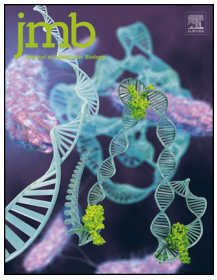

Untying a Protein Knot by Circular Permutation

Ya-Chu Chuang [1,2], I-Chen Hu [2], Ping-Chiang Lyu [2] and Shang-Te Danny Hsu [1,3,]

*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 - **Institute of Biochemical Sciences,** National Taiwan University, Taipei 10617, Taiwan*

***Correspondence to Ping-Chiang Lyu and Shang-Te Danny Hsu:** S.-T. Danny Hsu is to be contacted at: Institute of Biological Chemistry, Academia Sinica, Taipei 11529, Taiwan. pclyu@mx.nthu.edu.tw, sthsu@gate.sinica.edu.tw*
https://doi.org/10.1016/j.jmb.2019.01.005
Edited by Sheena Radford

Abstract

Topologically knotted proteins are tantalizing examples of how polypeptide chains can explore complex free energy landscapes to efficiently attain defined knotted conformations. The evolution trails of protein knots, however, remain elusive. We used circular permutation to change an evolutionally conserved topologically knotted SPOUT RNA methyltransferase into an unknotted form. The unknotted variant adopted the same three-dimensional structure and oligomeric state as its knotted parent, but its folding stability was markedly reduced with accelerated folding kinetics and its ligand binding was abrogated. Our findings support the hypothesis that the universally conserved knotted topology of the SPOUT superfamily evolved from unknotted forms through circular permutation under selection pressure for folding robustness and, more importantly, for functional requirements associated with the knotted structural element.
© 2019 Elsevier Ltd. All rights reserved.

Topologically knotted proteins are exquisite examples of how evolution can fold and thread intricate protein knots with efficiency and precision [1]. To date, over 1000 knotted protein structures have been deposited in the protein databank [2]. The SPOUT RNA methyltransferase superfamily is one of the best-known examples of knotted protein structures. SPOUT proteins exist in all kingdoms of life. All reported SPOUT protein structures contain a trefoil (3_1) knotted structural element that binds to the cofactor *S*-adenosyl-L-homocysteine (SAH), which acts as a donor for methyltransfer reactions [3]. The knotted cofactor-binding motif provides conformational constraints to activate the enzymatic activity.

YibK from *Haemophilus influenzae*, also known as TrmL, specifically modifies tRNA by 2′-*O*-methylation of the wobble nucleotide. It is also one of the best-studied knotted proteins in terms of the mechanisms underlying the folding and knotting [4,5]. YibK is a homodimeric α/β protein with a long C-terminal helix that threads through the knotting loop that binds the SAH co-factor, rendering it one of the most deeply knotted proteins known [6]. YibK can fold and knot spontaneously without the aid of molecular chaperones [7]. It forms several well-defined folding intermediates along its bifurcated folding pathway [4]. Correct folding and dimerization of YibK are essential for cofactor binding [8]. Remarkably, there is good experimental evidence to suggest that YibK remains topologically knotted under highly denaturing conditions [9], despite the lack of appreciable secondary structure content [10] and conformational compaction [11]. That raises the question of whether the knotting occurs co-translationally or post-translationally through a different folding pathway than the ones deduced from *in vitro* refolding experiments.

Bioinformatics survey of the SPOUT superfamily revealed that the trefoil knotted element is universally conserved. It was postulated that the trefoil protein knot may have evolved through circular permutation (CP) of unknotted parent proteins [3]. A circularly permuted protein pair has the same amino acid composition, but the termini are shifted such that the segment that forms the C-terminus in one protein is shifted to the N-terminus in the other protein by becoming concatenated to the original N-terminus of the second protein [12].

© 2019 Elsevier Ltd. All rights reserved.

Proteins that arise naturally as a result of CP adopt three-dimensional (3D) structures identical to those of their parent proteins. Engineered CP has been used to study the effect of reshuffling long-range native contacts in one-dimensional sequential terms, known as the contact order [13], on the folding kinetics while maintaining the nature and number of native contacts in 3D space. The folding kinetics of topologically knotted proteins are generally slower and more complex than those of the small globular proteins that exhibit two-state folding behaviors. All of the knotted proteins that have been experimentally examined so far exhibit highly populated kinetic folding intermediates, which are thought to be rate limiting [14–20]. We asked whether it is possible to generate an unknotted CP variant of YibK with the same 3D spatial arrangement of structural elements as the parent protein (i.e., the same native contacts, but different sequential order) and how CP and the consequential unknotting impact on the folding and function of YibK.

Using the structure-based CP predicting algorithm CPDB [21], we introduced nine CP sites into the knotted region of YibK to either open the knotting loop or shift the C-terminal helix to become the new N-terminus and thereby alleviate the need for threading (Supplementary information, Table S1 and Fig. S1). When we expressed each of the nine CP variants in *Escherichia coli*, CP82 (opening the threading loop at residue position 82) was expressed predominantly in the soluble fraction that can be purified using the same procedure used to purify the wild-type YibK (WT). The far-UV circular dichroism (CD) spectrum of CP82 is very similar to that of WT, indicating that the secondary structure content was preserved despite the reorganized sequential order (Fig. 1A). Size-exclusion chromatography-coupled multiple angle light scattering (SEC-MALS) confirmed the dimeric quaternary structures of both WT and CP82 with the expected molecular weights (Fig. 1B). Nevertheless, CP82 had a significantly reduced melting temperature (T_m; $T_{m,WT}$ = 64.5 °C *versus* $T_{m,CP82}$ = 50.0 °C; Table S2), as revealed by the CD signals at 222 nm as a function of temperature (Fig. 1C) and differential scanning calorimetry (DSC; Fig. S2). Urea-based equilibrium unfolding analysis by intrinsic fluorescence and far-UV CD spectroscopy revealed that compared with those for WT, the transition urea concentrations ($[D]_{50\%}$) for CP82 were reduced by 2.94 ± 0.01 and 1.54 ± 0.05 M, respectively (Fig. 1D). Fluorescence and CD data yielded the same unfolding isotherms for WT; however, there was a clear difference for CP82, with fluorescence data yielding a much lower stability than CD data, suggesting a loss of folding cooperativity in CP82. That was corroborated by the presence of an additional DSC profile of CP82 that was not detected by CD-based thermo-unfolding (Fig. S2). Further optimizations of refolding parameters (Supplementary information) enabled the recovery of CP118, CP124, CP128, and CP132 in their soluble forms, albeit with lower abundances and solubility than CP82. Those CP variants exhibited marginally different far-UV CD spectra that indicated reduced secondary structure content (Fig. S3A) with much lower melting temperatures (reduced T_m by 14–21 °C; Table S2) and lower thermo-unfolding cooperativity (Fig. S3B). Those results demonstrated that we could obtain unknotted CP variants of YibK with different reductions in folding stability.

To investigate how CP perturbs the structure of YibK, we solved the crystal structure of CP82 to a resolution of 1.65 Å (deposited under the PDB accession code: 6AHW; Fig. 2 and Table S3). CP82 adopted the same 3D structure and dimeric assembly as WT, with an overall backbone Cα atomic positional RMSD of 0.40 Å for the dimer as a whole. The only discernable structural difference lies in α2 and α3 (α4 and α5 in CP82) and the N- and C-termini (Fig. 2C). Structural analysis indicated a reduction in the dimer interface area (from ca. 1600 Å^2 per chain in WT to ca. 1440 Å^2 per chain in CP82; Fig. S4 and Table S4), which we ascribed to differences in inter-molecular hydrogen bonding and salt bridges around the periphery of the stable hydrophobic dimer interface (Fig. S5 and Table S5). We confirmed the quaternary structure of CP82

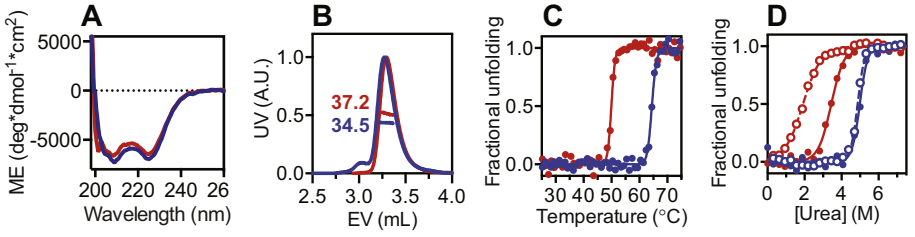

Fig. 1. Comparisons of structures and folding of WT (blue) and CP82 (red). (A) Far-UV CD spectra and (B) SEC-MALS profiles showing the same secondary and quaternary structures, respectively. The corresponding molecular weight (MW) distributions are shown as horizontal lines in the elution peaks, indicated with the averaged MW values in kDa on the left. Normalized fractions of unfolding derived from far-UV CD thermo-unfolding (C) and equilibrium urea unfolding (D). Intrinsic fluorescence and far-UV CD data points in panel D are shown in open and filled circles, respectively.

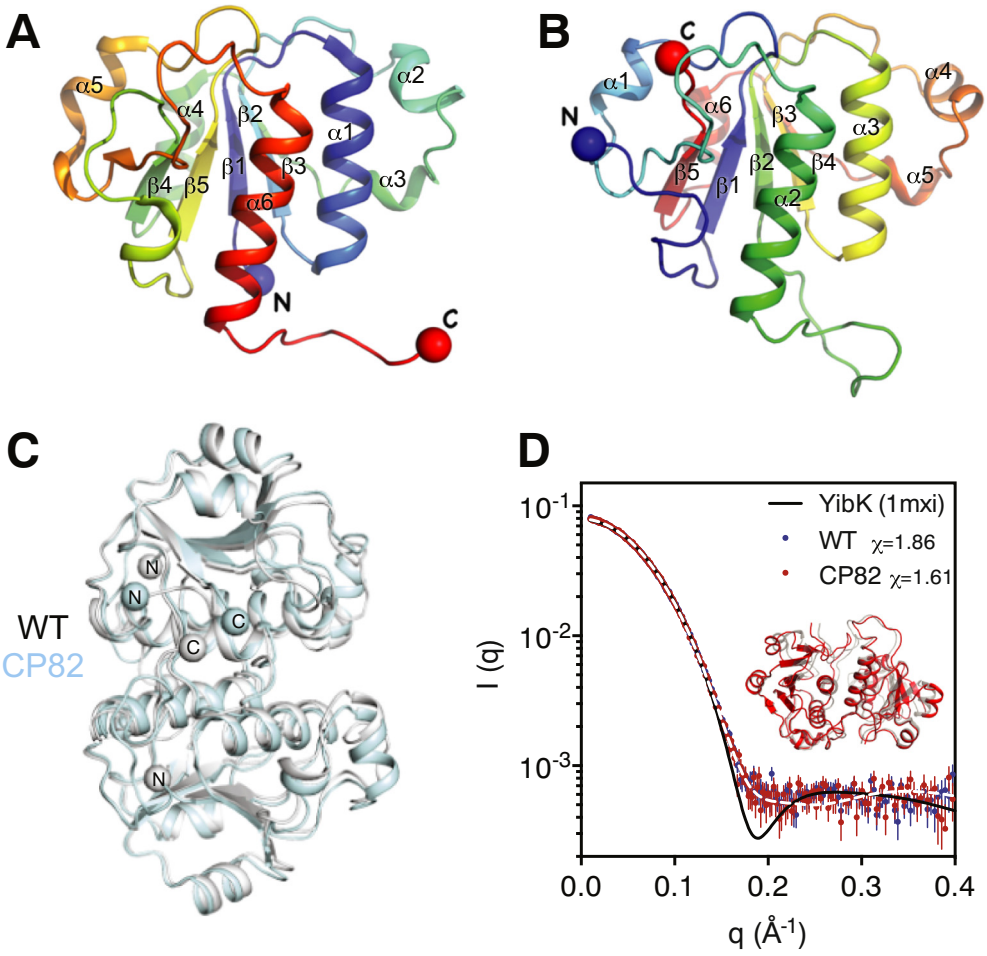

Fig. 2. Crystal structures of WT (A) and CP82 (B) shown in cartoon representations color-ramped in a reversed rainbow scale from the N-termini to the C-termini, indicated by blue and red spheres, respectively. The identities of individual secondary structure elements are indicated. (C) Superposition of the structures of dimeric WT in light gray and CP82 in light cyan. The Cα atoms of the N- and C-termini are shown in sphere and indicated accordingly. (D) SAXS profiles of WT (blue) and CP82 (red) superimposed with the back-calculated profile based on the WT structure (black line; the corresponding fitting χ-values for WT and CP82 are indicated on the upper right corner) and that of a model derived from MultiFoXS (white dashed line). Inset: Superposition of the crystal structure of CP82 (transparent gray) and the MultiFoXS-derived model (red).

in solution by SEC-coupled small-angle X-ray scattering (SEC-SAXS). The SAXS profiles and the corresponding Kratky plots of WT and CP82 were superimposable (Figs. 2D and S4), yielding identical radii gyration (R_g) of 21.2 ± 0.1 Å based on Guinier approximation (Fig. S6); the theoretical value based on the crystal structure of WT (PDB entry: 1MXI) is 19.2 Å. To account for the deviations between the experimental SAXS profile and back-calculated SAXS profile based on the WT crystal structure in the mid-q range (momentum transfer, q ~ 0.2), we generated a model using MultiFoXS [22] that yielded a theoretical SAXS profile that showed much improved match with the experiment profiles (Fig. 2D). The RMSD between the model and the crystal structure of CP82 was 1.63 Å, primarily due to marginal translational movements between the two (inset in Fig. 2D).

Despite the loop opening as a result of the introduction of the CP site on the threading loop, the SAH-interacting residues in apo CP82 were aligned in the same spatial arrangements as those in SAH-bound WT. Superposition of the two structures indicated that the SAH-binding pocket in CP82 is poised to form an extensive hydrogen bonding network with the adenine and ribose moieties of SAH. The only exceptions were the backbone carbonyl oxygen of Leu160 in CP82 (Leu78 in WT) that was displaced by ca. 0.5 Å from the 2′ hydroxyl oxygen of SAH, and the side-chain hydroxyl oxygen of Ser130 (Ser50 in WT) that was pointing away from, instead of toward, the solvent exposed hence more flexible amino oxygen of the homocysteine moiety of SAH (Fig. 3). To examine how CP may impact the SAH-binding capacity of CP82, which is functionally required for the tRNA methylation activity of YibK and

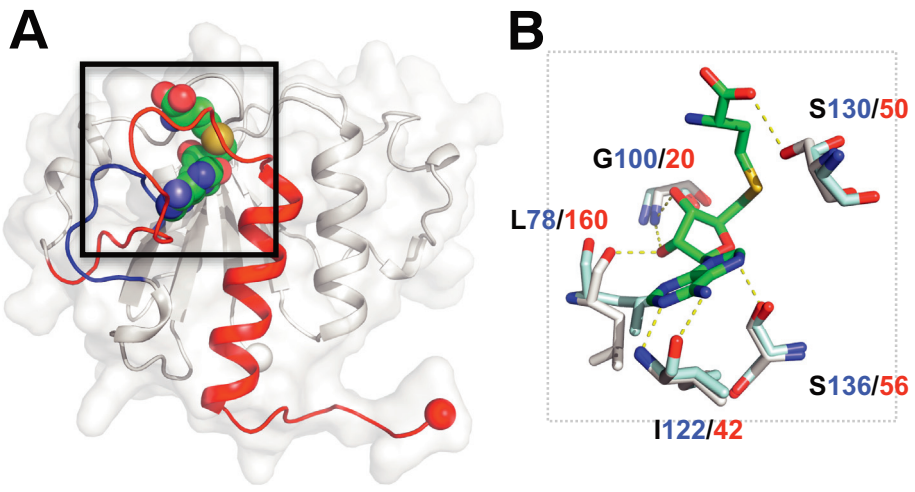

Fig. 3. Structural comparison of the substrate-binding motifs of WT and CP82. (A) The knotted structural element is involved in SAH recognition. The knotting loop and the C-terminal threading helix are colored in blue and red, respectively. SAH is shown as spheres with carbon, nitrogen, oxygen, and sulfur atoms colored green, blue, red, and gold, respectively. (B) Stick representation of the structural alignment of the SAH-interacting residues in WT (light gray) and the corresponding residues in CP82 (light cyan). Hydrogen bonds mediating SAH binding are shown as gold dashed lines. SAH is colored in the same scheme used in panel A.

all other SPOUT family members, we performed isothermal titration calorimetry to characterize SAH binding to WT and CP82. Consistent with previous findings, WT exhibited an SAH dissociation constant (K_d) of 24.5 ± 1.0 μM [8], whereas CP82 showed no appreciable binding to SAH, indicating that while the crystallographic analysis revealed only marginal structural perturbations within the SAH-binding motif, the loop opening (and unknotting) resulted in profound loss of SAH-binding affinity (Fig. S7).

To understand the loss of the SAH-binding affinity of CP82, we performed comparative native folding dynamics analysis of WT and CP82 using hydrogen–deuterium exchange mass spectroscopy (HDX-MS). Structural mapping of the differential deuterium uptakes of backbone amide groups showed that the most affected regions were the SAH-binding motif and the C-terminal residues of WT, which were moved to the N-terminus in CP82 (Fig. 4). In fact, almost half of CP82, particularly at the dimer interface, showed markedly enhanced HDX, indicating pronounced local fluctuations that were not reflected in the crystallographic study [23]. In keeping with the equilibrium thermochemical unfolding analyses (Fig. 1D), the HDX-MS analysis underscored the stabilizing effect of the knotted structural motif, but with much more detailed structural information, as in the case of the trefoil-knotted ornithine transcarbamylases [23]. It also suggested that the loop opening caused not only the immediately consequential unknotting but also long-range perturbations to the dimer interface, echoing the earlier mutational study of YibK that demonstrated the causal relationship between dimer interfacial disruption and the loss of SAH-binding capacity.

We next sought to investigate the impact of CP on the folding kinetics of YibK. On the one hand, the topological rearrangement increases the contact order of CP82 was to 12.3% compared 10.7% for WT, which is expected to lower the folding kinetics of CP82 relatively to that of WT [13]. On the other hand, untying the knotted topology of the unknotted CP82 may simplify and accelerate the folding kinetics compared to that of WT, which exhibits a complex bifurcated pathways linking the denatured state and two distinct monomeric folding intermediates followed by the formation of a native-like monomeric intermediate before assembling into a native dimer [4,5]. We carried out stopped-flow fluorescence measurements of CP82 and observed three distinct unfolding phases and four refolding phases (Fig. 5 and Table S6). While such a multiphasic folding behavior was similar to that of WT, most of microscopic reaction rate constants of CP82 were faster than their counterparts in WT with significantly larger *m*-values (note that an assumption was made that the unfolding and refolding arms could be paired in ascending order for CP82 in the same way as it was analyzed for WT. Consequently, the fastest refolding arm did not have a corresponding unfolding arm for chevron plot analysis). Complete delineation of the folding pathways of CP82 was limited in part by the exceedingly small amplitudes of the two fast refolding phases (Fig. S13). While it remains to be seen as to how CP-induced untying of the knotted topology of YibK remodels the associated free energy landscape, it is conceivable that CP effectively alleviates the topological frustration in the associated free energy landscape manifested in the accelerated the folding kinetics between the individual states.

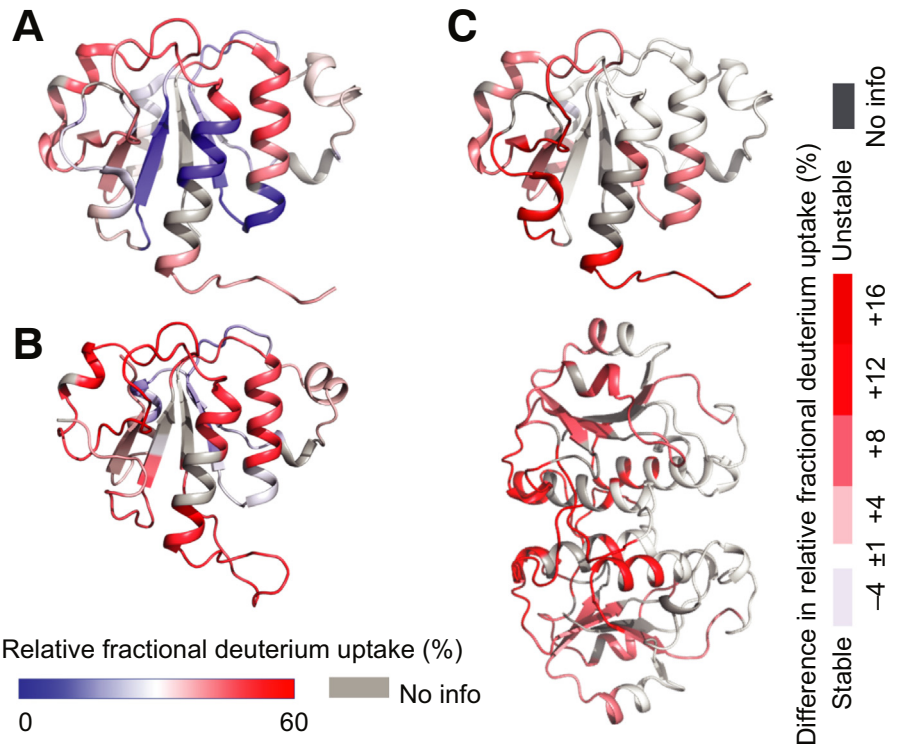

Fig. 4. HDX-MS analyses of WT (A) and CP82 (B). Structural mapping of the relative deuterium uptakes after incubation in deuterated buffer for 1 h at 25 °C. The hydrophobic core and the dimer interface (facing out of plane) were relatively more stable in WT (indicated by the blue patches), whereas the corresponding regions in CP82 were more susceptible to HDX. (C) Relative deuterium uptake of CP82 with respect to WT by subtracting the results in panel B from those in panel A. The dimeric structure of YibK is shown in the lower panel to highlight the long-range perturbations to the dimer interface. Dark red coloring corresponds to more destabilization caused by CP. The lack of common peptides between WT and CP82 in some regions, especially those within the dimer interface and the central β-sheet, precluded the ability to derive the relative deuterium uptakes. These regions are highlighted in dark gray (cf. Supporting Information and Figs. S8–S12 for details).

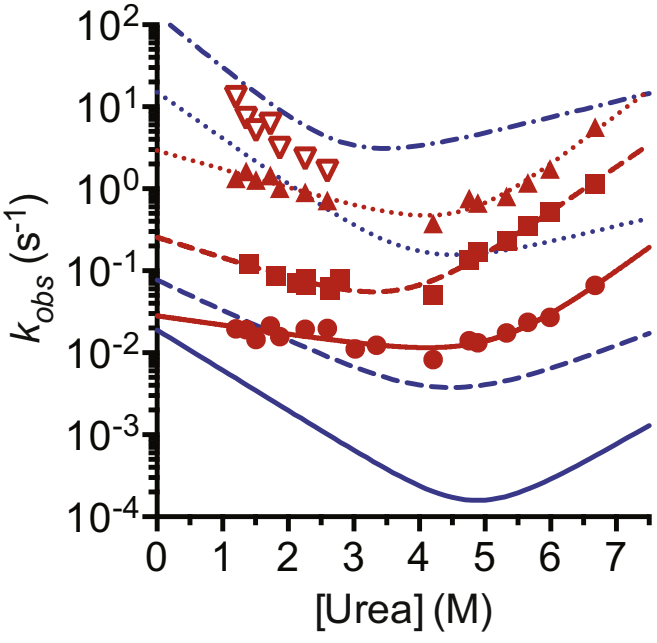

Fig. 5. Folding kinetic analysis of CP82 as a function of urea concentration. CP82 exhibited three unfolding kinetic phases and four refolding kinetic phases. The unfolding and refolding arms were paired in ascending orders (except for the fastest refolding phase) and fit to a two-state folding model to generate the chevron plots in solid, dashed, and dotted red lines. For comparison, the previously reported folding kinetic phases of WT were shown in solid, dashed, dotted, and dash-dotted blue lines in ascending order.

In summary, we reconfigured the sequence of the trefoil-knotted YibK from *H. influenzae* to create unknotted CP variants that can fold into the same 3D structure as YibK by forming the same native contacts as those in the parent protein. Our results suggest that the knotted structural element in the SPOUT superfamily is not essential in terms of proper protein structure and folding. Nevertheless, the knotted element provides enhanced thermo-chemical stability that might be evolutionarily advantageous. While untying the trefoil knot significantly accelerates the folding kinetics, opening the knotting loop abrogates the substrate binding capacity of the protein, possibly by removing the topological constraint required to maintain the strained knotting loop, which in turn bends the substrate into a strained, activated configuration. Our findings underscore the thermodynamic, kinetics and functional importance of the universally conserved protein knot in the SPOUT superfamily [3]. Considering the existence of other highly conserved knotted proteins, such as ubiquitin C-terminal hydrolases in eukaryotes, that utilize their knotted structural elements for substrate recognition and also for additional mechanostability [24], it is conceivable that protein knots were evolutionarily selected because of the synergistic relationship between folding stability and functionality despite the apparent frustrations across their free energy landscapes.

Accession numbers

The X-ray crystal structure has been deposited under the PDB accession code: 6AHW

Acknowledgments

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. We also thank Dr. Meng-Ru Ho of the Biophysics Facility, Ms. Hui-Ling Shih and Dr. Kai-Fa Huang of the Protein Crystalization Facility, Institute of Biological Chemistry, Academia Sinica, and members of the Hsu laboratory for assisting data collection. We acknowledge the National Synchrotron Radiation Research Center, Taiwan, for use of the BL13B beamline for protein crystallography and the BL23A beamline for SEC-SAXS data collection.

Funding Sources: No competing financial interests have been declared. This work was supported by the Ministry of Science and Technology, Taiwan (105-2113-M-001-005, 106-2113-M-001-004, and 107-2628-M-001-005-MY3 for S.-T.D.H. and 106-2311-B-007-004-MY3 for P.C.L.) and Academia Sinica, Taiwan.

Appendix A. Supplementary data

Supporting Information includes the Materials and Methods and additional information regarding the structural, folding, and ligand-binding assays. Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2019.01.005.

Received 17 November 2018;
Received in revised form 2 January 2019;
Accepted 2 January 2019
Available online 9 January 2019

Keywords:
knotted protein;
SPOUT RNA methyltransferase;
protein folding;
circular permutation

Abbreviations used:
SAH, *S*-adenosyl-L-homocysteine; CP, circular permutation; 3D, three-dimensional; WT, wild type; CD, circular dichroism; SEC-SAXS, size-exclusion chromatography-coupled small-angle X-ray scattering; HDX-MS, hydrogen–deuterium exchange mass spectroscopy; SEC-MALS, size-exclusion chromatography-coupled multiple angle light scattering; DSC, differential scanning calorimetry.

References

[1] S.E. Jackson, A. Suma, C. Micheletti, How to fold intricately: using theory and experiments to unravel the properties of knotted proteins, Curr. Opin. Struct. Biol. 42 (2017) 6–14.
[2] M. Jamroz, W. Niemyska, E.J. Rawdon, A. Stasiak, K.C. Millett, P. Sulkowski, et al., KnotProt: a database of proteins with knots and slipknots, Nucleic Acids Res. 43 (2015) D306–D314.
[3] K.L. Tkaczuk, S. Dunin-Horkawicz, E. Purta, J.M. Bujnicki, Structural and evolutionary bioinformatics of the SPOUT superfamily of methyltransferases, BMC Bioinforma. 8 (2007) 73.
[4] A.L. Mallam, S.E. Jackson, Probing nature's knots: the folding pathway of a knotted homodimeric protein, J. Mol. Biol. 359 (2006) 1420–1436.
[5] A.L. Mallam, E.R. Morris, S.E. Jackson, Exploring knotting mechanisms in protein folding, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 18740–18745.
[6] K. Lim, H. Zhang, A. Tempczyk, W. Krajewski, N. Bonander, J. Toedt, et al., Structure of the YibK methyltransferase from *Haemophilus influenzae* (HI0766): a cofactor bound at a site formed by a knot, Proteins 51 (2003) 56–67.
[7] A.L. Mallam, S.E. Jackson, Knot formation in newly translated proteins is spontaneous and accelerated by chaperonins, Nat. Chem. Biol. 8 (2012) 147–153.
[8] A.L. Mallam, S.E. Jackson, The dimerization of an α/β-knotted protein is essential for structure and function, Structure 15 (2007) 111–122.
[9] A.L. Mallam, J.M. Rogers, S.E. Jackson, Experimental detection of knotted conformations in denatured proteins, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 8189–8194.

[10] S.J. Hsieh, A.L. Mallam, S.E. Jackson, S.T.D. Hsu, Backbone ^{1}H, ^{13}C and ^{15}N assignments of YibK and avariant containing a unique cysteine residue at C-terminus in 8 M urea-denatured states [corrected], Biomol. NMR Assign. 8 (2014) 439–442.
[11] P.M. Shih, I. Wang, Y.T.C. Lee, S.J. Hsieh, S.Y. Chen, L.W. Wang, et al., Random-coil behavior of chemically denatured topologically knotted proteins revealed by small-angle X-ray scattering, J. Phys. Chem. B 119 (2015) 5437–5443.
[12] W.C. Lo, C.C. Lee, C.Y. Lee, P.C. Lyu, CPDB: a database of circular permutation in proteins, Nucleic Acids Res. 37 (2009) D328–D332.
[13] K.W. Plaxco, K.T. Simons, D. Baker, Contact order, transition state placement and the refolding rates of single domain proteins, J. Mol. Biol. 277 (1998) 985–994.
[14] A.L. Mallam, S.E. Jackson, A comparison of the folding of two knotted proteins: YbeA and YibK, J. Mol. Biol. 366 (2007) 650–665.
[15] F.I. Andersson, D.G. Pina, A.L. Mallam, G. Blaser, S.E. Jackson, Untangling the folding mechanism of the 5_2-knotted protein UCH-L3, FEBS J. 276 (2009) 2625–2635.
[16] I. Wang, S.Y. Chen, S.T.D. Hsu, Unraveling the folding mechanism of the smallest knotted protein, MJ0366, J. Phys. Chem. B 119 (2015) 4359–4370.
[17] I. Wang, S.Y. Chen, S.T.D. Hsu, Folding analysis of the most complex Stevedore's protein knot, Sci. Rep. 6 (2016), 31514.
[18] L.W. Wang, Y.N. Liu, P.C. Lyu, S.E. Jackson, S.T.D. Hsu, Comparative analysis of the folding dynamics and kinetics of an engineered knotted protein and its variants derived from HP0242 of *Helicobacter pylori*, J. Phys. Condens. Matter 27 (2015).
[19] S.C. Lou, S. Wetzel, H. Zhang, E.W. Crone, Y.T. Lee, S.E. Jackson, et al., 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 (2016) 2507–2520.
[20] Y.C. Lee, C.Y. Chang, S.Y. Chen, Y.R. Pan, M.R. Ho, S.T.D. Hsu, Entropic stabilization of a deubiquitinase provides conformational plasticity and slow unfolding kinetics beneficial for functioning on the proteasome, Sci. Rep. 7 (2017) 45174.
[21] W.C. Lo, L.F. Wang, Y.Y. Liu, T. Dai, J.K. Hwang, P.C. Lyu, CPred: a web server for predicting viable circular permutations in proteins, Nucleic Acids Res. 40 (2012) W232–W237.
[22] D. Schneidman-Duhovny, M. Hammel, J.A. Tainer, A. Sali, FoXS, FoXSDock and MultiFoXS: single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles, Nucleic Acids Res. 44 (2016) W424–W429.
[23] M.K. Sriramoju, T.J. Yang, S.T.D. Hsu, Comparative folding analyses of unknotted versus trefoil-knotted ornithine trans-carbamylases suggest stabilizing effects of protein knots, Biochem. Biophys. Res. Commun. 503 (2018) 822–829.
[24] M.K. Sriramoju, Y. Chen, Y.C. Lee, S.T.D. Hsu, Topologically knotted deubiquitinases exhibit unprecedented mechanostability to withstand the proteolysis by an AAA+ protease, Sci. Rep. 8 (2018) 7076.