfolding intermediate.[18] In the case of the SH3 domain, native-state HX in conjunction with mass spectrometry or NMR have revealed partially unfolded conformations of domains from many different proteins,[38] but optical methods have failed to detect any intermediate during refolding. Here, optical methods have shown that a partially unfolded conformation accumulates during the unfolding of the SH3 domain of PI3 kinase. The same intermediate is shown to form only after the rate-limiting transition state during refolding in GdnHCl. HX experiments in progress show clearly the presence of the unfolding intermediate, and the observed HX-monitored rate of unfolding of I to U at a low concentration of GdnHCl matches that predicted by the extrapolation of the fluorescence-monitored unfolding rates (our unpublished results). The HX experiments in conjunction with mass spectrometry are being used to obtain structural information about the intermediate, and to investigate the important question of whether the unfolding pathway of the protein is the same in low and high concentrations of denaturant.

Materials and methods

Buffers and reagents

All buffers and reagents were of ultra-pure grade, and 20 mM sodium phosphate (from Sigma) was used as the buffer at pH 7.2. GdnHCl and urea were from USB Corporation.

Cloning and purification of the SH3 domain of PI3 kinase

The gene encoding the SH3 domain of PI3 kinase, cloned in the pUC57 vector between the BamHI and EcoRI restriction sites, was obtained from GenScript Corporation. It was sub-cloned into the pET22b(+) expression

vector (from Novagen). The construct was confirmed by DNA sequencing.

Escherichia coli BL21(DE3) cells were transformed with the pET22b(+) construct containing the gene encoding the SH3 domain. A colony was grown in 100 ml of LB medium until the absorbance reached a value of 0.7, at 600 nm. A 25 mL sample of this culture was transferred to 500 mL of terrific broth, and grown for 3 h. Protein expression was induced by adding isopropylthiogalcto-pyranoside (IPTG) to a final concentration of 40 μM, and the culture was grown for another 9 h. Cells were pelleted by centrifugation at 6000 rpm for 20 min (1 L of culture was processed at a time). The pellet was suspended in 15 mL of lysis buffer (16 mM Na_2HPO_4, 4 mM NaH_2PO_4 150 mM NaCl, 1% (v/v) Triton X-100, 0.2 mM EDTA, 0.2 mM paramethyl sulfonyl fluoride, and one tablet of Protease Inhibitor (from Roche)). After sonication for 15 min, the lysate was centrifuged for 15 min at 17,000 rpm. The supernatant was loaded onto a Sephadex G-50 column equilibrated with buffer A (50 mM trIs, 20 mM NaCl, 0.1% (v/v) β-mercaptoethanol, 0.2 mM EDTA, pH 8). Fractions (10 mL) were collected, checked by SDS-PAGE, and fractions containing the SH3 domain were pooled. The protein was then loaded on to a DEAE Sepharose anion-exchange column, and eluted with a 0 – 0.5 M NaCl gradient using an AKTA Basic 10 HPLC. The purity of the protein was checked by SDS-PAGE, and electrospray ionization mass spectrometry (ESI-MS). The protein was> 98% pure. The mass obtained (9363.8 Da) was that expected for the SH3 domain, whose sequence is:

> SAEGYQYRALYDYKKEREEDIDLHLGDILTVNKGSL-VALGFSDGQEAKPEEIGWLNGYNETTGERGDFPGTY-VEYIGRKKISP

About 10–15% of the SH3 domain molecules had a mass of 9276.8 Da, indicating that the N-terminal serine residue had been cleaved off. The typical yield was 100 mg of pure SH3 domain per liter of growth culture.

The protein was also expressed as a GST-fusion protein using the pGEX-4T-2 plasmid. Thrombin cleavage of this fusion protein yielded a homogeneous SH3 domain, but with three extra residues preceding Ser1 in the SH3 domain sequence. This protein was identical with that used in an earlier study.[40] The folding and unfolding kinetics (rates as well as amplitudes) of this protein in GdnHCl as well as in urea were found to be identical in all respects with those reported here, including the occurrence of a burst phase change during fluorescence-monitored unfolding in high concentrations of denaturant. Hence, no aspect of the kinetics reported here can be attributed to the 10–15% heterogeneity of the protein used in this study. It should be noted that the pGEX-4T-2 expression system yielded 10-to 20-fold less protein than that obtained using the pET22b(+) expression system; hence the latter was used for expressing protein.

Equilibrium unfolding experiments

All refolding and unfolding experiments were carried out using the SFM-4 mixing module and its optical system (MOS-200) from Biologic, unless otherwise mentioned. Equilibrium unfolding was done at pH 7.2, by incubating the protein (~10 μM) in different concentrations of GdnHCl for 4 h at 25 °C. The unfolding was monitored by the change in fluorescence at 300 nm upon excitation at 268 nm. A 1 cm pathlength cuvette was used. The excitation slit-width was 1 mm, and a band-pass filter of 300

±10 nm was used for measuring the fluorescence emission. Equilibrium unfolding in urea was carried out by incubating the protein (~30 μM) in different concentrations of urea for about 8 h at 25 °C. Unfolding was monitored either by measuring the change in fluorescence at 300 nm using the MOS-450 module from Biologic, or by measuring the change in ellipticity at 222 nm using a Jasco-J720 spectropolarimeter. A 0.2 cm pathlength cuvette was used for the far-UV circular dichroism (CD) measurements and the excitation band width was set to 1 nm.

Kinetic unfolding experiments

The kinetic unfolding experiments were done by diluting the native protein at pH 7.2, to different concentrations of GdnHCl or urea (final protein concentration of 15–20 μM). Unfolding was monitored by the change in fluorescence at 300 nm using a band-pass filter of 300±10 nm. The excitation wavelength was 268 nm with a slit-width of 1 mm, and the mixing dead time was 10.8 ms.

In addition to the stopped-flow measurements of the kinetics of unfolding, manual mixing measurements of the kinetics of unfolding were done for all concentrations of denaturant in which the observed rate of unfolding was< 0.027 s^{-1}. In these manual mixing experiments, fluorescence changes during unfolding were measured using the MOS-450 optical system (Biologic).

Kinetic unfolding experiments that monitored the change in ellipticity at 222 nm were done with the Jasco-J720 spectropolarimeter. Unfolding was achieved by diluting the native protein at pH 7.2 in different concentrations of urea (final protein concentration 30 μM). A 0.2 cm pathlength cuvette was used, and the excitation band width was 1 nm.

Kinetic refolding experiments

The protein was unfolded in unfolding buffer containing either 6 M urea or 4 M GdnHCl at pH 7.2. Refolding was initiated by diluting the denaturant to different concentrations below the midpoint of the equilibrium unfolding transition, using an SFM-4 stopped-flow machine with a dead time of 10.8 ms. Refolding was monitored by the change in fluorescence at 300 nm, using a band-pass filter of 300±10 nm. The excitation wavelength was 268 nm, and the excitation slit width was 1 nm. The final protein concentration was 15–20 μM.

Interrupted unfolding experiments

Interrupted unfolding experiments were done with an SFM-400 machine (Biologic). Native protein (125 μL of at pH 7.2) was mixed with 125 μL of unfolding buffer containing 6 M GdnHCl so that the concentration of GdnHCl was 3 M. The mixture was incubated in a 190 μL loop (220 μL, inter-mixer volume) for different lengths of time from 100 ms to 50 s. Refolding was initiated by mixing 120 μL of the incubated mixture with 240 μL of refolding buffer so that the final concentration GdnHCl was 1 M. The mixing dead time was 19 ms, and the final protein concentration was 10–15 μM. The refolding kinetics was monitored by the change in fluorescence at 300 nm, using a band-pass filter of 300±5 nm. The excitation wavelength was 268 nm, and the excitation band width was 5 nm.

Analysis of equilibrium and pre-equilibrium unfolding data

The far-UV CD and fluorescence-monitored equilibrium unfolding transitions were analyzed together using a three-state N ↔ I ↔ U model.[49,54] For such an unfolding mechanism, the observed optical signal, y_{obs}, is given by:

$$y_{obs} = \frac{y_N + (y_N + (y_U - y_N)Z_I)e^{\frac{-\Delta G_{NI}}{RT}} + y_U e^{\frac{-\Delta G_{NU}}{RT}}}{1 + e^{\frac{-\Delta G_{NI}}{RT}} + e^{\frac{-\Delta G_{NU}}{RT}}} \quad (1)$$

where y_N and y_U are the optical signals of the native and unfolded proteins, respectively. ΔG_{NI} and ΔG_{NU} (=ΔG_{NI}+ ΔG_{IU}) are the free energies of unfolding of N to I, and to U, respectively, and ΔG_{IU} is the free energy of unfolding of I to U. Z_I is a measure of the optical signal of I, y_I, relative to y_N and y_U, and is given by:

$$Z_I = \frac{y_I - y_N}{y_U - y_N} \quad (2)$$

The fractional change in signal, f_{app}, at a particular concentration of denaturant, [D], is determined from y_{obs}, by:

$$f_{app} = \frac{y_{obs} - (y_N + m_N[D])}{(y_U + m_U[D]) - (y_N + m_N[D])} \quad (3)$$

In Eq. (3), it is assumed that the optical signals characteristic of N and U have linear dependence on [D], given by the slopes m_N and m_U, respectively.

For the three-state model, f_{app} is given by:

$$f_{app} = \frac{Z_I * e^{\frac{-\Delta G_{NI}}{RT}} + e^{\frac{-\Delta G_{NU}}{RT}}}{1 + e^{\frac{-\Delta G_{NI}}{RT}} + e^{\frac{-\Delta G_{NU}}{RT}}} \quad (4)$$

ΔG_{NI}, ΔG_{IU} and ΔG_{NU} are assumed to have linear dependence on [D], given by the slopes m^E_{NI}, m^E_{IU} and m^E_{NU}, respectively:

$$\Delta G_{NI} = \Delta G_{NI}(H_2O) + m^E_{NI}[D] \quad (5)$$

$$\Delta G_{IU} = \Delta G_{IU}(H_2O) + m^E_{IU}[D] \quad (6)$$

$$\Delta G_{NU} = \Delta G_{NU}(H_2O) + m^E_{NU}[D] \quad (7)$$

m^E_{NI}, m^E_{IU} and m^E_{NU} (=m^E_{NI}+ m^E_{IU}) represent the preferential free energies of interaction with denaturant of N in comparison to that of I, of I in comparison to that of U, and of N in comparison to that of U, respectively.

The curves traced by the $t = 0$ points of the kinetic unfolding traces were fit to a two-state N ↔ I model by:

$$y_{obs} = \frac{y_N + y_I e^{\frac{-\Delta G_{NI}}{RT}}}{1 + e^{\frac{-\Delta G_{NI}}{RT}}} \quad (8)$$

Analysis of kinetic unfolding and folding data

All the kinetic unfolding traces were fit to single-exponential equations. The kinetic refolding traces were fit to either a single-exponential or a double-exponential equation depending upon the denaturant concentration.

The three-state model used for analysis of the kinetic data was the same as that used for the equilibrium data:

$$N \underset{k_{IN}}{\overset{k_{NI}}{\rightleftharpoons}} I \underset{k_{UI}}{\overset{k_{IU}}{\rightleftharpoons}} U$$
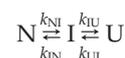
The analysis becomes simple when k_{NI}, $k_{IN} > k_{IU}$ for all concentrations of denaturant. Then during unfolding, N transforms rapidly to I so that a fast pre-equilibrium is established between N and I, defined by the equilibrium constant:

$$K_{NI} = k_{NI}/k_{IN}$$

then I unfolds to U in the observed unfolding phase. In high concentrations of denaturant, the observed rate constant of formation of U is given by:

$$\lambda_u = \frac{K_{NI}}{1 + K_{NI}} \cdot k_{IU} \qquad (9a)$$

At low concentrations of denaturant, when $k_{NI} \ll k_{IN}$, K_{NI} will be $\ll 1$, and λ_u (which represents an effective two-state, N to U unfolding rate constant under conditions when I is not populated, but which is not an elementary rate constant) is given by:

$$\lambda_u = K_{NI} \cdot k_{IU} \qquad (9b)$$

Equation (9) can be used to express λ_u in terms of $\Delta G^{\ddagger}_{IU}$, the free energy of activation of the I → U transition and ΔG_{NI}. A is the pre-exponential factor, which is a constant:

$$\lambda_u = \frac{e^{\frac{-\Delta G_{NI}}{RT}}}{1 + e^{\frac{-\Delta G_{NI}}{RT}}} \cdot A.e^{\frac{-\Delta G^{\ddagger}_{IU}}{RT}} \qquad (10)$$

Since $K_{NU} = \lambda_u / \lambda_f$ and $K_{NU} = K_{NI} \cdot K_{IU} = K_{NI} \cdot k_{IU}/k_{UI}$, and since $\lambda_u = K_{NI} \cdot k_{IU}$ (Eq. (9b)), the overall observed rate constant of formation of N is given by:

$$\lambda_f = k_{UI} \qquad (11)$$

Equation (11) is applicable when the U to I transition is the rate-limiting step during folding.

When λ_f is expressed in terms of $\Delta G^{\ddagger}_{UI}$, the free energy of activation of the U → I transition, Eq. (11) becomes:

$$\lambda_f = A.e^{\frac{-\Delta G^{\ddagger}_{UI}}{RT}} \qquad (12)$$

where A is the pre-exponential factor, which is a constant.

It should be noted that the three-state mechanism can be solved analytically to obtain expressions for the fast (λ_{fast}) and slow (λ_{slow}) observed rate constants in terms of k_{NI}, k_{IN}, k_{IU}, and When $\lambda_{fast} \gg \lambda_{slow}$, and k_{NI}, $k_{IN} \gg k_{IU}$, k_{UI}, λ_{slow} is found to be equal to the sum of λ_u and λ_f, as defined by Eqs. (9) and (11).[55]

In using Eqs. (10) and (12) to analyze the denaturant dependence of λ_u and λ_f, $\Delta G^{\ddagger}_{IU}$ and $\Delta G^{\ddagger}_{UI}$, like ΔG_{NI}, are assumed to have linear dependence on [D], which are given by the slopes m^k_{IU} and m^k_{UI}, respectively:

$$\Delta G^{\ddagger}_{IU} = \Delta G^{\ddagger}_{IU}(H_2O) + m^k_{IU}[D] \qquad (13)$$

$$\Delta G^{\ddagger}_{UI} = \Delta G^{\ddagger}_{UI}(H_2O) + m^k_{UI}[D] \qquad (14)$$

where m^k_{IU} and m^k_{UI} represent the preferential free energies of interaction with denaturant, of TS1 in comparison to that of I, and to that of U, respectively. TS1 is the transition state preceding I during folding.

Acknowledgements

References

1. Jackson, S. E. & Fersht, A. R. (1991). Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry*, **30**, 10428–10435.
2. Fersht, A. R. (1995). Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications. *Proc. Natl Acad. Sci. USA*, **92**, 10869–11087.
3. Bai, Y. (2003). Hidden intermediates and the Levinthal paradox in the folding of small proteins. *Biochem. Biophys. Res. Commun.* **305**, 785–788.
4. Korzhnev, D. M., Salvatella, X., Vendruscolo, M., Di Nardo, A. A., Davidson, A. R., Dobson, C. M. & Kay, L. E. (2004). Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature*, **430**, 586–590.
5. Khorasanizadeh, S., Peters, I. D. & Roder, H. (1996). Evidence for a three-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues. *Nat. Struct. Biol.* **3**, 193–205.
6. Houliston, R. S., Liu, C., Singh, L. M. & Meiering, E. M. (2002). pH and urea dependence of amide hydrogen-deuterium exchange rates in the beta-trefoil protein hisactophilin. *Biochemistry*, **41**, 1182–1194.
7. Gorski, S. A., Capaldi, A. P., Kleanthous, C. & Radford, S. E. (2001). Acidic conditions stabilise intermediates populated during the folding of Im7 and Im9. *J. Mol. Biol.* **312**, 849–863.
8. Chiti, F., Taddei, N., Webster, P., Hamada, D., Fiaschi, T., Ramponi, G. & Dobson, C. M. (1999). Acceleration of the folding of acylphosphatase by stabilization of local secondary structure. *Nat. Struct. Biol.* **6**, 380–387.
9. Wani, A. H. & Udgaonkar, J. B. (2006). HX-ESI-MS and optical studies of the unfolding of thioredoxin indicate stabilization of a partially unfolded, aggregation-competent intermediate at low pH. *Biochemistry*, **45**, 11226–11238.
10. Borgia, A., Bonivento, D., Travaglini-Allocatelli, C., Di Matteo, A. & Brunori, M. (2006). Unveiling a hidden folding intermediate in c-type cytochromes by protein engineering. *J. Biol. Chem.* **281**, 9331–9336.
11. Jonsson, T., Waldburger, C. D. & Sauer, R. T. (1996). Nonlinear free energy relationships in Arc repressor unfolding imply the existence of unstable, native-like folding intermediates. *Biochemistry*, **35**, 4795–4802.
12. Zaidi, F. N., Nath, U. & Udgaonkar, J. B. (1997). Multiple intermediates and transition states during protein unfolding. *Nat. Struct. Biol.* **4**, 1016–1024.
13. Juneja, J. & Udgaonkar, J. B. (2002). Characterization of the unfolding of ribonuclease a by a pulsed hydrogen exchange study: evidence for competing pathways for unfolding. *Biochemistry*, **41**, 2641–2654.
14. Sanchez, I. E. & Kiefhaber, T. (2003). Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. *J. Mol. Biol.* **325**, 367–376.
15. Sridevi, K. & Udgaonkar, J. B. (2002). Unfolding rates

of barstar determined in native and low denaturant conditions indicate the presence of intermediates. *Biochemistry*, **41**, 1568–1578.

16. Udgaonkar, J. B. (2008). Multiple routes and structural heterogeneity in protein folding. *Annu. Rev. Biophys.* **37**, 489–510.

17. Bai, Y., Sosnick, T. R., Mayne, L. & Englander, S. W. (1995). Protein folding intermediates: native-state hydrogen exchange. *Science*, **269**, 192–197.

18. Cellitti, J., Bernstein, R. & Marqusee, S. (2007). Exploring subdomain cooperativity in T4 lysozyme II: uncovering the C-terminal subdomain as a hidden intermediate in the kinetic folding pathway. *Protein Sci.* **16**, 852–862.

19. Rami, B. R. & Udgaonkar, J. B. (2001). pH-jump-induced folding and unfolding studies of barstar: evidence for multiple folding and unfolding pathways. *Biochemistry*, **40**, 15267–15279.

20. Roy, M. & Jennings, P. A. (2003). Real-time NMR kinetic studies provide global and residue-specific information on the non-cooperative unfolding of the beta-trefoil protein, interleukin-1beta. *J. Mol. Biol.* **328**, 693–703.

21. Patra, A. K. & Udgaonkar, J. B. (2007). Characterization of the folding and unfolding reactions of single-chain monellin: evidence for multiple intermediates and competing pathways. *Biochemistry*, **46**, 11727–11743.

22. Kumar, S., Mohanty, S. K. & Udgaonkar, J. B. (2007). Mechanism of formation of amyloid protofibrils of barstar from soluble oligomers: evidence for multiple steps and lateral association coupled to conformational conversion. *J. Mol. Biol.* **367**, 1186–1204.

23. Jackson, S. E. (1998). How do small single-domain proteins fold? *Fold. Des.* **3**, R81–R91.

24. Kuriyan, J. & Cowburn, D. (1997). Modular peptide recognition domains in eukaryotic signaling. *Annu. Rev. Biophys. Biomol. Struct.* **26**, 259–288.

25. Cohen, G. B., Ren, R. & Baltimore, D. (1995). Modular binding domains in signal transduction proteins. *Cell*, **80**, 237–248.

26. Mayer, B. J. (2001). SH3 domains: complexity in moderation. *J. Cell. Sci.* **114**, 1253–1263.

27. Musacchio, A. (2002). How SH3 domains recognize proline. *Adv. Protein Chem.* **61**, 211–268.

28. Koch, C. A., Anderson, D., Moran, M. F., Ellis, C. & Pawson, T. (1991). SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. *Science*, **252**, 668–674.

29. Northey, J. G., Di Nardo, A. A. & Davidson, A. R. (2002). Hydrophobic core packing in the SH3 domain folding transition state. *Nat. Struct. Biol.* **9**, 126–130.

30. Viguera, A. R., Serrano, L. & Wilmanns, M. (1996). Different folding transition states may result in the same native structure. *Nat. Struct. Biol.* **3**, 874–880.

31. Lindorff-Larsen, K., Vendruscolo, M., Paci, E. & Dobson, C. M. (2004). Transition states for protein folding have native topologies despite high structural variability. *Nat. Struct. Mol. Biol.* **11**, 443–449.

32. Zarrine-Afsar, A., Larson, S. M. & Davidson, A. R. (2005). The family feud: do proteins with similar structures fold via the same pathway? *Curr. Opin. Struct. Biol.* **15**, 42–49.

33. Korzhnev, D. M., Neudecker, P., Zarrine-Afsar, A., Davidson, A. R. & Kay, L. E. (2006). Abp1p and Fyn SH3 domains fold through similar low-populated intermediate states. *Biochemistry*, **45**, 10175–10183.

34. Borreguero, J. M., Ding, F., Buldyrev, S. V., Stanley, H. E. & Dokholyan, N. V. (2004). Multiple folding pathways of the SH3 domain. *Biophys. J.* **87**, 521–533.

35. Cheung, M. S., Garcia, A. E. & Onuchic, J. N. (2002). Protein folding mediated by solvation: water expulsion and formation of the hydrophobic core occur after the structural collapse. *Proc. Natl Acad. Sci. USA*, **99**, 685–690.

36. Shea, J. E., Onuchic, J. N. & Brooks, C. L., 3rd (2002). Probing the folding free energy landscape of the Src-SH3 protein domain. *Proc. Natl Acad. Sci. USA*, **99**, 16064–16068.

37. Mittermaier, A. & Kay, L. E. (2006). New tools provide new insights in NMR studies of protein dynamics. *Science*, **312**, 224–228.

38. Wales, T. E. & Engen, J. R. (2006). Partial unfolding of diverse SH3 domains on a wide timescale. *J. Mol. Biol.* **357**, 1592–1604.

39. Booker, G. W., Gout, I., Downing, A. K., Driscoll, P. C., Boyd, J., Waterfield, M. D. & Campbell, I. D. (1993). Solution structure and ligand-binding site of the SH3 domain of the p85 alpha subunit of phosphatidylinositol 3-kinase. *Cell*, **73**, 813–822.

40. Guijarro, J. I., Morton, C. J., Plaxco, K. W., Campbell, I. D. & Dobson, C. M. (1998). Folding kinetics of the SH3 domain of PI3 kinase by real-time NMR combined with optical spectroscopy. *J. Mol. Biol.* **276**, 657–667.

41. Otzen, D. E. & Oliveberg, M. (2002). Burst-phase expansion of native protein prior to global unfolding in SDS. *J. Mol. Biol.* **315**, 1231–1240.

42. Nakatani, H., Maki, K., Saeki, K., Aizawa, T., Demura, M., Kawano, K. *et al.* (2007). Equilibrium and kinetics of the folding and unfolding of canine milk lysozyme. *Biochemistry*, **46**, 5238–5251.

43. Galani, D., Fersht, A. K. & Perrett, S. (2002). Folding of the yeast prion protein Ure2: kinetic evidence for folding and unfolding intermediates. *J. Mol. Biol.* **315**, 213–227.

44. Zhang, H. J., Sheng, X. R., Niu, W. D., Pan, X. M. & Zhou, J. M. (1998). Evidence for at least two native forms of rabbit muscle adenylate kinase in equilibrium in aqueous solution. *J. Biol. Chem.* **237**, 7448–7456.

45. Agashe, V. R., Schmid, F. X. & Udgaonkar, J. B. (1997). Thermodynamics of the complex protein unfolding reaction of barstar. *Biochemistry*, **36**, 12288–12295.

46. Bhuyan, A. K. & Udgaonkar, J. B. (2001). Folding of horse cytochrome c in reduced state. *J. Mol. Biol.* **312**, 1135–1160.

47. Pradeep, L. & Udgaonkar, J. B. (2007). Diffusional barrier in the unfolding of a small protein. *J. Mol. Biol.* **366**, 1016–1028.

48. Park, S. H., O'Neil, K. T. & Roder, H. (1997). An early intermediate in the folding reaction of the B1 domain of protein G contains a native-like core. *Biochemistry*, **36**, 14277–14283.

49. Saeki, K., Arai, M., Yoda, T., Nakao, M. & Kuwajima, K. (2004). Localized nature of the transition-state structure in goat α-lactalbumin folding. *J. Mol. Biol.* **341**, 589–604.

50. Sadqi, M., Casares, S., Abril, M. A., Lopez-Mayorga, O., Conejero-Lara, F. & Freire, E. (1999). The native state conformational ensemble of the SH3 domain from alpha-spectrin. *Biochemistry*, **38**, 8899–8906.

51. Wright, C. F., Steward, A. & Clarke, J. (2004). Thermodynamic characterisation of two transition states along parallel protein folding pathways. *J. Mol. Biol.* **338**, 445–451.

52. Lam, A. R., Borreguero, J. M., Ding, F., Dokholyan, N. V., Buldyrev, S. V., Stanley, H. E. & Shakhnovich, E.

(2007). Parallel folding pathways in the SH3 domain protein. *J. Mol. Biol.* **373**, 1348–1360.

53. Guo, W., Lampoudi, S. & Shea, J. E. (2003). Posttransition state desolvation of the hydrophobic core of the src-SH3 protein domain. *Biophys. J.* **85**, 61–69.

54. Nath, U. & Udgaonkar, J. B. (1995). Perturbation of a tertiary hydrogen bond in barstar by mutagenesis of the sole His residue to Gln leads to accumulation of at least one equilibrium folding intermediate. *Biochemistry*, **34**, 1702–1713.

55. Hagerman, P. L. & Baldwin, R. L. (1976). A quantitative treatment of the kinetics of the folding transition of ribonuclease A. *Biochemistry*, **15**, 1462–1473.