

Probing Nature's Knots: The Folding Pathway of a Knotted Homodimeric Protein

Anna L. Mallam and Sophie E. Jackson*

Chemistry Department
Lensfield Road, Cambridge CB2
1EW, UK

The homodimeric protein YibK from *Haemophilus influenzae* belongs to a recently discovered superfamily of knotted proteins that has brought about a new protein-folding conundrum. Members of the α/β -knot clan form deep trefoil knots in their native backbone structure, a topological feature that is currently unexplained in the protein-folding field. To help solve the puzzle of how a polypeptide chain can efficiently knot itself, the folding kinetics of YibK have been studied extensively and the results are reported here. Folding was monitored using probes for changes in both secondary and tertiary structure, and the monomer–dimer equilibrium was perturbed with a variety of solution conditions to allow characterisation of otherwise inaccessible states. Multiphasic kinetics were observed in the unfolding and refolding reactions of YibK, and under conditions where the dimer is favoured, dissociation and association were rate-limiting, respectively. A folding model consistent with all kinetic data is proposed: YibK appears to fold *via* two parallel pathways, partitioned by proline isomerisation events, to two distinct monomeric intermediates. These form a common third intermediate that is able to fold to native dimer. Kinetic simulations suggest that all intermediates are on-pathway. These results provide the valuable groundwork required to further understand how Nature codes for knot formation.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: topological knot; protein folding; chevron plot; multistate kinetics; parallel pathway

*Corresponding author

Introduction

A group of proteins that boast a unique knot deep in their backbone structure have brought to light a new protein-folding problem. The α/β -knot superfamily of methyltransferases (MTases) are characterised by a distinct fold involving the formation of a deep trefoil knot by the backbone polypeptide chain.¹ It is unclear how such knots are able to form during folding, and this impressive topological display remains unexplained by current protein-folding models.^{2–4} To date, a number of α/β -knot superfamily structures have been solved; the most recent include TrmH from *Aquifex aeolicus*, AviRb from *Streptomyces viridochromogenes* and TrmH from *Thermus thermophilus*;^{5–7}

all form homodimers in the native state. Knots have also been identified in proteins other than those belonging to the α/β -knot superfamily of MTases. Examples include the deep figure-of-eight knot in the plant protein acetohydroxy acid isomeroreductase³ and, most recently, the deep trefoil knot contained in the chromophore-binding domain of *Deinococcus radiodurans* phytochrome.⁸

This report focuses on the folding of YibK from *Haemophilus influenzae*. YibK is a 160 amino acid residue protein belonging to the SpoU family of MTases, a subfamily of the α/β -knot superfamily. Crystallographic studies have showed that it possesses a deep trefoil knot at its C terminus, formed by the threading of the last 40 residues (121–160) through a knotting loop of approximately 39 residues (81–120) (Figure 1).⁹ Like other α/β -knot superfamily members, YibK is homodimeric. The dimer interface involves the N-terminal and C-terminal α helices ($\alpha 1$ and $\alpha 5$), and consists of two closely packed monomers arranged in a parallel fashion (Figure 1(b)). Previous studies on YibK have

Abbreviations used: MTase, methyltransferases; SEC, size-exclusion chromatography.

E-mail address of the corresponding author:
sej13@cam.ac.uk

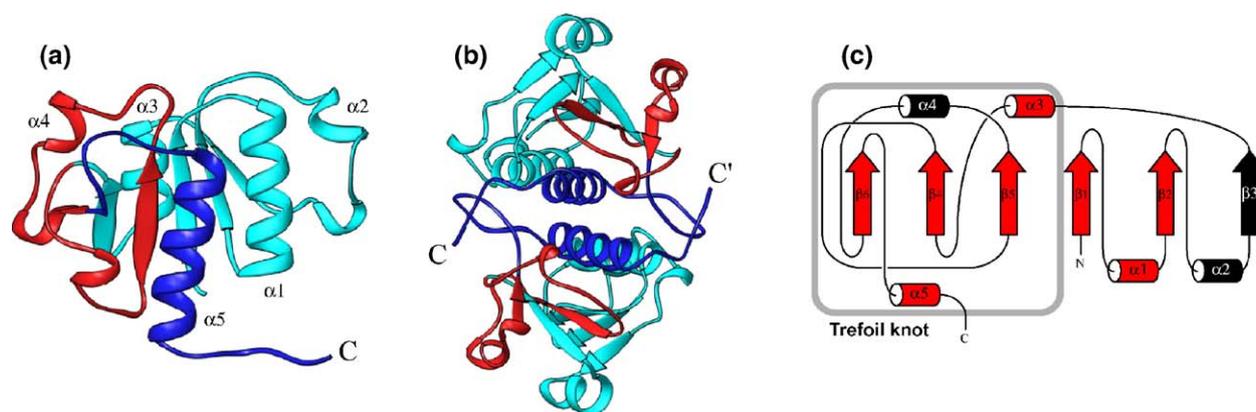


Figure 1. Structure of YibK from *Haemophilus influenzae*. (a) Ribbon diagram of a monomer subunit coloured to highlight the deep trefoil knot at the C terminus according to definitions given by Nureki *et al.*⁴⁰ The knotting loop is coloured red (residues 81–120), while the knotted chain appears dark blue (residues 121–160). (b) Dimeric YibK coloured as in (a). YibK dimerises in a parallel fashion, with $\alpha 1$ and $\alpha 5$ forming the majority of the dimer interface. Ribbon diagrams were generated using Ribbons.⁵¹ (c) Topological diagram of YibK showing the trefoil knot. Structural elements common to the α/β -knot superfamily of methyltransferases are highlighted in red.

shown that, despite its complicated knotted topology, it is able to fold efficiently and reversibly *in vitro* without the aid of molecular chaperones.¹⁰ The mechanism of folding has been determined under equilibrium conditions; remarkably, YibK appears to behave similarly to many unknotted dimers, and folds *via* a partially unfolded monomeric equilibrium intermediate.¹⁰

Small, easily manipulated, monomeric proteins are often the focus of research on protein folding and stability.¹¹ However, an increasing number of studies are now concentrating on the folding pathways of larger, dimeric protein systems. The folding and association steps required to form a dimer can occur by a variety of different mechanisms.¹² Some homodimers show kinetics under most conditions that can be described by a relatively simple cooperative two-state model, and fold to the dimeric native state in the absence of any detectable intermediate. This suggests that all native-like interactions are formed concurrently, and folding and association are simultaneous events. Examples of dimers that fold in this way include P22 Arc repressor^{13,14} and ORF56 from *Sulfolobus islandicus*.¹⁵ Others, such as the dimerisation domain of *Escherichia coli* Trp repressor,¹⁶ FIS¹⁷ and the H2A/H2B histone heterodimer,¹⁸ undergo near diffusion-limited association to form a dimeric intermediate that undergoes a slower unimolecular folding step to form native dimer. Monomeric intermediates have also been observed during folding to a dimer. In contrast to the behaviour of the dimerisation domain alone, full-length *E. coli* Trp repressor forms a burst phase ensemble of partially folded monomers,^{19,20} which then quickly dimerise and undergo isomerisation to the native state *via* three parallel channels.²¹ SecA from *E. coli*, one of the largest dimeric proteins to be characterised, undergoes very rapid dimerisation near the diffusion limit. However, this is preceded by the

formation of a monomeric intermediate observed on the microsecond time-scale. SecA goes on to fold *via* two parallel channels with sequential intermediates.²² The dimeric β -barrel domain E2C from human papillomavirus forms a stable, compact monomeric intermediate in the first 100 ms of folding, which then proceeds through two parallel channels to the native dimer.²³

Here, we present an extensive kinetic analysis of the YibK folding pathway. Folding is monitored with probes of both secondary and tertiary structure, and a variety of solution conditions are used to perturb the monomer–dimer equilibrium and allow characterisation of otherwise inaccessible states. Interrupted refolding experiments are used to map the time-course of folding intermediates along the refolding pathway, and a folding model consistent with the observed data is proposed.

Results

Previous studies on the thermodynamic denaturation of YibK revealed a stable, monomeric intermediate with appreciable secondary and tertiary structure populated under equilibrium conditions.¹⁰ Here, we present an extensive analysis of the unfolding and refolding kinetics of YibK. The chemical denaturant urea was used to perturb the equilibrium, and stopped-flow and manual mixing techniques were employed to measure folding rate constants using both intrinsic protein fluorescence and far-UV CD as probes of changes in tertiary and secondary structure, respectively. Interrupted refolding experiments were used to establish the presence of intermediate states along the refolding pathway, while double-jump refolding was used to investigate the presence of multiple unfolded states. A folding model was developed that is consistent with all the observed data.

Kinetic studies on YibK at pH 7.5

Earlier thermodynamic experiments on YibK were undertaken at pH 7.5 in a buffer containing salt and glycerol, and kinetic studies were performed under the same buffer conditions. At this pH, YibK remains dimeric at all experimental concentrations of protein (Figure 2(a)).¹⁰ Typical traces for single-jump unfolding and refolding as monitored by protein

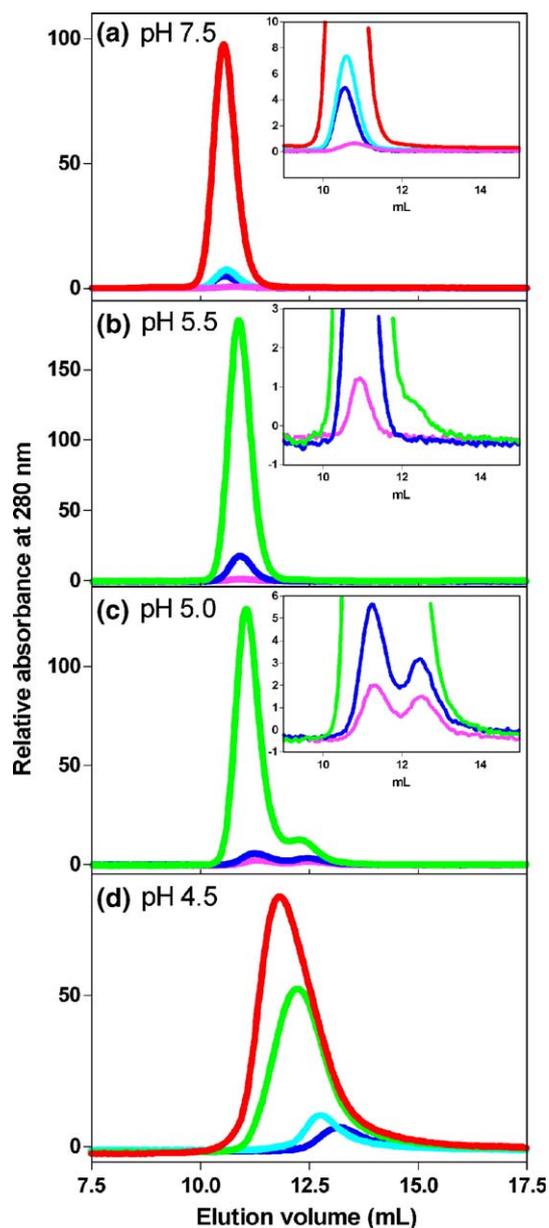


Figure 2. Determination of the oligomeric state of YibK by size-exclusion chromatography. Elution profiles at room temperature for (a) pH 7.5, (b) pH 5.5, (c) pH 5.0 and (d) pH 4.5 for 100 μ M (red), 50 μ M (green), 10 μ M (light blue), 5 μ M (dark blue) and 0.5 μ M (pink) concentrations of protein. Conditions for (a) are 50 mM Tris-HCl, 200 mM KCl, 10% (v/v) glycerol, 1 mM DTT and for (b)–(d) are 50 mM sodium acetate, 150 mM KCl, 1 mM DTT. An example of a calibration curve for YibK can be found elsewhere.¹⁰

fluorescence and far-UV CD are shown in Figures 3 and 4, respectively. Probes of secondary and tertiary structure both displayed an unfolding reaction that was best described by a slow, single, first-order exponential (Figure 3(a) and (b)). In contrast, refolding traces were best fit to a first-order reaction with four exponentials (Figure 4). There was no apparent burst phase in the unfolding or refolding reactions (data not shown). Here, the kinetic phases corresponding to the four observed rates have been numbered 1, 2, 3 and 4 in order from fastest to slowest, and colour-coded, appearing as red, green, light blue and dark blue, respectively. The protein concentration dependence of the four refolding phases was investigated (Figure 4(a), inset (2)), and all appear independent of protein concentration within experimental error. The [urea] dependence of unfolding and refolding rate constants at pH 7.5 was examined using both fluorescence and far-UV CD. The natural logarithm of the rate constants is shown in Figure 5(a) in the form of a chevron plot. The rate constants measured from fluorescence and far-UV CD experiments were in good agreement; therefore, only fluorescence was used to monitor folding in subsequent experiments, due to its enhanced sensitivity. Double-jump unfolding experiments were used to detect faster unfolding phases corresponding to non-native species populated on the refolding pathway.^{24–26} In these experiments, refolding was allowed for short amounts of time before unfolding was initiated to various final concentrations of urea. The results are shown in Figure 3(c)–(e). While the slow unfolding rate is limiting in the unfolding reaction for native dimeric YibK (shown dark blue in Figure 5(a)), a further three unfolding phases were detected during unfolding from partially refolded states. The [urea] dependence of the natural logarithm of the rate constants of each of these additional phases is shown in Figure 5(a), and each one joins smoothly to a refolding arm on the chevron plot. The m_{k_f} and m_{k_u} values were calculated for all four phases at 1 μ M YibK using equation (7), along with the corresponding unfolding and refolding rate constants in the absence of denaturant. The results of these fits are summarised in Table 1.

The pH dependence of the oligomeric state of YibK

YibK is prone to aggregation under certain buffer conditions, but remains soluble when buffered using 50 mM sodium acetate in the pH range 4.5–5.5 for protein concentrations up to at least 100 μ M (data not shown). Samples within this pH range were examined using analytical size-exclusion chromatography (SEC). The results are shown in Figure 2. At pH 7.5 and pH 5.5, YibK elutes in a single peak at a volume of 10.6 ml for all protein concentrations examined. This corresponds to a relative elution volume of 0.20 and a molecular mass of 36.7 kDa.¹⁰ This is close to the calculated mass of 36.803 kDa for a YibK dimer. In contrast, at

pH 4.5, YibK elutes in a single peak at a volume that is protein concentration dependent. For concentrations of protein below 10 μM , YibK has an elution volume of approximately 12.7 ml (Figure 2(d)) corresponding to a relative elution volume of 0.32 and a molecular mass of 17.3 kDa. This is near to the mass of 18.401 kDa expected for a YibK monomer. The elution volume moves closer to the

value expected for a YibK dimer with increasing concentrations of protein (Figure 2(d)). Analytical SEC traces for YibK at pH 5.0 show two peaks at 11.0 ml and 12.3 ml, likely to correspond to YibK monomer and dimer, respectively, and consistent with the presence of a slow monomer–dimer equilibrium (Figure 2(c)).

Equilibrium denaturation curves were measured using fluorescence at pH 5.5 and pH 4.5 over a 400-fold and a 200-fold change in protein concentration, respectively. The results are shown in Figure 6, and data measured at pH 7.5 are included for comparison. At pH 5.5, the equilibrium denaturation profiles exhibit a protein concentration dependence similar to that seen at pH 7.5; the slopes and midpoints increase with increasing protein concentration, consistent with the presence of a monomeric equilibrium intermediate.^{10,27} The trend is more pronounced at pH 5.5 (Figure 6(b)) compared to at pH 7.5 (Figure 6(a)), suggesting that the two transitions, $\text{N}_2 \leftrightarrow 2\text{I}$ and $\text{I} \leftrightarrow \text{D}$, are more separated in urea concentration at this pH. Accordingly, the data were globally fit to a three-state denaturation model involving a monomeric intermediate, and the results of this fit are shown in Table 2 and Figure 6(a) and (b). The $m_{\text{N}_2 \leftrightarrow 2\text{I}}$ and $m_{\text{I} \leftrightarrow \text{D}}$ values at pH 5.5 agree well with those obtained at pH 7.5; 2.0 kcal mol⁻¹ M⁻¹ and 1.5 kcal mol⁻¹ M⁻¹ compared to 1.8 kcal mol⁻¹ M⁻¹ and 1.5 kcal mol⁻¹ M⁻¹, respectively. The free energy of unfolding of dimeric YibK, $\Delta G_{\text{H}_2\text{O}}^{\text{N}_2 \leftrightarrow 2\text{D}}$, at pH 5.5 is significantly less than that at pH 7.5; 18.4 kcal mol⁻¹ compared to 31.9 kcal mol⁻¹.

At pH 4.5, the denaturation curves show no protein concentration dependence at 5 μM YibK and below (Figure 6(c)), and fit well to a two-state monomer denaturation model (equation (2)). The

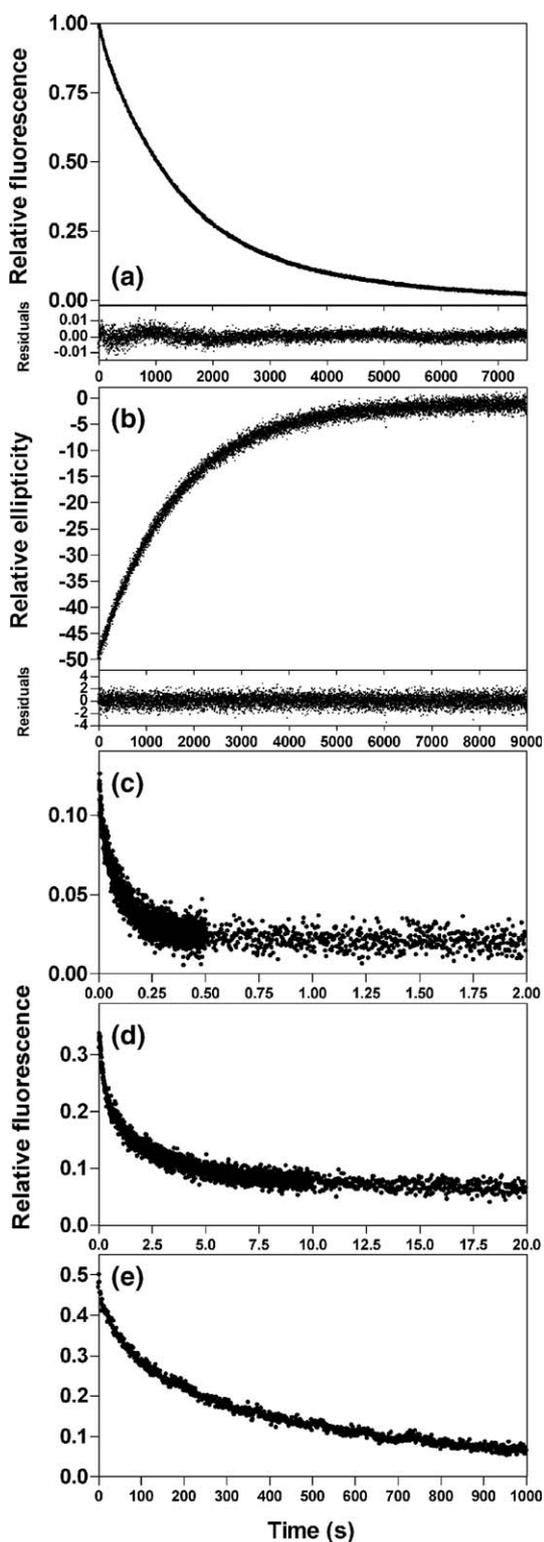


Figure 3. (a) and (b) YibK single-jump unfolding kinetic traces at pH 7.5. (a) Unfolding measured by fluorescence at 6.25 M final concentration of urea and 1 μM YibK. The trace is normalised relative to a denatured signal of 0 and a native dimer signal of 1. Residuals are for a fit of the trace to a first-order reaction with a single exponential. (b) Unfolding measured by far-UV CD at 6.5 M urea and 15 μM YibK. The trace is normalised relative to a denatured signal of 0 and a native dimer signal of -50. Residuals are for a fit of the trace to a first-order reaction with a single exponential. (c)–(e) Interrupted-refolding fluorescence kinetic traces at pH 7.5. Refolding was initiated by dilution to 1 M urea for various delay times before returning to unfolding conditions of 7.15 M urea and 1 μM protein. Traces are normalised as in (a). Delays were used that would populate the refolding species according to the refolding rate constants observed during single-jump experiments. Unfolding traces are shown after (c) 200 ms refolding delay, (d) 5 s refolding delay and (e) 20 s refolding delay. (c) and (d) were measured using a stopped-flow apparatus; (e) was measured using manual mixing techniques in a fluorimeter and fluorescence emission at 319 nm. Traces were fit to equation (3) with the required number of exponentials. Conditions: 25 °C, 50 mM Tris–HCl (pH 7.5), 200 mM KCl, 10% (v/v) glycerol, 1 mM DTT.

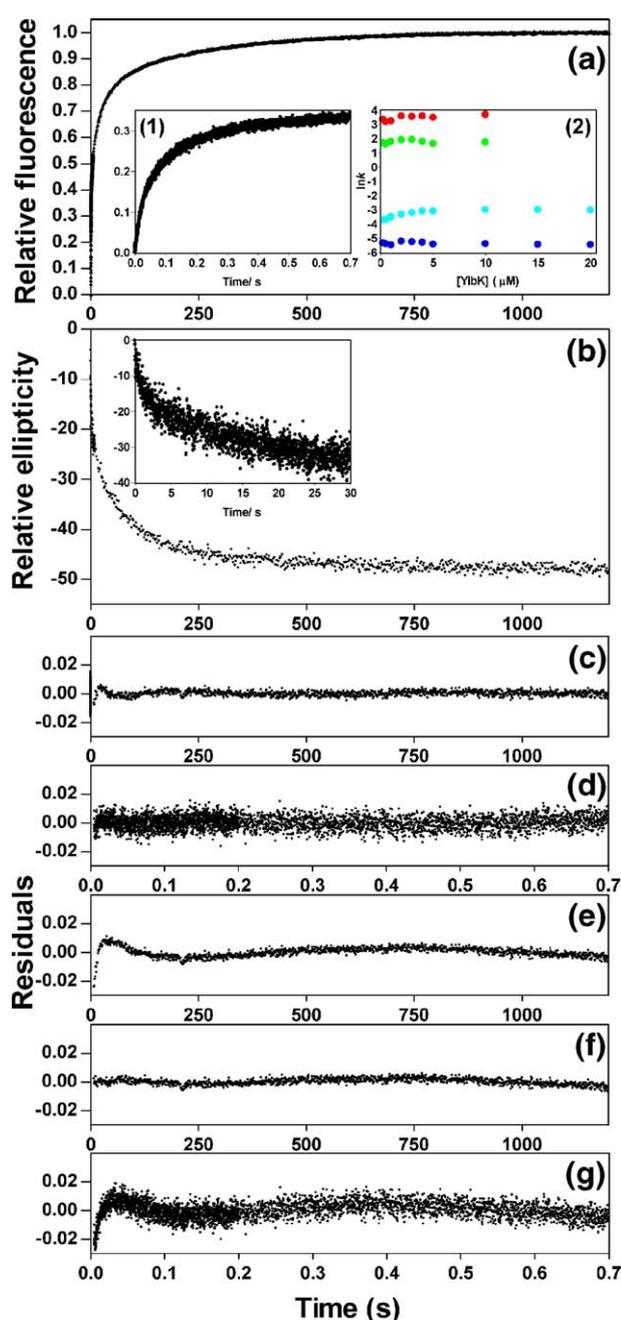


Figure 4. YibK refolding kinetic traces at pH 7.5. (a) Refolding measured by fluorescence at 1 M urea and 10 μ M YibK. Inset (1) shows an expanded view of the first 0.7 s. Inset (2) shows the [YibK] dependence of the four refolding phases, which are coloured red, green, light blue and dark blue in order from fastest to slowest, respectively. Symbols are larger than the errors in the rate constants. (b) Refolding measured by far-UV CD at 3.5 M urea and 15 μ M YibK. The inset shows an expanded view of the first 30 s. Traces are normalised as in Figure 3. Residuals of the fit of the trace in (a) to: (c) A first-order reaction with four exponentials; (d) shows an expanded view of the first 0.7 s. (e) A first-order reaction with two exponentials plus a second-order reaction. (f) A first-order reaction with three exponentials plus a second-order reaction. (g) A first-order reaction with three exponentials with an expanded view of the first 0.7 s. Conditions were as described for Figure 3.

results of this fit are shown in Table 2. The $m_{N \rightarrow D}$ value and $\Delta G_{H_2O}^{N \rightarrow D}$ for 5 μ M and 0.5 μ M YibK agree within error and have an average value of 1.9 kcal mol⁻¹ M⁻¹ and 2.6 kcal mol⁻¹, respectively. Deviation from this is seen at 100 μ M YibK, where the two-state monomer unfolding model can no longer be used to describe the denaturation profile (Figure 6(c)). This suggests that there is significant dimer present at this protein concentration at pH 4.5. Therefore, both SEC and equilibrium denaturation data are consistent with the presence of dimeric YibK at pH 5.5, and a protein concentration dependent ensemble of species at pH 4.5, which becomes predominantly monomeric at low (<10 μ M) concentrations of YibK.

Far-UV CD and fluorescence spectra were recorded for identical concentrations of YibK at both pH 5.5 and pH 4.5 (Figure 7(a) and (b)). Along with a decrease in signal, a shift in emission maximum to 333 nm is observed in the fluorescence spectrum for YibK at pH 4.5, compared to a maximum of 328 nm at pH 5.5. Far-UV CD spectra of YibK at pH 5.5 and pH 4.5 were analysed using the CDSSTR analysis program on the DichroWeb online circular dichroism analysis website to give their secondary structure content.^{28–32} The results are shown in Table 3, and are compared to those calculated from the crystal structure of YibK using PROMOTIF.³³ Analysis of the far-UV CD spectrum at pH 5.5 gives a secondary structure content that agrees well with that expected from the crystal structure (Table 3). In contrast, examination of the spectra at pH 4.5 indicates a loss of helical structure under these conditions; the analysis suggests that only 16 % of the structure is helical compared to 36 % and 35 % at pH 5.5 and from the crystal structure, respectively. Fluorescence and far-UV CD spectra together are consistent with a loss in both YibK tertiary and secondary structure at pH 4.5 relative to pH 5.5.

Kinetic studies on YibK at pH 4.5 and pH 5.5

Unfolding and refolding kinetics at pH 5.5 and pH 4.5 for both 1 μ M and 5 μ M final protein concentrations were monitored using fluorescence, and no obvious burst phase reaction was observed at either pH (data not shown). At pH 5.5, refolding traces showed four phases, similar to pH 7.5. The three fastest phases were best fit to a first-order reaction with three exponentials (equation (3)). The slowest refolding phase, however, was best fit to a second-order reaction (equation (5)) to give an apparent rate constant, k_{app} (data not shown). Unfolding traces were best described by a first-order reaction with a single exponential, comparable to pH 7.5. The [urea] dependence of the natural logarithm of the rate constants at pH 5.5 for both 1 μ M and 5 μ M final YibK concentrations are shown in Figure 5(b). Unlike the rate constants for the three fastest refolding phases, k_{app} displays an obvious dependence on protein concentration, as the refolding arm on the chevron plot for 5 μ M is above that for 1 μ M

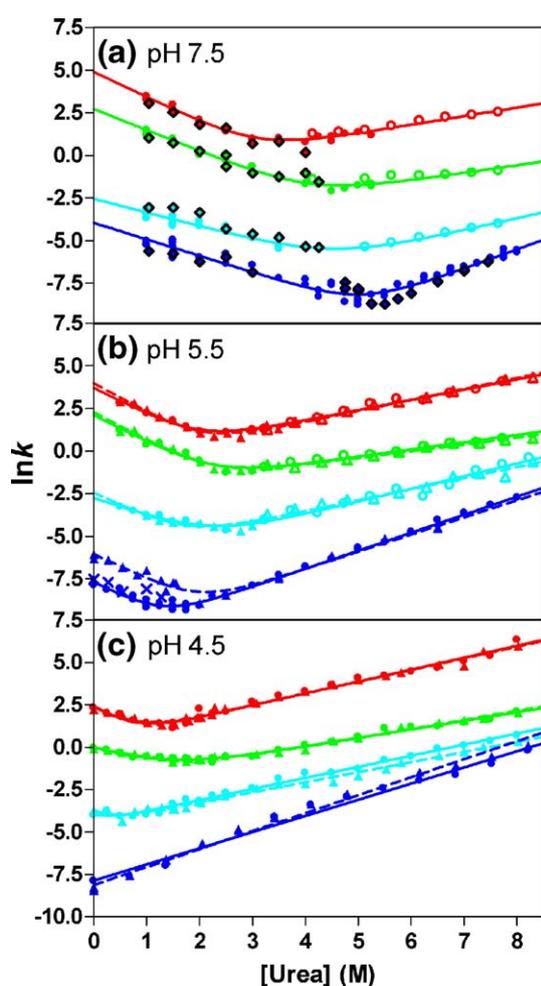


Figure 5. Chevron plots for YibK kinetics at (a) pH 7.5, (b) pH 5.5 and (c) pH 4.5 for 1 μ M (circles and crosses), 5 μ M (triangles) and 15 μ M (diamonds) protein. Single-jump experiments are represented by filled symbols, double-jump experiments at pH 7.5 and unfolding from 2.5–2.75 M urea at pH 5.5 by open symbols, and pH 4.5–5.5 jump experiments by crosses. Phases are coloured as in Figure 4(a), inset (2). All rate constants were measured using fluorescence, except those shown as diamonds with a black outline, which were measured using far-UV CD. All points represent rate constants calculated from a fit to a first-order reaction (equation (3)), except for those on the refolding arm of the slowest phase at pH 5.5, which are calculated from fits to a second-order reaction (equation (5)). The lines denote the fit of each phase to equation (7), or to equation (8) for the slowest unfolding phase at pH 4.5, for 1 μ M (continuous line) and 5 μ M (broken line) protein. Conditions: 25 $^{\circ}$ C, 50 mM Tris-HCl, 200 mM KCl, 10% (v/v) glycerol, 1 mM DTT in (a) and 50 mM sodium acetate, 1 mM DTT in (b) and (c).

(Figure 5(b)). This is expected for a second-order reaction, and equation (6) describes the relationship between k_{app} and k_{2nd} for a second-order process. The slowest (dark blue) phase in the chevron plot at pH 5.5 is therefore assigned to a dimerisation reaction involving the collision of two YibK monomeric species on the basis of its second-order nature and protein concentration dependence.

Figure 5(b) shows a midpoint for the slowest (dark blue) phase at a lower concentration of denaturant than the three faster phases. Unfolding was therefore initiated from 2.5 M and 2.75 M urea for 1 μ M and 5 μ M YibK, respectively; at these concentrations of urea, the species corresponding to the fastest three refolding phases should be populated. The resulting unfolding traces were best fit to a first-order reaction with four exponentials, revealing three further unfolding phases at pH 5.5. The [urea] dependence of the natural logarithm of the rate constants of these phases is shown in Figure 5(b), and they are all protein concentration independent. The unfolding arms of these phases join with the refolding arms of the three fastest refolding phases. The m_{k_f} and m_{k_u} values and rate constants in the absence of denaturant for all four phases were calculated using equation (7), and the values for 1 μ M YibK are shown in Table 1.

At pH 4.5, refolding traces were best fit to a first-order reaction with three exponentials and unfolding traces were described by a first-order reaction with four exponentials (data not shown). A chevron plot showing the [urea] dependence of the unfolding and refolding rate constants at pH 4.5 for both 1 μ M and 5 μ M final YibK concentrations is displayed in Figure 5(c). The arms of the three fastest unfolding phases join smoothly with those of the three refolding phases, while the slowest (dark blue) unfolding phase is present in the absence of denaturant and has no corresponding refolding phase. This is consistent with the assignment of the dark blue refolding phase to dimerisation, as SEC and equilibrium denaturation data at pH 4.5 suggest that there should be little or no dimer at concentrations of protein below 10 μ M, and so a refolding reaction corresponding to dimerisation would not be expected. An unfolding phase involving dissociation of the dimer is observed, as initial concentrations of protein are high enough to expect some dimer to be populated at pH 4.5 (11 μ M and 55 μ M for final YibK concentrations of 1 μ M and 5 μ M, respectively), and the protein is diluted during the unfolding reaction. All phases appear protein concentration independent, as the rate constants for both concentrations of YibK are the same within error (Figure 5(c)). The m_{k_f} and m_{k_u} values, and unfolding and refolding rate constants of the three fastest phases in the absence of denaturant were calculated using equation (7), while the slowest unfolding phase was analysed using equation (8). The results of the fits for 1 μ M protein are summarised in Table 1.

YibK pH-jump kinetics

The pH dependence of SEC and equilibrium denaturation curves for YibK suggest that while the protein remains dimeric at pH 5.5, it exists predominantly in a monomeric form at pH 4.5, at least for concentrations of YibK below 10 μ M. Rapidly changing solution conditions between pH 4.5 and pH 5.5 should therefore provide information on the kinetic transition between monomeric and dimeric

Table 1. Kinetic parameters for the unfolding and refolding of YibK at pH 7.5, pH 5.5 and pH 4.5 for 1 μ M final protein concentration

Phase	Colour	pH	$k_f^{\text{H}_2\text{O}}$ (s^{-1}) ^a	$k_u^{\text{H}_2\text{O}}$ (s^{-1})	m_{k_f} ($\text{kcal mol}^{-1} \text{M}^{-1}$)	m_{k_u} ($\text{kcal mol}^{-1} \text{M}^{-1}$)	$m_{k_{\text{kin}}}$ ($\text{kcal mol}^{-1} \text{M}^{-1}$) ^b	$\Delta G_{\text{H}_2\text{O}}^{\text{kin}}$ (kcal mol^{-1}) ^c
1	Red	7.5	133 \pm 22	0.30 \pm 0.06	0.87 \pm 0.06	0.30 \pm 0.02	1.2 \pm 0.1	3.6 \pm 0.2
		5.5	40.1 \pm 6.4	0.55 \pm 0.08	0.94 \pm 0.09	0.36 \pm 0.02	1.3 \pm 0.1	2.5 \pm 0.1
		4.5	10.1 \pm 2.1	1.4 \pm 0.2	1.2 \pm 0.3	0.42 \pm 0.02	1.6 \pm 0.3	1.2 \pm 0.1
2	Green	7.5	15.1 \pm 2.3	1.5(\pm 0.7) $\times 10^{-2}$	0.73 \pm 0.05	0.27 \pm 0.04	1.0 \pm 0.1	4.1 \pm 0.3
		5.5	8.1 \pm 1.6	0.09 \pm 0.02	0.94 \pm 0.10	0.25 \pm 0.02	1.2 \pm 0.1	2.7 \pm 0.2
		4.5	0.86 \pm 0.08	0.14 \pm 0.02	0.6 \pm 0.1	0.30 \pm 0.01	0.9 \pm 0.1	1.1 \pm 0.1
3	Light blue	7.5	7.7(\pm 1.1) $\times 10^{-2}$	9.0(\pm 7) $\times 10^{-5}$	0.48 \pm 0.05	0.42 \pm 0.08	0.9 \pm 0.1	4.0 \pm 0.10
		5.5	6(\pm 2) $\times 10^{-2}$	1.3(\pm 0.4) $\times 10^{-3}$	0.62 \pm 0.15	0.44 \pm 0.03	1.1 \pm 0.2	2.3 \pm 0.3
		4.5	1.3(\pm 0.4) $\times 10^{-2}$	1.2(\pm 0.1) $\times 10^{-2}$	–	0.39 \pm 0.01	–	0.06 \pm 0.2
4	Dark blue	7.5	1.9(\pm 0.3) $\times 10^{-2}$	4.9(\pm 2.0) $\times 10^{-7}$	0.57 \pm 0.03	0.67 \pm 0.03	1.2 \pm 0.1	14.0 \pm 0.3
		5.5	4.4(\pm 0.4) $\times 10^{-4}$	1.5(\pm 0.1) $\times 10^{-5}$	0.95 \pm 0.08	0.63 \pm 0.01	1.6 \pm 0.1	9.8 \pm 0.1
		4.5	–	3.8(\pm 0.7) $\times 10^{-4}$	–	0.57 \pm 0.02	–	–

Analyses were performed with Prism, version 4 (GraphPad Software) using equation (7) or equation (8). Errors quoted are the standard errors calculated by the fitting program. m_{k_f} and hence $m_{k_{\text{kin}}}$ is not quoted for phase 3 at pH 4.5 as the error from the fit was too large.

^a All rates are first-order, except for phase 4 at pH 5.5, where $k_{\text{app}}^{\text{H}_2\text{O}} = P_1 k_{2\text{nd}}^{\text{H}_2\text{O}}$.

^b $m_{k_{\text{kin}}} = m_{k_f} + m_{k_u}$.

^c $\Delta G_{\text{H}_2\text{O}}^{\text{kin}} = -RT \ln(k_u^{\text{H}_2\text{O}}/k_f^{\text{H}_2\text{O}})$ except for phase 4, where $\Delta G_{\text{H}_2\text{O}}^{\text{kin}} = -RT \ln(2k_u^{\text{H}_2\text{O}}/k_{2\text{nd}}^{\text{H}_2\text{O}})$.

species. Manual mixing techniques were used to jump YibK from pH 4.5 to pH 5.5 over an 80-fold change in protein concentration, and fluorescence was used to monitor any changes in tertiary structure. A typical kinetic trace is shown in Figure 7(c), and there is no sign of a burst phase during the dead-time of manual mixing. Traces fit well to a second-order reaction with a single exponential (Figure 7(d)), and a protein concentration dependence of the resulting second-order rate constant was observed (Figure 7(c), inset). For concentrations of YibK of 7.5 μ M and below, k_{app} varies linearly with protein concentration, as expected for a second-order reaction (equation (6)). The values for k_{app} below 7.5 μ M were fit to equation (6) to give a $k_{2\text{nd}}^{\text{H}_2\text{O}}$ of $4.7 \times 10^2 \text{ s}^{-1} \text{ M}^{-1}$ for the slowest refolding phase of YibK at pH 5.5. Above 7.5 μ M, k_{app} shows a deviation from linearity, consistent with a change in rate-determining step for this phase at higher concentrations of protein.

The [urea] dependence of the second-order transition between pH 4.5 and pH 5.5 was investigated at 1 μ M final concentration of YibK, and the natural logarithm of k_{app} is plotted in Figure 5(b). These points coincide with the slowest phase on the chevron plot of kinetics at pH 5.5, consistent with the assignment of the dark blue phase to a dimerisation step.

Together, the [urea]-jump and pH-jump results provide strong evidence that the slowest phase (dark blue) corresponds to a transition between monomeric and dimeric species on the YibK folding pathway, and accordingly this phase has been assigned to a dimerisation reaction.

YibK interrupted refolding experiments at pH 7.5

Single-jump unfolding experiments on YibK at pH 7.5 show that only one slow phase is observed when unfolding from the native dimer (Figure 3(a) and (b)). However, three further faster phases are

observed when unfolding from states populated after short refolding times (Figure 3(c)–(e)). Interrupted refolding experiments were carried out at pH 7.5 using a large range of refolding aging times to determine the time-course of the population of different species along the YibK refolding pathway, and formation of the native dimer. After various lengths of refolding time, the fraction of each species present is directly proportional to the amplitude of the corresponding unfolding reaction.^{26,34} Interrupted refolding experiments were undertaken for refolding in both 1.04 M urea and 2.5 M urea. Unfolding was initiated after various refolding delay times in 7.67 M urea and a final protein concentration of 1 μ M. Unfolding traces were fit to a first-order reaction with four exponentials, and the resulting amplitudes are shown in Figure 8. The unfolding rate constants observed in the interrupted refolding experiments agreed with the appropriate rate constants on the chevron plot and did not vary with refolding [urea], demonstrating that under both moderate and strong refolding conditions the same intermediate species are populated (data not shown).

Figure 8(a) and (b) show that there is no lag in the formation of species corresponding to phase 1 (red) or phase 2 (green), consistent with these reactions occurring in parallel. These species accumulate during the first 10 s of the refolding reaction, depending upon the concentration of urea, before their population decays. In contrast, the species corresponding to phase 4 (dimer formation, dark blue) shows an obvious lag in its formation before increasing in population to dominate the refolding ensemble. The species corresponding to phase 3 (light blue) shows a shorter lag in its formation, and accumulates during the refolding reaction between 10 s and 100 s, depending upon the concentration of urea, before decaying.

The presence of an essential, on-pathway sequential intermediate in the folding mechanism of a protein gives rise to two hallmarks; the absence of

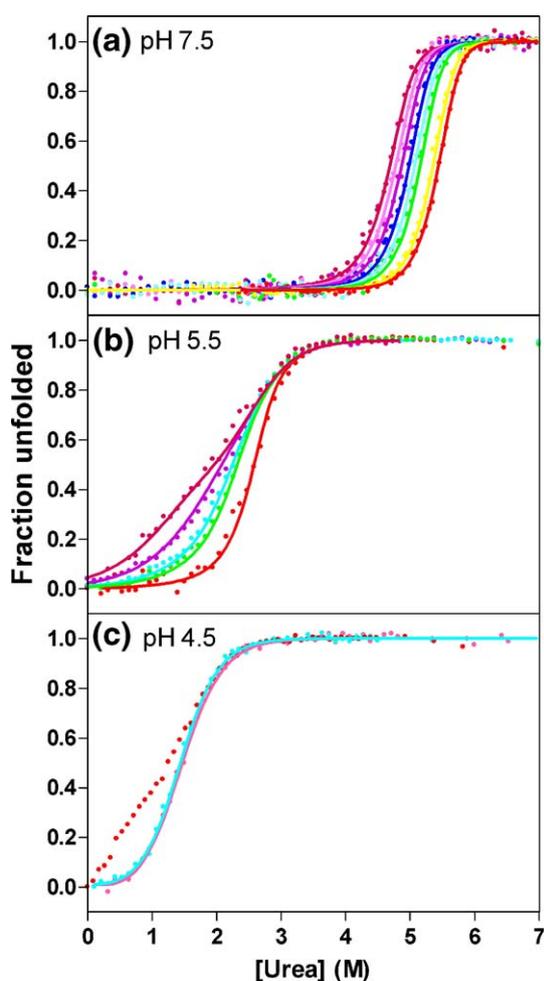


Figure 6. YibK equilibrium denaturation profiles at (a) pH 7.5, (b) pH 5.5 and (c) pH 4.5 for 100 μM (red), 50 μM (yellow), 10 μM (green), 5 μM (light blue), 2.5 μM (dark blue), 1 μM (purple), 0.5 μM (pink) and 0.25 μM (dark red) protein concentrations, as measured by fluorescence emission at 319 nm. Data have been normalised for ease of comparison, and continuous lines represent the global fit to a three-state dimer denaturation model with a monomeric intermediate¹⁰ for (a) and (b), and the separate fits to a two-state monomer denaturation model (equation (2)) in (c). Denaturation data at 100 μM YibK in (c) are not fit to a model. Conditions were as described for Figure 5.

fast track formation of the native state protein that would bypass formation of the intermediate, and a distinct lag phase in formation of the native state, during which time the intermediate accumulates.^{34,35} The lag phase observed in the accumulation of native dimer and in the formation of the species corresponding to phase 3 during YibK refolding is consistent with the presence of an essential intermediate preceding their formation. In addition to this, the species corresponding to phase 1 and phase 2 accumulate in parallel during refolding before decaying. Various models of the YibK folding pathway are possible, taking these constraints into account (Figure 8(c) and 9). Kinetic

simulations of these models were performed using KINSIM,³⁶ and the appropriate rate constants from the chevron plot at pH 7.5 (Figure 5(a)) to replicate the population of species present during the refolding reaction. The results of the simulations are shown in Figures 8 and 9. The model shown in Figure 8(c), involving intermediates from two parallel pathways (I_1 and I_2) folding *via* a third sequential intermediate (I_3) to form native dimer (N_2), is most consistent with the interrupted refolding data. Simulations of other models involving off-pathway species and alternative pathways do not describe the data well (Figure 9).

Investigating the denatured state of YibK: interrupted unfolding experiments

Interrupted refolding experiments show that the model most consistent with the kinetic data involves YibK folding *via* parallel pathways (Figure 8(c)). Each YibK monomer contains ten proline residues with one, Pro34, existing in the *cis* conformation in the native state. Upon unfolding, a large number of molecules will isomerise around this peptidyl-prolyl bond to the energetically more favourable *trans* conformation, forming a predominantly non-native ensemble of conformers.³⁷ Interrupted unfolding experiments were used to probe whether non-native-like peptidyl-prolyl bonds in the denatured state cause the parallel pathways to I_1 and I_2 . These require the protein to be unfolded faster than isomerisation events take place, therefore acid unfolding of YibK from pH 7.5 to pH 1.5 was used, as it occurs in less than 25 ms (data not shown). After various delay times, the protein was diluted back to refolding conditions at pH 4.5. Refolding at this pH was easiest to follow using just a stopped-flow mixing apparatus, and acid refolding back to pH 7.5 caused protein aggregation (data not shown). The resulting refolding traces were fit to a first-order reaction with three exponentials using the rate constants on the chevron plot (Figure 5(c)) and the amplitudes from the fits are shown in Figure 10. Figure 10 shows that after very short delay times, before any peptidyl-prolyl bond has had chance to isomerise from the native isomeric form, all molecules fold *via* I_2 . The population of molecules folding *via* I_2 decreases with increasing unfolding delay time with a rate constant of 0.05 s^{-1} , while the population of molecules folding *via* I_1 increases from zero at the same rate. This suggests that folding *via* I_2 occurs from a denatured state where all peptidyl-prolyl bonds are in their native state conformation, and folding *via* I_1 occurs from a denatured state where one or more proline residues has isomerised to a non-native-like conformation.

Discussion

The homodimeric protein YibK is one of an extraordinary group of proteins that fold to form a trefoil knot deep in their native backbone structure.

Table 2. Thermodynamic parameters for YibK fluorescence equilibrium unfolding data at pH 7.5, pH 5.5 and pH 4.5

pH	[YibK] (μM)	Y_1	$\Delta G_{\text{H}_2\text{O}}^{\text{N}_2 \leftrightarrow 2\text{I}}$ (kcal mol^{-1})	$m_{\text{N}_2 \leftrightarrow 2\text{I}}$ ($\text{kcal mol}^{-1} \text{M}^{-1}$)	$\Delta G_{\text{H}_2\text{O}}^{\text{I} \leftrightarrow \text{D}}$ (kcal mol^{-1})	$m_{\text{I} \leftrightarrow \text{D}}$ ($\text{kcal mol}^{-1} \text{M}^{-1}$)	$\Delta G_{\text{H}_2\text{O}}^{\text{N}_2 \leftrightarrow 2\text{D}}$ ^a (kcal mol^{-1})	$m_{\text{N}_2 \leftrightarrow 2\text{D}}$ ^b ($\text{kcal mol}^{-1} \text{M}^{-1}$)
7.5 ^c	–	0.61±0.04	18.9±0.4	1.8±0.1	6.5±0.2	1.5±0.05	31.9±1.2	4.9±0.3
5.5 ^d	–	0.39±0.01	11.0±0.03	2.0±0.02	3.7±0.01	1.5±0.01	18.4±0.04	5.0±0.03
4.5 ^e	0.5	–	–	–	2.5±0.1	1.8±0.1	–	–
4.5 ^e	5	–	–	–	2.7±0.1	1.9±0.1	–	–

Data were analysed using the non-linear least-squares fitting program Prism, version 4. Errors quoted are the standard errors calculated by the fitting program.

^a $\Delta G_{\text{H}_2\text{O}}^{\text{N}_2 \leftrightarrow 2\text{D}} = \Delta G_{\text{H}_2\text{O}}^{\text{N}_2 \leftrightarrow 2\text{I}} + 2\Delta G_{\text{H}_2\text{O}}^{\text{I} \leftrightarrow \text{D}}$.

^b $m_{\text{N}_2 \leftrightarrow 2\text{D}} = m_{\text{N}_2 \leftrightarrow 2\text{I}} + 2m_{\text{I} \leftrightarrow \text{D}}$.

^c Taken from Mallam & Jackson.¹⁰

^d YibK denaturation data collected at pH 5.5 for various concentrations of protein were globally fit to a three-state dimer denaturation model with a monomeric intermediate. Y_1 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.

^e Denaturation data collected at pH 4.5 were fit separately for each concentration of protein to a two-state monomer denaturation model (equation (2)). Data collected for 100 μM YibK were not fit to any model. At this pH, $m_{\text{I} \leftrightarrow \text{D}}$ and $\Delta G_{\text{H}_2\text{O}}^{\text{I} \leftrightarrow \text{D}}$ are equivalent to $m_{\text{N} \leftrightarrow \text{D}}$ and $\Delta G_{\text{H}_2\text{O}}^{\text{N} \leftrightarrow \text{D}}$.

The mechanism of formation of such a knot remains unexplained. To complement the earlier thermodynamic characterisation undertaken on YibK,¹⁰ we present a comprehensive analysis of the unfolding and refolding kinetics under various solution conditions, and propose a folding mechanism consistent with the observed data.

The oligomeric state of YibK is pH-dependent

Previous studies on YibK have reported that the protein exists exclusively in its dimeric state.^{9,10} Indeed, all current members of the α/β -knot superfamily of MTases have only ever been observed as dimers.^{6,7,38–40} SEC and equilibrium denaturation studies have been used here to investigate the effect of pH on the oligomeric state of YibK. While YibK remains dimeric at pH 7.5 and pH 5.5, SEC data are consistent with the existence of a slow equilibrium between monomer and dimer at pH 5.0, and a predominantly monomeric ensemble at pH 4.5 at concentrations of protein below 10 μM (Figure 2). Equilibrium denaturation studies at pH 4.5 and pH 5.5 agree with this observation (Figure 6(b) and (c), and Table 2). The denaturation profiles at pH 4.5 for concentrations of protein of 5 μM and below are protein concentration-independent, and are described well by a two-state monomer equilibrium model. A deviation from this is seen at 100 μM YibK (Figure 6(c)), agreeing with SEC data that the dimer can be populated at pH 4.5 at high concentrations of YibK. The equilibrium denaturation profiles at pH 5.5 exhibit a dependence on protein concentration similar to that observed at pH 7.5; specifically, an increase in the apparent m value and transition midpoint with increasing concentration of protein, and consequently are globally fit to a three-state dimer equilibrium denaturation model involving a monomeric intermediate.^{10,27} This trend is more obvious at pH 5.5, indicative of a larger separation of [urea] midpoints for the $\text{N}_2 \leftrightarrow 2\text{I}$ and $\text{I} \leftrightarrow \text{D}$ transitions in the denaturation profiles. The $m_{\text{N}_2 \leftrightarrow 2\text{D}}$ value calculated from this fit of 5.0 kcal

$\text{mol}^{-1} \text{M}^{-1}$ agrees well with the value of 4.9 kcal $\text{mol}^{-1} \text{M}^{-1}$ calculated from data at pH 7.5 (Table 2), providing evidence that the same equilibrium denaturation mechanism is occurring at both pH values. $\Delta G_{\text{H}_2\text{O}}^{\text{N}_2 \leftrightarrow 2\text{D}}$ at pH 5.5 is significantly less than at pH 7.5; 18.4 kcal mol^{-1} compared to 31.9 kcal mol^{-1} , respectively (Table 2). This decrease in stability may be due to the absence of stabilising agents in the buffer at pH 5.5 or to the protonation of side-chains at lower pH that will alter electrostatic interactions within the protein.

A pH dependence on the oligomeric state has been observed for other dimeric proteins, for example, *Trp* repressor from *E. coli* populates an ensemble of partially folded monomers at low pH.²¹ The effects of pH on the stability and dimerisation of dimeric procaspase-3 have been investigated, and the protein becomes mostly monomeric at pH 4.0.⁴¹ In this case, it was thought that acidic residues at the interface caused the dissociation at low pH. YibK has a glutamate residue present in its dimer interface at position 143 that forms an intramolecular salt-bridge with Arg146, and it is possible that Glu143 is responsible for the YibK pH-dependent dissociation.

Four phases are observed in the refolding and unfolding kinetics of YibK

The folding kinetics of YibK have been studied at pH 7.5 using both intrinsic protein fluorescence and far-UV CD as probes of changes in tertiary and secondary structure, respectively. Single-jump refolding experiments using both probes show four refolding phases, all of which appear independent of the concentration of protein at this pH (Figure 4(a), inset (2)). Only one slow unfolding phase is observed during single-jump unfolding from native dimer (Figure 3(a) and (b)) suggesting that this reaction is rate-limiting on the unfolding pathway. The good agreement between fluorescence and far-UV CD data suggest that all four phases are associated with changes in secondary and tertiary structure during folding. Interrupted refolding

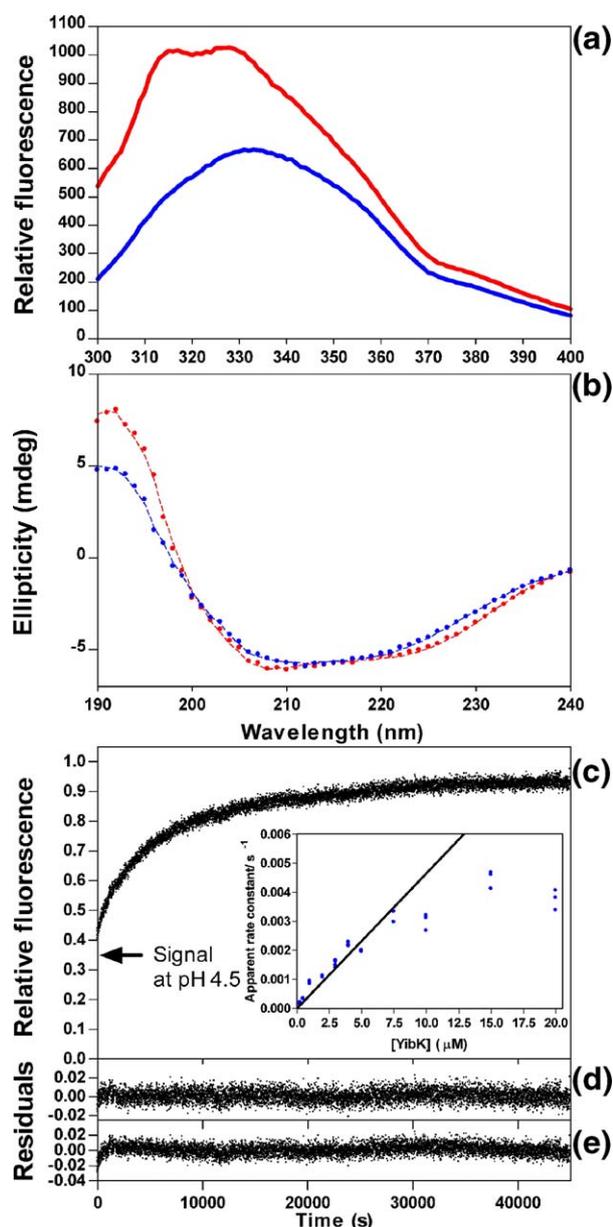


Figure 7. (a) Fluorescence and (b) far-UV CD spectra for YibK at pH 5.5 (red) and pH 4.5 (blue). Conditions: 25 °C, 10 mM sodium acetate, 2.5 μM YibK, 0.1 cm pathlength cuvette. The broken lines in (b) represent the analysis performed by the CDSSTR program on the DichroWeb online circular dichroism analysis website.^{28–32} (c) A typical pH 4.5–5.5 jump trace as measured by fluorescence at 319 nm. Final conditions were 25 °C, 50 mM sodium acetate (pH 5.5), 0.25 μM YibK. The arrow shows the signal expected at pH 4.5. Residuals of the fit of the trace shown in (c) to: (d) a second-order reaction with a single-exponential (equation (5)), and (e) a first-order reaction with one exponential and linear drift. The inset in (c) shows the protein concentration dependence of the apparent rate constant calculated by the fit of the pH-jump data to equation (5). The linear region of this graph was used to calculate a $k_{2\text{nd}}^{\text{H}_2\text{O}}$ value of $4.7 \times 10^2 \text{ M}^{-1} \text{ s}^{-1}$ for the dimerisation step at pH 5.5 from equation (6).

experiments, where YibK was allowed to refold for certain amounts of time before unfolding was initiated, showed a further three faster unfolding phases in addition to the slowest unfolding phase seen in the single-jump experiments (Figures 5(a) and 3(c)–3(e)). These phases are seen only when intermediate species along the refolding pathway are populated.

The [urea] dependence of the unfolding and refolding rate constants has been investigated (Figure 5(a)). All four phases have an unfolding and refolding arm that join smoothly on the chevron plot, suggesting that they each correspond to a reversible reaction.^{34,42}

Kinetic studies have also been undertaken at pH 5.5, where YibK remains dimeric, and at pH 4.5, where a predominantly monomeric ensemble exists (Figure 5(b) and (c)). As at pH 7.5, single-jump kinetic experiments at pH 5.5 show four refolding phases and one slow unfolding phase (Figure 5(b)). However, unlike the three fastest phases (1–3), the slowest refolding phase at pH 5.5 displays a dependence on the concentration of protein, and is described by a second-order reaction. Unfolding from concentrations of urea between 2.5 M and 2.75 M at pH 5.5, where non-native species are populated, allows the detection and characterisation of the three faster unfolding phases shown on the chevron plot.

Single-jump experiments at pH 4.5 show three refolding phases and four unfolding phases, all concentration-independent. The observation of all four unfolding phases during single-jump unfolding experiments implies that intermediate species are present at pH 4.5, as well as a small amount of dimer at the initial concentrations of protein used. Unfolding experiments to 1 μM YibK begin from a protein concentration of 11 μM , where it is likely that a small amount of dimer exists at pH 4.5 (Figure 2(d)). m_{k_1} and m_{k_2} values calculated at pH 5.5 and pH 4.5 agree well with those calculated from the kinetic data at pH 7.5 (Table 1). The total m -value for phase 1 is 1.2 kcal mol⁻¹ M⁻¹, 1.3 kcal mol⁻¹ M⁻¹ and 1.6 kcal mol⁻¹ M⁻¹ at pH 7.5, pH 5.5 and pH 4.5, respectively, and for phase 2 is 1.0 kcal mol⁻¹ M⁻¹, 1.2 kcal mol⁻¹ M⁻¹ and 0.9 kcal mol⁻¹ M⁻¹ at pH 7.5, pH 5.5 and pH

Table 3. Secondary structure analysis of YibK far-UV CD spectra and crystal structure

Conditions	Secondary structure summary (%)		
	β -Strand	Helix ^a	Other ^b
pH 5.5 ^c	14	36	50
pH 4.5 ^c	22	16	62
Crystal structure ^d	16	35	49

^a Includes α -helix and 3_{10} -helix motifs.

^b Includes turns and unordered regions.

^c Secondary structure content from analysis of the far-UV CD spectra shown in Figure 9 using the CDSSTR program on the DichroWeb online circular dichroism analysis website.^{28–32}

^d Secondary structure content computed from the crystal structure of YibK (PDB code 1MXI) using the PROMOTIF program.³³

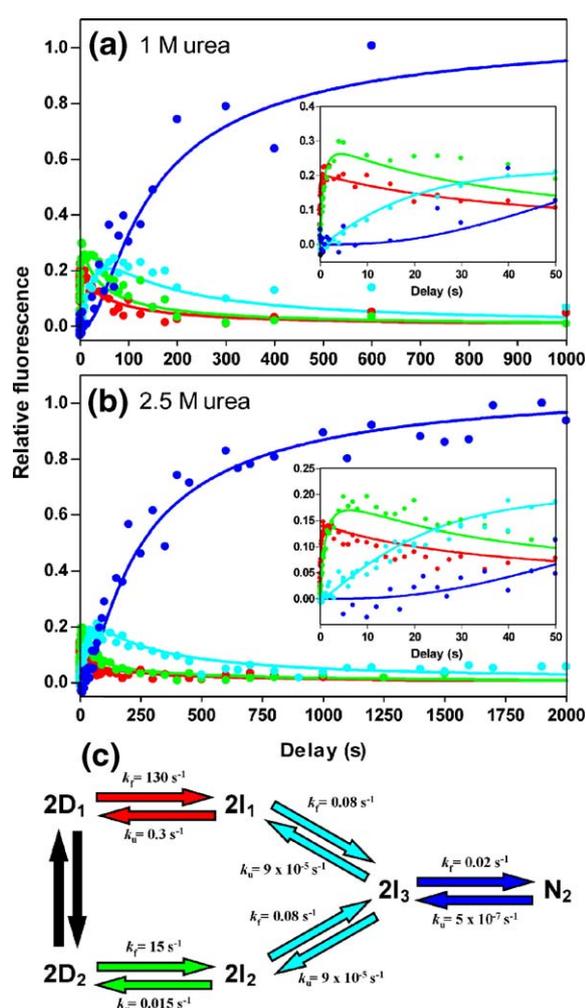


Figure 8. Relative amplitudes of the four YibK unfolding phases seen during interrupted refolding experiments at pH 7.5 after refolding to (a) 1.04 M urea and (b) 2.5 M urea for various delay times followed by unfolding to 7.67 M final concentration of urea and 1 μM final concentration of protein. The insets show an expanded view for delay times up to 50 s. The amplitudes are coloured according to their corresponding phase shown in Figure 5(a). (c) The folding pathway of YibK most consistent with all experimental data. The rate constants are for buffer at 25 $^{\circ}\text{C}$, pH 7.5, and the arrows are coloured to match their corresponding phase as in Figure 5(a). The continuous lines in (a) and (b) represent the KINSIM simulation of the time-course of intermediates and native dimer using the mechanism shown in (c) and the appropriate rate constants from the chevron plot (Figure 5(a)). Conditions were as described for Figure 3.

4.5, respectively. This is very suggestive that the folding mechanisms at pH 7.5, pH 5.5 and pH 4.5 are the same.

The slowest refolding phase corresponds to dimerisation

The slowest refolding phase at pH 5.5 exhibits a protein concentration dependence, and is best

described by a second-order reaction (Figure 5(b)). Consequently, this phase has been assigned to a dimerisation reaction on the folding pathway. The absence of an equivalent slow refolding phase at pH 4.5, where there is negligible dimer at concentrations of protein of 5 μM and below, is consistent with this assignment. Protein concentration dependent refolding reactions have been seen for other dimeric systems, such as the yeast (*Saccharomyces cerevisiae*) prion protein Ure2, P22 Arc repressor and ORF56 from *S. islandicus*, and have been assigned to dimerisation in each case.^{14,15,43} The kinetics of YibK dimer formation have been monitored directly using pH-jump experiments from pH 4.5 to pH 5.5. The [urea] dependence of the protein concentration dependent second-order rate constant observed at 1 μM YibK overlaps with the slow (dark blue) phase seen using [urea]-jump experiments at pH 5.5. This confirms that the slow refolding reaction corresponds to dimerisation, and a $k_{2\text{nd}}^{\text{H}_2\text{O}}$ of $4.7 \times 10^2 \text{ s}^{-1} \text{ M}^{-1}$ at pH 5.5 has been calculated (Figure 7(c), inset). Since only the dark blue phase is seen during unfolding kinetics under dimeric conditions, dissociation must be rate limiting on the unfolding pathway at pH 7.5 and pH 5.5.

The m -value for phase 4 is 1.2 kcal mol⁻¹ M⁻¹ and 1.6 kcal mol⁻¹ M⁻¹ at pH 7.5 and pH 5.5, respectively (Table 1). This is in reasonable agreement with the $m_{\text{N}_2 \leftrightarrow 2\text{I}}$ value of 1.8 kcal mol⁻¹ M⁻¹ and 2.0 kcal mol⁻¹ M⁻¹ seen at equilibrium for pH 7.5 and pH 5.5, respectively, calculated from the global fit to a three-state denaturation model with a monomeric intermediate (Table 2). The $\Delta G_{\text{H}_2\text{O}}$ calculated for phase 4 shown in Table 1 also agree well with the equilibrium $\Delta G_{\text{H}_2\text{O}}^{\text{N}_2 \leftrightarrow 2\text{I}}$, 14 kcal mol⁻¹ and 10 kcal mol⁻¹ compared to 19 kcal mol⁻¹ and 11 kcal mol⁻¹ for pH 7.5 and pH 5.5, respectively. This agreement suggests that the slowest dark blue refolding phase could be due to the folding and dimerisation of the partially unfolded monomeric intermediate seen in YibK equilibrium denaturation studies.¹⁰

In contrast to pH 5.5, the dimerisation phase at pH 7.5 does not exhibit an obvious protein concentration dependence, and fits well to a first-order reaction (Figure 4). This implies that dimerisation is limited by a conformational change at pH 7.5, rather than a collisional event, and hence becomes a first-order process. The deviation from linearity observed for k_{app} at pH 5.5 at a concentration of YibK above 7.5 μM is consistent with the observation that dimerisation is not always rate limiting, and suggests a change in rate-limiting step for phase 4 at higher concentrations of protein at pH 5.5 (Figure 7(c), inset). A similar deviation from linearity has been seen in other dimeric systems, such as the P22 Arc repressor and ORF56 from *S. islandicus*.^{14,15} Interestingly, the rate constant for dimerisation on the folding pathway of YibK at pH 5.5 is much smaller than that observed for other dimeric systems,^{14-16,21-23} and is six orders of magnitude below the diffusion

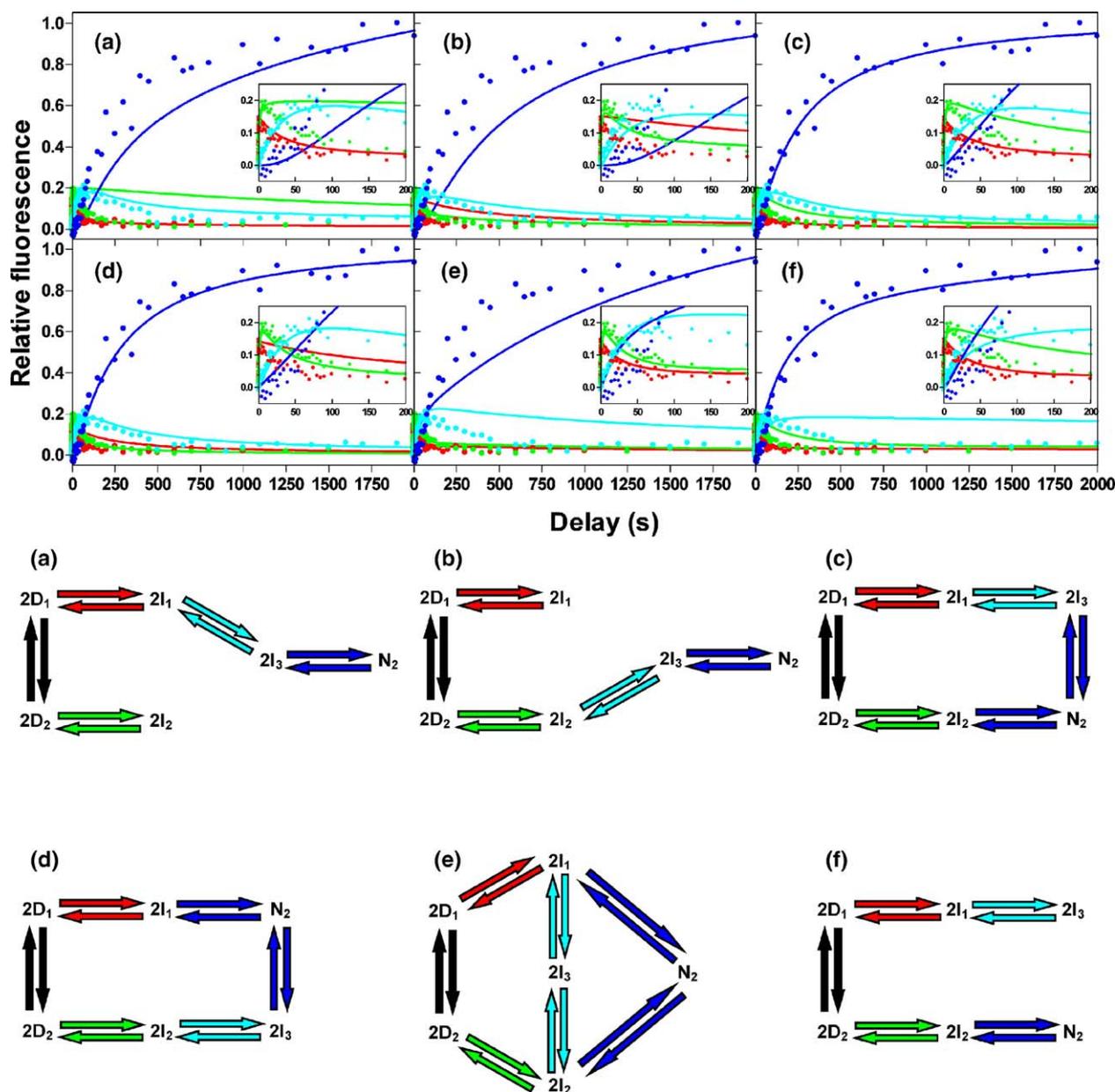


Figure 9. Kinetic simulations performed using KINSIM to model the time-course of the population of intermediates during the refolding of YibK to 2.5 M urea compared to the data shown in Figure 8(b). The continuous lines in graphs (a)–(f) represent the simulations performed using the mechanisms shown in models (a)–(f), respectively. Graphs and mechanisms are coloured as described for Figure 8.

limit of 10^8 – 10^9 $M^{-1} s^{-1}$. This implies that association is not diffusion limited.

YibK folds via a sequential mechanism involving parallel pathways

Kinetic [urea]-jump and pH-jump experiments have led to the assignment of phase 4, the slow, dark blue phase on the chevron plot, to a reaction involving the dimerisation of YibK. The protein concentration independence of the three fastest refolding phases at all pH values suggests that they correspond to the formation of monomeric species. If all phases related to on-pathway, sequen-

tial intermediates, then the total $m_{N_2 \rightarrow 2D}$ value and $\Delta G_{H_2O}^{N_2 \rightarrow 2D}$ for all phases together at pH 7.5 would be $7.4 \text{ kcal mol}^{-1} M^{-1}$ and $37.4 \text{ kcal mol}^{-1}$, respectively, calculated using the following equations:

$$\Delta G_{H_2O}^{N_2 \rightarrow 2D} = 2[\Delta G_{H_2O}^{\text{Phase 1}} + \Delta G_{H_2O}^{\text{Phase 2}} + \Delta G_{H_2O}^{\text{Phase 3}}] + \Delta G_{H_2O}^{\text{Phase 4}}$$

and

$$m_{N_2 \rightarrow 2D} = 2[m^{\text{Phase 1}} + m^{\text{Phase 2}} + m^{\text{Phase 3}}] + m^{\text{Phase 4}}$$

These values are significantly higher than those calculated from equilibrium studies, which give an

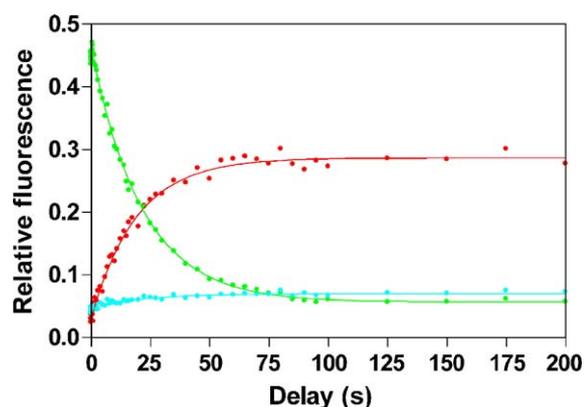


Figure 10. Relative amplitudes of the three refolding phases seen at pH 4.5 from interrupted unfolding traces for various unfolding delay times. Acid unfolding was initiated by jumping from pH 7.5 to pH 1.5, before dilution to refolding conditions at pH 4.5. Refolding traces were fit to a first-order reaction with three exponentials using the appropriate rate constants from the chevron plot at pH 4.5. Amplitudes are coloured according to their corresponding phase shown in Figure 5(c), and the continuous lines represent the fit to a first-order single exponential, which gave a rate constant of 0.05 s^{-1} .

$m_{N_2 \rightarrow 2D}$ value of $4.9 \text{ kcal mol}^{-1} \text{ M}^{-1}$ and a $\Delta G_{\text{H}_2\text{O}}^{N_2 \rightarrow 2D}$ of $31.9 \text{ kcal mol}^{-1}$ at pH 7.5 (Table 2).¹⁰ Therefore, it is likely that at least two phases correspond to reactions occurring in parallel.³⁴

Interrupted refolding experiments were performed to follow the time-course of the population of intermediates and the formation of native dimeric YibK along the refolding pathway at pH 7.5. The population of each species present after various refolding times is proportional to the amplitude of the corresponding unfolding phase. The rate constants of unfolding to 7.67 M urea were the same after interrupted refolding to both 1 M and 2.5 M urea, suggesting that the same intermediate species are present under strong and moderate refolding conditions. The population of the two fastest forming species (phase 1 and 2, red and green, respectively) increases with no observable lag, suggesting that they correspond to parallel pathways. A lag is seen in the formation of the species corresponding to the light blue phase (phase 3) and in the formation of native dimer (phase 4, dark blue) (Figure 8). This is consistent with them being sequential species on the YibK folding pathway, preceded by an obligatory intermediate; there is no fast track route to their formation.^{26,35} Simulations performed with KINSIM using these constraints and the rate constants for refolding and unfolding for each phase obtained from the chevron plot show that the folding model shown in Figure 8(c) is the most consistent with the data (Figure 8); two intermediates, I_1 and I_2 , formed by parallel pathways fold via a third obligatory sequential monomeric intermediate, I_3 , to form native dimer, N_2 . Simulations not including I_3 as an obligatory intermediate in the formation of N_2 or

involving I_1 , I_2 or I_3 as off-pathway species do not describe the data well (Figure 9). Complex mechanisms involving parallel and sequential intermediates have been observed for other dimeric systems, for example, full-length *Trp* Repressor and SecA, both from *E. coli*.^{21,22}

The total kinetic $m_{N_2 \rightarrow 2D}$ and $\Delta G_{\text{H}_2\text{O}}^{N_2 \rightarrow 2D}$ for folding at pH 7.5 via the mechanism shown in Figure 8(c) is $5.4 \text{ kcal mol}^{-1} \text{ M}^{-1}$ and $29.2 \text{ kcal mol}^{-1}$ for the red pathway through I_1 , and $5.0 \text{ kcal mol}^{-1} \text{ M}^{-1}$ and $30.2 \text{ kcal mol}^{-1}$ for the green pathway through I_2 , respectively. These values agree well with the equilibrium $m_{N_2 \rightarrow 2D}$ and $\Delta G_{\text{H}_2\text{O}}^{N_2 \rightarrow 2D}$ values at pH 7.5 of $4.9 \text{ kcal mol}^{-1} \text{ M}^{-1}$ and $31.9 \text{ kcal mol}^{-1}$, respectively (Table 2). The $m_{I_3 \rightarrow D}$ value for the intermediate I_3 is $2.1 \text{ kcal mol}^{-1} \text{ M}^{-1}$ and $1.9 \text{ kcal mol}^{-1} \text{ M}^{-1}$, and the $\Delta G_{\text{H}_2\text{O}}^{I_3 \rightarrow D}$ is $7.6 \text{ kcal mol}^{-1}$ and $8.1 \text{ kcal mol}^{-1}$ calculated for the pathway through I_1 and I_2 , respectively, at pH 7.5. This is similar to the $m_{I \rightarrow D}$ value and $\Delta G_{\text{H}_2\text{O}}^{I \rightarrow D}$ of $1.5 \text{ kcal mol}^{-1} \text{ M}^{-1}$ and $6.5 \text{ kcal mol}^{-1}$ calculated for the equilibrium monomeric intermediate at pH 7.5 (Table 2),¹⁰ suggesting that I_3 may be this intermediate.

The nature of the native state of YibK at pH 4.5

Fluorescence and far-UV CD spectra of YibK show changes in tertiary and secondary structure, respectively, at pH 4.5 relative to pH 5.5 (Figure 7(a) and (b), and Table 3). The loss of helical far-UV CD signal along with a shift to the red in fluorescence emission maximum at pH 4.5 suggest a loss of both tertiary and secondary structure at this pH relative to that of native dimeric YibK. A similar loss in secondary and tertiary structure was seen in the monomeric intermediate observed during equilibrium unfolding at pH 7.5.¹⁰

Four phases are detected during the unfolding of YibK at pH 4.5 (Figure 5(c)). If the model shown in Figure 8(c) for the YibK folding pathway is correct, then I_1 , I_2 , I_3 and N_2 must all be present at pH 4.5 at initial experimental concentrations of protein (11 μM and 55 μM for final concentrations of YibK of 1 μM and 5 μM , respectively) to account for the observation of four unfolding phases; if only I_3 and N_2 were present, their unfolding would be rate-limiting and only two unfolding phases would be seen. Therefore, the unfolding data at pH 4.5 suggest that YibK exists at this pH as an equilibrium ensemble of I_1 , I_2 , I_3 and N_2 for concentrations of protein over 10 μM .

Heterogeneity in the denatured state of YibK gives rise to parallel pathways

YibK has ten proline residues, with one, Pro34, adopting the *cis* conformation in the native structure. This *cis* peptidyl-prolyl bond will isomerise to the energetically more favourable *trans* conformation in the majority of molecules upon unfolding, resulting in a predominantly denatured ensemble of non-native-like proline isomers.^{37,44} Interrupted acid unfolding experiments showed that molecules unfolded for only a short amount of time (25 ms) fold

only *via* I_2 (Figure 10). During these short unfolding delays, peptidyl-prolyl bonds will not have had chance to isomerise, and so folding to I_2 occurs from only a denatured state with all proline isomers in a native-like conformation. The population of molecules folding *via* I_1 increased from zero with increasing time spent in the denatured state with a rate constant of 0.05 s^{-1} , while the population of molecules folding *via* I_2 decreased (Figure 10). This rate constant is within the range observed for proline isomerisation reactions,^{44–46} suggesting that folding to I_1 occurs from a denatured state with non-native-like proline isomers. I_1 and I_2 each have a well-defined unfolding rate constant and slightly different m_{k_f} and m_{k_u} values, suggesting that they are structurally distinct. Also, the folding rate constant of I_1 at all pH values is much faster than the rate constant of 0.05 s^{-1} observed for the proline isomerisation in the acid-denatured state of YibK. This implies that non-native-like proline isomers in the denatured state block folding to I_2 and cause faster folding to a structurally different intermediate I_1 . This is different to what has been observed for a number of proteins where non-native like proline isomers in the denatured state cause a slower parallel reaction to the same species, but limited by a proline isomerisation.^{44,47}

Conclusions

The folding mechanism of YibK has been studied at various pH values using unfolding and refolding single-jump and double-jump experiments. Four reversible folding phases have been observed corresponding to changes in both secondary and tertiary structure. A folding model for YibK consistent with the kinetic data has been proposed: two different intermediates from parallel pathways fold *via* a third sequential monomeric intermediate to form native dimer in a slow rate-limiting dimerisation reaction. All intermediates appear to be structurally distinct and on-pathway, and the parallel channels arise from heterogeneity in the denatured state as a result of proline isomerisation. These findings pave the way for more detailed studies into the role of the deep trefoil knot in the YibK folding mechanism.

Materials and Methods

Molecular biology grade urea was purchased from BDH Laboratory Supplies. All other chemicals were of analytical grade and were purchased from Sigma or Melford Laboratories. Millipore-filtered, double-deionised water was used throughout. YibK purification and expression is described elsewhere.¹⁰

Buffers

All experiments carried out at pH 7.5 were performed in a buffer of 50 mM Tris-HCl, 200mM KCl, 10 % (v/v) glycerol, 1 mM DTT. Experiments undertaken at pH 5.5,

pH 5.0 and pH 4.5 used 50 mM sodium acetate, 1 mM DTT buffered to the appropriate pH, except for far-UV CD scans for analysis by the CDSSTR program on the DichroWeb website, which were performed in 10 mM sodium acetate. Aggregation assays showed that YibK remained soluble under all conditions used (data not shown).⁴⁸

Size-exclusion chromatography

SEC methods for YibK are described elsewhere.¹⁰ YibK samples at various concentrations of protein between 0.5 μM and 100 μM , pre-equilibrated for 3 h in buffer containing 150 mM KCl at pH 5.5, pH 5.0 or pH 4.5, were injected (100 μl) onto an analytical gel-filtration column equilibrated in the same buffer. The relative elution volume was compared to that of molecular mass standards.

Spectroscopic measurements

All measurements were taken using a thermostatically controlled cuvette or cell at 25 °C. For fluorescence studies, an excitation wavelength of 280 nm was used in all experiments. An SLM-Amico Bowman series 2 luminescence spectrometer with a 1 cm path-length cuvette was used for manual mixing kinetic traces and equilibrium denaturation experiments. Fluorescence was monitored at 319 nm with a band pass of 4 nm for both excitation and emission. Far-UV CD spectra and measurements during manual mixing kinetic experiments were acquired with a Jasco J-720 spectropolarimeter. Scans were taken between 190 nm and 240 nm at a scan rate of 1 nm s^{-1} with 40 accumulations using a 0.1 cm path-length cuvette. A 0.3 cm path-length cuvette was used in kinetic experiments, and the change in signal was monitored at 225 nm with a band pass of 5 nm. Rapid mixing fluorescence data were collected using an Applied Photophysics SX.18MV stopped-flow fluorimeter with no cut-off filter, while far-UV CD data were measured using an Applied Photophysics Π^* -180 instrument.

Equilibrium denaturation experiments

Equilibrium denaturation experiments at pH 5.5 and pH 4.5 were performed as described elsewhere.¹⁰ Samples were left for at least 12 h to equilibrate, after which no change in spectroscopic signal was seen.

Measurement and analysis of far-UV CD spectra

Far-UV CD spectra for submission to the CDSSTR analysis program^{28,30,31} on the DichroWeb online circular dichroism analysis website^{29,32} were measured at a protein concentration of 2.5 μM at both pH 5.5 and pH 4.5.

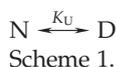
Kinetic unfolding and refolding experiments using fluorescence and far-UV CD

For kinetic unfolding experiments at all pH values, YibK in buffer was diluted 1:10 (v/v) to various concentrations of urea in buffer and the required final concentration of protein. Unfolding was then monitored over an appropriate time period. Kinetic refolding data were collected by taking YibK unfolded in 7.9 M urea for experiments done at pH 7.5, and 5.5 M urea for experiments done at pH 5.5 and pH 4.5, and diluting 1:10 (v/v) to various urea refolding

buffer conditions. Rapid mixing stopped-flow apparatus was used for kinetic traces shorter than 300 s, while manual mixing was used for traces where data were collected for more than 300 s. YibK unfolded in 32 mM HCl (pH 1.5) was diluted sixfold to either pH 5.5 or pH 4.5 to measure refolding rate constants in the absence of urea; dilution to pH 7.5 resulted in aggregation and so a refolding rate constant at 0 M urea was not measured. Interrupted refolding experiments at pH 7.5 were performed by diluting completely unfolded YibK in 6.25 M urea sixfold into final refolding conditions of either 1.04 M urea or 2.5 M urea. After various times, refolding was interrupted by dilution sixfold to the desired unfolding urea concentration and a final concentration of protein of 1 μ M. Interrupted unfolding experiments were performed by diluting YibK in buffer at pH 7.5 sixfold into final conditions of 32 mM HCl (pH 1.5) for various times before refolding was initiated by sixfold dilution to pH 4.5. A pH-jump was initiated by dilution of YibK at pH 4.5 sixfold to final conditions of 50 mM sodium acetate (pH 5.5), 1 mM DTT. At least three traces were averaged for each experiment.

Data analysis

All data analysis was performed using the non-linear, least-squares fitting program Prism, version 4 (GraphPad Software). Equilibrium unfolding curves at pH 5.5 were fit to a three-state dimer denaturation model involving a monomeric intermediate, and analysis has been described elsewhere.¹⁰ Normalised denaturation data collected at pH 4.5 for final concentrations of YibK of 0.5 μ M and 5 μ M were fit to a two-state monomer denaturation model:⁴⁹



$$\Delta G_{N \rightarrow D} = \Delta G_{N \rightarrow D}^{\text{H}_2\text{O}} - m_{N \rightarrow D}[\text{urea}] \quad (1)$$

$$[D] = \frac{([D] + [N]) \exp(\{m_{N \rightarrow D}[\text{urea}] - \Delta G_{N \rightarrow D}^{\text{H}_2\text{O}}\} / RT)}{1 + \exp(\{m_{N \rightarrow D}[\text{urea}] - \Delta G_{N \rightarrow D}^{\text{H}_2\text{O}}\} / RT)} \quad (2)$$

where N is a monomeric species and D is the denatured state of a YibK monomer.

All kinetic traces, except those for the slowest phase observed at pH 5.5, were fit individually to a first-order reaction with the required number of exponentials:

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

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_i is the first-order rate constant for each phase. The slow refolding phase observed at pH 5.5, and the traces from pH 4.5–5.5 jump experiments were fit to a second-order reaction described by the following model:

$$2I \xrightleftharpoons{k_{2\text{nd}}} N_2 \quad d[N_2]/dt = k_{2\text{nd}}[I]^2 \quad (4)$$

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

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

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 $k_{2\text{nd}}$ as follows:

$$k_{\text{app}} = P_t k_{2\text{nd}} \quad (6)$$

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 the concentration of urea is assumed to be linear,^{42,50} and each phase on the chevron plots was fit to:

$$\ln k_{\text{obs}} = \ln(k_f^{\text{H}_2\text{O}} \exp(-m_{k_f}[\text{urea}]) + k_u^{\text{H}_2\text{O}} \exp(m_{k_u}[\text{urea}])) \quad (7)$$

where k_{obs} is the observed rate constant, $k_f^{\text{H}_2\text{O}}$ and $k_u^{\text{H}_2\text{O}}$ are the refolding and unfolding rate constants for each phase in water, respectively, and m_{k_f} and m_{k_u} are constants of proportionality.^{42,50} Phase 4 at pH 4.5 had only an unfolding arm on the chevron plot, and was fit to:

$$\ln k_u = \ln k_u^{\text{H}_2\text{O}} + m_{k_u}[\text{urea}] \quad (8)$$

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

Kinetic simulations

Kinetic simulations to model the time-course of species present during refolding *via* various possible folding mechanisms for YibK were performed using KINSIM³⁶ and the rate constants from the chevron plot.

Acknowledgements

A.L.M. holds an MRC PhD studentship and the work was funded, in part, by the Welton Foundation.

References

- Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S. *et al.* (2004). The Pfam protein families database. *Nucl. Acids Res.* **32**, D138–D141.
- Mansfield, M. L. (1997). Fit to be tied. *Nature Struct. Biol.* **4**, 166–167.
- Taylor, W. R. (2000). A deeply knotted protein structure and how it might fold. *Nature*, **406**, 916–919.
- Taylor, W. R. & Lin, K. (2003). Protein knots: a tangled problem. *Nature*, **421**, 25.
- Mosbacher, T. G., Bechthold, A. & Schulz, G. E. (2005). Structure and function of the antibiotic resistance-mediating methyltransferase AviRb from *Streptomyces viridochromogenes*. *J. Mol. Biol.* **345**, 535–545.
- Nureki, O., Watanabe, K., Fukai, S., Ishii, R., Endo, Y., Hori, H. & 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. & Batey, R. T. (2005). Structure of a class II TrmH tRNA-modifying enzyme from *Aquifex aeolicus*. *Acta Crystallogr. sect. F*, **61**, 722–728.

8. Wagner, J. R., Brunzelle, J. S., Forest, K. T. & Vierstra, R. D. (2005). A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature*, **438**, 325–331.
9. Lim, K., Zhang, H., Tempczyk, A., Krajewski, W., Bonander, N., Toedt, J. *et al.* (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.
10. Mallam, A. L. & Jackson, S. E. (2005). Folding studies on a knotted protein. *J. Mol. Biol.* **346**, 1409–1421.
11. Jackson, S. E. (1998). How do small single-domain proteins fold? *Fold. Des.* **3**, 81–91.
12. Jaenicke, R. & Lilie, H. (2000). Folding and association of oligomeric and multimeric proteins. *Advan. Protein Chem.* **53**, 329–401.
13. Milla, M. E., Brown, B. M., Waldburger, C. D. & Sauer, R. T. (1995). P22 Arc repressor: transition state properties inferred from mutational effects on the rates of protein unfolding and refolding. *Biochemistry*, **34**, 13914–13919.
14. Milla, M. E. & Sauer, R. T. (1994). P22 Arc repressor: folding kinetics of a single-domain, dimeric protein. *Biochemistry*, **33**, 1125–1133.
15. Zeeb, M., Lipps, G., Lilie, H. & Balbach, J. (2004). Folding and association of an extremely stable dimeric protein from *Sulfolobus islandicus*. *J. Mol. Biol.* **336**, 227–240.
16. Gloss, L. M. & Matthews, C. R. (1998). Mechanism of folding of the dimeric core domain of *Escherichia coli* trp repressor: a nearly diffusion-limited reaction leads to the formation of an on-pathway dimeric intermediate. *Biochemistry*, **37**, 15990–15999.
17. Topping, T. B., Hoch, D. A. & Gloss, L. M. (2004). Folding mechanism of FIS, the intertwined, dimeric factor for inversion stimulation. *J. Mol. Biol.* **335**, 1065–1081.
18. Placek, B. J. & Gloss, L. M. (2005). Three-state kinetic folding mechanism of the H2A/H2B histone heterodimer: the N-terminal tails affect the transition state between a dimeric intermediate and the native dimer. *J. Mol. Biol.* **345**, 827–836.
19. Gittelman, M. S. & Matthews, C. R. (1990). Folding and stability of trp aporepressor from *Escherichia coli*. *Biochemistry*, **29**, 7011–7020.
20. Mann, C. J. & Matthews, C. R. (1993). Structure and stability of an early folding intermediate of *Escherichia coli* trp aporepressor measured by far-UV stopped-flow circular dichroism and 8-anilino-1-naphthalene sulfonate binding. *Biochemistry*, **32**, 5282–5290.
21. Gloss, L. M., Simler, B. R. & Matthews, C. R. (2001). Rough energy landscapes in protein folding: dimeric *E. coli* Trp repressor folds through three parallel channels. *J. Mol. Biol.* **312**, 1121–1134.
22. Doyle, S. M., Bilsel, O. & Teschke, C. M. (2004). SecA folding kinetics: a large dimeric protein rapidly forms multiple native states. *J. Mol. Biol.* **341**, 199–214.
23. de Prat-Gay, G., Nadra, A. D., Corrales-Izquierdo, F. J., Alonso, L. G., Ferreira, D. U. & Mok, Y. K. (2005). The folding mechanism of a dimeric beta-barrel domain. *J. Mol. Biol.* **351**, 672–682.
24. Bilsel, O., Zitzewitz, J. A., Bowers, K. E. & Matthews, C. R. (1999). Folding mechanism of the alpha-subunit of tryptophan synthase, an alpha/beta barrel protein: global analysis highlights the interconversion of multiple native, intermediate, and unfolded forms through parallel channels. *Biochemistry*, **38**, 1018–1029.
25. Kiefhaber, T. (1995). Kinetic traps in lysozyme folding. *Proc. Natl Acad. Sci. USA*, **92**, 9029–9033.
26. Schmid, F. X. (1983). Mechanism of folding of ribonuclease A. Slow refolding is a sequential reaction via structural intermediates. *Biochemistry*, **22**, 4690–4696.
27. Hobart, S. A., Meinhold, D. W., Osuna, R. & Colon, W. (2002). From two-state to three-state: the effect of the P61A mutation on the dynamics and stability of the factor for inversion stimulation results in an altered equilibrium denaturation mechanism. *Biochemistry*, **41**, 13744–13754.
28. Compton, L. A. & Johnson, W. C., Jr (1986). Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal. Biochem.* **155**, 155–167.
29. Lobley, A., Whitmore, L. & Wallace, B. A. (2002). DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics*, **18**, 211–212.
30. Manavalan, P. & Johnson, W. C., Jr (1987). Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal. Biochem.* **167**, 76–85.
31. Sreerama, N. & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **287**, 252–260.
32. Whitmore, L. & Wallace, B. A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucl. Acids Res.* **32**, W668–W673.
33. Hutchinson, E. G. & Thornton, J. M. (1996). PROMO-TIF—a program to identify and analyze structural motifs in proteins. *Protein Sci.* **5**, 212–220.
34. Wallace, L. A. & Matthews, C. R. (2002). Sequential vs. parallel protein-folding mechanisms: experimental tests for complex folding reactions. *Biophys. Chem.* **101–102**, 113–131.
35. Heidary, D. K., O'Neill, J. C., Jr, Roy, M. & Jennings, P. A. (2000). An essential intermediate in the folding of dihydrofolate reductase. *Proc. Natl Acad. Sci. USA*, **97**, 5866–5870.
36. Dang, Q. & Frieden, C. (1997). New PC versions of the kinetic-simulation and fitting programs, KINSIM and FITSIM. *Trends Biochem. Sci.* **22**, 317.
37. Brandts, J. F., Halvorson, H. R. & Brennan, M. (1975). Consideration of the Possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry*, **14**, 4953–4963.
38. Ahn, H. J., Kim, H. W., Yoon, H. J., Lee, B. I., Suh, S. W. & Yang, J. K. (2003). Crystal structure of tRNA (m1G37)methyltransferase: insights into tRNA recognition. *EMBO J.* **22**, 2593–2603.
39. Elkins, P. A., Watts, J. M., Zalacain, M., van Thiel, A., Vitazka, P. R., Redlak, M. *et al.* (2003). Insights into catalysis by a knotted TrmD tRNA methyltransferase. *J. Mol. Biol.* **333**, 931–949.
40. Nureki, O., Shirouzu, M., Hashimoto, K., Ishitani, R., Terada, T., Tamakoshi, M. *et al.* (2002). An enzyme with a deep trefoil knot for the active-site architecture. *Acta Crystallog. sect. D*, **58**, 1129–1137.
41. Bose, K. & Clark, A. C. (2005). pH effects on the stability and dimerization of procaspase-3. *Protein Sci.* **14**, 24–36.
42. Tanford, C. (1970). Protein denaturation: C. Part. Theoretical models for the mechanism of denaturation. *Advan. Protein Chem.* **24**, 1–95.
43. Galani, D., Fersht, A. R. & Perrett, S. (2002). Folding of

- the yeast prion protein Ure2: kinetic evidence for folding and unfolding intermediates. *J. Mol. Biol.* **315**, 213–227.
44. Schmid, F. X. (1993). Prolyl isomerase: enzymatic catalysis of slow protein-folding reactions. *Annu. Rev. Biophys. Biomol. Struct.* **22**, 123–142.
 45. Grathwohl, C. & Wuthrich, K. (1981). NMR studies of the rates of proline cis/trans isomerisation in oligopeptides. *Biopolymers*, **20**, 2623–2633.
 46. Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M. & Fischer, G. (1998). Side-chain effects on peptidyl-prolyl cis/trans isomerisation. *J. Mol. Biol.* **279**, 449–460.
 47. Schmid, F. X., Mayr, L. M., Mucke, M. & Schonbrunner, E. R. (1993). Prolyl isomerases: role in protein folding. *Advan. Protein Chem.* **44**, 25–66.
 48. Bondos, S. E. & Bicknell, A. (2002). Detection and prevention of protein aggregation before, during, and after purification. *Anal. Biochem.* **316**, 223–231.
 49. Pace, C. N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280.
 50. Tanford, C. (1968). Protein denaturation. A. Part, Characterisation of the denatured state. B. Part, The transition from native to denatured state. *Advan. Protein Chem.* **23**, 121–282.
 51. Carson, M. (1997). Ribbons. *Methods Enzymol.* **277**, 493–505.

Edited by F. Schmid

(Received 27 January 2006; received in revised form 11 April 2006; accepted 13 April 2006)
Available online 2 May 2006