View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Structure
Article

# The Dimerization of an α/β-Knotted Protein Is Essential for Structure and Function

Anna L. Mallam[1] and Sophie E. Jackson[1,*]
[1] Chemistry Department, Lensfield Road, Cambridge CB2 1EW, United Kingdom
*Correspondence: sej13@cam.ac.uk
DOI 10.1016/j.str.2006.11.007

## SUMMARY

**α/β-Knotted proteins are an extraordinary example of biological self-assembly; they contain a deep topological trefoil knot formed by the backbone polypeptide chain. Evidence suggests that all are dimeric and function as methyltransferases, and the deep knot forms part of the active site. We investigated the significance of the dimeric structure of the α/β-knot protein, YibK, from *Haemophilus influenzae* by the design and engineering of monomeric versions of the protein, followed by examination of their structural, functional, stability, and kinetic folding properties. Monomeric forms of YibK display similar characteristics to an intermediate species populated during the formation of the wild-type dimer. However, a notable loss in structure involving disruption to the active site, rendering it incapable of cofactor binding, is observed in monomeric YibK. Thus, dimerization is vital for preservation of the native structure and, therefore, activity of the protein.**

## INTRODUCTION

Due to the apparent complexities involved, it was thought highly improbable, if not completely impossible, that a chain of amino acids could "knot" itself to form a functional protein. It was somewhat surprising, therefore, when proteins possessing this entirely unexpected structural property were recently identified (Taylor and Lin, 2003). Most contain a deep trefoil knot (Nureki et al., 2004; Taylor and Lin, 2003; Wagner et al., 2005), but a protein with an intricate figure-of-eight knot has been observed (Taylor, 2000), as has a knotted structure with five projected crossings (Virnau et al., 2006). Over 30 knotted proteins have been recognized in the Protein Data Bank, and hundreds more are predicted. Determining the structural and functional significance of the unusual knotted topology, as well as how these proteins knot and fold, represents an important new challenge.

Many of the knotted proteins discovered to date are structurally related and belong to the α/β-knot superfamily (Ahn et al., 2003; Bateman et al., 2004). Proteins in this clan share some common characteristics: all possess a deep trefoil knot in their backbone topology, and there is evidence to suggest that all are dimeric and function as methyltransferases (MTases) (Ahn et al., 2003; Elkins et al., 2003; Forouhar et al., 2003; Lim et al., 2003; Michel et al., 2002; Mosbacher et al., 2005; Nureki et al., 2002, 2004; Pleshe et al., 2005; Zarembinski et al., 2003). Functional studies on α/β-knot superfamily members have shown that the knotted region of the protein forms the S-adenosylmethionine (AdoMet)-binding crevice, the cofactor involved in the methylation process, and those enzymes fully characterized are all involved in the methylation of tRNA (Ahn et al., 2003; Elkins et al., 2003; Mosbacher et al., 2005; Nureki et al., 2004; Watanabe et al., 2005). Although the cofactor binding site is not always situated directly at the dimer interface, dimerization of the knotted domains is thought to be important for MTase function (Elkins et al., 2003; Nureki et al., 2004; Watanabe et al., 2005). This study aims to investigate in detail the role of dimerization in maintaining the structure and function of the α/β-knotted protein, YibK, from *Haemophilus influenzae*.

YibK is a 160 residue homodimer, described as an SpoU-type MTase due to the presence of three characteristic sequence motifs (Anantharaman et al., 2002). It is one of the smallest knotted proteins to be identified to date and has a deep trefoil knot in its structure formed by the threading of the last 40 residues of the polypeptide chain through a loop of approximately 39 residues (Lim et al., 2003) (Figure 1A). Although its biological substrate is unknown, YibK displays the catalytic fold common to all knotted MTases. Furthermore, its crystal structure has been solved with the bound cofactor, AdoHcy, the product of AdoMet after methyl-group transfer to the substrate has taken place (Lim et al., 2003), indicating that YibK most likely functions as an MTase (Figure 1C). The behavior of the YibK dimer as it folds under thermodynamic and kinetic control has been studied extensively (Mallam and Jackson, 2005, 2006a), and in both cases, a monomeric intermediate species of considerable stability and structure is populated during the folding process. YibK is, therefore, an ideal candidate for investigations into the role of dimerization in α/β-knotted proteins by the engineering of a stable monomeric variant.

A variety of approaches have been used in the past to create stable monomeric species that are incapable of associating to their native oligomeric states. Many are based on rational mutations made from the analysis of quaternary contacts from a known three-dimensional crystal

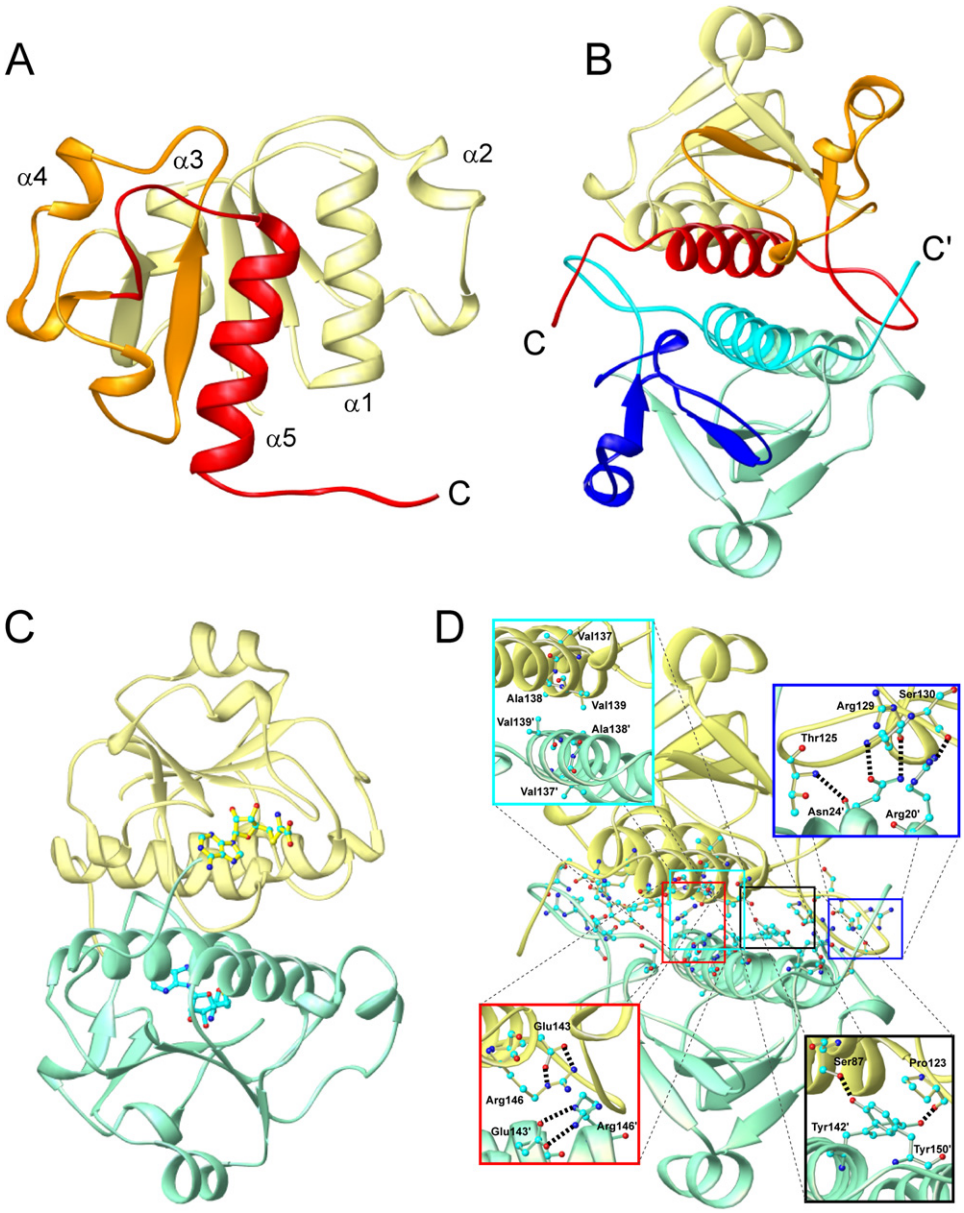


**Figure 1. Structure of YibK from *H. influenzae***

(A) Ribbon diagram of a monomer subunit (PDB code 1MXI) colored to highlight the deep trefoil knot at the C terminus, according to definitions given by Nureki et al. (2002). The knotting loop is colored orange (residues 81–120), while the knotted chain appears red (residues 121–160).

(B) Structure of dimeric YibK. One subunit is colored as in (A), while the other is shown in shades of blue.

(C) The binding site of AdoHcy in wild-type dimeric YibK. The two monomeric subunits of YibK are shown in light yellow and light blue, while AdoHcy molecules are shown as ball-and-stick models. The crystal structure contains one AdoHcy binding site per monomer, located in the knotted region of the protein.

(D) Areas of the YibK dimer interface targeted by mutagenesis. The prime (′) specifies a residue from the other subunit. Residues of interest are shown as ball-and-stick models, and thick, black dashed lines depict intermolecular-hydrogen bonds. Residues Arg20, Asn24, Ser87, and Tyr142, outlined in dark blue and black, respectively, were mutated to disrupt intermolecular hydrogen-bonding interactions. Residue Val139 was mutated to a bulky, charged residue to disrupt the hydrophobic core of the interface, and is outlined in light blue. Glu143, highlighted in red, was targeted by mutagenesis to create unfavorable electrostatic interactions between subunits. Protein structures were generated with Ribbons (Carson, 1997).

structure. In many cases, single-point mutations are used to disrupt the association of a protein interface (Beernink and Tolan, 1996; Sano et al., 1997; Shao et al., 1997), or a combination of deletions and mutations (Thoma et al., 2000). Alternative techniques involve the careful manipulation of interfacial loop regions (Borchert et al., 1994; Dickason and Huston, 1996; Mossing and Sauer, 1990) or modification of the peptide backbone using chemical

**Table 1. Analysis of YibK Wild-Type and Mutant Fluorescence Denaturation Data**

| Mutant | Wild-Type | E143A | E143K | V139R | R20A/N24A/ S87A | R20A/N24A/ S87A/Y142F | R20A/N24A/ S87A/Y142F/ E143K | R20D/N24A/ S87A/V139R/ Y142F/E143K[d] |
|---|---|---|---|---|---|---|---|---|
| $Y_I$ | 0.61 | 0.66 | 0.60 | 0.46 | 0.65 | 0.58 | 0.54 | — |
| $\Delta G_{H_2O}^{N_2 \leftrightarrow 2I}$ (kcal mol$^{-1}$) | 18.9 | 10.5 | 7.1 | 4.7 | 13.2 | 11.6 | 5.5 | — |
| $m_{N_2 \leftrightarrow 2I}$ (kcal mol$^{-1}$ M$^{-1}$) | 1.80 | 1.50 | 1.45 | 1.20 | 1.51 | 1.50 | 1.20 | — |
| $\Delta G_{H_2O}^{I \leftrightarrow D}$ (kcal mol$^{-1}$) | 6.5 | 8.6 | 6.6 | 6.9 | 8.9 | 8.9 | 6.9 | 5.2 (3.9) |
| $m_{I \leftrightarrow D}$ (kcal mol$^{-1}$ M$^{-1}$) | 1.53 | 2.01 | 1.65 | 1.72 | 1.88 | 1.92 | 1.56 | 1.42 (1.1) |
| $\Delta G_{H_2O}^{N_2 \leftrightarrow 2D}$ (kcal mol$^{-1}$)[a] | 31.9 | 27.7 | 20.2 | 18.5 | 31.1 | 29.3 | 19.3 | — |
| $m_{N_2 \leftrightarrow 2D}$ (kcal mol$^{-1}$ M$^{-1}$)[b] | 4.9 | 5.5 | 4.8 | 4.6 | 5.3 | 5.3 | 4.3 | — |
| $K_D^{N_2 \rightarrow 2I}$ (μM)[c] | $1.4 \times 10^{-8}$ | 0.02 | 6.3 | 360 | $2 \times 10^{-4}$ | $3 \times 10^{-3}$ | 93 | — |
| % of monomers present as dimer at 1 μM protein | 100 | 90 | 20 | 1 | 99 | 96 | 2 | 0 |

Fitting errors are not quoted, as they are unrealistically small, a consequence of the global analysis, and do not reflect the true experimental error, which is estimated to be ± 5% for all parameters. Data for wild-type YibK were taken from Mallam and Jackson (2006a). $Y_I$ is the spectroscopic signal of the monomeric intermediate relative to a signal of 0 for a native monomeric subunit in a dimer and 1 for a denatured monomer.
[a] $\Delta G_{H_2O}^{N_2 \leftrightarrow 2D} = \Delta G_{H_2O}^{N_2 \leftrightarrow 2I} + 2\Delta G_{H_2O}^{I \leftrightarrow D}$.
[b] $m_{N_2 \leftrightarrow 2D} = m_{N_2 \leftrightarrow 2I} + 2m_{I \leftrightarrow D}$.
[c] $K_D^{N_2 \rightarrow 2I}$ is the constant for dissociation of dimer to monomeric intermediate.
[d] Values in parenthesis were calculated using far-UV CD data measured at 225 nm.

synthesis (Rajarathnam et al., 1994) to produce stable monomeric variants of dimeric proteins. The ability of a number of these mutants to retain functionality has been studied; however, their structural and stability properties are often not examined in great detail, and few have been characterized fully to determine their relationship to intermediates observed during the folding of the wild-type oligomeric species.

The present study uses protein-engineering techniques to disrupt interactions at the dimeric interface of YibK and represents the first, to our knowledge, attempt to create a monomeric version of a dimeric, knotted protein. Structural and functional characterization of the resulting monomeric proteins allows important insights into the structure-function relationship of dimerization in α/β-knot MTases to be gained, and stability and folding experiments enable comparisons between monomeric YibK and intermediates observed during the folding of the dimeric protein to be made. Significantly, dimerization of the protein appears essential to maintain the integrity of the cofactor-binding pocket.

## RESULTS

### Mutant Design

The crystal structure of dimeric YibK indicates that association of monomeric subunits involves a variety of interactions. The two monomers are closely packed, and both α1 and α5 helices participate in dimer formation (Figure 1B). Areas involving favorable intermolecular hydrogen-bonding, hydrophobic and electrostatic interactions were identified as targets for mutagenesis (Figure 1D). Residues Arg20, Asn24, Ser87, and Tyr142 all form particularly short intermolecular hydrogen bonds (the distance between electronegative atoms is less than 2.9 Å) with residues Ser130′, Arg129′ and Thr125′, Tyr150′, and Pro123′, respectively (the prime [′] specifies a residue from the other subunit), and were all targeted by mutagenesis to remove their hydrogen-bonding capabilities (Figure 1D). The hydrophobic component of the dimer interface consists of residues Leu21, Ala138, Val139, and Tyr142 (Lim et al., 2003); Val139 was targeted by mutagenesis and altered to a bulky, charged residue (Figure 1D).

Previous studies suggest that electrostatic interactions play an important role in the dimerization of YibK, and association between monomers weakens with decreasing pH (Mallam and Jackson, 2006a). An intramolecular salt bridge exists at the dimer interface, formed by residues Glu143 and Arg146, that projects toward a counterpart ion pair in the other subunit, Glu143′ and Arg146′ (Lim et al., 2003) (Figure 1D). Glu143 was chosen as a mutagenesis target, and was altered to either an alanine (neutral) or a lysine (positive) residue to remove the salt bridge and disrupt electrostatic interactions between monomeric subunits.

Seven mutants were constructed in total, and these are listed in Table 1. Previous work has shown that the energy involved in association of two YibK-equilibrium monomeric intermediates is considerable—approximately 19 kcal mol$^{-1}$ (Mallam and Jackson, 2005). With this in mind, additional quintuple and sextuple mutants were made to disrupt more than one type of interaction.

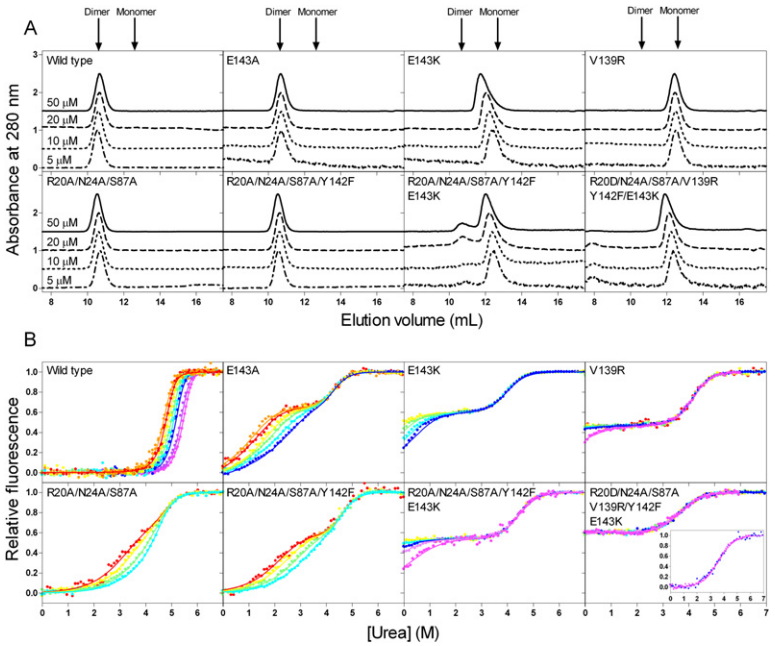


**Figure 2. Determination of the Oligomeric State and Stability of YibK Mutants**

(A) SEC elution profiles for 50, 20, 10, and 5 μM protein, displayed from top to bottom, respectively. Absorbance signal is normalized against protein concentration. The arrows indicate the expected elution volume for YibK monomer and dimer. A calibration curve has previously been shown (Mallam and Jackson, 2005). Conditions: room temperature in 50 mM Tris-HCl (pH 7.5), 200 mM KCl, 10 % (v/v) glycerol, 1 mM DTT.

(B) YibK mutant denaturation profiles for 100 (pink), 50 (dark purple), 20 (light purple), 10 (dark blue), 5 (light blue), 2.5 (green), 1 (yellow), 0.5 (red), and 0.25 μM (orange) protein, monitored by fluorescence emission 319 nm. Data are normalized relative to a folded monomer subunit in a dimer signal of 0 and a denatured monomer signal of 1. Continuous lines represent the global fit to a three-state dimer-denaturation model with a monomeric intermediate (Equation 1), except for the sextuple mutant, where data were globally fit to a two-state monomer-denaturation model (Equation 2). Inset for the sextuple mutant shows denaturation curves measured by far-UV CD signal at 225 nm. Conditions: 25°C in 50 mM Tris-HCl (pH 7.5), 200 mM KCl, 10% (v/v) glycerol, 1 mM DTT.

## The Oligomeric State and Stability of the YibK Mutants

The oligomeric state, stability, and degree of structure retained relative to wild-type YibK of the mutants was examined using size-exclusion chromatography (SEC) and fluorescence-equilibrium denaturation experiments. SEC was undertaken over a range of protein concentrations between 5 and 50 μM, and the results are shown in Figure 2A. The mutants R20A/N24A/S87A, R20A/N24A/S87A/Y142F, and E143A elute with protein concentration-independent peaks at an elution volume of 10.6 ml, corresponding to a molecular mass of 36.7 kDa, very similar to that expected for a YibK dimer of 36.8 kDa. In contrast, E143K eluted with a protein concentration-dependent peak; the elution volume for 20 μM protein was 12.0 ml, corresponding to a molecular mass of 22.2 kDa, much closer to the mass of 18.4 kDa for a YibK monomer. Similarly, two protein concentration-dependent elution peaks, at approximately 10.7 and 12.3 ml, were seen for the quintuple mutant, likely to correspond to dimeric and monomeric protein, respectively. Finally, V139R and R20D/N24A/S87A/V139R/Y142F/E143K elute with single peaks that are relatively protein concentration independent, and at a volume of 12.4 and 12.2 ml, corresponding to molecular weights of 19.2 and 20.7 kDa, respectively; these mutants appear predominantly monomeric at all concentrations of protein studied.

Equilibrium denaturation studies using the chemical denaturant urea were performed on all mutants over at least a 10-fold change in protein concentration, and results are shown in Figure 2B, along with data for wild-type protein for comparison. All mutants displayed significantly different denaturation profiles to wild-type YibK, and, with the exception of the sextuple mutant, all profiles were biphasic. The following analysis refers to all mutants except the sextuple mutant. Equilibrium unfolding transitions observed at lower urea concentrations were protein concentration dependent, consistent with mutants unfolding via a three-state dimer-denaturation model involving a monomeric intermediate (Mallam and Jackson, 2005). Data for each mutant were globally fit to this model across all concentrations of protein, and the results are summarized in Table 1. The parameter $\Delta G^{N_2 \leftrightarrow 2I}_{H_2O}$, the free energy change corresponding to the unfolding of a dimer molecule to two monomeric intermediates, is an indication of dimer stability for each of the mutants. Values range from 4.7 to 13.2 kcal $mol^{-1}$, compared to a value of 18.9 kcal $mol^{-1}$ for the wild-type protein. The most significant destabilization relative to wild-type YibK occurred in V139R and the quintuple mutant, where it is predicted that only 1%–2% of monomer molecules exist as dimers at 1 μM protein, compared to 100% for the wild-type protein (Table 1). $\Delta G^{I \leftrightarrow D}_{H_2O}$ and $m_{I \leftrightarrow D}$ values relate to the stability and structure loss upon unfolding of the monomeric species observed during the equilibrium unfolding of the mutants, respectively, and remain relatively unchanged compared with those of the wild-type protein (Table 1); this suggests that only dimeric structure and stability was notably disrupted by the mutations, not the stability of the equilibrium monomeric intermediate.

 ©2007 Elsevier Ltd All rights reserved

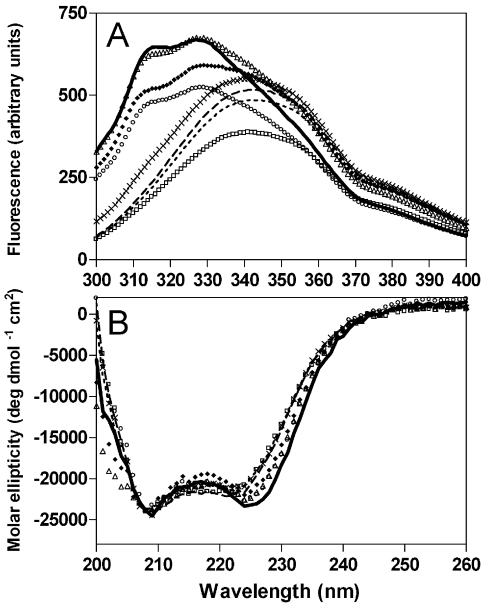


**Figure 3. Native Spectra for Wild-Type and Mutant YibK**

(A) Fluorescence and (B) far-UV CD scans for wild-type protein (solid black line), E143A (open circles), E143K (crosses), V139R (long-dashed line), R20A/N24A/S87A (open triangles), R20A/N24A/S87A/Y142F (filled diamonds), R20A/N24A/S87A/Y142F/E143K (short-dashed line), and R20D/N24A/S87A/V139R/Y142F/E143K (open squares) at 5 μM are shown. Conditions: 25°C in 50 mM Tris-HCl (pH 7.5), 200 mM KCl, 10 % (v/v) glycerol, 1 mM DTT.

Together, SEC and equilibrium denaturation data indicate that the mutants R20A/N24A/S87A, R20A/N24A/S87A/Y142F, and E143A are predominantly dimeric, E143K exists as an equilibrium ensemble of dimeric and monomeric species, and V139R and R20A/N24A/S87A/Y142F/E143K are predominantly monomeric at 1 μM protein (Figure 2 and Table 1).

In contrast to other mutants, equilibrium-denaturation profiles for the sextuple mutant are protein concentration-independent at all concentrations of protein studied. Unfolding occurs with a single transition that can be described by a two-state monomer-denaturation model (Figure 2B and Table 1). Equilibrium-denaturation profiles for R20D/N24A/S87A/V139R/Y142F/E143K were additionally measured using far-UV circular dichroism (CD) (Figure 2B, inset). These data, which monitor loss of secondary structure on unfolding, agree well with the fluorescence-denaturation profiles, demonstrating that a global unfolding event is being monitored. The protein-concentration independence of SEC and equilibrium-denaturation data indicates that the sextuple mutant remains completely monomeric at all experimental concentrations of protein examined.

The *m* values obtained from the analysis of thermodynamic unfolding data are useful parameters that relate to the amount of solvent-accessible surface area (SASA) exposed during unfolding, which in turn can be used to assess the degree of structure in a protein (Myers et al., 1995). The values shown in Table 1 are an indication of the amount of tertiary structure retained by the mutant proteins relative to wild-type YibK. The *m* value corresponding to complete unfolding of native dimer to two unfolded monomers, $m_{N_2 \leftrightarrow 2D}$, is similar for all dimeric and partially dimeric mutants, indicating that all lose comparable amounts of structure when unfolding from their native dimeric state. The *m* value predicted for dissociation of a YibK dimer into two fully folded, native-like monomers is 0.4 kcal mol$^{-1}$ M$^{-1}$ (Mallam and Jackson, 2005). This is substantially less than the experimental $m_{N_2 \leftrightarrow 2I}$ values measured for wild-type protein and dimeric mutants, which range from 1.2 to 1.8 kcal mol$^{-1}$ M$^{-1}$, suggesting that each monomer has partially unfolded upon dissociation to form the intermediate state. Likewise, the *m* value for unfolding of a fully folded YibK monomeric subunit in a dimer was predicted to be between 2.0 and 2.5 kcal mol M$^{-1}$ (Mallam and Jackson, 2005). The *m* value for unfolding of the sextuple monomeric mutant is considerably smaller than this, again indicating some structure has been lost relative to a fully folded YibK monomer in the dimer (Table 1). Fluorescence and far-UV CD spectra for YibK mutants provide further evidence for this: the decrease in native fluorescence signal and a red shift in the emission maximum observed for those YibK mutants most monomeric in nature is consistent with a decrease in tertiary structure relative to wild-type protein, while a reduction in the far-UV CD signal at 225 nm suggests a loss in secondary structure (Figure 3).

## Folding Kinetics of Selected YibK Mutants

The relationship between E143A, V139R, quintuple, and sextuple mutants and folding intermediates identified during the formation of native dimer in previous studies on the wild-type protein was examined using fluorescence kinetic-folding experiments. Several monomeric intermediates are observed during the folding of wild-type dimeric YibK, and four reversible folding phases are seen at pH 7.5 (Mallam and Jackson, 2006a). The urea-concentration dependence of the unfolding- and refolding-rate constants observed for each mutant during single-jump experiments was investigated at 1 μM protein, and V-shaped plots of the natural logarithm of the rate constants versus denaturant concentration are shown in Figure 4. Resulting kinetic phases are colored according to their similarity to those observed for wild-type protein, and appear red, green, and light blue in order from fastest to slowest, respectively. A protein concentration-dependent refolding phase was observed for E143A and is colored dark blue (Figure 4B). This phase is likely to correspond to a dimerization reaction. Rate constants for all other phases were protein concentration independent (data not shown). Double-jump unfolding experiments, where YibK mutants were allowed to refold for various amounts of time before unfolding was initiated, were used to detect additional faster unfolding phases from intermediates populated along the refolding pathway. Values of $m_{k_f}$ and $m_{k_u}$ were calculated for each mutant for all phases at 1 μM YibK, along with the corresponding unfolding- and refolding-rate constants

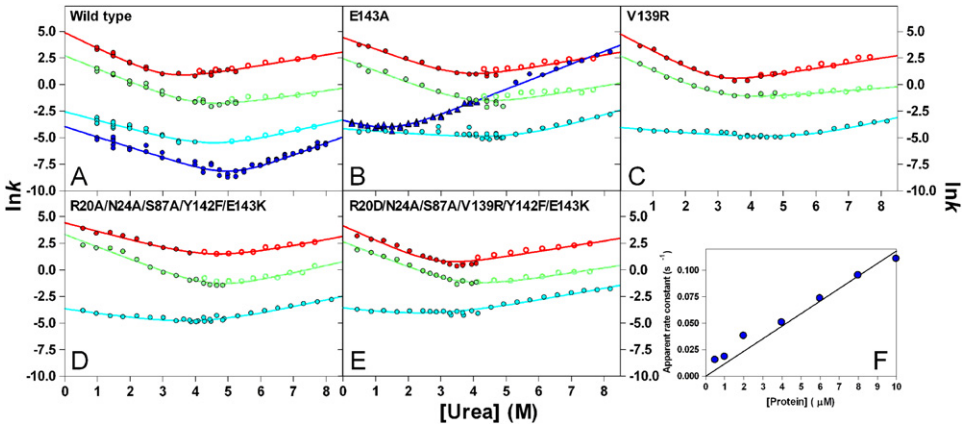


**Figure 4. Kinetics of Selected YibK Mutants at 1 μM Protein**

(A–E) V-shaped plots of the natural logarithm of rate constants observed during folding and unfolding at various concentrations of urea. Rate constants from single-jump and double-jump experiments monitored using stopped-flow apparatus are represented by filled and open circles, respectively. Single-jump rate constants measured at 319 nm on a fluorimeter using manual mixing techniques are shown as filled triangles. Phases are colored red, green, and light blue in order from fastest to slowest, respectively, and the phases that correspond to dimerization are shown in dark blue. Continuous lines represent the fit of each phase to a two-state model (Equation 7). All symbols represent rate constants calculated from a fit to a first-order reaction, except those on the refolding arm of the dimerization phase for E143A, which represent apparent rate constants calculated from a fit to a second-order reaction. A chevron plot for wild-type protein (Mallam and Jackson, 2006a) is included for comparison. Conditions were as described for Figure 3.

(F) The protein-concentration dependence for the dimerization phase for E143A at 0.75 M urea (the solid line represents the fit of the apparent rate constants to Equation 6).

in the absence of denaturant. These results are summarized in Table 2.

The dimeric mutant E143A displays folding kinetics similar to wild-type YibK, and four reversible folding phases are observed. The $m_{kin}$ values, which relate to the SASA change associated with an observed kinetic step, calculated for the dark-blue dimerization phase for both proteins, are in excellent agreement, indicating that the SASA buried is similar (Table 2). However, this phase is significantly destabilized in E143A relative to wild-type YibK, and the $\Delta G^{kin}_{H_2O}$ of 9.5 kcal mol$^{-1}$ observed is considerably lower than the value of 14.0 kcal mol$^{-1}$ for the wild-type protein. No equivalent dimerization phase is seen for monomeric and predominantly monomeric mutants; they display three reversible phases only (Figure 4).

In previous work, extensive interrupted-refolding studies were undertaken on dimeric wild-type YibK to learn more about its folding mechanism (Mallam and Jackson, 2006a). Similar experiments were carried out on the quintuple and sextuple mutants, which are monomeric at 1 μM protein (Table 1). This allowed the time course for intermediates involved in the refolding reaction to be followed, as the population of any species present after various durations of refolding is proportional to the amplitude of the corresponding unfolding reaction (Schmid, 1983; Wallace and Matthews, 2002). The resulting unfolding amplitudes for the three phases observed after various refolding periods are shown in Figures 5A and 5B. The time course of refolding species was similar for both mutants; the populations of the species corresponding to the two fastest phases (red and green) increased in parallel with no observable lag, while a lag was seen in the formation of the species corresponding to the light-blue phase before its population escalated to dominate the refolding ensemble, indicating that its formation is preceded by an obligatory intermediate (Heidary et al., 2000). The folding mechanism shown in Figure 5C involving three on-pathway species, $I_1$, $I_2$, and $I_3$, best describes the interrupted-refolding data, and simulations of the population of species present during refolding via this mechanism are shown in Figures 5A and 5B. Mechanisms involving either of the species corresponding to the fastest two phases, $I_1$ and $I_2$, as off-pathway intermediates do not describe the interrupted-refolding data well (data not shown). The mechanism shown in Figure 5C involves $I_1$ and $I_2$ folding to a third species $I_3$. This is very similar to the proposed folding pathway of wild-type protein (Figure 5D). During the folding of wild-type protein, the species corresponding to the light-blue phase, $I_3$, folds to native dimer, $N_2$. This can be compared to the monomeric mutants, where the population of $I_3$ does not decay and no detectable dimer is formed.

In summary, kinetic experiments indicate that dimeric mutants of YibK fold in a manner similar to that of wild-type protein, but with a significantly destabilized dimerization phase, while monomeric mutants display a strong resemblance to the monomeric intermediate $I_3$ observed on the wild-type folding pathway.

## Affinity of S-Adenosyl Homocysteine for Dimeric and Monomeric Forms of YibK

Since the physiological substrate of YibK is not yet known, the ability of mutants of YibK to retain their MTase function

**Table 2. Kinetic Parameters for the Unfolding and Refolding of Selected YibK Mutants at pH 7.5 and 1 μM Final Protein Concentration**

| Phase | Color | Mutant[a] | $k_f^{H_2O}$ ($s^{-1}$)[b] | $k_u^{H_2O}$ ($s^{-1}$) | $m_{k_f}$ (kcal $mol^{-1}$ $M^{-1}$) | $m_{k_u}$ (kcal $mol^{-1}$ $M^{-1}$) | $m_{kin}$ (kcal $mol^{-1}$ $M^{-1}$)[c] | $\Delta G^{kin}_{H_2O}$ (kcal $mol^{-1}$)[d] |
|---|---|---|---|---|---|---|---|---|
| 1 | Red | Wild-type | 133 ± 22 | 0.30 ± 0.06 | 0.87 ± 0.06 | 0.30 ± 0.02 | 1.2 ± 0.1 | 3.6 ± 0.2 |
| | | E143A | 84 ± 16 | 0.22 ± 0.09 | 0.66 ± 0.07 | 0.22 ± 0.04 | 0.9 ± 0.1 | 3.5 ± 0.5 |
| | | V139R | 118 ± 30 | 0.26 ± 0.06 | 0.94 ± 0.09 | 0.29 ± 0.02 | 1.2 ± 0.1 | 3.6 ± 0.3 |
| | | Quintuple | 84 ± 10 | 0.17 ± 0.09 | 0.48 ± 0.04 | 0.34 ± 0.05 | 0.8 ± 0.1 | 3.7 ± 0.5 |
| | | Sextuple | 64 ± 16 | 0.30 ± 0.09 | 0.80 ± 0.1 | 0.29 ± 0.03 | 1.1 ± 0.1 | 3.2 ± 0.4 |
| 2 | Green | Wild-type | 15.1 ± 2.3 | 1.5 (± 0.7) × $10^{-2}$ | 0.73 ± 0.05 | 0.27 ± 0.04 | 1.0 ± 0.1 | 4.1 ± 0.3 |
| | | E143A | 11 ± 3 | 1.5 (± 1) × $10^{-2}$ | 0.67± 0.09 | 0.30 ± 0.07 | 1.0 ± 0.1 | 3.9 ± 0.7 |
| | | V139R | 15 ± 2 | 0.11 ± 0.02 | 0.80 ± 0.05 | 0.14 ± 0.02 | 0.9 ± 0.05 | 2.9 ± 0.2 |
| | | Quintuple | 28 ± 5 | 6.4 (± 4) × $10^{-3}$ | 0.68 ± 0.05 | 0.40 ± 0.06 | 1.1 ± 0.1 | 5.0 ± 0.7 |
| | | Sextuple | 14 ± 2 | 2.9 (± 1) × $10^{-2}$ | 0.71 ± 0.04 | 0.28 ± 0.04 | 1.0 ± 0.1 | 3.7 ± 0.4 |
| 3 | Light blue | Wild-type | 7.7 (± 1.1) × $10^{-2}$ | 9.0 (± 7) × $10^{-5}$ | 0.48 ± 0.05 | 0.42 ± 0.08 | 0.9 ± 0.1 | 4.0 ± 0.1 |
| | | E143A | 1.6 (± 0.2) × $10^{-2}$ | 3.1 (± 3) × $10^{-5}$ | 0.13 ± 0.03 | 0.55 ± 0.08 | 0.7 ± 0.1 | 3.7 ± 1 |
| | | V139R | 1.7 (± 0.2) × $10^{-2}$ | 6.8 (± 6) × $10^{-5}$ | 0.15 ± 0.03 | 0.45 ± 0.07 | 0.6 ± 0.1 | 3.3 ± 0.9 |
| | | Quintuple | 2.6 (± 0.3) × $10^{-2}$ | 3.0 (± 1) × $10^{-4}$ | 0.26 ± 0.05 | 0.39 ± 0.04 | 0.7 ± 0.1 | 2.6 ± 0.4 |
| | | Sextuple | 2.6 (± 0.4) × $10^{-2}$ | 2.0 (± 0.6) × $10^{-3}$ | 0.24 ± 0.07 | 0.33 ± 0.03 | 0.6 ± 0.1 | 1.5 ± 0.3 |
| 4 | Dark blue | Wild-type | 1.9 (± 0.3) × $10^{-2}$ | 4.9 (± 2.0) × $10^{-7}$ | 0.57 ± 0.03 | 0.67 ± 0.03 | 1.2 ± 0.1 | 14.0 ± 0.3 |
| | | E143A | 3.3 (± 0.3) × $10^{-2}$ | 1.7 (± 0.1) × $10^{-3}$ | 0.53 ± 0.07 | 0.7 ± 0.01 | 1.2 ± 0.1 | 9.5 ± 0.1 |

Errors quoted are the standard errors calculated by the fitting program. $k_f^{H_2O}$ and $k_u^{H_2O}$ are the rate constants for refolding and unfolding, respectively, in the absence of denaturant; $m_{k_f}$ and $m_{k_u}$ are the kinetic refolding and unfolding $m$ values, respectively.
[a] Quintuple and sextuple mutants are R20A/N24A/S87A/Y142F/E143K and R20D/N24A/S87A/V139R/Y142F/E143K, respectively.
[b] All refolding rates are first order, except for phase 4 where $k_{app}^{H_2O}$ is quoted. $k_{app}^{H_2O} = P_t k_{2nd}^{H_2O}$, where $P_t$ is the concentration of protein.
[c] $m_{kin} = m_{k_f} + m_{k_u}$.
[d] $\Delta G^{kin}_{H_2O} = -RT\ln(k_u^{H_2O}/k_f^{H_2O})$ except for phase 4 where $\Delta G^{kin}_{H_2O} = -RT\ln(2k_u^{H_2O}/k_{2nd}^{H_2O})$.

was determined by AdoHcy affinity studies. The binding of AdoHcy is more straightforward to examine than that of AdoMet, as the latter is unstable in vitro (Hoffman, 1986). The AdoHcy cofactor binding site, shown in Figure 1C, consists of a pocket formed by two loops of the knot, residues 80–85 and 102–105, and the loop connecting β6 and α5 (Lim et al., 2003). Binding of AdoHcy was measured by isothermal titration calorimetry (ITC); results are shown in Figure 6 and summarized in Table 3. Mutants that are predominantly dimeric, as determined by SEC and equilibrium-denaturation experiments, display a similar affinity for AdoHcy as wild-type YibK, and only a small increase in their dissociation constant ($K_D$), relative to that of wild-type protein, is observed. In contrast, much weaker binding to AdoHcy was seen for the quintuple mutant and V139R, both largely monomeric, and no observable binding was seen for the completely monomeric sextuple mutant. ITC experiments, therefore, indicate that AdoHcy is only able to bind to dimeric forms of YibK. The stoichiometry of the observed AdoHcy binding event was approximately 0.5 for wild-type protein and all mutants, indicating that only one AdoHcy molecule binds to each YibK dimer.

## DISCUSSION

The α/β-knot superfamily of homodimeric MTases is an extraordinary group of proteins that contain a deep trefoil knot in their backbone topology. Before their discovery, it was thought that knot formation in proteins would be impossible, due to the apparent complications involved; it is still not obvious how, during the process of protein folding, a substantial length of polypeptide chain manages to spontaneously thread itself through a loop. Based on sequence classification, α/β-knot proteins can be divided into four distinct families, known as SpoU, TrmD, YbeA, and AF2226 (Anantharaman et al., 2002). Several mutational studies on SpoU- and TrmD-like proteins have shed light on the functional role of the knotted region in their structure, and have shown that it forms the cofactor-binding pocket and the active site (Elkins et al., 2003; Mosbacher et al., 2005; Nureki et al., 2004; Watanabe et al., 2005). In this study, the purpose of dimerization in the SpoU α/β-knotted protein, YibK, of *H. influenzae* has been examined by the generation of a monomeric version of the protein.

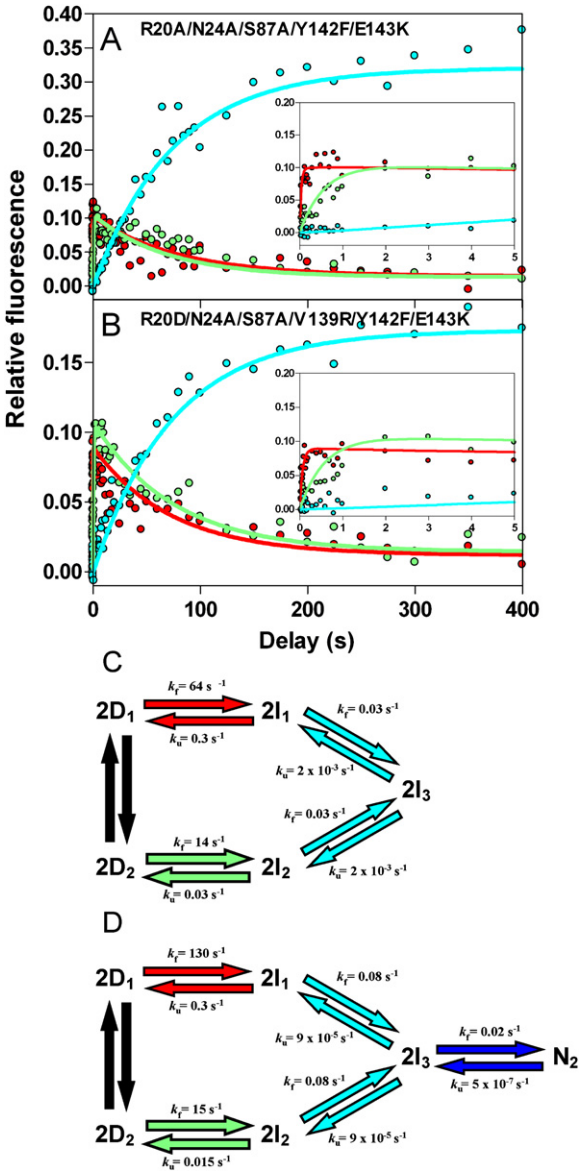


**Figure 5. Determining the Folding Mechanism of Monomeric YibK**

(A and B) Relative amplitudes of the three unfolding reactions seen during interrupted-refolding experiments on (A) the quintuple and (B) the sextuple YibK mutants after refolding at 1 M urea, and subsequent unfolding at 7.7 M urea and a final concentration of protein of 1 μM. Insets show an expanded view for delay times up to 5 s. Amplitudes are colored according to their corresponding phase shown in Figure 4.

(C) The folding mechanism of YibK monomeric mutants most consistent with all experimental data. Rate constants are shown for the sextuple mutant in buffer at 25°C, and arrows are colored according to their corresponding phase in Figure 4. The continuous lines in (A) and (B) represent simulations of the time course of intermediates folding via the mechanism shown in (C).

(D) The folding mechanism proposed for wild-type YibK dimer (Mallam and Jackson, 2006a). Conditions were as described for Figure 3.

Six disruptive mutations were necessary to render YibK entirely monomeric at all experimental protein concentrations examined. That such extensive disruptions were required is perhaps not surprising, as thermodynamic folding studies on the wild-type protein showed that the interaction between two monomeric intermediates was strong, some 18.9 kcal $mol^{-1}$ (Mallam and Jackson, 2005). Furthermore, other α/β-knotted proteins have only ever been observed as dimers, indicating that subunits of proteins in this superfamily are not easily dissociated (Ahn et al., 2003; Elkins et al., 2003; Mallam and Jackson, 2005, 2006b; Mosbacher et al., 2005; Nureki et al., 2004). Dimer stability was significantly reduced in all of the YibK mutants engineered. The sequence conservation of residues targeted for mutagenesis within the SpoU subfamily of α/β-knotted proteins is shown in Figure 7. The mutated residues Arg20, Tyr142, and Glu143 are all highly conserved, and the residue corresponding to position 139 in the YibK sequence is always hydrophobic, suggesting that amino acids at these positions have been preserved for dimer stability in SpoU-like knotted proteins.

A notable loss in secondary and tertiary protein structure in monomeric forms of YibK demonstrates that dimerization of the protein is essential for maintaining native-like structure. Furthermore, upon monomerization, YibK is unable to bind the MTase cofactor, AdoHcy. Residues Leu78, Gly100, Ile122, Met131, and Ser136 have been identified as those that form hydrogen bonds with AdoHcy in its bound state (Figure 7) (Lim et al., 2003) and are independent of the residues mutated in this study, which are not located in the AdoHcy-binding pocket. Consequently, it is reasonable to assume that the loss of affinity observed for monomeric YibK is caused by the disruption of the structure of the binding pocket upon monomerization of the protein, and not by the removal of direct AdoHcy-binding interactions. Dimerization is, therefore, crucial to maintain the integrity of the cofactor binding site. Since YibK would not be able to act as an MTase in the absence of a bound cofactor molecule, it follows that dimerization is necessary for the preservation of the function of the protein. This demonstrates the biological significance of the strong dimerization observed in α/β-knotted proteins.

Studies on other α/β-knotted proteins have postulated that dimerization is important for MTase function (Ahn et al., 2003; Elkins et al., 2003; Nureki et al., 2004; Watanabe et al., 2005). Residues potentially crucial to the MTase activity of the SpoU knotted homodimer, TrmH, from *Thermus thermophilus*, a protein closely related to YibK, have been identified (Nureki et al., 2004; Watanabe et al., 2005). Nureki and coworkers proposed a novel RNA-dependent methylation mechanism for TrmH, and suggested that dimerization was critical for tRNA binding and methylation catalysis, as one monomer subunit binds AdoMet, while the other serves as a tRNA-binding site (Nureki et al., 2004; Watanabe et al., 2005). Extensive mutational analysis performed on the knotted protein, TrmD, from *Escherichia coli* led to the suggestion that formation of a homodimer was required for activity. This conclusion was based on the observation that mutations made outside the catalytic region but at the dimer interface led to inactivation of the protein (Elkins et al., 2003). The results presented here demonstrate the structural importance of

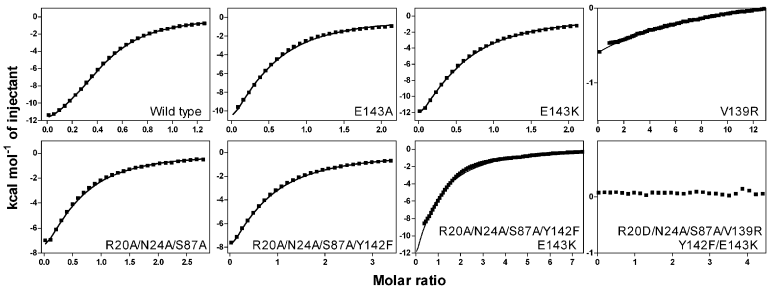


**Figure 6. Affinity of AdoHcy for YibK Wild-Type and Mutant Proteins**

The continuous line represents the fit of ITC data to a single-site binding model using the Origin software package (MicroCal Inc.). Conditions were 50 mM Tris-HCl (pH 7.5), 200 mM KCl, 10% glycerol (v/v), 1 mM β-mercaptoethanol. Data have been corrected for the heat of dilution.

the dimerization observed in YibK. However, there is evidence to suggest that these findings may be applied to other α/β-knot proteins. Folding studies recently undertaken on the knotted homodimer, YbeA, from *E. coli* established that, like YibK, YbeA unfolds via a thermodynamic and kinetic monomeric intermediate that has lost significant structure relative to the native monomeric subunit in the dimer, implying that dimerization is also essential to maintain native structure and, perhaps, function in this α/β-knot protein (Mallam and Jackson, 2006b).

The stoichiometry of the observed cofactor-binding event suggests that only one AdoHcy molecule binds to each YibK dimer. We note that this is in contrast to the crystal structure that shows one AdoHcy moiety bound to each monomer subunit. A possible explanation for this discrepancy is that binding of one AdoHcy molecule causes a conformational change that prevents observable binding of a second AdoHcy unit; Lim and coworkers reported a small conformational change involving five loop residues upon AdoHcy binding (Lim et al., 2003). Cocrystallization of AdoHcy was achieved by soaking a YibK protein crystal in a high-concentration solution of cofactor over a long period of time, conditions potentially sufficient to cause the equilibrium to favor a higher binding stoichiometry.

Upon monomerization of YibK, a species is formed that has similar secondary and tertiary structure (as judged by agreement of the $m_{I\leftrightarrow D}$ values for monomeric YibK and wild-type protein) to the equilibrium monomeric intermediate observed during wild-type unfolding. Kinetic characterization of the YibK mutants allowed their folding pathway to be compared to that of the wild-type dimer, which forms by a complex kinetic mechanism involving two different intermediates ($I_1$ and $I_2$) from parallel pathways folding via a third sequential monomeric intermediate ($I_3$) to form native dimer ($N_2$) in a slow, rate-limiting dimerization reaction (Mallam and Jackson, 2006a). The folding mechanisms for monomeric mutants of YibK and wild-type dimer appear very similar, except that, during folding of the wild-type protein, $I_3$ is an intermediate that precedes formation of the native dimer, $N_2$. This agreement validates the proposed wild-type dimer folding pathway, and suggests that monomeric YibK is an excellent model for the folding intermediate, $I_3$, observed during wild-type folding. The comparable stability and *m* values for the kinetic folding of the monomeric mutants to the values for wild-type $I_3$ adds further weight to this argument (Table 2). Furthermore, monomeric mutants similar to $I_3$ display little or no binding to AdoHcy, implying that the cofactor binding site is formed during the final folding step ($2I_3 \leftrightarrow N_2$). Importantly, while the kinetic phase corresponding to dimerization in the dimeric mutant, E143A, is considerably destabilized relative to that for wild-type protein, the other three phases remain the same,

**Table 3. Thermodynamic Parameters for the Binding of AdoHcy to YibK Wild-Type and Mutant Proteins**

| Mutant | Binding Stoichiometry ([AdoMet]/[YibK]) | $K_D$ (μM) | $\Delta G_b$ (kcal mol$^{-1}$)[a] |
|---|---|---|---|
| Wild-type | 0.44 ± 0.002 | 26 ± 0.5 | −6.3 ± 0.01 |
| E143A | 0.41 ± 0.03 | 55 ± 3.4 | −5.8 ± 0.04 |
| E143K | 0.46 ± 0.02 | 61 ± 3.0 | −5.8 ± 0.03 |
| V139R | 0.5[b] | 8130 ± 3600 | −2.9 ± 0.26 |
| R20A/N24A/S87A | 0.51 ± 0.05 | 86 ± 5.4 | −5.6 ± 0.04 |
| R20A/N24A/S87A/Y142F | 0.49 ± 0.02 | 65 ± 2 | −5.7 ± 0.02 |
| R20A/N24A/S87A/Y142F/E143K | 0.5[b] | 581 ± 19 | −4.4 ± 0.02 |
| R20A/N24A/S87A/V139R/Y142F/E143K | No binding | No binding | No binding |

ITC data were analyzed using Origin version 7, and the errors quoted are the standard errors calculated by the fitting program. "No binding" indicates that no binding was observed. Concentration of protein in the ITC cell varied between 110 and 330 μM.

[a] The free energy of binding was calculated using $\Delta G_b = -RT\ln(1/K_D)$.

[b] The binding stoichiometry was fixed to 0.5 as suggested by Turnbull and Daranas (2003) to allow a more accurate determination of $K_D$ and, hence, $\Delta G_b$, in low-affinity systems.

 ©2007 Elsevier Ltd All rights reserved 

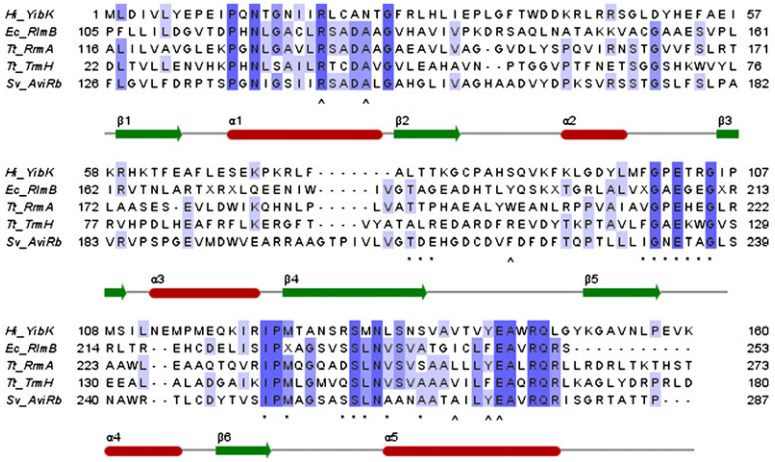


**Figure 7. Multiple Sequence Alignment of SpoU Family Proteins**

All proteins are known to contain a deep trefoil knot in their backbone structure. The alignment was performed using ClustalW (Chenna et al., 2003), and the figure was generated using Jalview (Clamp et al., 2004). Proteins are listed according to the species name followed by the gene: Hi, *Haemophilus influenzae*; Ec, *Escherichia coli*; Tt, *Thermus thermophilus*; Sv, *Streptomyces viridochromogenes*. YibK residues marked "∧" have been targeted by mutagenesis in this study, and '*' indicates residues buried by AdoHcy binding (Lim et al., 2003).

demonstrating that dimeric interactions have been selectively disrupted. An intermediate that could correspond to a fully folded YibK monomer is never observed during the folding of wild-type or any of the YibK mutants; native-like monomer subunits can only exist when accompanied by the formation of quaternary interactions in the dimeric structure. It is interesting to note that the folding kinetics of monomeric mutants of YibK are slow in comparison with the kinetics of other stable monomeric species that have been engineered from dimeric proteins. Examples are the folding of tryptophan repressor monomer and a monomeric form of phage λ repressor, both of which occur on a submillisecond timescale (Huang and Oas, 1995; Shao et al., 1997). It is possible that the development of the knot in YibK is responsible for its slow folding; however, it is necessary to establish the kinetic step corresponding to knot formation in order to confirm this.

## Conclusions

The intriguing deep trefoil knots found in the backbone topology of α/β-knotted MTases demonstrate that nature has evolved mechanisms not only to successfully fold protein chains, but to knot them as well. Discovering how and why such knots occur represents a fundamental and exciting challenge in structural biology. A characteristic of all α/β-knotted proteins is their existence as homodimers. In this study, the construction of mutants that disrupt the dimer interface of YibK has allowed the effects of dimerization on structure and function of this knotted protein to be examined directly. Thermodynamic and kinetic characterization of mutant proteins provided a convenient way of confirming their oligomeric state and assessing their relative structures, stability and folding pathways. Results clearly show that activation of this enzyme occurs upon dimerization, and monomerization of the protein leads to a loss of both structure and function. Consequently, the knotted topology alone is insufficient to maintain the active conformation of the cofactor binding site in YibK, and additional stability is required from dimerization of the protein. This demonstrates that, while the knot assembly may be advantageous in terms of constricting movement in the active-site region, dimerization is also essential to preserve the correct active-site structure. The conservation of many residues targeted by mutagenesis in this study, as well as the observation of partially folded monomeric species in other α/β-knot proteins, suggest that these findings may be applicable to other knotted homodimers.

## EXPERIMENTAL PROCEDURES

### Materials

Molecular biology-grade urea was purchased from BDH Laboratory Supplies. Point mutations were introduced into the gene encoding for YibK wild-type protein using the QuikChange Site-Directed Mutagenesis Kit (Strategene). A series of site-directed mutagenesis reactions were performed to obtain triple, quadruple, quintuple, and sextuple mutants. Mutant proteins were expressed and purified as described for wild-type YibK (Mallam and Jackson, 2005), with the following modifications: protein was incubated postinduction for 16 hr at 25°C, apart from the sextuple mutant, which was incubated for 16 hr at 15°C. Additionally, the step involving an SP-sepharose cation-exchange column was performed using a buffer of 50 mM phosphate (pH 6.7), 125 mM KCl, 5 % glycerol (v/v), 1 mM DTT. The identity of mutants was confirmed by DNA sequencing and mass spectrometry. All experiments were performed in a buffer of 50 mM Tris-HCl (pH 7.5), 200 mM KCl, 10 % glycerol (v/v), 1 mM DTT, except for the ITC experiments where β-mercaptoethanol replaced DTT as the reducing agent. All protein concentrations are in monomer units.

### Mutant Characterization

SEC was performed on an ÄKTA FPLC system using a Superdex 75 10/300 GL analytical gel filtration column, as described previously (Mallam and Jackson, 2005). All spectroscopic measurements were taken using a thermostatically controlled cuvette or cell at 25°C. For fluorescence studies, data were collected with an SLM-Amico Bowman series 2 luminescence spectrometer with an excitation wavelength of 280 nm (4 nm band pass) with a 1 cm path-length cuvette. Fluorescence was monitored at 319 nm (4 nm band pass) for manual-mixing kinetic experiments on E143A, while scans between 310 and 350 nm were recorded for equilibrium-denaturation experiments. Far-UV CD spectra were acquired with an Applied Photophysics Chirascan, and scans were taken between 200 and 260 nm at a scan rate of 1 nm $s^{-1}$ using a 0.1 cm path-length cuvette and a bandwidth of 1 nm. For equilibrium denaturation, the change in far-UV CD signal was monitored at 225 nm. Rapid-mixing fluorescence data were collected using an Applied Photophysics SX.18MV stopped-flow fluorimeter with no cut-off filter.

Equilibrium-denaturation experiments on YibK mutants were performed using the same methods as for wild-type protein, and these are described in detail elsewhere (Mallam and Jackson, 2005). Samples were left for at least 1 hr to equilibrate, after which no change in spectroscopic signal was seen. The reversibility of unfolding in urea for all mutants was confirmed using fluorescence and far-UV CD. Kinetic unfolding and refolding experiments using fluorescence were undertaken on selected mutants of YibK by the same methods described for wild-type protein (Mallam and Jackson, 2006a) and were performed at a final concentration of protein of 1 μM unless otherwise stated. There was no observable burst phase for any mutant, as all amplitude change was accounted for by the kinetic traces. The dimerization phase for E143A was measured by refolding protein unfolded in 3 M urea.

### Data Analysis

All data analysis was performed using the nonlinear, least squares-fitting program, Prism version 4 (GraphPad Software). Mutant equilibrium-denaturation data measured at 319 nm, with the exception of that for the sextuple mutant, were globally fit over all concentrations of protein to a three-state dimer denaturation model involving a monomeric intermediate:

$$Y_{rel} = Y_N\left(\frac{2P_tF_I^2}{K_1}\right) + Y_I(F_I) + Y_D(K_2F_I), \quad (1)$$

where: $Y_{rel}$ is the normalized spectral signal, $Y_N$, $Y_I$, and $Y_D$ are the spectroscopic signals of the native, intermediate, and denatured state, respectively; $P_t$ is the total protein concentration in terms of monomer, $F_I$ represents the fraction of monomeric subunits involved in the intermediate state, and $K_1$ and $K_2$ are the equilibrium constants for the first and second transitions, respectively.

Equilibrium-unfolding data measured for the sextuple mutant were globally fit to a two-state monomer-denaturation model:

$$[D] = \frac{([D]+[N])\exp\left(\left\{m_{N\leftrightarrow D}[\text{urea}] - \Delta G^{H_2O}_{N\leftrightarrow D}\right\}RT\right)}{1+\exp\left(\left\{m_{N\leftrightarrow D}[\text{urea}] - \Delta G^{H_2O}_{N\leftrightarrow D}\right\}/RT\right)}, \quad (2)$$

where N is a folded monomeric species and D is a denatured monomer.

These models have been described in detail elsewhere (Mallam and Jackson, 2005, 2006a).

All kinetic traces, except those for the protein concentration-dependent phase observed for E143A, were fit individually to a first-order reaction with the required number of exponentials:

$$Y(t) = Y_{Native} + \sum_{i=1}^{N} Y_i\exp(-k_{1st}t), \quad (3)$$

where $Y_{(t)}$ is the signal at time $t$, $Y_{Native}$ is the signal expected for fully folded native protein, $Y_i$ is the amplitude change corresponding to a given kinetic phase, and $K_{1st}$ is the first-order rate constant. The protein concentration-dependent traces observed during the refolding kinetics of E143A were fit to a second-order reaction described by the following model:

$$2I \overset{k_{2nd}}{\leftrightarrow} N_2 \; d[N_2]/dt = k_{2nd}[I]^2, \quad (4)$$

where $K_{2nd}$ is the bimolecular folding rate constant. The differential equation can be solved to give:

$$Y(t) = Y_{t=0} + Y_i(k_{app}t)/(1+k_{app}t), \quad (5)$$

where $Y_{t=0}$ is the signal at time $t = 0$ and $k_{app}$ is the apparent rate constant. The apparent rate constant is related to $k_{2nd}$ as follows:

$$k_{app} = P_tk_{2nd}, \quad (6)$$

where $P_t$ is the concentration of protein in terms of monomer.

The dependence of the natural logarithm of the unfolding- and refolding-rate constants on urea concentration is assumed to be linear (Tanford, 1968, 1970), and each phase on the chevron plots was fit to:

$$\ln k_{obs} = \ln\left(k_f^{H_2O}\exp(-m_{k_f}[\text{urea}]) + k_u^{H_2O}\exp(m_{k_u}[\text{urea}])\right), \quad (7)$$

where $k_{obs}$ is the observed rate constant, $k_f^{H_2O}$ and $k_u^{H_2O}$ are the refolding- and unfolding-rate constants for each phase in water, and $m_{k_f}$ and $m_{k_u}$ are constants of proportionality.

Traces from interrupted refolding and unfolding experiments for different delay times were globally fit to Equation 3, with values for the first-order unfolding-rate constants shared throughout all datasets.

Kinetic simulations to model the time course of species present during refolding of YibK monomeric mutants via various possible folding mechanisms were performed with the numerical simulation program KINSIM (Dang and Frieden, 1997) and the rate constants from the chevron plots.

### ITC

ITC was performed with a MicroCal VP-ITC instrument (MicroCal Inc. Northampton, MA). AdoHcy at an appropriate concentration was injected into a 2.5 ml cell containing protein. Parallel experiments were carried out injecting AdoHcy into buffer alone to correct data for the heat of dilution in subsequent data analysis using Origin (MicroCal Inc.). Protein and AdoHcy concentrations were determined spectrophotometrically.

## ACKNOWLEDGMENTS

The authors thank E. Coulstock for assistance with ITC experiments. A.L.M was supported by a Medical Research Council studentship. The work was funded in part by the Welton Foundation.

Received: October 11, 2006
Revised: November 24, 2006
Accepted: November 29, 2006
Published: January 16, 2007

## 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.

Anantharaman, V., Koonin, E.V., and Aravind, L. (2002). SPOUT: a class of methyltransferases that includes spoU and trmD RNA methylase superfamilies, and novel superfamilies of predicted prokaryotic RNA methylases. J. Mol. Microbiol. Biotechnol. *4*, 71–75.

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.

Beernink, P.T., and Tolan, D.R. (1996). Disruption of the aldolase A tetramer into catalytically active monomers. Proc. Natl. Acad. Sci. USA *93*, 5374–5379.

Borchert, T.V., Abagyan, R., Jaenicke, R., and Wierenga, R.K. (1994). Design, creation, and characterization of a stable, monomeric triosephosphate isomerase. Proc. Natl. Acad. Sci. U.S.A. *91*, 1515–1518.

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

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D. (2003). Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. *31*, 3497–3500.

Clamp, M., Cuff, J., Searle, S.M., and Barton, G.J. (2004). The Jalview Java alignment editor. Bioinformatics *20*, 426–427.

Dang, Q., and Frieden, C. (1997). New PC versions of the kinetic-simulation and fitting programs, KINSIM and FITSIM. Trends Biochem. Sci. *22*, 317.

 ©2007 Elsevier Ltd All rights reserved 

Dickason, R.R., and Huston, D.P. (1996). Creation of a biologically active interleukin-5 monomer. Nature *379*, 652–655.

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.

Forouhar, F., Shen, J., Xiao, R., Acton, T.B., Montelione, G.T., and Tong, L. (2003). Functional assignment based on structural analysis: crystal structure of the yggJ protein (HI0303) of *Haemophilus influenzae* reveals an RNA methyltransferase with a deep trefoil knot. Proteins. *53*, 329–332.

Heidary, D.K., O'Neill, J.C., Jr., Roy, M., and Jennings, P.A. (2000). An essential intermediate in the folding of dihydrofolate reductase. Proc. Natl. Acad. Sci. U.S.A. *97*, 5866–5870.

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

Huang, G.S., and Oas, T.G. (1995). Submillisecond folding of monomeric lambda repressor. Proc. Natl. Acad. Sci. USA *92*, 6878–6882.

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. *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. (2006a). 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. (2006b). A comparison of the folding of two knotted proteins: YbeA and YibK. J. Mol. Biol., in press.

Michel, G., Sauve, V., Larocque, R., Li, Y., Matte, A., and Cygler, M. (2002). The structure of the RlmB 23S rRNA methyltransferase reveals a new methyltransferase fold with a unique knot. Structure *10*, 1303–1315.

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.

Mossing, M.C., and Sauer, R.T. (1990). Stable, monomeric variants of lambda Cro obtained by insertion of a designed beta-hairpin sequence. Science *250*, 1712–1715.

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., Shirouzu, M., Hashimoto, K., Ishitani, R., Terada, T., Tamakoshi, M., Oshima, T., Chijimatsu, M., Takio, K., Vassylyev, D.G., et al. (2002). An enzyme with a deep trefoil knot for the active-site architecture. Acta Crystallogr. D Biol. Crystallogr. *58*, 1129–1137.

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.

Pleshe, E., Truesdell, J., and Batey, R.T. (2005). Structure of a class II TrmH tRNA-modifying enzyme from *Aquifex aeolicus*. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. *61*, 722–728.

Rajarathnam, K., Sykes, B.D., Kay, C.M., Dewald, B., Geiser, T., Baggiolini, M., and Clark-Lewis, I. (1994). Neutrophil activation by monomeric interleukin-8. Science *264*, 90–92.

Sano, T., Vajda, S., Smith, C.L., and Cantor, C.R. (1997). Engineering subunit association of multisubunit proteins: a dimeric streptavidin. Proc. Natl. Acad. Sci. U.S.A. *94*, 6153–6158.

Schmid, F.X. (1983). Mechanism of folding of ribonuclease A: slow refolding is a sequential reaction via structural intermediates. Biochemistry *22*, 4690–4696.

Shao, X., Hensley, P., and Matthews, C.R. (1997). Construction and characterization of monomeric tryptophan repressor: a model for an early intermediate in the folding of a dimeric protein. Biochemistry *36*, 9941–9949.

Tanford, C. (1968). Protein denaturation, part A: characterisation of the denatured state. Part B: the transition from native to denatured state. Adv. Protein Chem. *23*, 121–282.

Tanford, C. (1970). Protein denaturation, part C: theoretical models for the mechanism of denaturation. Adv. Protein Chem. *24*, 1–95.

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

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

Thoma, R., Hennig, M., Sterner, R., and Kirschner, K. (2000). Structure and function of mutationally generated monomers of dimeric phosphoribosylanthranilate isomerase from *Thermotoga maritima*. Structure *8*, 265–276.

Turnbull, W.B., and Daranas, A.H. (2003). On the value of c: can low affinity systems be studied by isothermal titration calorimetry? J. Am. Chem. Soc. *125*, 14859–14866.

Virnau, P., Mirny, L.A., and Kardar, M. (2006). Intricate knots in proteins: function and evolution. PLoS Comput. Biol. *2*, e122, Epub 2006 Jul 28.

Wagner, J.R., Brunzelle, J.S., Forest, K.T., and Vierstra, R.D. (2005). A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. Nature *438*, 325–331.

Wallace, L.A., and Matthews, C.R. (2002). Sequential vs. parallel protein-folding mechanisms: experimental tests for complex folding reactions. Biophys. Chem. *101–102*, 113–131.

Watanabe, K., Nureki, O., Fukai, S., Ishii, R., Okamoto, H., Yokoyama, S., Endo, Y., and Hori, H. (2005). Roles of conserved amino acid sequence motifs in the SpoU (TrmH) RNA methyltransferase family. J. Biol. Chem. *280*, 10368–10377.

Zarembinski, T.I., Kim, Y., Peterson, K., Christendat, D., Dharamsi, A., Arrowsmith, C.H., Edwards, A.M., and Joachimiak, A. (2003). Deep trefoil knot implicated in RNA binding found in an archaebacterial protein. Proteins. *50*, 177–183.

 ©2007 Elsevier Ltd All rights reserved