A Dry Transition State More Compact Than the Native State Is Stabilized by Non-Native Interactions during the Unfolding of a Small Protein

Sreemantee Sen, Rama Reddy Goluguri, and Jayant B. Udgaonkar*

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

Supporting Information

ABSTRACT: Defining the role of non-native interactions in directing the course of protein folding or unfolding reactions has been a difficult challenge. In particular, the extent to which such interactions play a productive role by stabilizing the structures of transition states (TSs) found on the folding and unfolding pathways of proteins is not known. On the contrary, it is thought that the TSs are expanded forms of the N state stabilized by native interactions, and it is not known whether non-native interactions can modulate TS structure. In this study of the unfolding of the SH3 domain of PI3 kinase using a microsecond mixing methodology, partial non-native structure formation is shown to occur initially during unfolding. The TS of this partial “folding during unfolding” reaction is more compact than the N state; the apparent rate constant of Trp53 burial during this reaction decreases with an increase in denaturant concentration. Kinetic studies of the unfolding of mutant variants suggest that the unusually compact TS is stabilized by interactions not present in N and that these non-native interactions are hydrophobic in nature. It was determined that mutation could be used to tune the degree of compaction in the TS.

Understanding the origin of the barriers that slow folding and unfolding on the free energy landscape of a protein is a major goal in protein folding studies. In particular, it is important to determine the nature of the interactions that determine the stability of the TS that the first free energy barrier encountered by a protein during unfolding, because it is this barrier that plays a pivotal role in maintaining the integrity of the structure of the protein. It has been suggested that this barrier arises because of the loss of tight packing interactions.1,2 Although TSs possessing this characteristic property of a DMG have not yet been identified, unfolding intermediates with DMG-like properties have been studied.2-5

The TSs of folding of many proteins have been characterized indirectly by the use of $\phi$-value analysis,6-9 and such studies have revealed that the dominant TS on a folding pathway is invariably an expanded form of the N state. $\phi$-Value analysis suggests that the TS has N-like interactions that are partially formed,10 and it appears to resemble a WMG in structure.10 However, the interpretation of $\phi$-values in terms of partially formed N-like structure has been questioned.11 Moreover, the observation of noncanonical $\psi$-values7,8,12,13 for some proteins and noncanonical $\psi$-values in the case of Protein L14 suggests that TSs might be stabilized also by non-native interactions. The nature of non-native interactions cannot, however, be probed readily by $\phi$-value analysis;13,15 consequently, they are largely ignored, especially because of the prevalence of native-centric views of protein folding.

Non-native interactions have also been identified in folding17-20 and unfolding21-25 intermediates, as well as in the U state.23-23 They may facilitate folding,26 possibly by restricting the conformational space that has to be sampled during folding. Intermediates with non-native structure may have function,27 but they are also likely to be prone to aggregation and to initiate aggregation.28 It is, however, not known whether the TSs that precede and follow an intermediate with non-native structure also possess the non-native structure. It is therefore important to determine how non-native interactions can stabilize and modulate TS structure13,15,29 and play a productive role in protein folding or unfolding.

SH3 domains constitute a family of proteins that have been used widely as model proteins for protein folding studies. Not surprisingly then, several SH3 domains have been shown to fold or unfold via intermediates with non-native structure.20,22-24,30 SH3 domains are predominantly $\beta$-sheet proteins, but an intermediate with $\alpha$-helical structure has been shown to form transiently on the folding pathways of the Src,30 Fym,20 and PI3K20 SH3 domains. In the case of the PI3K SH3 domain, a WMG-like intermediate, $I_U$, has been shown to form early, within a few milliseconds, before the rate-limiting step of unfolding.31 $I_U$ has been shown to possess non-native structure, in which Trp53 that is solvent-exposed in N has become partially buried.22 The PI3K SH3 domain is therefore a good model system for determining whether non-native interactions are present in a TS preceding a non-native intermediate and how they might play a role in modulating the structure and stability of the TS.

To obtain structural information about $I_U$ as well as about TS1 (the TS defining the N to $I_U$ transition), and in particular about the residues that might be involved in non-native interactions with Trp53, the effects of three mutations on the structure as well as on the rate of formation of $I_U$ were studied. In each of these three mutant variants, a polar residue in the

Received: April 28, 2017
Revised: June 17, 2017
Published: July 6, 2017

DOI: 10.1021/acs.biochem.7b00388
Biochemistry 2017, 56, 3699–3703
nSrc loop (Figure 1a), which had been implicated by a MD simulation to interact with Trp53 in IU, was replaced by a more hydrophobic residue. As expected, the N31M, K32I, and S34V mutations did not affect the fluorescence spectrum of N (Figure 1b), indicating that Trp53 is as fully hydrated in the mutant variants as it is in the wild-type (wt) protein. The three mutations also did not affect the apparent rate constant of unfolding of U, measured in the millisecond to second time domain, or its dependence on GdnHCl concentration (Figure 1d). The mutations did, however, affect the extent to which the fluorescence of Trp53 increased in the submillisecond time domain, before the observable decrease in fluorescence on the millisecond to second time scale (Figure 2 and Figure S1). The initial hyperfluorescence of Trp53 is a signature of the formation of IU,\(^22\) and Figure 2 shows that the intensity of the fluorescence spectrum of IU decreased upon mutation of the protein, and that the extent of the decrease depended on the site of the mutation. The spectra of the wt and three mutant variants all showed the same 5 nm blue shift with respect to the fluorescence spectrum of N, indicating that none of the three mutations had affected the solvent accessibility of Trp53 in IU. It is possible that the quantum yield of Trp53 in IU was affected by the mutations, but it is more likely that each of the mutations destabilized IU and that the decrease in the fluorescence intensity of Trp53 upon mutation was because IU was populated to a lesser extent.

The Tyr fluorescence of IU is known to be the same as that of U,\(^11\) for the wt protein. Figure S2 shows that this is also true for the three mutant variants. At high (>4 M) GdnHCl concentrations, the entire increase in Tyr fluorescence upon complete unfolding occurred in the submillisecond time domain. Because IU forms on this time scale, this result indicated that the Tyr fluorescence of IU is same as that of U. At GdnHCl concentrations in the range of 2–3 M, IU was not fully
Communication

populated, and analysis (see the text of the Supporting Information) of the dependence of the Tyr fluorescence of IU on denaturant concentration yielded the value of the free energy of unfolding of N to IU (ΔG_{NI}) for all four protein variants (Table S1), as well as the dependencies of ΔG_{NI} on GdnHCl concentration (m_{NI}). It is seen that the extent to which a mutation decreased the fluorescence intensity of Trp53 in IU (Figure 2, inset) correlated well with the extent to which the mutations destabilized IU (Table S1). This correlation indicated that the mutations affected the fluorescence spectrum of IU because the mutations affected the extent to which IU was transiently populated.

To monitor the formation of IU, kinetic unfolding experiments in the submillisecond time domain were performed (Figure 3 and Figure S3), using a home-built, continuous flow microsecond mixer. For the wt protein as well as for the mutant variants, it was seen that the kinetic traces, obtained from continuous flow mixing, captured the entire increase in fluorescence signal of the native protein to IU. The kinetic traces before and after the break in the x-axis were acquired using continuous flow and stopped flow mixing, respectively. The kinetic traces from continuous flow mixing were found to merge with the kinetic traces obtained using stopped flow mixing. Each trace was normalized to a value of 1 for the fluorescence signal of the native protein. The dashed and dotted lines represent the fluorescence signals of the native protein and of unfolded protein in 5.5 M GdnHCl, respectively. The solid lines through the data fits to a double-exponential equation.

Figure 3. Kinetic traces of unfolding of the PI3K SH3 domain. GdnHCl-induced unfolding, using continuous flow and stopped flow mixing, was monitored at 320 nm upon excitation at 295 nm for the (a) wt, (b) N31M, (c) K32I, and (d) S34V PI3K SH3 domains. The data were obtained by diluting the native protein to 3 (green), 3.6 (blue), 4 (cyan), and 4.5 (red) M GdnHCl. The kinetic traces before and after the break in the x-axis were acquired using continuous flow and stopped flow mixing, respectively. The kinetic traces from continuous flow mixing were found to merge with the kinetic traces obtained using stopped flow mixing. Each trace was normalized to a value of 1 for the fluorescence signal of the native protein. The dashed and dotted lines represent the fluorescence signals of the native protein and of unfolded protein in 5.5 M GdnHCl, respectively. The solid lines through the data fits to a double-exponential equation.

For the wt protein as well as for the mutant variants, it was seen that the kinetic traces, obtained from continuous flow mixing, captured the entire increase in Trp53 fluorescence, which occurred in the burst phase of stopped flow mixing (Figure 3). In all cases, the kinetic traces in the submillisecond time domain could be described by a single-exponential equation, and they merged into the kinetic traces in the millisecond to second time domain. The logarithm of the apparent rate constant of unfolding in the submillisecond time domain was found to decrease linearly with an increase in GdnHCl concentration (Figure 4a). This was surprising because unfolding rate constants are expected to increase with an increase in the concentration of the denaturant. This observation indicated that the increase in the fluorescence of Trp53, which occurred in the submillisecond time domain, originated from a transient “folding-like” reaction occurring during unfolding. The observation that the apparent rate constant of this reaction decreased with an increase in the concentration of GdnHCl indicated that the SASA decreased during the transition from N to the TS (TS1) preceding IU (Figure 4a). TS1 must therefore be more compact than N, as the change in the SASA would correlate to a change in R_g (see below).

TSs on the free energy landscapes of protein folding are invariably expanded and solvated relative to the N state. A ϕ-value analysis done on more than 800 mutant variants of 24 proteins, with simple kinetic behavior, showed a conserved ϕ-value of 0.24, which signifies that TSs on the protein folding landscape are expanded forms of N and are relatively poorly structured compared to N. TS1 is therefore very unusual. It is more compact than N and must therefore be (like N) a dry state that precedes the WMG, IU, which is more expanded and solvated than N. The folding-like behavior could arise due to the Trp burial process being an event that leads to a gain in structure, as Trp53 is solvent-exposed in N.

In this study, the effects of three mutations on the stability and unfolding kinetics were studied. None of these mutations, N31M, K32I, or S34V, affected the stability of N. Hence, residues N31, K32, and S34 are unlikely to be involved in any stabilizing interactions in N. Nevertheless, two of these mutations, K32I and S34V, resulted in the transition from N to IU being accelerated (Figure 4a and Table S1). Hence, these two mutations must stabilize TS1. This result suggested that residues 32 and 34 are engaged in interactions in TS1 that are not present in N. It is these non-native interactions that stabilize TS1. The observation here that a TS encountered during unfolding is stabilized by non-native interactions is unique; until now, only TSs encountered on folding pathways have been shown to be stabilized by non-native interactions.

The observation that a TS on the unfolding pathway of the PI3K SH3 domain is more compact than N, due to stabilizing, non-native interactions, is remarkable. These interactions in TS1 appear to be hydrophobic in nature because TS1 is stabilized by mutations that introduce more hydrophobic residues at positions 32 and 34. It is likely that the non-native cluster, in which I32 and V34 participate in TS1, is the same cluster that partially buries Trp53.

The observation that the K32I and S34V mutations resulted in a decrease in the logarithm of the apparent rate constants (λ_{NI}) of the N to IU transition, with an increase in GdnHCl concentration, steeper than that seen for the wt protein and for the N31M variant (Figure 4a) indicates that more surface area is buried in TS1 of the K32I and S34V variants than in that of the wt and N31M variants. The kinetic m values (m_{NI}) listed in Table S1, and the β_{NI} values calculated using these kinetic m values and m_{NI} indicated that the extent of the decrease in solvent-exposed surface area during the N to TS1 transition is only 3%–9% of the decrease in the solvent-exposed surface area that occurs during the U to N transition. The extents of compaction in TS1 with respect to N for the different protein variants, as measured by the fractional decrease in R_g during the
Communication

Figure 4. Stabilization of TS1. (a) Apparent rate constants of Trp burial. The dependencies of the observed rate constants of Trp burial on GdnHCl concentration are shown for wt (blue diamonds), N31M (magenta triangles), K32I (orange circles), and S34V (green squares) PI3K SH3 domains. The solid lines through the data are fits to the equation log φ = log φ_{NI} + m_{NI} [GdnHCl] (the values of 2.303RT ln φ_{NI} and φ_{NI} are listed in Table S1). The error bars, representing the spread in the data obtained from two independent experiments, indicate that the average error (∼5%) in measurement of the rate constants is significantly smaller than the differences in rate constants between different protein variants (except for the wt and N31M variants). (b) ΔΔG_{NT} plotted vs ΔΔG_{NI}. The slope of the straight line through the data is −1.5.

N to TS1 transition, using the known values for R_g of the N and U states,20 are listed in Table S1.

A previous MD simulation of denaturant-induced unfolding of the PI3K SH3 domain indicated that I_U is preceded by a DMG in which Trp53 is partially buried, as it is in I_U.32 This study shows that I_U is preceded by a TS (TS1) that is more compact than N and which is therefore likely to be dry, with no penetration of water into its structure. Nevertheless, TS1 possesses non-native structure, and it is likely that to accommodate this non-native structure, packing interactions would have to break in at least some parts of the structure of TS1. Hence, TS1 appears to be a DMG.

Thus, non-native interactions are present in both TS1 and I_U. In TS1, the non-native interactions involve residues 32 and 34 presumably forming a hydrophobic cluster in which Trp53 may participate. These non-native interactions stabilize TS1 (Figure 4a and Table S1). In I_U, Trp53 is known to become partially buried in a hydrophobic cluster,22,32 which may be the same as that found in TS1. Certainly, the earlier MD simulations had indicated that residues at positions 32 and 34 interact with Trp53 in I_U,32 but the mutations that stabilize TS1 destabilize I_U (Table S1). This anticorrelation (Figure 4b) could be the result of TS1 being DMG-like, while I_U is WMG-like,31 or it could be because the stabilizing hydrophobic cluster in TS1 is different from that in I_U. The overall average q-value, with I_U as a reference, of the region where the mutations were made (Figure 1a) is −1.5 (Figure 4b), which indicates that non-native structure associated with residues 32 and 34 is present in TS1, causing the first free energy barrier on the unfolding pathway of the protein to come down. It is interesting to note that the apparent rate constants of unfolding of I_U, which is the rate-limiting step in unfolding, are not affected by any of the three mutations (Figure 1d). This result suggests that the non-native structure accommodating the partial burial of Trp53 remains present in TS2, the dominant TS of unfolding.

MD simulations have suggested that non-native interactions play minimal roles during protein folding37 and that obligatory folding intermediates would have few if any such interactions.35 The demonstration, here, of the presence of obligatory intermediates and TSs with non-native interactions, on the unfolding pathway of the PI3K SH3 domain, reveals the importance of examining carefully unfolding transition paths by MD simulations.

SH3 domains have a conserved Trp residue (Trp53 in the PI3K SH3 domain) in their sequences. Previous studies have shown that Trp53 plays a critical role in the binding of the PI3K SH3 domain to other proteins and is therefore important for its function.38 Hence, it is necessary for Trp53 to be solvent-exposed in N for the function of the protein. Trp53 becomes buried early during unfolding and may remain buried until late during unfolding. Proteins can aggregate upon partial unfolding,39 and it is likely that the burial of the hydrophobic Trp53 in partially unfolded TSs and intermediates makes the protein less prone to aggregation.

The microsecond mixing experiments performed to monitor the formation of the intermediate, I_U, with non-natively buried Trp53, could characterize this step as a transient folding event happening on the unfolding pathway of the protein. Moreover, the TS encountered during the formation of this non-native intermediate was shown to possess stabilizing, non-native interactions, presumably hydrophobic in nature, with the penetration of water into the core unlikely to have occurred. These results are significant because a TS more compact than the N state, which is stabilized by non-native interactions, has not been seen before.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.7b00388.

Detailed experimental procedures, supplementary figures, and a supplementary table (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: jayant@ncbs.res.in. Telephone: 91-80-23666150. Fax: 91-80-23636662.

ORCID

Sreemantee Sen: 0000-0001-6713-815X
Jayant B. Udgaonkar: 0000-0002-7005-224X
Funding
J.B.U. is a recipient of a J. C. Bose National Research Fellowship from the Government of India. This work was funded by the Tata Institute of Fundamental Research and by the Department of Science and Technology, Government of India.

Notes
The authors declare no competing financial interest.

• ACKNOWLEDGMENTS

We thank M. K. Mathew, Raghavan Varadarajan, and members of our laboratory for discussions.

• ABBREVIATIONS

TS, transition state; N, native; DMG, dry molten globule; WMG, wet molten globule; U, unfolded; MD, molecular dynamics; SASA, solvent-accessible surface area; $R_g$, radius of gyration.

• REFERENCES