Supporting Information

Mallam et al. 10.1073/pnas.0912161107

SI Text

Supporting Materials and Methods. ***Creation of circularized species.*** The terminal cysteine mutants of YibK and YbeA in their linear reduced form, $YibK_{SH}^{SH}$ and $YbeA_{SH}^{SH}$, respectively, were shown to have the same structural and functional properties as the wild-type proteins: they are homodimeric as determined by analytical size-exclusion chromatography (SEC), contain the correct secondary structure as shown by far-ultraviolet circular dichroism (far-UV CD) spectra and have similar binding affinities for the methyltransferase cofactor S-adenosylhomocysteine (AdoHcy) (Fig. 2). $YibK_{SH}^{SH}$ and $YbeA_{SH}^{SH}$ were fully reduced [5 mM tris(2-carboxyethyl) phosphine (TCEP)] in either native [50 mM Tris-HCl (pH 7.5), 200 mM KCl, 10% (vol/vol) glycerol] or denaturing [0.1 M Tris-HCl (pH 8.0), 6 M GdmCl] conditions. The reducing agent was then removed by buffer exchange using a HiPrep 26/10 desalting column (GE Healthcare), and the proteins were left to oxidize at 25 or 4 °C (to minimize aggregation) for denaturing and native conditions, respectively. The progress of disulfide-bond formation was monitored by the addition of 5, 5′-dithio-bis(2-nitrobenzoic acid) in the presence of 6 M GdmCl, 0.1 M Tris-HCl (pH 8.0), 1 mM EDTA, and measurement of absorption at 412 nm, otherwise known as Ellman's assay (1). For both conditions, approximately half of the free thiols react within 10 min, followed by a slower disappearance of the remaining sulfhydryl groups. These observed rates of disulfide-bond formation are consistent with faster formation of circularized monomeric molecules and dimeric species with a single disulfide bond, followed by a slower intra- and intermolecular oxidation of these dimeric molecules with remaining free thiols. This model is supported by analysis of the oligomeric species present immediately after the reducing agent is removed, and the observation of an approximately equal ratio of monomeric to oligomeric disulfide-bonded species after oxidation is complete. We have considered the possibility that oxidation can occur in transiently folded species even under highly denaturing conditions. The presence of 6 M GdmCl shifts the folded-unfolded equilibrium far in favor of the denatured protein. Although partially folded species can be transiently populated, analysis of unfolding and refolding rate constants of $YibK_{SH}^{SH}$ and $YbeA_{SH}^{SH}$ demonstrates that such species are rare and very short lived. The observed rate of disulfide-bond formation and the ratio of monomeric: oligomeric species formed strongly suggests that the majority of molecules are circularized in their fully denatured states. When the assay indicated that there were no free sulfhydryls remaining in solution, monomeric and multimeric disulfide-bonded species were separated using denaturing analytical size-exclusion chromatography with a Superdex 75 10/300 column (GE Healthcare) in 8 M urea, 50 mM Tris-HCl (pH 7.5), 200 mM KCl, 10% (vol/vol) glycerol. In addition, this step allowed the circularized proteins to be buffer exchanged into an appropriate urea denaturant for subsequent biophysical characterization. The presence of isolated monomeric circularized protein was verified by mass spectrometry before and after treatment with iodoacetamide, and using the Measure-iT thiol assay kit (Invitrogen). Both techniques confirmed the complete absence of free thiols.

Protein characterization. Experiments on circularized proteins under nonreducing conditions were carried out in 50 mM Tris-HCl (pH 7.5), 200 mM KCl, 10% (vol/vol) glycerol. Experiments in reducing conditions were undertaken in the same buffer but with the addition of 10 mM DTT, except for isothermal titration calorimetry (ITC) experiments where 1 mM TCEP was used. The presence of DTT does not affect the fluorescence signal or the kinetic folding rate constants of the wild-type proteins. The shape of ITC binding isotherms is dependent upon the concentrations of ligand and protein used in the experiment (2); ITC experiments were therefore performed at approximately the same concentration of protein to allow for a more accurate comparison. Analysis of urea denaturation profiles was performed as previously described (3). In these experiments, the change in average emission wavelength was used to monitor unfolding. This ensured that a global unfolding event was measured and reduced the signal-to-noise ratio in the unfolding transition. A global analysis of the kinetic traces for each protein was undertaken with refolding or unfolding kinetic transients at different concentrations of urea for each observable phase considered together and fit to a single equation (4, 5). This assumes a linear dependence of the natural logarithm of the folding and unfolding rate constants on urea concentration (6, 7), an assumption shown to be valid by analysis of separate refolding traces that display no obvious rollover. The exponential reaction for each observable phase at different concentrations of urea outside the transition region was globally fit to

$$Y(t) = Y_{\text{native}} + Y_1 \exp(-k_{\text{obs}}t), \qquad \textbf{[S1]}$$

where k_{obs} is equal to $k_{\text{f}}^{\text{H}_2\text{O}} \exp(-m_{k_{\text{f}}}[\text{urea}])$ or $k_{\text{u}}^{\text{H}_2\text{O}} \exp(m_{k_{\text{u}}}[\text{urea}])$ for refolding and unfolding reactions, respectively, and Y_1 is the corresponding fluorescence amplitude change. The parameters $k_{\text{f}}^{\text{H}_2\text{O}}$ and $k_{\text{u}}^{\text{H}_2\text{O}}$, the refolding and unfolding rate constants in the absence of denaturant, respectively, and $m_{k_{\text{f}}}$ and $m_{k_{\text{u}}}$ values, a measure of the response of the folding or unfolding rate constants to changing urea concentration, respectively, for each phase were shared throughout all datasets. Chevron plots were constructed using the equation (8)

$$\ln k_{\text{obs}} = \ln(k_{\text{f}}^{\text{H}_2\text{O}} \exp(-m_{k_{\text{f}}}[\text{urea}]) + k_{\text{u}}^{\text{H}_2\text{O}} \exp(m_{k_{\text{u}}}[\text{urea}])). \qquad \textbf{[2]}$$

For both wild-type YibK and YbeA, the kinetic phase assigned to the slow dimerization reaction appears protein concentration independent at pH 7.5 (5, 9). The data fit well to a first-order reaction and we do not observe protein concentration-dependent, second-order kinetics. We observed similar first-order kinetics for the mutants studied here. The kinetic step for the dimerization of YibK was assigned previously by performing kinetic experiments under conditions where the protein starts to populate a monomeric state (e.g., low pH) and was further verified by kinetic analysis of the folding of monomeric mutants (10). In an earlier study on YibK, we proposed that, at pH 7.5, dimerization is limited by a conformational change rather than a collision event and therefore becomes a first-order reaction (9). The dimerization step for YbeA was assigned using interrupted refolding experiments and by analysis of m values (5). In contrast to previous studies, here we have calculated the change in free energy associated with the dimerization of YibK and YbeA by assuming that it is rate-limited by a first-order process (Table S3).

To investigate the effect of the equilibration time and temperature of the denatured state on the refolding efficiencies of wild-type YibK and YbeA, samples of denatured protein [0.1 M Tris-HCl (pH 8.0), 6 M GdmCl, 5 mM TCEP] were incubated at temperatures ranging from 25 to 90 °C for durations of 3 to 96 h. Samples were then cooled to room temperature, fully

reduced by the addition of DTT (5 mM) and refolded by buffer exchange or dilution to refolding conditions [50 mM Tris-HCl (pH 7.5), 200 mM KCl, 10% (vol/vol) glycerol, 1 mM TCEP]. The refolding efficiency was assessed by quantifying the presence of native secondary, tertiary, and quaternary structure measured using far-UV CD, fluorescence, and SEC, respectively. The samples were examined for any degradation using SDS-PAGE and mass spectrometry.

Structural figures were made using PyMOL. Knot pictures were generated using KnotPlot.

1. Ellman GL (1959) Tissue sulfhydryl groups *Arch Biochem Biophys* 82:70–77.
2. Turnbull WB, Daranas AH (2003) On the value of c: Can low affinity systems be studied by isothermal titration calorimetry? *J Am Chem Soc* 125:14859–14866.
3. Mallam AL, Jackson SE (2005) Folding studies on a knotted protein *J Mol Biol* 346:1409–1421.
4. Mallam AL, Morris ER, Jackson SE (2008) Exploring knotting mechanisms in protein folding *Proc Natl Acad Sci USA* 105:18740–18745.
5. Mallam AL, Jackson SE (2007) A comparison of the folding of two knotted proteins: YbeA and YibK *J Mol Biol* 366:650–665.
6. Tanford C (1970) Protein denaturation: Part C. Theoretical models for the mechanism of denaturation. *Adv Prot Chem* 24:1–95.
7. Tanford C (1968) Protein denaturation: Part A. Characterisation of the denatured state. Part B. The transition from native to denatured state. *Adv Prot Chem* 23:121–282.
8. Jackson SE, Fersht AR (1991) Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition *Biochemistry* 30:10428–10435.
9. Mallam AL, Jackson SE (2006) Probing nature's knots: The folding pathway of a knotted homodimeric protein *J Mol Biol* 359:1420–1436.
10. Mallam AL, Jackson SE (2007) The dimerization of an alpha/beta-knotted protein is essential for structure and function *Structure* 15:111–122.
11. Wallace LA, Matthews CR (2002) Sequential vs. parallel protein-folding mechanisms: Experimental tests for complex folding reactions *Biophys Chem* 101–102:113–131.

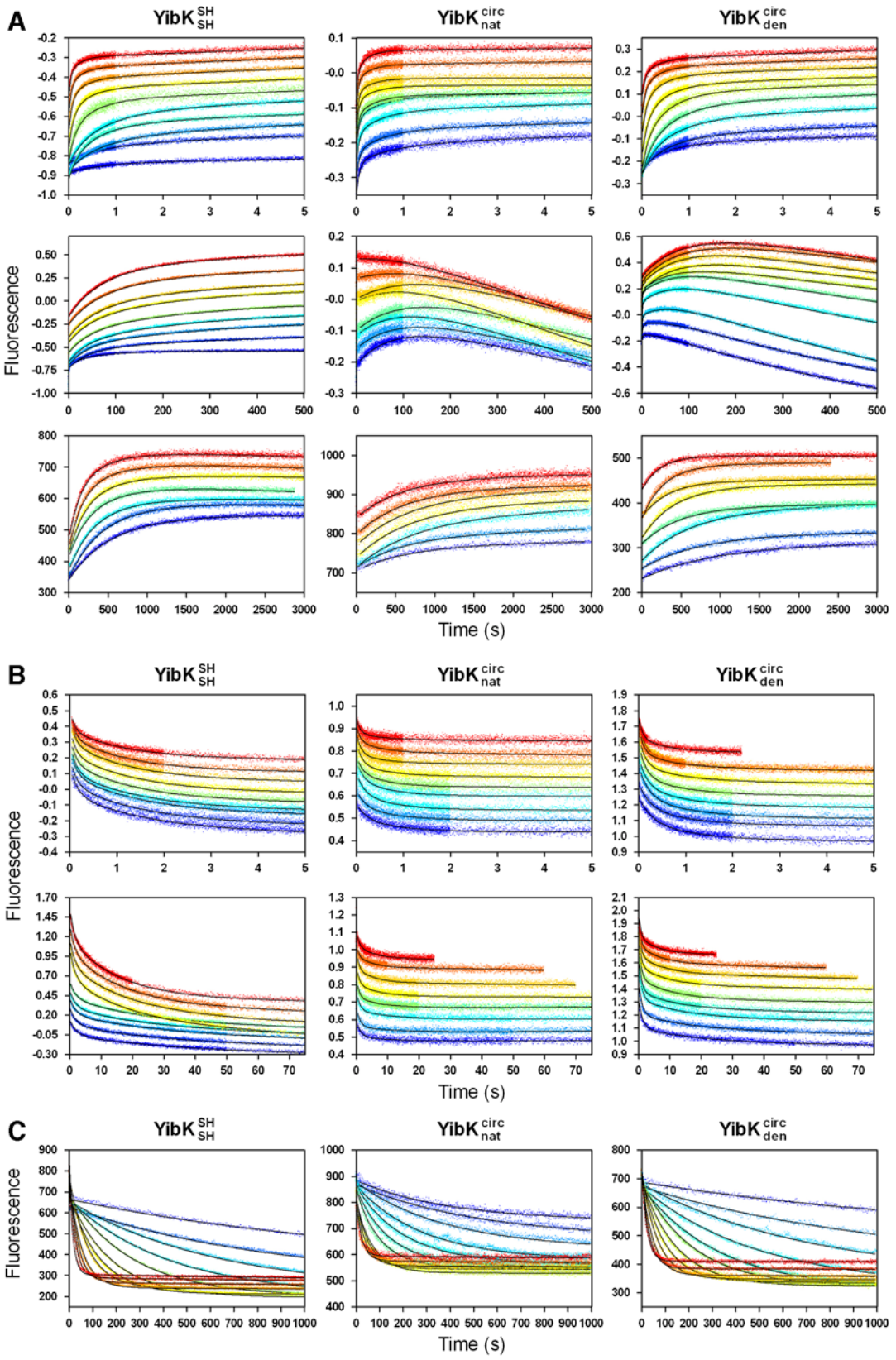

Fig. S1. Global analysis of the refolding and unfolding fluorescence kinetic data for linear and circularized $YibK^{SH}_{SH}$. Data were acquired at 1-μM protein using a combination of a stopped-flow apparatus and a fluorescence spectrometer. Solid black lines represent the global fit of the data to Eq. **1**. (*A*) Single-jump refolding transients (≈0.5–3.4 M final concentration of urea, red to blue). (*B*) Interrupted refolding transients (≈6.8–4.1 M final concentration of urea, red to blue). During these experiments the protein was allowed to refold for various delay times to populate intermediate states. Unfolding was then initiated and the change in fluorescence monitored. (*C*) Single-jump unfolding transients (≈7.1–3.6 M final concentration of urea, red to blue).

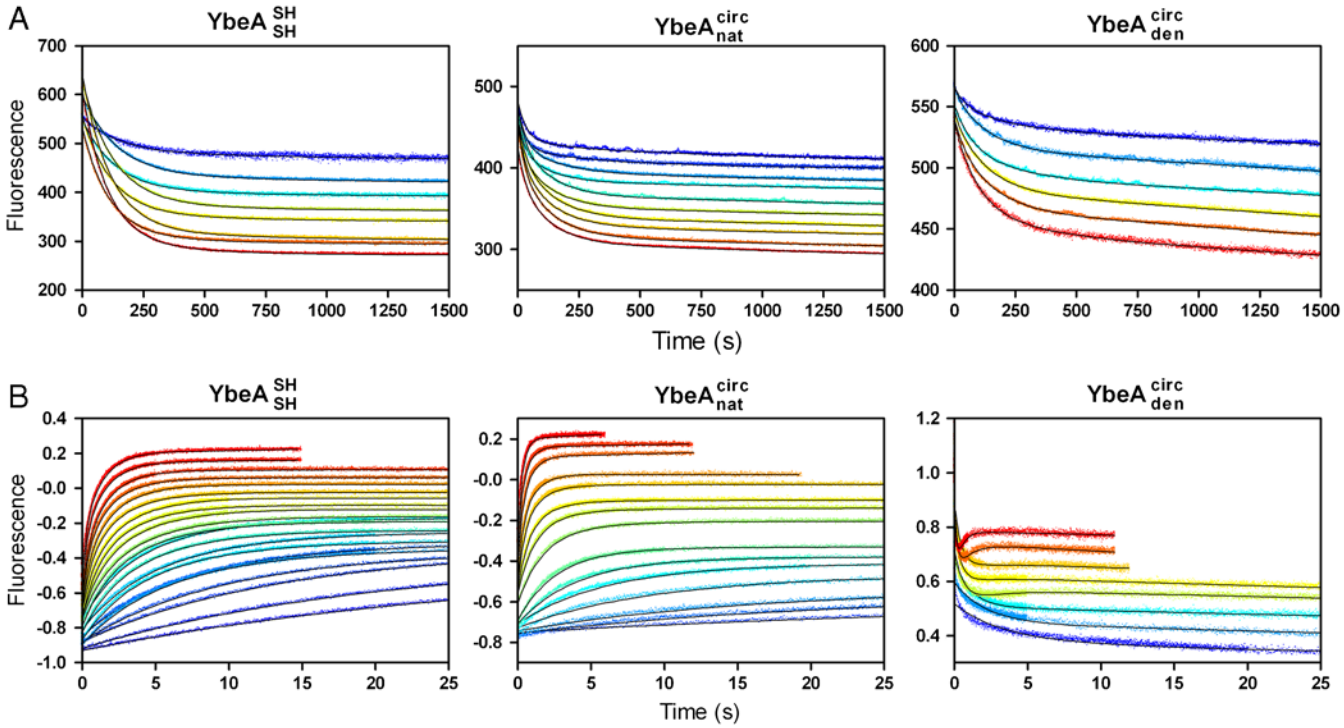

Fig. S2. Global analysis of the refolding and unfolding fluorescence kinetic data for linear and circularized YbeA$^{SH}_{SH}$. Data were acquired at 1-μM protein using a combination of a stopped-flow apparatus and a fluorescence spectrometer. Solid black lines represent the global fit of the data to Eq. **1**. (*A*) Single-jump refolding transients (≈0.3–1.8 M final concentration of urea, red to blue). (*B*) Single-jump unfolding transients (≈7.2–1.8 M final concentration of urea, red to blue).

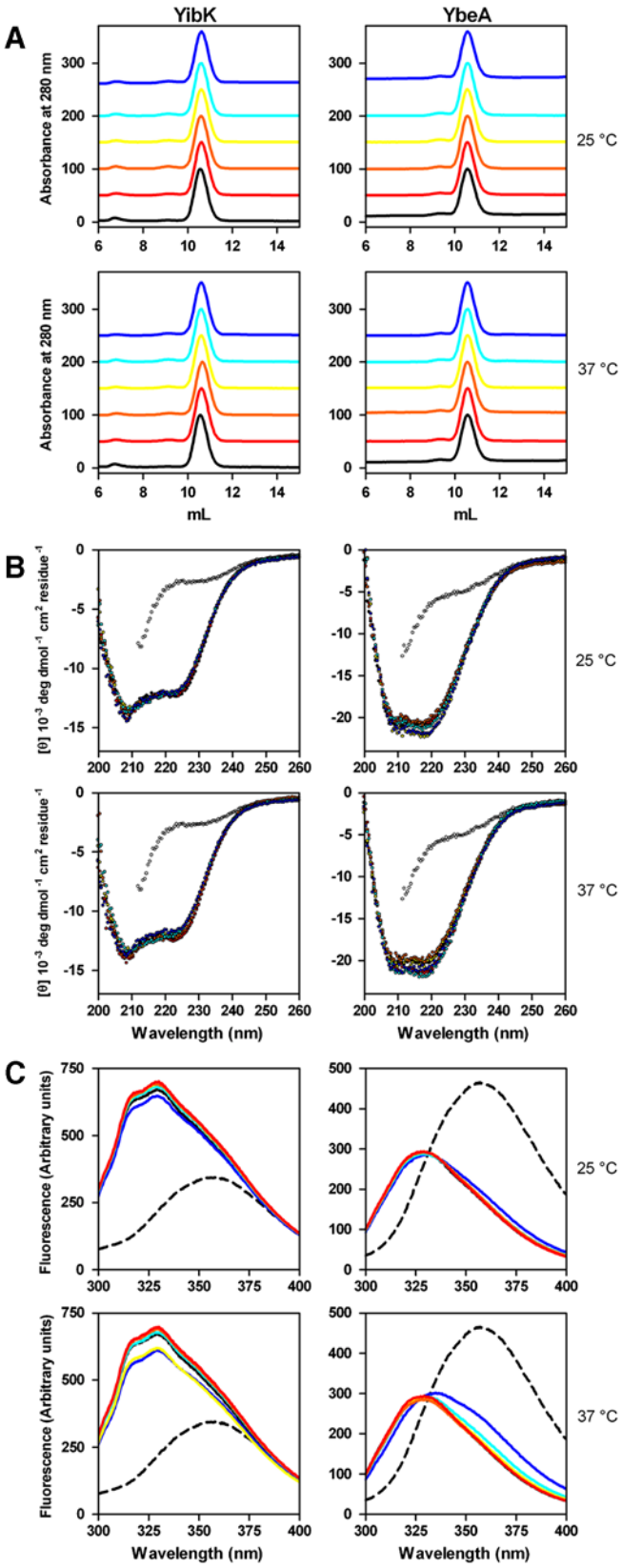

Fig. S3. The effect of the denatured-state equilibration time and temperature on the refolding efficiencies of wild-type YibK and YbeA. The ability of wild-type YibK (*Left*) and YbeA (*Right*) to refold to their native dimeric structures was assessed by (*A*) analytical size-exclusion chromatography, (*B*) far-UV CD, and (*C*) fluorescence spectra as probes of quaternary, secondary, and tertiary protein structure, respectively. Data are shown for refolded samples that were incubated in 6 M GdmCl at 25 °C and 37 °C for 0 h (solid black), 3 h (red), 6 h (orange), 9 h (yellow), 24 h (cyan), and 96 h (blue). The expected elution volume of dimeric YibK and YbeA in *A* is 10.6 mL. Spectra for unfolded proteins in 6 M GdmCl are included in *B* and *C* for comparison (far-UV CD, open black circles; fluorescence, black dashed line).

Table S1. Thermodynamic parameters for the binding of AdoHcy to linear and circularized knotted proteins

Protein	Binding stoichiometry, [AdoHcy]/[protein]	K_d, μM	ΔG_b, kcal mol^{-1}
YibK proteins			
YibK wild type	0.63	20 ± 2.0	−6.4 ± 0.1
$YibK^{SH}_{SH}$	0.5*	74 ± 2.2	−5.6 ± 0.1
$YibK^{circ}_{nat}$	0.5*	71 ± 2.6	−5.7 ± 0.1
$YibK^{circ}_{den}$	0.5*	59 ± 2.3	−5.8 ± 0.1
YbeA proteins			
YbeA wild type	0.17 ± 0.06	2.5 ± 0.3	−7.6 ± 0.1
$YbeA^{SH}_{SH}$	0.19 ± 0.03	2.4 ± 0.2	−7.7 ± 0.1
$YbeA^{circ}_{nat}$	0.17 ± 0.04	1.6 ± 0.2	−7.9 ± 0.1
$YbeA^{circ}_{den}$	0.23 ± 0.02	0.6 ± 0.1	−8.5 ± 0.1

ITC data were analyzed using Origin, version 7, and the errors quoted are the standard errors calculated by the fitting program. Conditions: 50 mM Tris-HCl (pH 7.5), 200 mM KCl, 10% glycerol, 25 °C. 1 mM TCEP was added to the buffer for linear proteins. Concentration of protein in the ITC cell varied between 16–25 μM for YibK proteins and 5–6 μM for YbeA proteins. The free energy of binding was calculated using $\Delta G_b = -RT\ln(1/K_d)$.

*Indicates that the binding stoichiometry was fixed to 0.5, the value calculated previously for wild-type YibK (10), as suggested by Turnbull and Daranas (2) to allow a more accurate determination of the value of K_d, and hence ΔG_b, in low affinity systems. This analysis assumes that all the protein is correctly folded and active, although the calculated values for K_d do not change significantly if the concentration of protein is varied ±15%.

Table S2. Characteristics of equilibrium denaturation profiles for linear and circularized knotted proteins at 1 μM calculated from a fit of the data to a two-state dimer denaturation model

Protein	$[D]_{50\%}$ (M)	m_{app}, kcal mol^{-1} M^{-1}
YibK proteins		
YibK wild type	4.8 ± 0.01	4.4 ± 0.10
$YibK^{SH}_{SH}$	3.2 ± 0.01	2.8 ± 0.06
$YibK^{circ}_{nat}$	3.6 ± 0.02	2.1 ± 0.05
$YibK^{circ}_{den}$	3.3 ± 0.02	2.0 ± 0.08
YbeA proteins		
YbeA wild type	2.7 ± 0.01	3.9 ± 0.01
$YbeA^{SH}_{SH}$	1.2 ± 0.01	2.3 ± 0.05
$YbeA^{circ}_{nat}$	1.4 ± 0.04	0.87 ± 0.03
$YbeA^{circ}_{den}$	1.4 ± 0.06	0.75 ± 0.03

Parameters are quoted with their standard errors. Urea denaturation profiles were analyzed as previously described (3). $[D]_{50\%}$ is the midpoint of the unfolding transition and m_{app} is the apparent m value, a measure of the slope of the transition. In dimer systems such as YibK and YbeA, that actually unfold under equilibrium conditions with a three-state mechanism involving a monomeric intermediate, m_{app} depends upon the concentration of protein and is therefore not equal to $m_{N_2 \leftrightarrow 2D}$ (3, 5). This analysis assumes that all protein molecules are able to correctly refold.

Table S3. Kinetic folding parameters for linear and circularized knotted proteins

Phase	Protein	$k_f^{H_2O}$, s^{-1}	$k_u^{H_2O}$, s^{-1}	m_{k_f}, kcal^{-1} mol^{-1} M^{-1}	m_{k_u}, kcal^{-1} mol^{-1} M^{-1}	m_{kin}, kcal^{-1} mol^{-1} M^{-1}	$\Delta G^{kin}_{H_2O}$, kcal mol^{-1}	$\Delta G^{kin,total}_{H_2O}$, kcal mol^{-1}
YibK proteins								
Very fast D^1 ↔ I^1	YibK wild type	81	0.60	0.69	0.24	0.93	2.9	—
	$YibK^{SH}_{SH}$	41	2.7	0.61	0.13	0.74	1.6	—
	$YibK^{circ}_{nat}$	36	0.76	0.14	0.26	0.40	2.3	—
	$YibK^{circ}_{den}$	28	0.54	0.38	0.28	0.66	2.3	—
Fast D^2 ↔ I^2	YibK wild type	10	0.11	0.62	0.11	0.73	2.7	—
	$YibK^{SH}_{SH}$	7.3	0.17	0.60	0.19	0.79	2.2	—
	$YibK^{circ}_{nat}$	7.1	0.39	0.35	0.15	0.50	1.7	—
	$YibK^{circ}_{den}$	5.5	0.36	0.34	0.13	0.47	1.6	—
Slow I^1 ↔ I^3 I^2 ↔ I^3	YibK wild type	9.5×10^2	5.6×10^4	0.52	0.27	0.79	3.0	—
	$YibK^{SH}_{SH}$	1.0×10^2	1.3×10^2	0.32	0.14	0.46	−0.2	—
	$YibK^{circ}_{nat}$	5.4×10^3	1.0×10^2	−0.21	0.27	0.06	−0.4	—
	$YibK^{circ}_{den}$	1.2×10^2	7.0×10^3	−0.16	0.28	0.12	0.3	—
Very slow 2I^3 ↔ N_2	YibK wild type	1.3×10^2	3.4×10^7	0.42	0.70	1.12	6.3	29.8
	$YibK^{SH}_{SH}$	3.1×10^3	1.8×10^5	0.26	0.68	0.94	3.1	13.4
	$YibK^{circ}_{nat}$	1.9×10^3	8.0×10^5	0.09	0.53	0.62	1.9	11.0
	$YibK^{circ}_{den}$	6.5×10^3	2.0×10^5	0.33	0.64	0.97	3.4	15.2
YbeA proteins								
Slow D ↔ I	YbeA wild type	0.20	1.8×10^3	0.86	0.41	1.27	2.8	—
	$YbeA^{SH}_{SH}$	1.1×10^2	4.6×10^3	0.20*	0.42	0.62	0.5	—
	$YbeA^{circ}_{nat}$	3.7×10^2	4.5×10^3	0.10*	0.51	0.61	1.2	—
	$YbeA^{circ}_{den}$	3.7×10^2	6.3×10^3	0.20*	0.46	0.66	1.0	—
Very slow 2I ↔ N_2	YbeA wild type	4.1×10^2	1.2×10^4	0.61	0.71	1.32	3.5	12.6
	$YbeA^{SH}_{SH}$	4.5×10^3	1.6×10^3	0.72*	0.64	1.36	0.61	2.2
	$YbeA^{circ}_{nat}$	8.7×10^3	2.6×10^3	0.10*	0.68	0.78	0.72	3.8
	$YbeA^{circ}_{den}$	9.5×10^3	9.3×10^4	0.29*	0.74	1.0	1.4	4.8

Data analysis was performed with Prism, version 4 (GraphPad Software); errors have not been shown for the sake of clarity, but are between 1–10%. A detailed description of the kinetic analysis, including the equations used, has been previously described (4, 5, 9). $k_f^{H_2O}$ and $k_u^{H_2O}$ are the rate constants for refolding and unfolding, respectively, in the absence of denaturant and are all first order. m_{k_f} and m_{k_u} are the kinetic refolding and unfolding m values, respectively.

*Indicates that the m_{k_f} values for linear and circularized YbeA proteins could not be determined to a satisfactory degree of accuracy as the refolding arm is too short. $m_{kin} = m_{k_f} + m_{k_u}$; $\Delta G^{kin}_{H_2O} = -RT \ln(k_u^{H_2O}/k_f^{H_2O})$; the total free-energy change associated with the dimeric protein, $\Delta G^{kin,total}_{H_2O} = 2(\sum_{all\ phases} \Delta G^{kin}_{H_2O})$ In previous studies on YibK and YbeA, the change in stability associated with the very slow dimerization reaction was calculated using an apparent second-order refolding rate constant. However, since for both wild-type proteins, this phase appears first order and protein-concentration independent at pH 7.5 (please see *SI Materials and Methods*), the stability change has been calculated here assuming a first-order process. It is important to note that it is difficult to deconvolute the separate energetic contributions of the cooperative association and folding that occurs during this dimerization step. The value calculated is likely to be an underrepresentation of the true $\Delta G^{kin}_{H_2O}$ associated with dimerization because, under the experimental conditions, association appears to be rate limited by a first-order conformational change. A negative m_{k_f} value was calculated for the Slow refolding phase of the circularized YibK proteins. This suggests that, for these proteins, the formation of this intermediate involves an increased exposure of solvent accessible surface area (11).