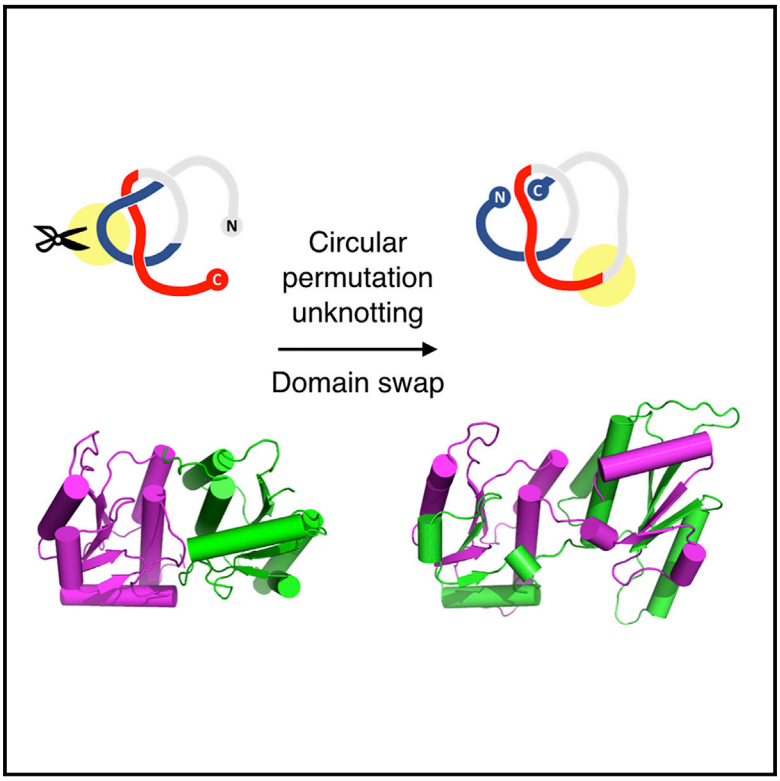
Article

Structure

Untying a Knotted SPOUT RNA Methyltransferase by Circular Permutation Results in a Domain-Swapped Dimer

Graphical Abstract



Highlights

- Circular permutation was introduced to untie the topological knot of YbeA
- The unknotted circular permutant forms a domain-swapped dimer with reduced stability
- Untying the conserved protein knot abrogates cofactor binding essential for function

Authors

Kuang-Ting Ko, I-Chen Hu, Kai-Fa Huang, Ping-Chiang Lyu, Shang-Te Danny Hsu

Correspondence

sthsu@gate.sinica.edu.tw

In Brief

An unknotted circular permutant of YbeA was generated to understand how the conserved knotted topology emerged from evolution. Induction of a circular permutation at the knotting loop relieved the mechanical strain essential for maintaining the cofactor in a high-energy configuration, and resulted in a domain-swapped dimer.

Ko et al., 2019, Structure 27, 1224–1233
August 6, 2019 © 2019 Elsevier Ltd.
https://doi.org/10.1016/j.str.2019.04.004

CellPress

CellPress

Structure
Article

Untying a Knotted SPOUT RNA Methyltransferase by Circular Permutation Results in a Domain-Swapped Dimer

Kuang-Ting Ko,[1,2] I-Chen Hu,[2] Kai-Fa Huang,[1] Ping-Chiang Lyu,[2,3] and Shang-Te Danny Hsu[1,4,5,*]
[1]Institute of Biological Chemistry, Academia Sinica, Taipei 11529, Taiwan
[2]Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu 30013, Taiwan
[3]Department of Medical Sciences, National Tsing Hua University, Hsinchu 30013, Taiwan
[4]Institute of Biochemical Sciences, National Taiwan University, Taipei 10617, Taiwan
[5]Lead Contact
*Correspondence: sthsu@gate.sinica.edu.tw
https://doi.org/10.1016/j.str.2019.04.004

SUMMARY

YbeA from *E. coli* is a trefoil-knotted SpoU-TrmD (SPOUT) RNA methyltransferase. While its knotted motif plays a key functional role, it is unclear how the knotted topology emerged from evolution. Here, we reverse-engineered an unknotted circular permutant (CP) of YbeA by introducing a new opening at the knotting loop. The resulting CP folded into an unexpected domain-swapped dimer. Untying the knotted loop abrogated its function, perturbed its folding stability and kinetics, and induced allosteric dynamic changes. We speculated that the knotted loop of YbeA is under tension to keep the cofactor in a high-energy configuration while keeping the threading C-terminal helix being knotted. Circular permutation released the mechanical strain thereby allowing the spring-loaded threading helix to flip, to relax, and to form a domain-swapped dimer. Being knotted may be the consequence of selection pressure for the unique structure-function relationship of the SPOUT superfamily that exists in all kingdoms of life.

INTRODUCTION

With more than 1,000 knotted structures being deposited in the database, topologically knotted proteins represent a small but significant fraction of the protein topological universe (Dabrowski-Tumanski et al., 2019; Dabrowski-Tumanski and Sulkowska, 2017; Jackson et al., 2017; Jarmolinska et al., 2018; Lim and Jackson, 2015). Unlike random entanglements of long, flexible polymer chains, protein knots adopt well-defined and compact three-dimensional structures. Even in their chemically denatured states, some knotted proteins retain their knotted topologies, while exhibiting the random coil-like characteristics as predicted for well-solvated linear polymers (Flory, 1953; Kohn et al., 2004; Shih et al., 2015). Experimental surveys of the mechanisms underlying the folding of a variety of topologically knotted proteins revealed the presence of highly populated kinetic folding intermediates in all cases (Andersson et al., 2009; Lee and Hsu, 2018; Lou et al., 2016; Mallam and Jackson, 2006; Mallam et al., 2008; Wang et al., 2015a, 2015b, 2016), which contrast the expectation for minimum frustration in the free energy landscapes of robust protein folding systems (Onuchic et al., 1997; Wolynes et al., 1995).

Although it seems counterintuitive for Nature to evolve functionally important proteins to fold into topologically knotted configurations from an energetic point of view, a large number of knotted proteins are enzymes that play key physiological roles. These include trefoil (3_1) knotted SPOUT RNA methyltransferases (Tkaczuk et al., 2007) and Gordian (5_2) knotted ubiquitin C-terminal hydrolases (UCHs) (Hsu, 2016), of which all family members are exclusively knotted. UCH and SPOUT members utilize their knotted structural elements to bind to their physiological substrate and cofactor, namely ubiquitin and S-adenosyl-L-methionine (SAM), respectively, attesting to the functional implications of these knotted motifs that may justify the apparent energetic costs associated with folding (Christian et al., 2016). Note that while UCHs are only present in eukaryotes as part of the ubiquitin proteasome system regulatory system (Hsu, 2016), SPOUT RNA methyltransferases exist in all kingdoms of life with a broad spectrum of sequence diversity. An early bioinformatics survey suggested that the last universal common ancestor (LUCA) of the SPOUT superfamily may be knotted, but the study also proposed that the knotted topology may evolve through circular permutation (CP) (Tkaczuk et al., 2007).

Naturally occurring CPs have been identified as a result of gene duplication (Goldenberg and Creighton, 1983; Lo et al., 2009). Artificial CPs have also been designed to fold into the same three-dimensional structures as their parent proteins (Chiang et al., 2006; Reitinger et al., 2010; Yu and Lutz, 2011). These natural and artificial CPs afford an opportunity to examine the relationship between folding topology—defined as the sequential and spatial arrangements of secondary structural elements—and folding kinetics assuming that the parent proteins and the CPs share the same native contacts and therefore free energies associated with their native states. The folding kinetics is expected to be slower for a protein variant that has more long-range native contacts than its counterpart that has more short-range native contacts, as quantitatively described

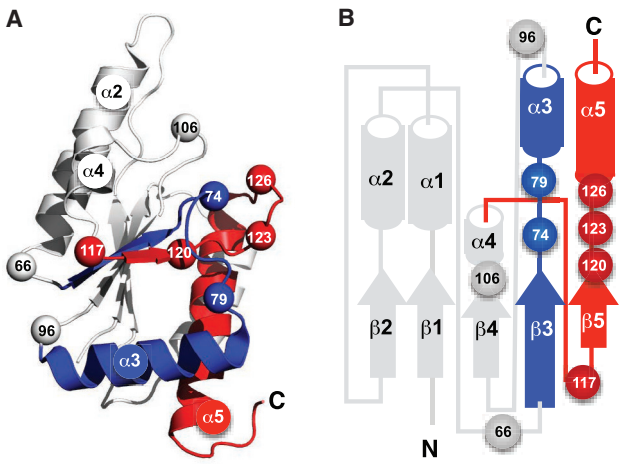


Figure 1. Structure of Trefoil-Knotted YbeA and CP Designs Generated in This Work

(A and B) (A) Cartoon representation and (B) topological representation of the crystal structure of YbeA with the knotting loop and the C-terminal threading loop shown in blue and red, respectively. The remaining part of YbeA is shown in light gray. The CP sites are indicated in spheres with their residue identities indicated.

by the concept of contact order (Plaxco et al., 1998). Nevertheless, there has neither been an example of a CP of a given knotted protein found in Nature, nor an engineered CP based on a knotted protein.

In this study, we successfully designed a well-folded, unknotted CP based on the well-characterized trefoil-knotted YbeA from *E. coli* (Mallam and Jackson, 2007a, 2012; Mallam et al., 2010). Unexpectedly, the CP was found to form a domain-swapped dimer. We quantitatively compared the global folding stability and kinetics of the CP and wild type (WT). We further examined the impact of the circular permutation on the folding dynamics and function of YbeA. Our results demonstrated the ability to reconfigure the topological arrangement of a knotted protein into an unknotted form in a reverse engineering approach. Our findings may shed light on how the knotted LUCA of the SPOUT superfamily emerged through evolution.

RESULTS

To untie the trefoil-knotted YbeA, we constructed nine CPs of YbeA based on the suggestions of the CP design server, CPred (Lo et al., 2012). The CP sites were individually introduced to nine different positions in close proximity to the knotted structural elements. These included residue positions 66, 74, 79, 96, 106, 117, 120, 123, and 126 (Figure 1). Of the nine CPs, only CP74, which has its CP site introduced to the knotting loop through which the long C-terminal helix threads, could be expressed in *E. coli* as a soluble recombinant protein, allowing subsequent purification steps to achieve homogeneity following the same procedures used to purify WT. All other CPs were expressed in inclusion bodies. We were unable to refold these CPs in sufficient quantities for detailed characterizations despite extensive tests of different refolding conditions. We therefore focused on the characterization of CP74 in the following studies.

The far-UV circular dichroism (CD) spectra of WT and CP74 indicated well-folded structure for both variants while that of CP74 appeared to contain more helical content compared with that of WT (Figure S1A). Meanwhile, the intrinsic fluorescence spectrum of CP74 exhibited a red shift of 10 nm—the maximum emission wavelength was shifted from 340 to 350 nm—accompanied by some increase in the emission fluorescence intensity compared with that of WT (Figure S1B), suggesting some structural perturbations around the five endogenous tryptophan residues, namely W15, W79, W91, W120, and W144 in WT, and W7, W19, W48, W72, and W100 in CP74. At a quaternary structure level, size-exclusion chromatography (SEC)-coupled multiangle light scattering yielded the monodispersed molecular weights of 30.16 ± 0.12 and 33.60 ± 0.06 kDa for WT and CP74, respectively, which were comparable with the expectation for a dimeric YbeA (34.68 kDa; Figure S1C).

Having established that CP74 shared comparable secondary, tertiary, and quaternary structures with those of WT, we sought to determine the crystal structure of CP74 to examine in atomic detail the structural differences between WT and CP74. The crystal structure of CP74 was determined and refined to a resolution of 1.75 Å with the aid of mercury-soaking derived single-wavelength anomalous diffraction (the structure was deposited in the PDB under accession number PDB: 5ZYO) (Figure 2A; Table 1). Remarkably, the crystal structure of CP74 revealed a domain-swapped dimer, of which residues 1–49 (corresponding to residues 74–121 in WT; Figure 2B) of one monomer flipped to form extensive intermolecular contacts with the other monomer, while maintaining the overall structure and environment around the swapped structural elements (Bennett et al., 1995): The positional root-mean-square deviation between the Cα atoms of the secondary structure elements of the WT monomer and the corresponding Cα atoms of the half of domain-swapped CP74 is 0.7 Å (Figure S2). The electron density corresponding to the linker residues that bridges the two segments of the same monomer was well resolved (see also Figure S3), and the resulting domain-swapped dimer had an extensive intermolecular surface area of ca. 4,600 $Å^2$ as opposed to ca. 1,100 $Å^2$ in WT. The latter was limited to a small contact surface between two globular monomers (Figure 2C). While the overall domain structure showed limited structural perturbation (Figure S3), CP74 underwent considerable conformational rearrangements in terms of the relative domain orientations and the positioning of the individual α helices. Super-positioning the two crystal structures with respect to one of the two long α helix at the dimer interface (α5 in WT and α3 in CP74; Figure 2D) showed how the two monomers were rearranged with respect to each other as a result of CP and the consequential domain swap. Such a structural alignment gave rise to the largest displacement between α3 in WT and α1 in CP ($\alpha1_{cp}$) by 14 Å, while several other structural elements, as well as the loops, also exhibited major conformational differences. These prominent structural differences also explained why initial attempts by molecular replacement based on the structure of WT failed to solve the phase problem of CP74.

To validate the conformational difference observed in the crystalline states, we carried out synchrotron-based SEC-coupled small angle X-ray scattering (SEC-SAXS) analyses to evaluate the solution structures and dynamics of WT and CP74 (Figure 3). The theoretical SAXS profiles of WT and CP74 were

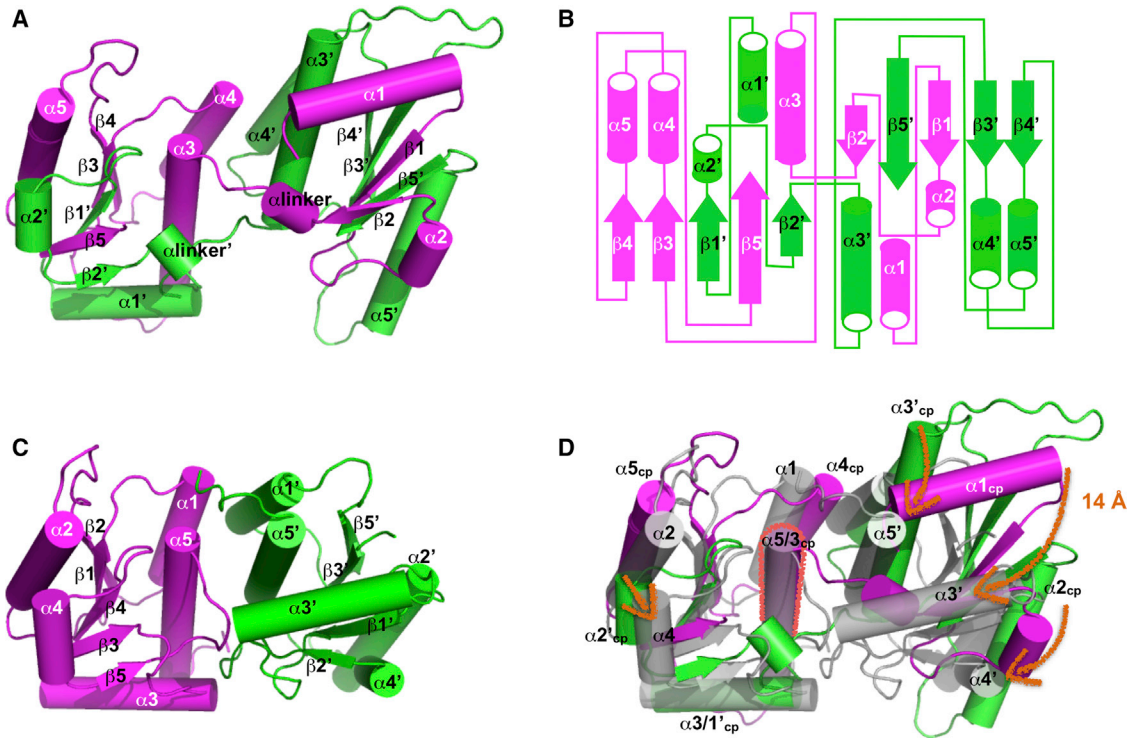


Figure 2. Structural and Topological Comparison of WT and CP74

(A and B) Cartoon representation (A) and topological representation (B) of the crystal structure of CP74 with one monomer shown in green and the other in purple to highlight the extent of domain swapping.

(C) Cartoon representation of the crystal structure of WT (PDB: 1NS5).

(D) Superposition of CP74 and WT (semi-transparent gray) with respect to the C-terminal helix α5 in WT and the corresponding α3 in CP74 (annotated as $\alpha3_{cp}$) highlighted by an orange contour. Pairwise displacements of individual secondary structure elements between WT and CP74 are indicated by orange arrows.

back-calculated from their crystal structures (Schneidman-Duhovny et al., 2016), which showed very good agreement with the experimental SAXS profiles (Blanchet and Svergun, 2013). Furthermore, the radii of gyration (R_g) derived from Guinier approximation showed a significantly larger dimension for CP74 compared with WT (21.8 ± 0.1 Å and 24.5 ± 0.1 Å for WT and CP74, respectively; inset in Figure 3A). Finally, the Kratky plots of WT and CP74 exhibited well-defined bell-shape profiles with some differences in the mid-q range ($0.1 < q < 0.2$), indicating well-folded structures with limited local dynamics and clear structural differences between WT and CP74 (Figure 3B). Collectively, our structural analyses underscored the impact of introducing a CP site on the knotting loop of YbeA that led to the formation of an unusual domain-swapped dimer with a well-defined three-dimensional structure. Due to the introduction of the CP site at the knotting loop, the cofactor-binding pocket underwent considerable conformational rearrangements, resulting in the loss of structural complementarity to form an appropriate hydrogen bonding network to accommodate SAM (Figure 4) (Koh et al., 2017). This was confirmed by isothermal titration calorimetry analyses, in which we used the more stable analog of SAM, S-adenosyl-L-homocysteine (SAH) as the ligand; SAH was previously used as an SAM analog to study SPOUT protein variants (Mallam and Jackson, 2007b, 2012). While the analysis of WT yielded a dissociation constant (K_d) of 4.4 ± 0.3 μM for SAH with a stoichiometry (N) of 1.02 ± 0.01 (Figure 4C), CP74 showed no SAH binding (Figure 4F), confirming that the opening of the knotting loop completely abrogated the substrate binding of YbeA, and that the knotted structural element plays an essential role for substrate recognition among the SPOUT family members.

To evaluate the impacts of CP on the folding stability and kinetics of YbeA, we compared the thermal stabilities of WT and CP74 by differential scanning calorimetry (DSC) and far-UV CD spectroscopy. DSC revealed a prominent unfolding event for CP74 at a melting temperature (T_m) of 59.3°C in addition to the one (51.3°C) that was similar to that of WT (52.8°C), indicating the presence of a stable folding intermediate during thermal unfolding (Figure 5A; Table S1). Similar to DSC, far-UV CD also revealed an additional thermal transition for CP74: A single T_m was observed for WT (50.7°C), while two T_m values were observed for CP74 (46.3°C and 54.2°C) (Figure 5B; Table S1), suggesting that CP74 has at least two distinct local unfolding events that are reminiscent of our previously findings on amyloidogenic human lysozyme variants (Dhulesia et al., 2010). We next compared the urea-induced equilibrium unfolding of WT and CP74 by monitoring the intrinsic fluorescence changes as a function of urea concentration. By fitting the isotherms to an apparent two-state unfolding model without considering the dimeric nature of YbeA, the results revealed a significant increase in the transition urea concentration for CP74 (from 2.85 M in WT to 3.80 M in CP74; Figure 5C; Table S1). Urea-induced unfolding of both WT and

Table 1. X-Ray Crystallography Data Collection and Refinement Statistics

	CP74 Hg Derivative	CP74 Native
Data Collection		
Space group	$P2_1$	$P2_1$
Wavelength (Å)	1.00928	1.00000
Resolution (Å)	30–2.16 (2.24–2.16)[a]	30–1.75 (1.81–1.75)
Unit cell dimensions		
a, b, c (Å)	47.57, 88.34, 80.89	47.46, 88.05, 81.18
α, β, γ (°)	90.0, 101.0, 90.0	90.0, 101.5, 90.0
Total observations	257,725	239,045
Unique reflections	35,173 (3,358)	65,452 (6,405)
Redundancy	7.3 (5.6)	3.7 (3.1)
Completeness (%)	99.5 (95.9)	99.5 (98.1)
I/σ(I)	16.5 (2.1)	34.8 (2.1)
R_{meas} (%)	9.9 (72.5)	3.9 (52.5)
Phasing statistics		
Number of sites	20	
Figure of merit	0.376	
Refinement		
Resolution (Å)		29.52–1.75
Reflections		62,103 (6,284)
R_{work}/R_{free}		0.1693/0.1947
Root-mean-square deviation		
Bond lengths (Å)		0.007
Bond angles (°)		1.38
Average B factor (Å^2)		
Protein		43.27
Water		46.41
Ramachandran plot, residues in (%)		
Favored regions		98.7
Allowed regions		100
Outliers		0

[a]Values in parentheses correspond to the highest-resolution shell.

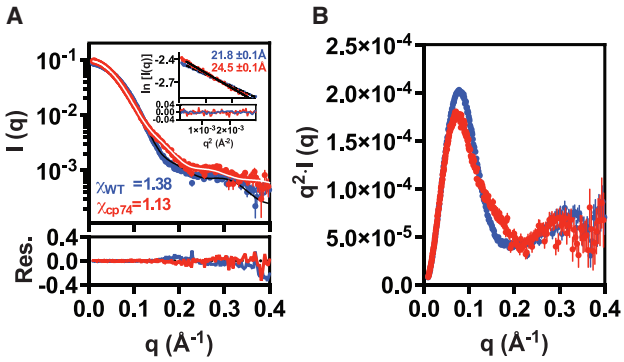


Figure 3. SAXS Analyses of WT and CP74

(A) SAXS profiles of WT (blue) and CP74 (red) were fit to the back-calculated SAXS profiles based on their respective crystal structures shown in black and white, respectively. The corresponding c values are indicated on the lower right corner and the fitting residues in logarithm scales are shown below. Inset: Guinier approximations of WT and CP74 were carried out to derive the radii of gyration of WT and CP74 as indicated next to the linear fits.

(B) Kratky plots of WT and CP74 with the same coloring scheme as in (A).

CP74 were reversible despite minor differences in the associated *m*-values of both variants (see also Figure S4).

We next asked how untying the topological knot and domain swapping may affect the folding kinetics of YbeA by carrying out stopped-flow intrinsic fluorescence measurements of CP74 (Figure 6). Reminiscent of the previously reported folding kinetics of WT, CP74 also exhibited two distinct kinetic phases (Mallam and Jackson, 2007a). While the faster kinetic phase of CP74 was similar to that of WT, the slower one was markedly slower than that of WT. Assuming that CP74 also unfolds via a monomeric folding intermediate as does WT, the slower intrinsic unfolding rate ($k_u = 7.1 \times 10^{-6}$ s^{-1}) of CP74 was ca. 20-fold slower than that of WT (Table 2). This is in line with the expectation for the need to disentangle the extensive intermolecular interactions as a result of domain swapping. In contrast, the intrinsic folding rate of the slower kinetic phase of CP74 ($k_f = 2.5 \times 10^{-2}$ s^{-1}) was comparable with that of WT suggesting robust dimer formation as a result of the countering effects between the release of topological constrain, i.e., unknotting, and complicated domain-swapping event. Once CP74 unfolds into a monomeric intermediate state, the folding and unfolding into the denatured state became much more robust than WT, which may be attributed to the effect of unknotting.

Finally, we employed hydrogen-deuterium exchange mass spectrometry (HDX-MS) to examine with greater spatial resolution the impacts of CP and unknotting on the folding dynamics of YbeA, manifested in the rates at which HDX occurred at an individual peptide level (Sriramoju et al., 2018b). Unlike the global folding analyses by far-UV CD and intrinsic fluorescence, which report on the overall changes in folding characteristics of a given protein, structural mapping of the observed relative deuterium uptakes revealed relatively fast HDX for both WT and CP74 (Figures 7A and 7B). Within 4 h, most of the backbone amide groups exhibited significant relative fractional deuterium uptakes of up to 50%, indicating abundant local dynamics for both WT and CP74. By comparing the fractional deuterium uptakes of the common peptides between WT and CP74 derived from random on-column pepsin digestion after HDX, i.e., differential HDX as defined by the relative fractional deuterium uptakes of WT subtracted by that of CP74 (Figure 7C), we observed significantly increased HDX within the CP site and its proximity, particularly within the dimer interface, while the regions distant from the CP site, e.g., α2 in WT and α5 in CP74 and part of the central β sheet, actually showed reduced HDX, i.e., attenuated local fluctuations, suggesting allosteric perturbations of folding dynamics within the relatively small protein structure (Figure 7).

DISCUSSION

Despite a large number of knotted protein structures being deposited in the database, the trail through which topologically knotted proteins emerged from evolution remains poorly defined. Except the ornithine transcarbamylases (OTCs) superfamily, of which some bacterial OTC homologs were found to contain a trefoil-knotted structural element, while the majority of the family members are unknotted (Sriramoju et al., 2018b), there is a lack

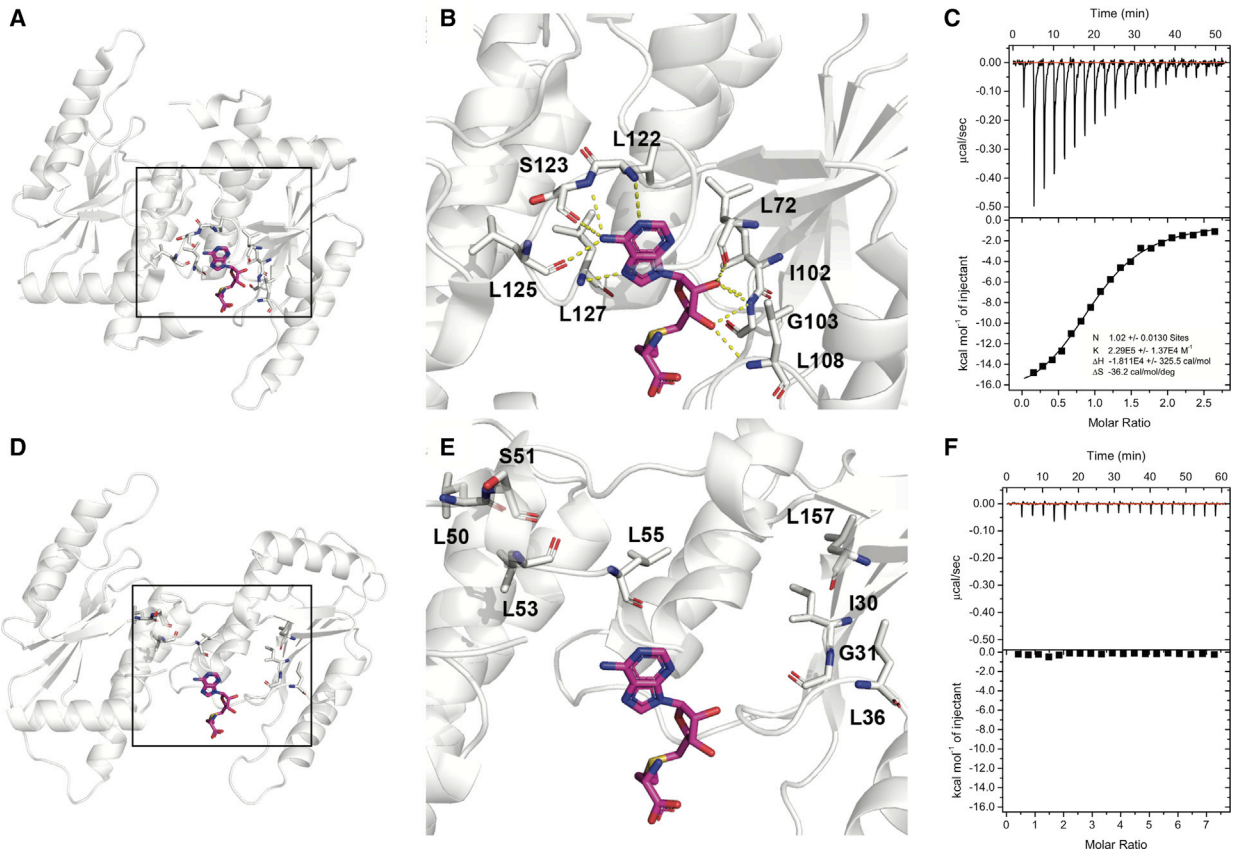


Figure 4. Impact of CP on Cofactor Binding of YbeA

(A) Cartoon representation of WT with cofactor, SAM, and its interacting residues shown in sticks. The carbon atoms shown in magenta and white for SAM and WT residues, respectively. Nitrogen, oxygen, and sulfur atoms are shown in blue, red and gold, respectively. The model was generated using the SAM-bound YbeA crystal structure (PDB: 5TWJ).

(B) Expanded view of SAM binding pocket of WT as indicated by the box in (A). The identities of the SAM interacting residues are indicated and the intermolecular hydrogen bonding shown in dashed yellow lines.

(C) The isothermal titration calorimetry (ITC) data of SAH titrating into WT as an SAM analog. The resulting data were fit to an one-site binding model and the associated thermodynamics parameters are shown within the panel.

(D–F) Cartoon representation of CP74 (D) and expanded view of the region corresponding to the SAM binding pocket in WT (E). The SAM molecule was maintained in the same position after super-positioning the structures of CP74 onto WT to illustrate how the corresponding SAM interacting residues in CP74 as indicated underwent conformational rearrangements that rendered the loss of SAM binding as demonstrated by the ITC analysis using SAH as an SAM analog as shown in (F).

of coexistence of knotted and unknotted members within a protein family, making it nontrivial to glean information from sequence comparison with predict a knotted protein structure (Potestio et al., 2010). One of the largest knotted protein families in the KnotProt database is the SPOUT RNA methyltransferase family (CATH superfamily 3.40.1280.10) with more than 120 reported crystal structures derived from all kingdoms of life (Dabrowski-Tumanski et al., 2019). An earlier bioinformatics survey proposed that the emergence of the trefoil-knotted structure in the SPOUT superfamily may predate the emergence of LUCA, which seems

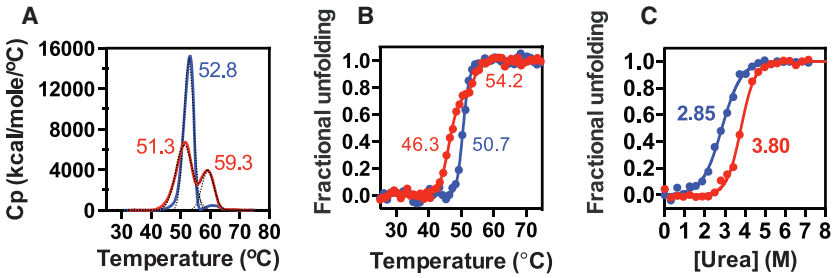


Figure 5. Comparative Analyses of Thermal and Chemical Folding Stabilities of WT and CP74

(A) DSC profiles of WT (blue) and CP74 (red).

(B) Fractional unfolding of WT (blue) and CP74 (red) as a function of temperature derived from far-UV CD spectroscopy.

(C) Fractional unfolding of WT (blue) and CP74 (red) as a function of urea concentration derived from intrinsic fluorescence spectroscopy. The values of individual transition points are indicated.

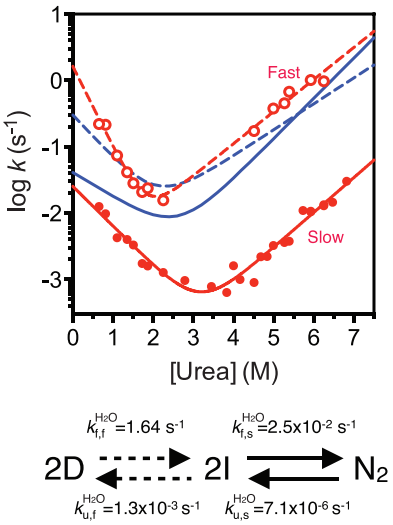


Figure 6. Comparative Analysis of Folding Kinetics of WT and CP74

CP74 exhibited two distinct kinetic phases, shown in filled and open circles, corresponding to the slow and fast phases, respectively. The slow and fast phases were fit to a two-state model indicated by solid and dashed red lines, respectively. For comparison, previously reported kinetic data of WT are shown in dashed lines in the same chevron plot. Proposed folding pathway and the corresponding microscopic reaction rates derived from the two-state fitting are indicated below. The intrinsic folding rate ($k_f^{H_2O}$) and unfolding rate ($k_u^{H_2O}$) of the slow and fast kinetic phases are indicated by subscripts, *s* and *f*, respectively (cf. Table 2).

counterintuitive since LUCA is considered to be the origin (Tkaczuk et al., 2007). The authors proposed that one of the potential mechanisms by which a knotted SPOUT protein evolves from an unknotted parent protein may be through CP. While naturally occurring and designed CPs have been reported in the literature, there is no report on a successful design of an unknotted CP based on a topologically knotted protein, or indeed any designed protein that unties a knotted topology while maintaining the overall structure, or vice versa; at least not until our very recent report of an engineered unknotted CP based on another SPOUT protein, YibK from *Haemophilus influenza* (Chuang et al., 2019).

In this work, we successfully untied a topologically knotted SPOUT protein, YbeA from *E. coli*, by introducing a CP site (CP74) that opens the knotting loop and closes the original N and C termini to reconfigure the trefoil-knotted topology into a trivial knot (unknotted). Unlike the unknotted CP of YibK, which retains most of the native contacts with minimum structural perturbations (Chuang et al., 2019), the crystal structure of the CP74 forms a domain-swapped dimer by exchanging a large segment from one domain to the other. Such a major conformational and topological rearrangement resulted in significant change of the overall molecular shape and the relative domain orientation, which were confirmed by SAXS analyses in solution state (Figure 3). In addition to the structural changes, CP74 exhibited reduced thermal and chemical stabilities (Figure 5; Table S1) and, more importantly, complete loss of SAH binding, which is essential for its RNA methyltransferase activity (Figure 4). Considering that most topologically knotted proteins, including SPOUT proteins and UCHs, bind to their natural substrates or cofactors via their knotted structural elements (Christian et al., 2016; Hsu, 2016; Ito et al., 2015), being knotted may not be the most energetically favorable solution to fold a polypeptide chain, but it may be driven by a specific functional requirement for which a unique knotted topology has been evolved.

While computational analyses predict that protein knots may confer stabilizing effects, the nature of stabilization may be context dependent. We recently showed that the Gordian-knotted human UCHs are remarkably resilient to withstand mechano-unfolding by the bacterial AAA+ protease, ClpXP, compared with the chemically and thermally much more stable model substrate, namely the green fluorescent protein (Sriramoju et al., 2018a). A similar observation has been reported for the smallest knotted protein, MJ0366 (San Martin et al., 2017), of which the biological function remains undefined (Thiruselvam et al., 2017). Indeed, many knotted proteins are implicated in key enzymatic reactions. Therefore, functional dynamics is expected to be present in topologically knotted enzymes. In the case of SPOUT proteins, the knotted structural element of TrmD has been shown to play an important role in straining the cofactor molecule in an unusually high-energy configuration for catalytic activation (Christian et al., 2016). To impose such a mechanic constraint, the knotted structural element of TrmD, and possibly all SPOUT family members, may be highly rigidified under tension. Indeed, comparison of the crystal structures of a selection of SPOUT protein-bound SAM and SAH, including that of YbeA (Koh et al., 2017) reveals the same highly bent, high-energy configurations, suggesting a conversed ligand binding mode associated with the knotted motifs (Figure S5).

The restrained local dynamics within the knotted loop of YbeA was confirmed by HDX-MS analysis (Figure 7A). In CP74, the loop opening resulted in an increase of local fluctuations within the CP site and concomitant allosteric dynamic changes on the opposite side of CP74, which showed substantial reduction in local

Table 2. Kinetics Parameters of YbeA Variants Derived from Stopped-Flow Fluorescence Measurements

	$k_f^{H_2O}(s^{-1})$	$k_u^{H_2O}(s^{-1})$	m_f (kcal mol^{-1} M^{-1})	m_u (kcal mol^{-1} M^{-1})	m_{kin}[a] (kcal mol^{-1} M^{-1})
CP74					
Slow	2.53 (±0.62) × 10^{-2}	7.09 (±3.0) × 10^{-6}	0.82 ± 0.09	0.72 ± 0.05	1.54 ± 0.10
Fast	1.64 ± 0.54	1.29 (±0.51) × 10^{-3}	0.72 ± 0.05	0.66 ± 0.04	1.38 ± 0.06
WT[b]					
Slow	4.1 (±0.02) × 10^{-2}	1.2 (±0.02) × 10^{-4}	0.61 ± 0.01	0.71 ± 0.002	1.32 ± 0.01
Fast	0.20 ± 0.003	1.8 (±0.02) × 10^{-3}	0.86 ± 0.01	0.41 ± 0.002	1.27 ± 0.01

[a]$m_{kin} = m_f + m_k$.

[b]The values correspond to WT at pH 7.5 as described previously (Mallam and Jackson, 2006).

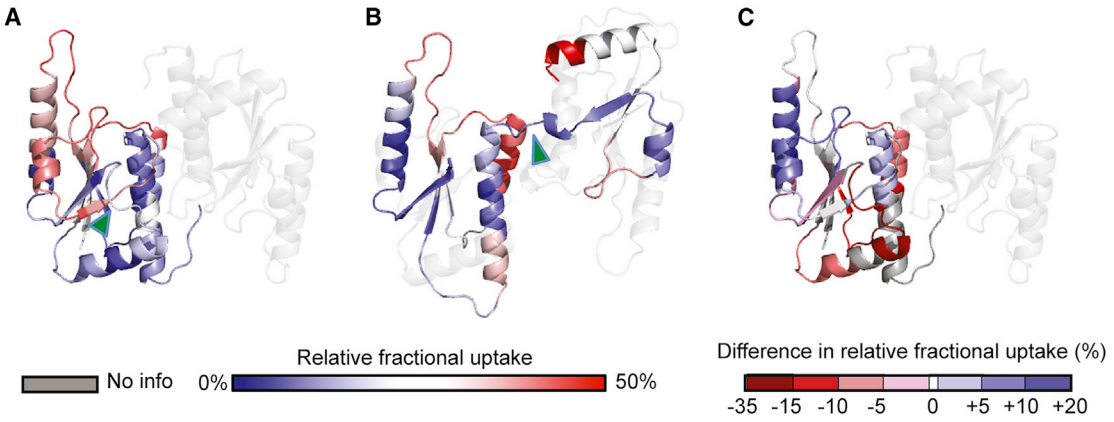


Figure 7. HDX-MS Analyses of WT and CP74 Revealed Allosteric Dynamic Changes on Topological Rearrangement

(A and B) Structural mapping of the residue-specific fractional deuterium uptakes in WT (A) and CP74 (B) after 10 min of HDX. The structures are color-ramped from blue to white to red, corresponding to low-to-high fractional deuterium uptake values as indicated by the color bar below. The coloring is applied to only one of the two molecules within the dimeric assembly to highlight the domain swap in CP74. The positions of the knotting loop in WT and that of the linker for domain-swapped segment in CP74 are indicated by green arrows.

(C) Relative fractional deuterium uptake of CP74 with respect to WT. Blue and red colorings correspond to less and more deuterium uptakes in CP74, which reflect reduction and increase in local fluctuations as a result of CP. Due to the lack of common peptides between WT and CP74, the differential deuterium uptakes in some regions cannot be defined, which are indicated by dark gray. The structural mapping is applied to the WT structure.

fluctuations that compensated the increase of dynamics within the CP site (Figures 7B and 7C). One could envisage that the knotting loop may resemble the holding bar sitting on top of a spring-loaded hammer of a mousetrap to prevent it from being snapped to the opposite direction. Once the holding bar is relieved from its holding position, the hammer that corresponds to the domain-swapped fragment will readily flip out to form extensive native-like intermolecular contacts with the other monomer (Figure 2A), a process that will take a lot longer than forming intramolecular native contacts as demonstrated by the slower refolding kinetics of CP74, compared with that of WT (Figure 6). In contrast to the accelerated folding kinetics of the CP of YibK from *H. influenzae*, which was attributed to the alleviation from the topological constrains (Chuang et al., 2019), the significantly slower unfolding kinetics of CP74 ($k_u^{H_2O}$ of the slower kinetic phase, associated with the intermediate-to-denature state transition) may be attributed to the need to break the extensive intermolecular interactions at the domain-swapped dimer interface that generate a much higher kinetic barrier for unfolding (Table 2).

According to the CP design and the resulting topological and structural rearrangements, we hereby propose a likely folding pathway for CP74 (Figure 8). First, the introduction of the CP

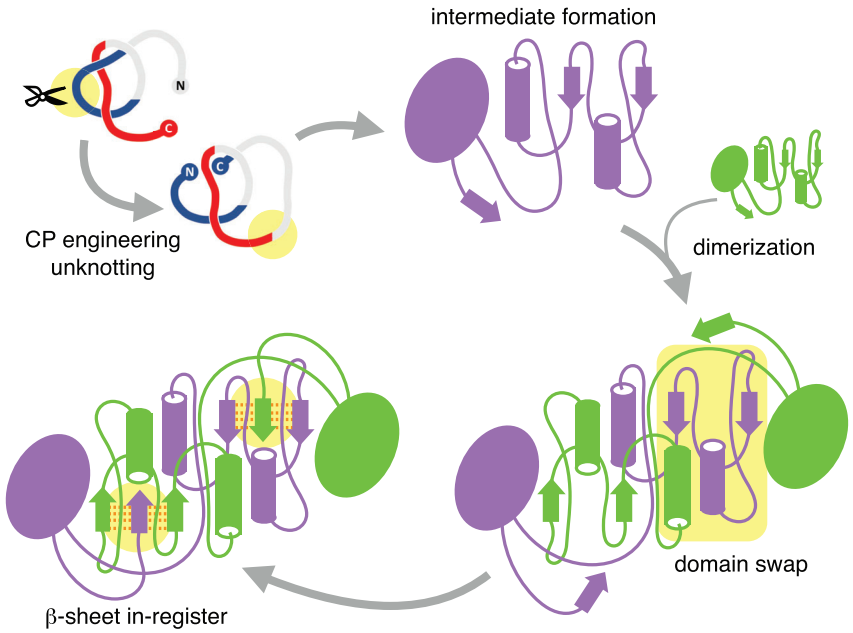


Figure 8. Proposed Folding Pathway of CP74

Regions that are involved in topological reconfiguration and/or folding events at individual steps are highlighted yellow.

site on the knotting loop and the closure of the original N and C termini reshuffled the topological arrangement of individual secondary structures, rendering the original α3 to become the new N-terminal α helix, α1, and the original β3 to become the new C-terminal β strand, β5 (Figure 2B). During the refolding process, the N- and C-terminal halves of CP74 fold independently as opposed to a much more cooperative folding process of WT, demonstrated by the two distinct (un)folding peaks (folding transitions) in the DSC profile of CP74 during thermal unfolding compared with a single transition of WT (Figure 5A). After extensive conformational searches and template matching, the two N-terminal halves of CP74 manage to form native-like intermolecular contacts with their counterpart C-terminal halves. At this point, the C-terminal β5 of CP74 can tug into the cleft between β1 and β2 to complete the domain-swapped dimer formation (Figure 8). While the validity of such a convoluted folding pathway to form a domain-swapped dimer of CP74 remains to be experimentally and/or computationally verified, the crystal structure of the domain-swapped CP74 dimer, the previously reported WT structure, and the quantitative thermodynamics and kinetics analyses of folding of CP74 and WT reported herein provide a vantage point for further studies.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Construct Designs of YbeA CPs
  - Purification of WT YbeA and CP Variants
  - Far-UV CD Spectroscopy
  - Intrinsic Fluorescence Spectroscopy
  - DSC
  - SEC-MALS
  - SEC-SAXS
  - Protein Crystallization, Data Collection, Processing and Refinement
  - ITC
  - Stopped-Flow Fluorescence Measurements
  - HDX-MS
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.str.2019.04.004.

ACKNOWLEDGMENTS

We thank Dr. Meng-Ru Ho of the Biophysics Facility, Ms Hui-Ling Shih of the Protein Crystallization Facility, Institute of Biological Chemistry, Academia Sinica, for assisting data collection and analyses. We thank Dr. Shu-Chuan (Chris) Jao of the Biophysics Core Facility, Department of Academic Affairs and Instrument Service at Academia Sinica for providing technical assistance and useful discussion of isothermal titration calorimetry experiments. We acknowledge the National Synchrotron Radiation Research Center, Taiwan, for the use of protein crystallography beamlines (15A1, 13B1, and 13C1) and SAXS beamline (23A). We also acknowledge the Taiwan Protein Project (AS-KPQ-105-TPP) for providing beamtime for part of our crystallography data collection. We thank Dr. Shu-Yu Lin and Mr. Ming-Jie Tsai of the Mass Spectrometry Core Facility, Academia Sinica, for assisting the HDX-MS data collection. This work was supported by the Ministry of Science and Technology, Taiwan (106-2113-M-001-004 and 107-2628-M-001-005-MY3 to S.-T.D.H and 106-2311-B-007-004-MY3 to P.-C.L), and Academia Sinica, Taiwan (intramural funding to S.-T.D.H.).

AUTHOR CONTRIBUTIONS

Conceptualization, P.-C.L. and S.-T.D.H.; Methodology, K.-T.K., K.-F.H., I.-C.H., and S.-T.D.H.; Investigation, K.-T.K., K.-F.H., and S.-T.D.H.; Formal Analysis, K.-T.K., K.-F.H., and S.-T.D.H.; Writing – Original Draft, K.-T.K. and S.-T.D.H.; Writing – Review & Editing, S.-T.D.H.; Funding Acquisition, P.-C.L. and S.-T.D.H.; Resources, P.-C.L. and S.-T.D.H.; Supervision, P.-C.L., I.-C.H., and S.-T.D.H.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 7, 2019
Revised: March 1, 2019
Accepted: April 5, 2019
Published: May 16, 2019

REFERENCES

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr. *68*, 352–367.

Ahn, H.J., Kim, H.W., Yoon, H.J., Lee, B.I., Suh, S.W., and Yang, J.K. (2003). Crystal structure of tRNA(m^1G37)methyltransferase: insights into tRNA recognition. EMBO J. *22*, 2593–2603.

Andersson, F.I., Pina, D.G., Mallam, A.L., Blaser, G., and Jackson, S.E. (2009). Untangling the folding mechanism of the 5_2-knotted protein UCH-L3. FEBS J. *276*, 2625–2635.

Bennett, M.J., Schlunegger, M.P., and Eisenberg, D. (1995). 3D Domain swapping: a mechanism for oligomer assembly. Protein Sci. *4*, 2455–2468.

Blanchet, C.E., and Svergun, D.I. (2013). Small-angle X-ray scattering on biological macromolecules and nanocomposites in solution. Annu. Rev. Phys. Chem. *64*, 37–54.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Chiang, J.J.H., Li, I., and Truong, K. (2006). Creation of circularly permutated yellow fluorescent proteins using fluorescence screening and a tandem fusion template. Biotechnol. Lett. *28*, 471–475.

Christian, T., Sakaguchi, R., Perlinska, A.P., Lahoud, G., Ito, T., Taylor, E.A., Yokoyama, S., Sulkowska, J.I., and Hou, Y.M. (2016). Methyl transfer by substrate signaling from a knotted protein fold. Nat. Struct. Mol. Biol. *23*, 941–948.

Chuang, Y.C., Hu, I.C., Lyu, P.C., and Hsu, S.-T.D. (2019). Untying a knotted protein by circular permutation. J. Mol. Biol. *431*, 857–863.

Dabrowski-Tumanski, P., Rubach, P., Goundaroulis, D., Dorier, J., Sulkowski, P., Millett, K.C., Rawdon, E.J., Stasiak, A., and Sulkowska, J.I. (2019). KnotProt 2.0: a database of proteins with knots and other entangled structures. Nucleic Acids Res. *47*, D367–D375.

Dabrowski-Tumanski, P., and Sulkowska, J. (2017). To tie or not to tie? That is the question. Polymers (Basel) *9*, 454.

Dhulesia, A., Cremades, N., Kumita, J.R., Hsu, S.-T.D., Mossuto, M.F., Dumoulin, M., Nietlispach, D., Akke, M., Salvatella, X., and Dobson, C.M. (2010). Local cooperative in an amyloidogenic state of human lysozyme observed at atomic resolution. J. Am. Chem. Soc. *132*, 15580–15588.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Flory, P.J. (1953). Principles of Polymer Chemistry (Cornell Univ. Press).

Franke, D., Petoukhov, M.V., Konarev, P.V., Panjkovich, A., Tuukkanen, A., Mertens, H.D.T., Kikhney, A.G., Hajizadeh, N.R., Franklin, J.M., Jeffries, C.M., et al. (2017). ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. J. Appl. Crystallogr. *50*, 1212–1225.

Goldenberg, D.P., and Creighton, T.E. (1983). Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor. J. Mol. Biol. *165*, 407–413.

Hsu, S.-T.D. (2016). Folding dynamics and structural basis of the enzyme mechanism of ubiquitin C-terminal hydroylases. In Understanding Enzymes - Function, Design, Engineering and Analysis, A. Svendsen, ed. (Pan Stanford Publishing), pp. 167–202.

Ito, T., Masuda, I., Yoshida, K., Goto-Ito, S., Sekine, S., Suh, S.W., Hou, Y.M., and Yokoyama, S. (2015). Structural basis for methyl-donor-dependent and sequence-specific binding to tRNA substrates by knotted methyltransferase TrmD. Proc. Natl. Acad. Sci. U S A *112*, E4197–E4205.

Jackson, S.E., Suma, A., and Micheletti, C. (2017). How to fold intricately: using theory and experiments to unravel the properties of knotted proteins. Curr. Opin. Struct. Biol. *42*, 6–14.

Jarmolinska, A.I., Perlinska, A.P., Runkel, R., Trefz, B., Ginn, H.M., Virnau, P., and Sulkowska, J.I. (2018). Proteins' knotty problems. J. Mol. Biol. *431*, 244–257.

Jeng, U.-S., Su, C.H., Su, C.-J., Liao, K.-F., Chuang, W.-T., Lai, Y.-H., Chang, J.-W., Chen, Y.-J., Huang, Y.-S., Lee, M.-T., et al. (2010). A small/wide-angle X-ray scattering instrument for structural characterization of air-liquid interfaces, thin films and bulk specimens. J. Appl. Crystallogr. *43*, 110–121.

Koh, C.S., Madireddy, R., Beane, T.J., Zamore, P.D., and Korostelev, A.A. (2017). Small methyltransferase RlmH assembles a composite active site to methylate a ribosomal pseudouridine. Sci. Rep. *7*, 969.

Kohn, J.E., Millett, I.S., Jacob, J., Zagrovic, B., Dillon, T.M., Cingel, N., Dothager, R.S., Seifert, S., Thiyagarajan, P., Sosnick, T.R., et al. (2004). Random-coil behavior and the dimensions of chemically unfolded proteins. Proc. Natl. Acad. Sci. U S A *101*, 12491–12496.

Lee, Y.C., and Hsu, S.-T.D. (2018). A natively monomeric deubiquitinase UCH-L1 forms highly dynamic but defined metastable oligomeric folding intermediates. J. Phys. Chem. Lett. *9*, 2433–2437.

Lim, N.C., and Jackson, S.E. (2015). Molecular knots in biology and chemistry. J. Phys. Condensed Matter *27*, 354101.

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

Lo, W.C., Lee, C.C., Lee, C.Y., and Lyu, P.C. (2009). CPDB: a database of circular permutation in proteins. Nucleic Acids Res. *37*, D328–D332.

Lo, W.C., Wang, L.F., Liu, Y.Y., Dai, T., Hwang, J.K., and Lyu, P.C. (2012). CPred: a web server for predicting viable circular permutations in proteins. Nucleic Acids Res. *40*, W232–W237.

Lou, S.C., Wetzel, S., Zhang, H., Crone, E.W., Lee, Y.T., Jackson, S.E., and Hsu, S.-T.D. (2016). The knotted protein UCH-L1 exhibits partially unfolded forms under native conditions that share common structural features with its kinetic folding intermediates. J. Mol. Biol. *428*, 2507–2520.

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

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

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

Mallam, A.L., and Jackson, S.E. (2012). Knot formation in newly translated proteins is spontaneous and accelerated by chaperonins. Nat. Chem. Biol. *8*, 147–153.

Mallam, A.L., Morris, E.R., and Jackson, S.E. (2008). Exploring knotting mechanisms in protein folding. Proc. Natl. Acad. Sci. U S A *105*, 18740–18745.

Mallam, A.L., Rogers, J.M., and Jackson, S.E. (2010). Experimental detection of knotted conformations in denatured proteins. Proc. Natl. Acad. Sci. U S A *107*, 8189–8194.

Murshudov, G.N., Skubak, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. *67*, 355–367.

Onuchic, J.N., Luthey-Schulten, Z., and Wolynes, P.G. (1997). Theory of protein folding: the energy landscape perspective. Annu. Rev. Phys. Chem. *48*, 545–600.

Plaxco, K.W., Simons, K.T., and Baker, D. (1998). Contact order, transition state placement and the refolding rates of single domain proteins. J. Mol. Biol. *277*, 985–994.

Potestio, R., Micheletti, C., and Orland, H. (2010). Knotted vs. unknotted proteins: evidence of knot-promoting loops. PLoS Comp. Biol. *6*, e1000864.

Reitinger, S., Yu, Y., Wicki, J., Ludwiczek, M., D'Angelo, I., Baturin, S., Okon, M., Strynadka, N.C., Lutz, S., Withers, S.G., et al. (2010). Circular permutation of *Bacillus circulans* xylanase: a kinetic and structural study. Biochemistry *49*, 2464–2474.

San Martin, A., Rodriguez-Aliaga, P., Molina, J.A., Martin, A., Bustamante, C., and Baez, M. (2017). Knots can impair protein degradation by ATP-dependent proteases. Proc. Natl. Acad. Sci. U S A *114*, 9864–9869.

Schneidman-Duhovny, D., Hammel, M., Tainer, J.A., and Sali, A. (2016). FoXS, FoXSDock and MultiFoXS: single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. Nucleic Acids Res. *44*, W424–W429.

Shih, P.-M., Wang, I., Lee, Y.-T.C., Hsieh, S.-J., Chen, S.-Y., Wang, L.-W., Huang, C.-T., Chien, C.-T., Chang, C.-Y., and Hsu, S.-T.D. (2015). Random-coil behavior of chemically denatured topologically knotted proteins revealed by small-angle X-ray scattering. J. Phys. Chem. B *119*, 5437–5443.

Sriramoju, M.K., Chen, Y., Lee, Y.C., and Hsu, S.-T.D. (2018a). Topologically knotted deubiquitinases exhibit unprecedented mechanostability to withstand the proteolysis by an AAA+ protease. Sci. Rep. *8*, 7076.

Sriramoju, M.K., Yang, T.J., and Hsu, S.-T.D. (2018b). Comparative folding analyses of unknotted versus trefoil-knotted ornithine transcarbamylases suggest stabilizing effects of protein knots. Biochem. Biophys. Res. Commun. *503*, 822–829.

Terwilliger, T.C., Adams, P.D., Read, R.J., McCoy, A.J., Moriarty, N.W., Grosse-Kunstleve, R.W., Afonine, P.V., Zwart, P.H., and Hung, L.W. (2009). Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. Acta Crystallogr. D Biol. Crystallogr. *65*, 582–601.

Thiruselvam, V., Kumarevel, T., Karthe, P., Kuramitsu, S., Yokoyama, S., and Ponnuswamy, M.N. (2017). Crystal structure analysis of a hypothetical protein (MJ0366) from *Methanocaldococcus jannaschii* revealed a novel topological arrangement of the knot fold. Biochem. Biophys. Res. Commun. *482*, 264–269.

Tkaczuk, K.L., Dunin-Horkawicz, S., Purta, E., and Bujnicki, J.M. (2007). Structural and evolutionary bioinformatics of the SPOUT superfamily of methyltransferases. BMC Bioinformatics *8*, 73.

Wang, I., Chen, S.-Y., and Hsu, S.-T.D. (2016). Folding analysis of the most complex Stevedore's protein knot. Sci. Rep. *6*, 31514.

Wang, I., Chen, S.Y., and Hsu, S.-T.D. (2015a). Unraveling the folding mechanism of the smallest knotted protein, MJ0366. J. Phys. Chem. B *119*, 4359–4370.

Wang, L.-W., Liu, Y.-N., Lyu, P.-C., Jackson, S.E., and Hsu, S.-T.D. (2015b). Comparative analysis of the folding dynamics and kinetics of an engineered knotted protein and its variants derived from HP0242 of *Helicobacter pylori*. J. Phys. Condensed Matter. *27*, 354106.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. *67*, 235–242.

Wolynes, P.G., Onuchic, J.N., and Thirumalai, D. (1995). Navigating the folding routes. Science *267*, 1619–1620.

Yu, Y., and Lutz, S. (2011). Circular permutation: a different way to engineer enzyme structure and function. Trends Biotechnol. *29*, 18–25.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial Strains		
E. coli DH5α	Yeatern Biotech	Cat#FYE678-80VL
E. coli BL21 (DE3)	Yeatern Biotech	Cat#FYE207-40VL
Chemicals, Peptides, and Recombinant Proteins		
YbeA-WT	(Mallam and Jackson, 2007a, 2007b)	N/A
YbeA-CP74	This paper	N/A
S-(5'-Adenosyl)-L-homocysteine	Sigma-Aldrich	A9384
Critical Commercial Assays		
Pierce™ BCA Protein Assay Kit	ThermoFisher	Cat#23225
Deposited Data		
YbeA-WT	(Koh et al., 2017)	PDB:1ns5
YbeA-WT with SAM	(Koh et al., 2017)	PDB:5twj
YbeA-CP74	This paper	PDB:5zyo
TrmD with SAM	(Ahn et al., 2003)	PDB:1uak
YibK with SAH	(Lim et al., 2003)	PDB:1mxi
Oligonucleotides		
Primer: ybeA-CTCGAG-ybeA Forward: 5'-CCGCTCGAGATGAAGCTGCAACTTGTCGC	This paper	N/A
Primer: ybeA-CTCGAG-ybeA Reverse: 5'- CCGCTCGAGCTCACGGTGATAAGGATGGT	This paper	N/A
Primer: CP74 Forward: 5'- GGAATTCCATATGA TTCCAGGCAAGCCCTGGGA	This paper	N/A
Primer: CP74 Reverse: 5'- CCCAAGCTTTTAAT CGAGGGTGACAATGCGG	This paper	N/A
Primer: CP-GS-linker Forward: 5'- CAACCATCCTTATC ACCGTGAGGGATCCATGAAGCTGCAACTTGTCGCCG	This paper	N/A
Primer: CP-GS-linker Reverse: 5'- CGGCGACAAGTTG CAGCTTCATGGATCCCTCACGGTGATAAGGATGGTTG	This paper	N/A
Recombinant DNA		
pET17b-YbeA	(Mallam and Jackson, 2007a, 2007b)	N/A
pET21b-YbeA-CP66	This paper	N/A
pET21b-YbeA-CP74	This paper	N/A
pET21b-YbeA-CP79	This paper	N/A
pET21b-YbeA-CP96	This paper	N/A
pET21b-YbeA-CP106	This paper	N/A
pET21b-YbeA-CP117	This paper	N/A
pET21b-YbeA-CP120	This paper	N/A
pET21b-YbeA-CP123	This paper	N/A
pET21b-YbeA-CP126	This paper	N/A
Software and Algorithms		
PyMOL	Schrodinger LLC	http://www.pymol.org
ASTRA 6.0	Wyatt Technology	http://www.wyatt.com/products/software/astra.html
ATSAS	(Franke et al., 2017)	http://www.embl-hamburg.de/biosaxs/atsas-online/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
FoXS	(Schneidman-Duhovny et al., 2016)	http://modbase.compbio.ucsf.edu/foxs/
PHENIX	(Adams et al., 2010)	http://www.phenix-online.org
REFINE	(Afonine et al., 2012)	http://www.phenix-online.org
CCP4	(Winn et al., 2011)	http://www.ccp4.ac.uk
REFMAC5	(Murshudov et al., 2011)	http://www.ccp4.ac.uk
Autosol	(Terwilliger et al., 2009)	http://www.phenix-online.org/
Coot	(Emsley and Cowtan, 2004)	http://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
MolProbity	(Chen et al., 2010)	http://molprobity.biochem.duke.edu/
Prism	Graphpad	http://www.graphpad.com/
PLGS	Waters	http://www.waters.com/
DynamX	Waters	http://www.waters.com/
Origin	OriginLab	http://www.originlab.com/
Other		
HiTrap™ SP HP 5 mL	GE Life Science	Cat#17115201
HiTrap™ Q HP 5 mL	GE Life Science	Cat#17115401
HiLoad 16/600 Superdex 75 pg	GE Life Science	Cat#28989333
Superdex 75 Increase 10/300 GL	GE Life Science	Cat#29148721

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Shang-Te Danny Hsu (sthsu@gate.sinica.edu.tw).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Plasmids of YbeA variants were transformed into in *E. coli* BL21 (DE3) followed by overnight growth at 37°C in 15 mL of LB broth with 100 μg/mL ampicillin. The overnight preculture was inoculated into 1 L of LB broth to allow further growth at 37°C until the cell culture optical density at 600 nm (OD_{600}) reached 0.6, at which point isopropyl β-D-1-thiogalactopyranoside (IPTG; 0.5 mM final concentration) was added to induce recombinant protein overexpression. After overexpression, the cells were harvested by centrifugation at 6,000 rpm for 20 min at 4°C, and the cell pellets were stored at −80°C until further use.

METHOD DETAILS

Construct Designs of YbeA CPs

CPred was used to suggest potential CP sites in YbeA (Lo et al., 2012). These included residues 66, 74, 79, 96, 106, 117, 120, 123, 126 to be the new N-termini. To generate the YbeA CP constructs, a tandem repeat DNA sequence was constructed of which two ybeA genes were concatenated via a hexanucleotide CTCGAG linker and an XhoI restriction site was introduced to the 5'- and 3'-ends to yield XhoI-*ybeA*-CTCGAG-*ybeA*-XhoI (Chiang et al., 2006). To obtain different YbeA CP variants, pairs of forward and backward primers were designed to obtain the desired CP constructs with NdeI, HindIII or EcoRI restriction sites at the 5'- and 3'-ends, respectively, for subcloning into the expression vector pET21b. Due to the introduction of the LE linker residue (translated from CTCGAG), the isoelectric points of YbeA CP variants were decreased from 8.7 to 7.8, thus reducing their binding affinity to the cation column. We therefore replaced the linker DNA sequence from CTCGAG to GGATCC by PCR to yield an electrostatic neutral amino acid composition of GS to help improve the yield of recombinant YbeA CP variants.

Purification of WT YbeA and CP Variants

Except for WT and CP74, which was expressed in soluble forms at 37°C, all other CPs were aggregated, which could not be refolded with sufficient quantities. To purify WT and CP74, the cell pellets were resuspended in SP buffer (50 mM sodium acetate (pH 5.8), 5% glycerol and 1 mM DTT) followed by sonication, and centrifugation at 13,000 rpm for 20 minutes at 4°C to remove cell debris. The supernatant was loaded to a HiTrap™ SP HP 5 mL cation-exchange column (GE Life Science, U. S. A.) in SP buffer, followed by elution with a linear gradient of NaCl (0-0.5 M) in SP buffer by using an ÄKTA Prime FPLC system (GE Life Science, U. S. A.). The target proteins were eluted by 0.1 M NaCl, collected and dialyzed against Q buffer (20 mM Tris-HCl (pH 7.5), 200 mM KCl, 10% (v/v) glycerol and 1 mM DTT) for three hours. Visible precipitants were removed by centrifugation prior to loading onto a HiTrap™

Q HP 5 mL anion-exchange column (GE Life Science, U. S. A.). Flow-through was collected, dialyzed against buffer A (50 mM Tris-HCl (pH 7.5), 200 mM KCl, 10% glycerol and 1 mM DTT), and concentrated to 5 mL to be loaded onto a Superdex 75 16/60 gel-filtration column (GE Life Science, U. S. A.) for SEC in buffer A using an ÄKTA Prime FPLC system. The purified proteins were pooled, aliquoted and flash-frozen by liquid nitrogen for storage at –30°C until further use. All following experiments were carried in buffer A unless otherwise specified.

Far-UV CD Spectroscopy

Far-UV CD spectra of YbeA variants were recorded using a CD spectrometer (J-815, JASCO, Japan) for native CD spectra and thermal unfolding.

Intrinsic Fluorescence Spectroscopy

Intrinsic fluorescence spectra of WT and CP74 and in different urea concentrations ranging between 0 and 8 M were collected with an excitation wavelength of 280 nm and the emission spectra between 300 and 450 nm using the Infinite M1000 PRO plate reader (TECAN, Switzerland) at 25°C. The equilibrium unfolding spectral series were subject to singular value decomposition to extract the thermodynamics parameters associated with the unfolding, as described previously (Wang et al., 2015a, 2015b, 2016).

DSC

Samples were subject to fully automated DSC measurements in Buffer A without DTT using MicroCal PEAQ-DSC (Malvern Panalytical, UK). The temperature was ramped from 10 to 90°C at a rate of 200°C/hr. The data were analyzed by using the Origin™ software (OriginLab, U. S. A.).

SEC-MALS

The apparent molecular weights of WT and CP74 were determined by SEC-MALS using a Wyatt Dawn HELEOS II multiangle light scattering detector (Wyatt Technology, U. S. A.) and a refractive index (RI) detector (Wyatt Technology, U. S. A.) coupled with BioSEC-3 HPLC analytical size-exclusion column (Agilent, U. S. A.) at a flow rate of 0.3 ml/min using an ÄKTApurifier UPC 10 FPLC system (GE Life Science, U.S.A.), as described previously (Lee and Hsu, 2018). The data were processed by ASTRA 6.0 software (Wyatt Technology, U. S. A.).

SEC-SAXS

SEC-SAXS data of WT and CP74 were collected using the BL23A beamline at the National Synchrotron Radiation Research Center (NSRRC) in Hsinchu, Taiwan, as described previously (Jeng et al., 2010; Lee and Hsu, 2018). The protein concentrations were set to 15 and 12.5 mg/mL for CP74 and WT respectively. The solvent subtraction was achieved by an in-house developed program. The SAXS data were analyzed by using the ATSAS package (Franke et al., 2017), and FoXS (Schneidman-Duhovny et al., 2016). The radii of gyration (R_g) were estimated using the Guinier approximation for the low-resolution scattering ($qR_g < 1.3$ for globular scatters): $I(q) = I(0)\exp(-q^2R_g^2/3)$, where $I(q)$ is the scattering intensity and $I(0)$ is the forward scattering intensity.

Protein Crystallization, Data Collection, Processing and Refinement

Freshly thawed aliquot of CP74 was concentrated to 12 mg/ml for an initial crystallization screening of ca. 500 conditions (Protein Crystallization Facility. Institute of Biological Chemistry, Academia Sinica). The crystallization conditions were manually refined to the final condition: 0.1M PIPES (pH 7.0) and 25% PEG8000. The crystals were grown at 20°C by mixing the protein aliquot with twice equilivant volume of crystallization buffer through the sitting-drop vapor-diffusion method. To solve the phase problem, the CP74 crystals were soaked in mother liquid added with 2-5 mM various mercury derivatives (Hampton Research, U.S.A.) at 20°C for 16 hours. The crystals were immediately flash-frozen in liquid nitrogen after dipping into cryoprotectant composed of 0.1 M PIPES (pH 7.0), 25 % (v/v) PEG8000 and 10 % (v/v) glycerol. All X-ray diffraction data were collected at beamlines 15A1, 13B1, or 13C1 of NSRRC (Hsinchu, Taiwan). The diffraction data were processed and scaled with *HKL2000* package. The resulting crystal has a space group of $P2_1$ with one asymmetric unit containing four CP74 molecules, with a solvent content of ca. 48%.

The crystal structure of CP74 was solved by the Hg-SAD phasing method using a 2.16-Å resolution data collected at the peak wavelength (1.00928Å). The positions of bound mercury were determined and then 77% of the model was automatically traced with *Autosol* (Terwilliger et al., 2009) within the *PHENIX* software suite (Adams et al., 2010). The model was further built and refined with *Coot* (Emsley and Cowtan, 2004) against a native data at 1.75-Å resolution to improve the quality of the structure. The computational refinements were conducted by *REFMAC5* (Murshudov et al., 2011) and *phenix.refine* (Afonine et al., 2012). Water molecules were located with *Coot*. Finally, the refinement converged on R_{work} and R_{free} of 0.169 and 0.195, respectively, using anisotropic temperature factors. The stereochemical quality of the refined structure was checked with *MolProbity* (Chen et al., 2010) with results tabulated in Table 1.

ITC

SAH binding to WT and CP74 was evaluated by ITC using a MicroCal iTC_{200} system (GE Life Sciences, U. S. A.) in Buffer A with 0.5 mM TCEP, replacing DTT. The protein concentration in the sample cell was set to 24.5 μM and 20 μM for WT and CP74, respectively. The concentration of SAH was set to 315 and 700 μM for WT and CP74, respectively. The protein concentrations

were confirmed by using the Pierce™ BCA Protein Assay Kit (Thermo Scientific, U.S.A). The concentration of SAH was confirmed by measuring the UV absorbance at 260 nm with an extinction coefficient at 260 nm (15400 M^{-1} cm^{-1}) using an UV-Vis spectrophotometer (V630, JASCO, Japan). The experiments were carried out at 25°C with a slower sample cell-stirring rate (650 rpm) to prevent sample aggregation due to agitation. The data were analyzed by using the Origin™ software (OriginLab, U. S. A.).

Stopped-Flow Fluorescence Measurements

The folding kinetics of CP74 was characterized at 25°C using a stopped-flow spectrometer by monitoring changes of its total intrinsic fluorescence using an excitation wavelength of 280 nm with a 320 nm cut-off filter (SX18 stopped-flow spectrometer, Applied Photophysics, UK) as described previously (Wang et al., 2015a; 2015b, 2016). The unfolding and refolding of CP74 were triggered by rapidly mixing the unfolding and refolding buffer with native and urea-denatured CP74, respectively, at an asymmetric mixing ratio of 10:1. The final protein concentration was 1.8 μM. The resulting kinetic traces were fit to double exponential to extract two observed reaction rates (k_{obs}) as a function of urea concentration. They were subsequently fit to the following equation to derive the intrinsic folding rate ($k_f^{H_2O}$) and unfolding rate ($k_u^{H_2O}$):

$$k_{obs} = k_f^{H_2O} e^{-\frac{m_f \cdot [urea]}{RT}} + k_u^{H_2O} e^{-\frac{m_u \cdot [urea]}{RT}}$$

where R is the gas constant, T the temperature in Kelvin, m_f and m_u the *m*-values associated with folding and unfolding, respectively, and [urea] the urea concentration. The non-linear regression was carried out by using Prism (Graphpad, U.S.A).

HDX-MS

HDX-MS measurements were carried out using a SYNAPT G2-Si HDMS system (Waters, U. S. A.). The protein concentrations were adjusted to 100 μM in buffer A without glycerol. The data collection was carried out in deuterated buffer (5 mM potassium phosphate in D_2O, adjusted with DCl to pD 7.0: the pH reading was 6.6 to account for the isotope effect) to trigger HDX for different durations: 0, 0.5, 10, 30, 60, and 240 min in triplicate. All HDX reactions were carried out at 25°C, followed by quenching at 0°C by adding equal amount of quench duffer (0.05 M potassium phosphate, pH 2.3 in H_2O). The protein samples were digested into peptide fragments using an immobilized pepsin digestion column (Enzymate BEH Pepsin, 2.1 x 30 mm, Waters, U. S. A.), which were trapped by a C18 trapping column (Acquity BEH VanGuard 1.7 μm, 2.1 x 5.0 mm, Waters) followed by separation of the peptides by an analytical C18 column (Acquity BEH C18: 1.7 μm, 1.0 x 100 mm, Waters, U. S. A.). All columns were temperature-controlled at 4°C. The data were processed by ProteinLynx Global Server (PLGS) (Waters, U. S. A.) to identify individual peptides. The time-dependent peptide assignments were achieved by DynamX (Waters, U. S. A.). The fractional deuterium uptakes were calculated by calculating the uptakes of individual peptides and the commonality of the overlapping regions in the coverage map to generate a heat map of relative deuterium uptake as a function of HDX time. A given peptide was selected when present in triplicate during the random on-column pepsin digestion process. The relative fractional deuterium uptake of a given peptide was calculated when it was commonly present in both triplicate experiments of WT and CP74 and the difference is defined as the deuterium uptake of CP74 minus that of WT for each peptide pair. As a result of the stringent selection criterion, the coverage of the common peptides between WT and CP74 was significantly less than the individual coverage. The fractional deuterium uptakes of individual peptides were mapped onto the crystal structures using PyMOL (Chuang et al., 2019; Sriramoju et al., 2018b).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistics of the X-ray crystallographic data processing, refinement and structure validation are summarized in Table 1. The thermal unfolding statistics derived from DSC was analyzed by Origin™ (OriginLab, U. S. A.). The thermal and chemical unfolding statistics derived from far-UV CD and intrinsic fluorescence were analyzed by Prism (Graphpad, U.S.A) to extract the associated physicochemical parameters as tabulated in Table S1. The kinetics parameters derived from stopped-flow fluorescence measurements were analyzed by using Prism (Graphpad, U.S.A), and the results were tabulated in Table 2.

DATA AND SOFTWARE AVAILABILITY

The accession number for the crystal structure of CP74 reported in this paper is Protein Data Bank (PDB): 5ZYO.