Molecular Cell
Short Article

Knotted Fusion Proteins Reveal Unexpected Possibilities in Protein Folding

Anna L. Mallam,[1] Shimobi C. Onuoha,[1] J. Günter Grossmann,[2] and Sophie E. Jackson[1,*]
[1]Chemistry Department, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK
[2]Molecular Biophysics Group, STFC Daresbury Laboratory, Warrington WA4 4AD, UK
*Correspondence: sej13@cam.ac.uk
DOI 10.1016/j.molcel.2008.03.019

SUMMARY

Proteins that contain a distinct knot in their native structure are impressive examples of biological self-organization. Although this topological complexity does not appear to cause a folding problem, the mechanisms by which such knotted proteins form are unknown. We found that the fusion of an additional protein domain to either the amino terminus, the carboxy terminus, or to both termini of two small knotted proteins did not affect their ability to knot. The multidomain constructs remained able to fold to structures previously thought unfeasible, some representing the deepest protein knots known. By examining the folding kinetics of these fusion proteins, we found evidence to suggest that knotting is not rate limiting during folding, but instead occurs in a denatured-like state. These studies offer experimental insights into when knot formation occurs in natural proteins and demonstrate that early folding events can lead to diverse and sometimes unexpected protein topologies.

INTRODUCTION

Haemophilus influenzae YibK and *Escherichia coli* YbeA are homodimeric α/β-knot methyltransferases (MTases) (Figures 1A–1C) (Lim et al., 2003; Mallam and Jackson, 2007a). Their backbone structure contains an unusual, deeply embedded trefoil knot that requires at least 40 amino acid residues to have threaded through a similarly sized loop. Tying equivalent knots on a human scale involves considerable skill and purpose, but a protein molecule possessing none of these attributes can manage a similar task with complete efficiency and reproducibility. Despite their intricacy, a significant number of "well-tied" protein knots have been discovered in recent years (Taylor, 2007; Taylor and Lin, 2003; Virnau et al., 2006; Yeates et al., 2007); as well as simple trefoils, a figure-of-eight knot (Taylor, 2000) and a knotted structure with five projected crossings (Virnau et al., 2006) have been observed. Both YibK and YbeA unfold spontaneously and reversibly upon addition of the chemical denaturant urea (Mallam and Jackson, 2005; 2007a). The folding of YibK is complicated, primarily due to its dimeric nature and the presence of multiple proline residues that cause heterogeneity in the unfolded state (Mallam and Jackson, 2006). Although YbeA appears to fold via a simpler pathway, similarities between the folding of YibK and YbeA imply that the mechanisms involved in knot formation in both proteins may be related (Mallam and Jackson, 2007a).

By constructing a set of proteins involving the fusion of the thermophilic protein *Archeoglobus fulgidus* ThiS to either YibK or YbeA at the amino terminus (ThiS-YibK and ThiS-YbeA), carboxy terminus (YibK-ThiS and YbeA-ThiS) or both termini (ThiS-YibK-ThiS and ThiS-YbeA-ThiS), we were able to probe the events that occur during the folding of a knotted structure. ThiS is a 91-residue monomeric protein that we used as a "molecular plug" to attempt to hinder the threading motions of the polypeptide chain or to prevent it from knotting altogether. Characterization of the resultant fusion proteins allowed us to conclude that the folding of α/β-knotted proteins likely propagates from a loosely knotted, denatured-like state, a result that has important implications for protein folding, simulation, and design.

RESULTS

Fusion Proteins Are Homodimeric and Knotted

The structural properties of the native-state fusion proteins were examined using small-angle X-ray scattering (SAXS), and both the radius of gyration (R_g) and the maximum particle dimension (D_{max}) were determined (see Table S1 available online). Particle shapes were generated ab initio from the experimental scattering data using the program GASBOR (Svergun et al., 2001), while rigid-body modeling of the individual domains to the experimental scattering data was performed using BUNCH (Petoukhov and Svergun, 2005) (Figure 1D; Figure S1). Since these are independent techniques, the good agreement between the structures obtained from both gives a high degree of confidence in a particular solution. The SAXS data are consistent with all fusion proteins existing as fully folded homodimeric entities, and the structures generated show the relative positions of the knotted and ThiS domains. Analytical gel filtration chromatography further confirmed the oligomeric state of the fusion proteins; all eluted as a single peak, indicating a homogenous dimer population at the concentrations of protein studied (Figure S2).

α/β-knot proteins such as YibK and YbeA bind the MTase cofactor S-adenosylmethionine (AdoMet) in a crevice formed by the knotted region of the protein (Ahn et al., 2003; Lim et al., 2003; Nureki et al., 2004). Functional studies have shown that affinity for AdoMet is highly sensitive to changes in the structure of this binding site, and often only a single mutation can cause

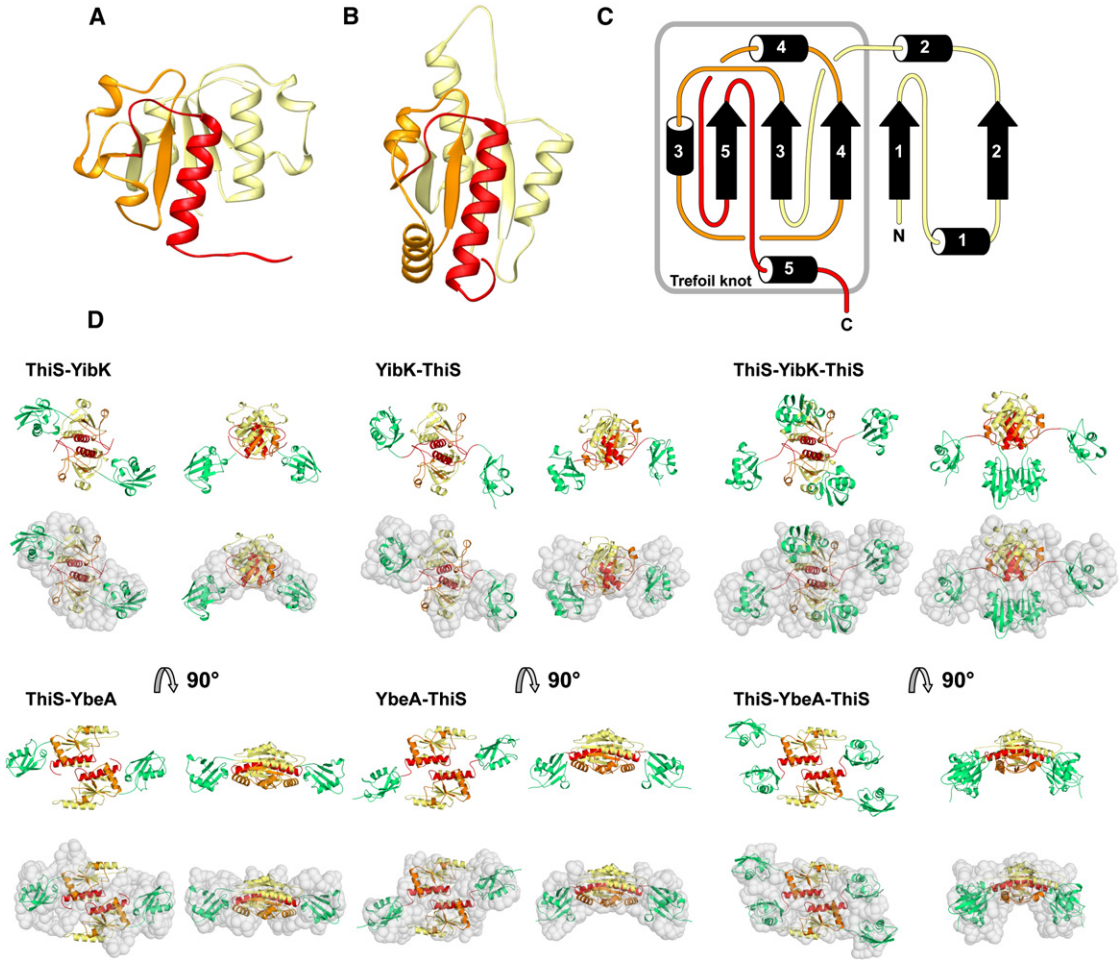


Figure 1. Structures of Knotted Proteins

(A and B) Monomeric subunits of *Haemophilus influenzae* YibK (160 residues) (A) and *Escherichia coli* YbeA (155 residues) (B).

(C) Topological diagram of an α/β-knot MTase. Numbers refer to secondary structure elements. The structures in (A)–(C) are colored to highlight knotting loop (orange) and the knotted chain (red).

(D) Knotted fusion proteins reconstructed from SAXS data. Results from ab initio modeling are represented as gray dummy residues and are superimposed onto structures obtained from rigid-body modeling. Knotted domains are colored according to (A) and (B), and ThiS domains are shown in green.

inactivity (Elkins et al., 2003; Mallam and Jackson, 2007b). We measured the affinity of S-adenosylhomocysteine (AdoHcy), the product of AdoMet after methyl-group transfer to the substrate has taken place, for the fusion proteins using isothermal titration calorimetry (ITC) to confirm the integrity of the cofactor-binding pocket and, therefore, the presence of the knotted structure (Figure 2); the binding of AdoHcy is more straightforward to examine than that of AdoMet, as the latter is unstable in vitro (Hoffman, 1986). With the exception of ThiS-YibK-ThiS, all fusion proteins displayed binding parameters comparable to those of the equivalent wild-type protein, demonstrating that the knotted region remains intact (Table S2). While binding is still observed, the affinity of AdoHcy for ThiS-YibK-ThiS appears somewhat reduced compared to wild-type YibK; a possible explanation is that the binding pocket is obstructed by the ThiS domains in some protein molecules, creating a population of inactive conformers. The dimeric nature of all the fusion proteins provides further evidence that the knotted structure is undisrupted, as this area additionally forms part of the dimer interface (Lim et al., 2003). Thus, regardless of our attempts to "plug" one or more termini, YibK and YbeA appear able to fold to their knotted, homodimeric native states.

Knotted Fusion Proteins Remain Stable and Fold Reversibly

We quantified the effect of additional ThiS domains on the stability of YibK and YbeA by using chemical denaturants to perturb the equilibrium between folded and unfolded states (Figure 3). Wild-type ThiS is a very stable protein and unfolds at much higher concentrations of guanidinium chloride (GdmCl) than either YibK or YbeA (Figure 3A). Denaturation profiles in GdmCl for the knotted fusion proteins were biphasic, consisting of transitions corresponding to the unfolding of the knotted and ThiS domains at low and high concentrations of GdmCl, respectively,

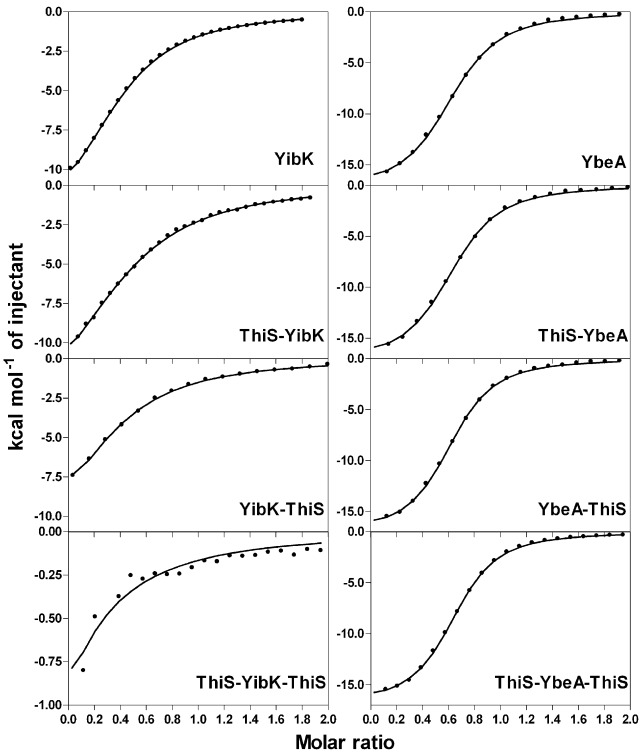


Figure 2. Affinity of AdoHcy for Wild-Type and Fusion-Knotted Proteins

The continuous line represents the fit of ITC data to a single-site binding model using the Origin software package (MicroCal Inc.). Data have been corrected for the heat of dilution.

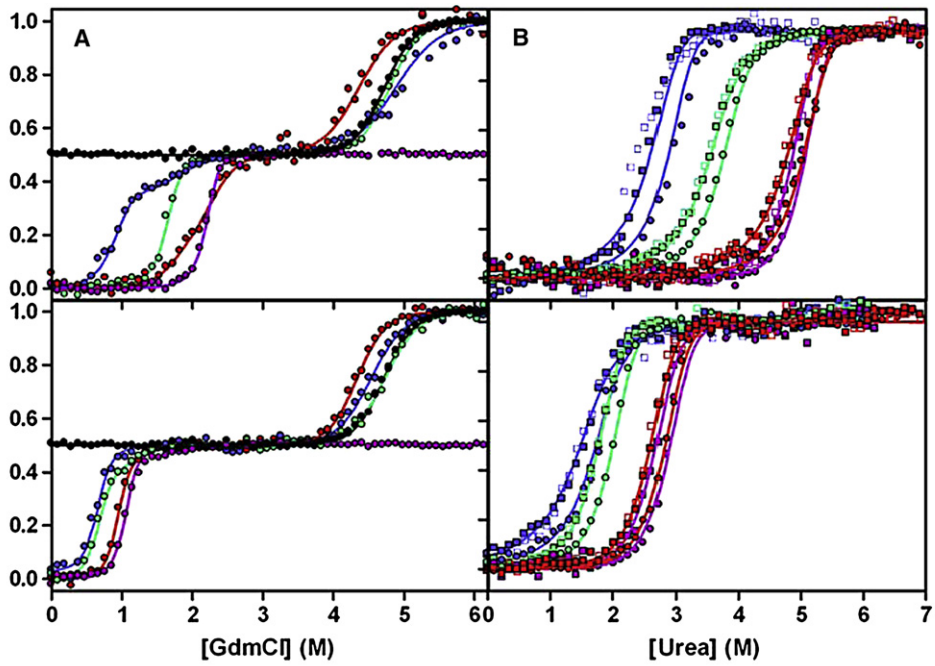
consistent with them behaving independently despite being linked. The separation of these transitions meant that it was possible to selectively unfold only the knotted protein domains. Far ultraviolet circular dichroism (far-UV CD) scans of the fusion proteins confirmed that the ThiS domain remained structured in 3.0 M GdmCl and in 6.0 M urea (Figure S3). Urea titrations were used to calculate an accurate thermodynamic stability of the YibK and YbeA domains in the fusion proteins (Figure 3B); the ThiS domains remained folded during these experiments, allowing the structure loss of the knotted domains to be monitored exclusively (Figure S3). The overlapping denaturation and renaturation urea profiles at equivalent concentrations of protein and the recovery of approximately 100% of native far-UV CD signal upon refolding (Figure S3) confirmed that the folding of the knotted domains was fully reversible in the presence of a folded ThiS domain fused to either or both termini. The free energies of unfolding ($\Delta G^{N_2 \leftrightarrow 2D}_{H_2O}$) of YibK and YbeA in the fusion proteins were calculated from the global fit of the data measured at different concentrations of protein to a three-state dimer denaturation model with a monomeric intermediate (Mallam and Jackson, 2005; 2007a) (Figure 3C). The knotted domains in YibK-ThiS and YbeA-ThiS had almost identical stabilities to the equivalent wild-type proteins. This result is significant, as it is the carboxy-terminal α-helix of YibK and YbeA that appears the most likely segment of chain to have undergone threading, and addition of a ThiS domain at this terminus might have been expected to prevent folding altogether. A pattern of stabilities was observed for the other fusion proteins, with any reduction in $\Delta G^{N_2 \leftrightarrow 2D}_{H_2O}$ mainly due to a destabilization of the dimer rather than the equilibrium monomeric intermediate. Knotted domains with a ThiS molecule fused to both termini were most destabilized relative to wild-type protein by 5–10 kcal mol^{-1}, while those with an amino-terminus ThiS were destabilized to a lesser extent by 4–5 kcal mol^{-1}. We suggest that much of this destabilization is caused by the unfavorable steric effect of a ThiS domain attached to the relatively solvent unexposed amino termini of YibK and YbeA.

Additional Protein Domains Do Not Alter the Folding Characteristics of a Knotted Protein

Wild-type YibK folds in urea with a complex kinetic mechanism involving four reversible folding phases: two different intermediates from parallel pathways fold via a third sequential monomeric intermediate to form native dimer in a slow rate-limiting dimerization reaction (Mallam and Jackson, 2006). YbeA exhibits two reversible urea-dependent kinetic phases with the dimer forming in the last step from a monomeric intermediate (Mallam and Jackson, 2007a). Kinetic experiments using urea to perturb the equilibrium were performed on the knotted fusion proteins; the fused ThiS domains remain folded during these measurements, as they are resistant to urea denaturation (Figure 3; Figure S3). Thus, the unfolding and refolding of the knotted domains can be examined while attached to one or more folded ThiS domains. There is a remarkable similarity between the kinetic folding data of the knotted fusion proteins and that of wild-type YibK and YbeA (Figure 4); the same number of reversible phases is observed, each with comparable kinetic *m* values that relate to the solvent-accessible surface area change upon folding (Myers et al., 1995) (Table S3). Crucially, fusing a ThiS domain to either or both termini of a knotted protein has surprisingly little effect on its folding rate; refolding rate constants in the absence of denaturant, other than those that correspond to dimerization, vary by less than 4-fold, equivalent to an energetic change of less than 1 kcal mol^{-1}.

DISCUSSION

That all our artificial multidomain constructs can knot and fold is an unexpected result; even the folding of ThiS-YibK-ThiS and ThiS-YbeA-ThiS can be described by kinetic parameters comparable to those of the wild-type proteins despite the fact that a considerably longer segment of chain has to be threaded through a loop. Indeed, these fusion proteins contain the most deeply embedded protein knots observed to date, as determined by the number of amino acids that can be removed from each side before the structure becomes unknotted (Kolesov et al., 2007), which is over 125. We estimate that a polypeptide "knotting loop" of at least 40 completely unstructured residues is necessary for a folded ThiS domain to thread through (see the Experimental Procedures). The fastest kinetic folding step of the wild-type and fusion proteins, indicated by the red and green phases from parallel pathways for YibK and the blue phase for YbeA (Figure 4), involves formation of a compact



C Analysis of fusion protein urea denaturation data.

Protein	Y_I	$\Delta G^{N_2 \leftrightarrow 2I}_{H_2O}$ (kcal mol^{-1})	$\Delta G^{I \leftrightarrow D}_{H_2O}$ (kcal mol^{-1})	$\Delta G^{N_2 \leftrightarrow 2D}_{H_2O}$ (kcal mol^{-1})
YibK	0.61 ± 0.1	18.9 ± 0.4	6.5 ± 0.2	31.9 ± 1.2
ThiS-YibK	0.49 ± 0.2	14.6 ± 0.1	5.6 ± 0.1	25.7 ± 0.1
YibK-ThiS	0.85 ± 0.2	17.4 ± 0.2	7.4 ± 0.1	32.2 ± 0.2
ThiS-YibK-ThiS	0.64 ± 0.1	13.7 ± 0.2	4.0 ± 0.1	21.7 ± 0.2
YbeA	0.32 ± 0.1	15.2 ± 0.1	2.5 ± 0.01	20.2 ± 0.1
ThiS-YbeA	0.74 ± 0.3	13.2 ± 0.2	1.5 ± 0.1	16.2 ± 0.2
YbeA-ThiS	0.65 ± 0.3	15.0 ± 0.1	2.4 ± 0.1	19.8 ± 0.1
ThiS-YbeA-ThiS	0.52 ± 0.1	10.5 ± 0.1	2.3 ± 0.1	15.1 ± 0.1

Data for each fusion protein at 1 μM and 5 μM protein were globally analyzed using the equation for a three-state dimer denaturation model involving a monomeric intermediate as described (Mallam and Jackson, 2005; 2007a). Y_I is the spectroscopic signal of the knotted protein monomeric intermediate relative to a signal of 0 for a native monomeric subunit in a dimer and 1 for a denatured monomer. $m_{N_2 \leftrightarrow 2I}$ and $m_{I \leftrightarrow D}$ values for fusion proteins, which relate to the solvent-accessible surface area change upon dissociation to the monomeric intermediate and upon unfolding of the monomeric intermediate, respectively, were fixed to those of the equivalent wild-type protein as calculated previously (Mallam and Jackson, 2005; 2007a).

$$\Delta G^{N_2 \leftrightarrow 2D}_{H_2O} = \Delta G^{N_2 \leftrightarrow 2I}_{H_2O} + 2\Delta G^{I \leftrightarrow D}_{H_2O}$$

Figure 3. Denaturation Profiles of Knotted Fusion Proteins

(A and B) Equilibrium denaturation of YibK (top) and YbeA (bottom) fusion proteins with an appending ThiS at the amino (green), carboxy (red), and both termini (blue) in GdmCl (A), measured by far-UV CD signal at 225 nm, and urea (B), measured by fluorescence emission at 319 nm. Denaturation profiles are also shown for wild-type knotted domains (purple) and ThiS (black). Protein concentrations are 5 μM (circles) and 1 μM (squares). Urea renaturation profiles are represented by open symbols. The continuous lines in (A) below 3 M GdmCl and those in (B) represent the best fit to a three-state dimer denaturation model with a monomeric intermediate, while those in (A) above 3 M GdmCl represent the fit to a two-state monomer denaturation model. For ease of comparison, data have been normalized relative to a signal of 0 for a folded fusion protein, in (A) to 0.5 for a folded ThiS domain and 1 for an unfolded fusion protein, and in (B) to 1 for a fusion protein with an unfolded knotted domain.

(C) Analysis of fusion protein urea denaturation data.

monomeric intermediate with significant structure (between 60%–77 % of a fully folded monomer, estimated using kinetic *m* values) (Mallam and Jackson, 2006; 2007a). We propose that there would no longer be a knotting loop large enough for a folded ThiS domain to pass through in this intermediate species, especially without a notable decrease in the corresponding rate constant, and, therefore, that the knot is already formed. In the simplest kinetic model, the initial folding step could take

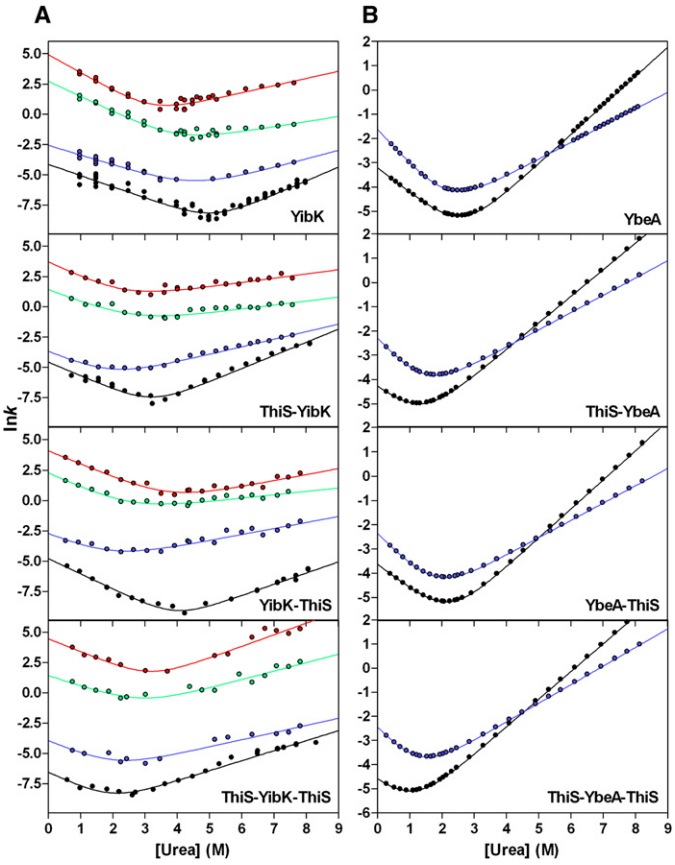


Figure 4. Urea Kinetics of Knotted Fusion Proteins at 1 μM Protein

(A and B) V-shaped plots of the natural logarithm of rate constants observed during folding and unfolding at various concentrations of urea for YibK (A) and YbeA (B) fusion proteins. Folding and unfolding data were measured using fluorescence and analyzed as described for wild-type proteins (Mallam and Jackson, 2006, 2007a). Equivalent kinetic phases are identically colored; the phases that correspond to dimerization are shown in black, while the red and green phases in (A) represent two species forming in parallel (Mallam and Jackson, 2006). Continuous lines denote the fit of each phase to a two-state folding model (Pace, 1986).

place from either a knotted or an unknotted denatured state. If the latter occurred, the "plugging" of the threading terminus with a ThiS domain should slow down or prevent folding. Instead, values of the rate constant in buffer for the red and green folding phases of YibK and the blue folding phase of YbeA are similar for the wild-type and fusion proteins (Table S3) and are consistent with folding taking place from a "preknotted" configuration. Consequently, we conclude that knotting is not the rate-limiting step during folding and, therefore, independent of the number of residues that must be fed through the loop. This is in contrast to what has been suggested previously, where it was thought likely that threading would be rate limiting (Taylor, 2007). Whether the knot becomes completely untied in the urea-denatured state of a knotted protein remains to be determined; we predict that a rapid pre-equilibrium exists between unknotted and loosely knotted conformations. Further research into the relevance of this concept to naturally unknotted proteins would be of interest, as polymer simulations demonstrate that a string of beads equivalent to approximately 600 residues has a 50 % probability of being knotted of its own accord (Virnau et al., 2005), and physical experiments involving the rotation of a confined string suggest that long flexible strings are almost certain to become knotted (Raymer and Smith, 2007); where no native-state knot is present, productive folding would only occur from an untangled chain. It is worth noting that it was speculated some time ago that protein knots might originate from a knotted denatured state (Mansfield, 1997).

We can use kinetic folding data for wild-type ThiS to consider how the knotted fusion proteins engineered here might knot and fold in the cellular environment. During synthesis, the elongating nascent polypeptide chain will initially be unknotted. Wild-type ThiS folds rapidly in vitro with a rate constant of 413 s^{-1}, equivalent to half the population of protein molecules forming the native state in the first 2 ms (Figure S4; Table S3). This is faster than the folding of wild-type YibK or YbeA and quicker than the rate of translation in *Escherichia coli* (approximately 40 amino acids per second). The ThiS domains in a fusion protein will, therefore, likely fold, whether in a cotranslational or posttranslational manner, before they have the opportunity to pass through a knotting loop. Consequently, it is probable that the knotting mechanism of ThiS-YibK-ThiS and ThiS-YbeA-ThiS in vivo involves a natively folded ThiS threading through a loop, similar to our in vitro kinetic experiments.

The results discussed here suggest that productive knotting occurs from a largely unfolded protein chain early on in the folding process; however, they do not identify the threading terminus or establish whether knotting is a cotranslational or a posttranslational event. One condition for knot formation must be sufficient elongation of the polypeptide chain to allow a knotting loop to develop. Moreover, cotranslational knotting of the polypeptide chain would require threading to take place from the amino terminus, as the threading of an *Escherichia coli* ribosome (approximately 200 Å in diameter [Schuwirth et al., 2005]) would necessitate an unfeasibly large knotting loop. If knotting of the polypeptide chain occurs in any naturally unknotted proteins of sufficient length, as suggested by both physical and simulation experiments (Virnau et al., 2005; Raymer and Smith, 2007), it could preclude successful folding. It is interesting to speculate that one role of molecular chaperones may be to prevent such unproductive knotting events in the cell.

The similar behavior of YibK and YbeA when involved in a fusion protein suggests that loose knotting in a denatured-like state might be a folding feature of all α/β-knot MTases. Furthermore, it appears that knot formation poses no problems for natural proteins, as it occurs early on in the folding process, before that of any significant structure. This explains the existence of other deeply embedded protein knots, many with additional amino or carboxy domains (Ahn et al., 2003; Elkins et al., 2003; Mosbacher et al., 2005; Taylor, 2000). Interestingly, there are no domain architectures for known sequences that contain internally positioned YibK-like or YbeA-like knotted domains similar to ThiS-YibK-ThiS and ThiS-YbeA-ThiS (Bateman et al., 2004).

More generally, it may be that complex protein topologies, such as trefoil knots, can only be anticipated by examining early folding events; interactions that occur before the formation of secondary and tertiary structural elements, perhaps very different from those present in the native state as proposed in a recent computational study on YibK (Wallin et al., 2007), could have a large effect on a protein's final structure. The process of three-dimensional domain swapping is another example of this, where the interactions present in highly intertwined oligomers have been found to form early on in the folding reaction (Hayes et al., 1999; Rousseau et al., 2001, 2003). Conversely, if something as surprising as a trefoil knot remains essentially silent during the folding process, it is possible that knotted folds will never be predicted. However, we anticipate that further study to establish the exact nature of the early-folding interactions required to productively knot a polypeptide chain will be instrumental to the advancement of protein simulation and design.

EXPERIMENTAL PROCEDURES

Plasmids and Proteins

The genetic sequences for YibK, YbeA, and ThiS were amplified by polymerase chain reaction (PCR) and introduced into pET-17b (Novagen) to encode for the appropriate multidomain protein with separate domains joined by a serine-glycine linker. Resultant fusion proteins were expressed and purified as described (Mallam and Jackson, 2005, 2007a) with the following modifications: ThiS-YibK-ThiS and ThiS-YbeA-ThiS were incubated postinduction for 16 hr at 25°C and purified by anion exchange chromatography in a buffer of 50 mM Tris-HCl (pH 8.4), 50 mM KCl, and 1 mM DTT followed by gel filtration chromatography. The expression rate of all fusion proteins was comparable to that of the equivalent wild-type protein. While less of YibK-ThiS and ThiS-YibK-ThiS appeared in the soluble fraction during purification, the solubility of all other fusion proteins was similar to wild-type, suggesting that this depends on the individual protein rather than the addition of a ThiS domain to a particular termini. The identity and purity of the fusion proteins were confirmed by SDS-PAGE and mass spectrometry. All experiments were performed in a buffer of 50 mM Tris-HCl (pH 7.5), 200 mM KCl, 10 % (v/v) glycerol, and 1 mM DTT at 25°C unless otherwise stated, except for the ITC experiments where β-mercaptoethanol replaced DTT as the reducing agent. All protein concentrations are in monomer units.

Solution X-Ray Scattering

Solution X-ray Scattering data were collected at station 2.1 of the Daresbury Synchrotron Radiation Source, UK, and processed with procedures described previously (Grossmann et al., 2002). Radius of gyration, R_g, forward scattering intensity, and the intra particle distance distribution function, $p(r)$, were evaluated with the indirect Fourier transform program GNOM (Svergun, 1992). The maximum linear dimension, D_{max}, of the particle was evaluated according to the characteristic of $p(r)$. In order to check the consistency of the results, radii of gyration were also determined from the very low-angle profiles by using the Guinier analysis based on the approximation $\ln I(q) = \ln I_0 - 4\pi^2 R_g^2 s^2/3$. At least ten particle shapes for each protein were restored from the experimental scattering profiles following ab initio procedures based on the simulated annealing algorithm using GASBOR (Svergun et al., 2001), and the most probable structure was found using DAMAVER (Volkov and Svergun, 2003). A homology-modeled structure of ThiS was generated using MODELER (Marti-Renom et al., 2000). This, along with the known crystal structures of YibK and YbeA, was used to perform rigid body modeling with the program BUNCH (Petoukhov and Svergun, 2005). For all fusion proteins, multiple repetitions of modeling were undertaken to check the reproducibility of a particular solution. Structures from rigid body modeling and ab initio were superimposed with the program SUPCOMB (Kozin and Svergun, 2001). Structural figures were made with PyMOL (www.pymol.org) and Ribbons (Carson, 1997).

Protein Characterization

Analytical gel filtration chromatography, ITC, and thermodynamic and kinetic folding experiments were carried out as described (Mallam and Jackson, 2005; 2006; 2007a; 2007b). Urea denaturation data were fit to a three-state dimer denaturation model involving a monomeric intermediate using the m values calculated previously for wild-type proteins. Individual ITC experiments were repeated at least twice. The amount of structure formed in the fastest kinetic phases was estimated by taking the ratio of the kinetic m value, m_{kin}, for this step to the m value for a fully folded monomer subunit, as estimated previously (Mallam and Jackson, 2005; 2007a). The size of the polypeptide knotting loop needed for a ThiS domain to pass through was estimated using the maximum linear dimension of ThiS, as calculated from SAXS data, and an average interresidue distance of 3.5 Å.

SUPPLEMENTAL DATA

The Supplemental Data include four figures and three tables and can be found with this article online at http://www.molecule.org/cgi/content/full/30/5/642/DC1/.

ACKNOWLEDGMENTS

We thank K. Geraki for assistance with SAXS data acquisition and A.G. Brown and J. Clarke for reading of the manuscript and critical discussion. A.L.M. is a Research Fellow (Title A) at St. John's College, Cambridge. S.C.O. was supported by a BBSRC studentship. The project was funded in part by the Leverhulme Trust. A.L.M. designed the study, undertook the major experimental work and data analysis, prepared the figures, and wrote the paper. S.C.O. carried out the SAXS experiments under supervision of J.G.G.; S.C.O. and A.L.M. performed the SAXS data analysis. All authors discussed the results and commented on the manuscript.

Received: November 20, 2007
Revised: February 22, 2008
Accepted: March 28, 2008
Published: June 5, 2008

REFERENCES

Ahn, H.J., Kim, H.W., Yoon, H.J., Lee, B.I., Suh, S.W., and Yang, J.K. (2003). Crystal structure of tRNA(m1G37)methyltransferase: insights into tRNA recognition. EMBO J. *22*, 2593–2603.

Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E.L., et al. (2004). The Pfam protein families database. Nucleic Acids Res. *32*, D138–D141.

Carson, M. (1997). Ribbons. Methods Enzymol. *277*, 493–505.

Elkins, P.A., Watts, J.M., Zalacain, M., van Thiel, A., Vitazka, P.R., Redlak, M., Andraos-Selim, C., Rastinejad, F., and Holmes, W.M. (2003). Insights into catalysis by a knotted TrmD tRNA methyltransferase. J. Mol. Biol. *333*, 931–949.

Grossmann, J.G., Hall, J.F., Kanbi, L.D., and Hasnain, S.S. (2002). The N-terminal extension of rusticyanin is not responsible for its acid stability. Biochemistry *41*, 3613–3619.

Hayes, M.V., Sessions, R.L.B., and Clarke, A.R. (1999). Engineered assembly of intertwined oligomers of an immunoglobulin domain. J. Mol. Biol. *285*, 1857–1867.

Hoffman, J.L. (1986). Chromatographic analysis of the chiral and covalent instability of S-adenosyl-L-methionine. Biochemistry *25*, 4444–4449.

Kolesov, G., Virnau, P., Kardar, M., and Mirny, L.A. (2007). Protein knot server: detection of knots in protein structures. Nucleic Acids Res. *35*, W425–W428.

Kozin, M.B., and Svergun, D.I. (2001). Automated matching of high- and low-resolution structural models. J. Appl. Cryst. *34*, 33–41.

Lim, K., Zhang, H., Tempczyk, A., Krajewski, W., Bonander, N., Toedt, J., Howard, A., Eisenstein, E., and Herzberg, O. (2003). Structure of the YibK methyltransferase from *Haemophilus influenzae* (HI0766): a cofactor bound at a site formed by a knot. Proteins Struct. Funct. Genet. *51*, 56–67.

Mallam, A.L., and Jackson, S.E. (2005). Folding studies on a knotted protein. J. Mol. Biol. *346*, 1409–1421.

Mallam, A.L., and Jackson, S.E. (2006). Probing Nature's knots: The folding pathway of a knotted homodimeric protein. J. Mol. Biol. *359*, 1420–1436.

Mallam, A.L., and Jackson, S.E. (2007a). A comparison of the folding of two knotted proteins: YbeA and YibK. J. Mol. Biol. *366*, 650–665.

Mallam, A.L., and Jackson, S.E. (2007b). The dimerization of an α/β-knotted protein is essential for structure and function. Structure *15*, 111–122.

Mansfield, M.L. (1997). Fit to be tied. Nat. Struct. Biol. *4*, 166–167.

Marti-Renom, M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F., and Sali, A. (2000). Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct. *29*, 291–325.

Mosbacher, T.G., Bechthold, A., and Schulz, G.E. (2005). Structure and function of the antibiotic resistance-mediating methyltransferase AviRb from *Streptomyces viridochromogenes*. J. Mol. Biol. *345*, 535–545.

Myers, J.K., Pace, C.N., and Scholtz, J.M. (1995). Denaturant *m* values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. Protein Sci. *4*, 2138–2148.

Nureki, O., Watanabe, K., Fukai, S., Ishii, R., Endo, Y., Hori, H., and Yokoyama, S. (2004). Deep knot structure for construction of active site and cofactor binding site of tRNA modification enzyme. Structure *12*, 593–602.

Pace, C.N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. Methods Enzymol. *131*, 266–280.

Petoukhov, M.V., and Svergun, D.I. (2005). Global rigid body modeling of macromolecular complexes against small-angle scattering data. Biophys. J. *89*, 1237–1250.

Raymer, D.M., and Smith, D.E. (2007). Spontaneous knotting of an agitated string. Proc. Natl. Acad. Sci. USA *104*, 16432–16437.

Rousseau, F., Schymkowitz, J.W.H., Wilkinson, H.R., and Itzhaki, L.S. (2001). Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by conserved proline residues. Proc. Natl. Acad. Sci. USA *98*, 5596–5601.

Rousseau, F., Schymkowitz, J.W.H., and Itzhaki, L.S. (2003). The unfolding story of three-dimensional domain swapping. Structure *11*, 243–251.

Schuwirth, B.S., Borovinskaya, M.A., Hau, C.W., Zhang, W., Vila-Sanjurjo, A., Holton, J.M., and Doudna Cate, J.H. (2005). Structures of the Bacterial Ribosome at 3.5 Å Resolution. Science *310*, 827–834.

Svergun, D.I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Cryst. *25*, 495–503.

Svergun, D.I., Petoukhov, M.V., and Koch, M.H. (2001). Determination of domain structure of proteins from X-ray solution scattering. Biophys. J. *80*, 2946–2953.

Taylor, W.R. (2000). A deeply knotted protein structure and how it might fold. Nature *406*, 916–919.

Taylor, W.R. (2007). Protein knots and fold complexity: some new twists. Comput. Biol. Chem. *31*, 151–162.

Taylor, W.R., and Lin, K. (2003). Protein knots: A tangled problem. Nature *421*, 25.

Virnau, P., Kantor, Y., and Kardar, M. (2005). Knots in globule and coil phases of a model polyethylene. J. Am. Chem. Soc. *127*, 15102–15106.

Virnau, P., Mirny, L.A., and Kardar, M. (2006). Intricate Knots in Proteins: Function and Evolution. PLoS Comput. Biol. *2*, e122. 10.1371/journal.pcbi.0020122.

Volkov, V.V., and Svergun, D.I. (2003). Uniqueness of *ab initio* shape determination in small-angle scattering. J. Appl. Cryst. *36*, 860–864.

Wallin, S., Zeldovich, K.B., and Shakhnovich, E.I. (2007). The folding mechanics of a knotted protein. J. Mol. Biol. *368*, 884–893.

Yeates, T.O., Norcross, T.S., and King, N.P. (2007). Knotted and topologically complex proteins as models for studying folding and stability. Curr. Opin. Chem. Biol. *11*, 1–9.