Supporting Information

Sułkowska *et al.* 10.1073/pnas.0805468105

SI Materials

In this supporting information, we present more details concerning the two proteins we consider: 1c9y (not knotted) and 1yh1 (knotted).

The Two Proteins. In the classical arginine biosynthetic pathway, the ormithine transcarbamylase (OTCase) enzyme first deacetylates *N*-acetylornithine to the L-ornithine and then forms citrulline through carbamylation. In the other pathway involving *N*-acetylornithine transcarbamylase (AOTCase), the process is reversed: *N*-acetylornithine is first carbamylated to the *N*-acetylocitrulline and then deacetylated to the citrulline.

The folding, thermal, and mechanical properties of these two proteins have not been compared up to now, mostly because the structure of AOTCase has not been known until recently and because the presence of the knot makes experiments harder to interpret. However, it has been determined that both substrates *N*-acetylcitrulline (1yh1) (1) and L-norvaline (1c9y) (2) obey Michaelis–Menten kinetics (3). It has also been found (3) that affinity of the human OTCase to ornithine is 10 times greater than the affinity of AOTCase to *N*-acetylornithine, whereas the affinity for carbamyl phosphate is approximately five times smaller. The thermal stability was measured only for the OTCase 1c9y (by measuring the temperature at which 50% of the enzyme activity is lost) was determined to be 56 ± 1 °C (4).

It must noted that there exists another member of the transcarbamoylase family, SOTCase (extracted from *Bacillus fragilis* argF′, 1sj1), which also contains a knot. The rmsd between two structures, based on 280 equivalent C_α positions, is ≈1.4 Å (1). There are also structures that are similar to the human OTCase 1c9y like *Escherichia coli* ATCase (PDB ID code 1ekx) with rmsd 1.7 Å based on 262 equivalent C_α atoms (1). The superpositions of these 4 proteins with their substrates were carried out in ref. 5, where only slight differences between corresponding pairs of the knotted and unknotted proteins were found. We have also checked that the properties of 1js1 and 1ekx in the model are nearly identical to those of 1yh1 and 1c9y, respectively. For this reason, our analysis is focused on the 1yh1 and 1c9y proteins.

Typical Trajectory of a Knot in 1yh1 under Stretching by a Constant Force. We now analyze an example of the constant force pathway for 1yh1 for $F > 1.9$ ε/Å in more detail (Fig. S1). The right end of the knot is initially located close to β-strand L (248–252). The left end of the knot is located at the sharp turn involving glycine (170), and it is rather hard to force this end to leave this turn. For this reason, when a constant force is applied, the right end of the knot starts to move: as the protein backbone is being pulled out of the loop, the sequential location of this end decreases. Interestingly, the decrease is linear in time (Fig. S1 *Upper*) and it involves motion of the right end across several pinning centers comprising the turn at His-237, located between β-strand K (232–236), the small helix H2 (238–244), and the sharp turn at Pro-210. These centers are seen as accumulations of points along the tilted interval in Fig. S1 *Upper*. Nonetheless, these pinning sites are weaker than the force that holds the left end at site 170 and, hence, it is only the right end that can slide. The situation changes when the right end reaches Gly-200 within a sharp turn at the end of a helix H1. Interestingly, the left end is now "pushed" out of site 170 by the right end and it starts to move to the left, expanding the knot region a bit and resulting in a translation of the whole knot to the left. This translational motion stops on the left at Gly-164 at the end of the long helix (147–164). At the same time the right end passes through a half-loop (with a sharp and rigid turn involving prolines at 181 and 183) between β-strand G (172–177) and helix H1 (185–199). Finally, the right end stops at site 175 and the knot becomes fully tightened. This scenario of events is consistent with the observation of the role of the sharp structural turns in the dynamics of knot's ends made in ref. 6.

We also analyzed in detail an example of a constant force pathway for 1yh1 for $F > 1.9$ ε/A (Fig. S1 *Upper*). The scenarios of events reported here are consistent with the prominent role of the sharp structural turns in the dynamics of knot's ends (see Ref. 6). In particular, when the knot is tightened, its right end moves across several pinning centers comprising the turn between β strand K and the small helix H2, and the sharp turn at Pro-210, slowing down at successive pinning centers. Each slowing down manifests itself as accumulation of points along the tilted interval in Fig. S1 *Upper*.

So far we have discussed the differences between 1c9y and 1yh1 as seen at the level of single stretching trajectories. These differences are also visible after averaging over many trajectories, as demonstrated in Fig. S1 *Lower*. In particular, we find that for $\tilde{F} = 2.0$, which allows for the full extension in both proteins, 1yh1 takes longer to unfold than 1c9y. However, for forces higher than 2.3 ε/A the differences in the averaged trajectories are minor. Fig. S1 *Lower* shows that to match the time scale of unfolding 1c9y at $\tilde{F} = 1.9$ in 1yh1 one has to enhance the value of $\tilde{F}$ to 2.2.

We enclose a video presentation of the stretching of the two proteins, as generated using our implementation of the Go-like model. The first animation presents the protein 1c9y, and the second, protein 1yh1. The process of tightening of the knot in the animation corresponds to Fig. 2 of the main text. The knotted region in 1yh1 and corresponding region in 1c9y are marked in green.

Threshold Force F*. We define F^* as the threshold force at which the free-energy barrier for the transition from the native to the unfolded state vanishes and the protein begins to unfold in a downhill manner. Unfolding is then essentially immediate, without any intermediate states (see Fig. 5 *Inset*). The force F^* is analogous to that found in simulations of ubiquitin (7) above which the unfolding times are short and distributed log-normally and below which they are substantially longer and distributed exponentially. For forces below F^*, the median unfolding times follow a trend, which in general is a superposition of exponential functions (7). For forces above F^*, unfolding times also decrease with an increasing force, but at a much slower rate.

For 1yh1 and 1c9y we find F^* of 3.2 and 2.5 ε/Å, respectively, as indicated in Fig. 7 by the arrows. The data shown in this figure are based on 300 trajectories for $F < 1.9$ ε/Å and 100 trajectories for $F > 1.9\varepsilon/Å$. The relative shift in the location of F^* is notable: F^* for the knotted 1yh1 is higher pointing to a higher stability.

Thermal Untying of the Knot in the 1yh1 Protein. Fig. 7 shows the knot untying process in a schematic way. Fig. S2 shows the corresponding conformations in more detail, with the original position of the knot along the backbone marked.

Backtracking. The process that involves a series of breaking and forming of the same group of contacts due to topological barrier is called backtracking (8, 9). Complete thermal unfolding of the

knotted proteins (i.e., unfolding to the trivial topology, with the knot untightened) would not be possible without such backtracking. An example of a backtracking due to knot topology is untying of the protein 1yh1 from the N-terminal. In this case the knot has to move along almost entire chain. The translocation of the knot across the backbone is correlated with refolding due to backtracking of a part of the structure, as seen in Fig. S3. *Lower* shows the number of contacts Q in domain b during unfolding. Fig. S3 *Upper* shows the number of contacts Q inside the domain a. In the native state the position of the knot is stabilized by contacts G+I and I+K (in domain b). These contacts periodically break (black line); however, until $2{,}800\tau$ the knot is localized in domain b, whereas in domain a all contacts keep breaking randomly. When the knot moves to domain a at $2{,}800\tau$, the periodic refolding of contacts A+E is observed (Fig. S3 *Upper*). Eventually, the knot slides off the chain through the terminus N.

1. Morizono H, *et al.* (2006) Acetylornithine transcarbamylase: A novel enzyme in arginine biosynthesis. *J Bacteriol* 188:2974–2982.
2. Morizono H, *et al.* (1997) Expression, purification and kinetic characterization of wild-type human ornithine transcarbamylase and a recurrent mutant that produces-late onset-hyperammonaemia. *Biochem J* 322:625–631.
3. Shi D, Moriziono H, Aoyagi M, Tuchman M, Allewell NM (2000) Crystal structure of human ornithine transcarbamylse complexed with carbamyl phosphate and l-norvaline at 1.9 a resolution. *Proteins Struct Funct Genet* 39:271–277.
4. Shi D, Morizono H, Yu X, Caldovic L (2005) Crystal structure of n-acetylornithine transcarbamylas from xanthomonas cmapestris: A novel enzyme in a new arginine biosynthetic pathway found in several eubacteria. *J Biol Chem* 280:14366–14369.
5. Shi D, Yu X, Roth L, Morizono H, Tuchman M, Allewell NM (2006) Structures of n-acetylornithine transcarbamoylase from xanthomonas campestris complexed with substrates and substrate analogs imply mechanisms for substrate binding and catalysis. *Proteins Struct Funct Genet* 64:532–542.
6. Sułkowska JI, Sułkowski P, Szymczak P, Cieplak M (2008) Tightening of knots in the proteins. *Phys Rev Lett* 100:58106.
7. Szymczak P, Cieplak M (2006) Stretching proteins in the force-clamp. *J Chem Phys* 18:L21–L28.
8. Gosavi S, Chavez LL, Jennings PA, Onuchic JN (2006) Topological frustration and the folding of interleukin-1beta. *J Mol Biol* 357:986–996.
9. Gosavi S, Whitford PC, Jennings PA, Onuchic JN (2008) Extracting function from a β-trefoil folding motif *Proc Natl Acad Sci USA* 105:10384–10389.

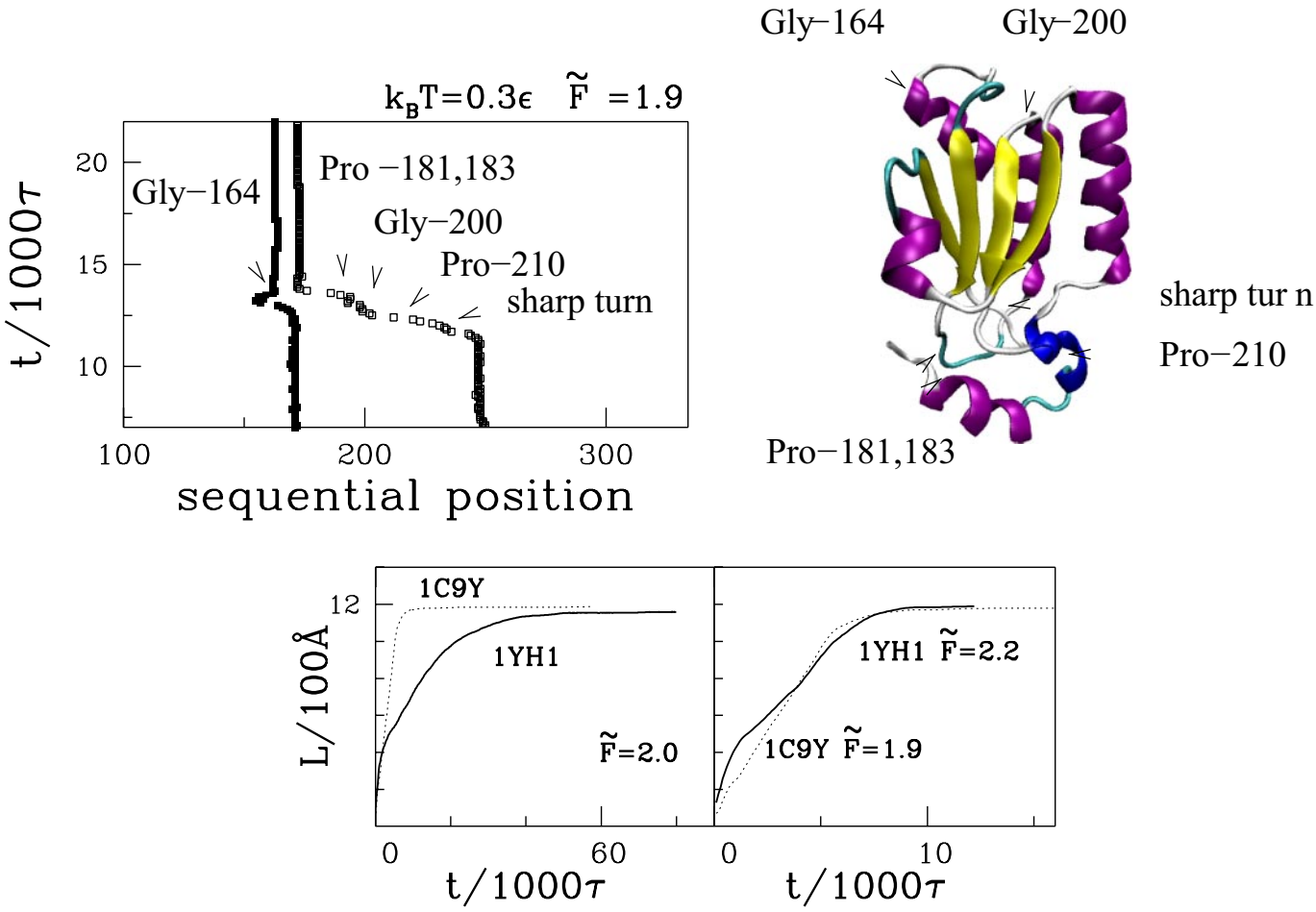


Fig. S1. Dynamics of the knot and the average end-to-end distance during stretching at constant force. (*Upper*) A typical trajectory of knot's ends locations for stretching at constant force (*Left*) and the corresponding region of the knot in 1yh1 shown in the diagrammatic representation (*Right*). The pinning centers are indicated on both sides of the figure. (*Lower*) The average end-to-end distance as a function of time for the forces indicated.

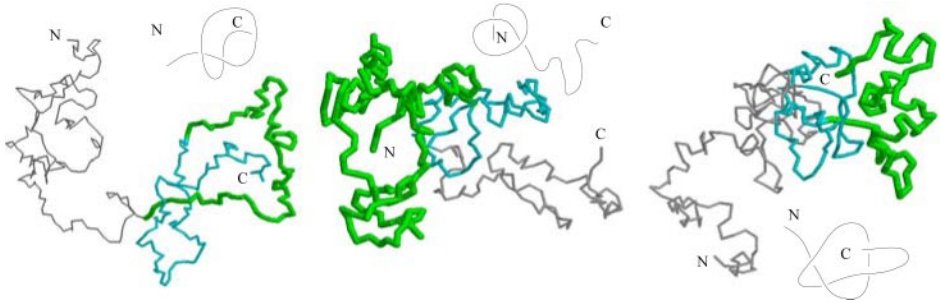


Fig. S2. Three possible ways of thermal untying of the knot. From the left to the right: simple from the C terminus, simple from the N terminus, and through formation of a slipknot. The cyan (medium thick) line indicates the native location of the knot, whereas the green (thick) line in the *Center* and *Left* shows the instantaneous position of the knot. (*Right*) The position of the slipknot is indicated by the combined lines of medium and large thickness.

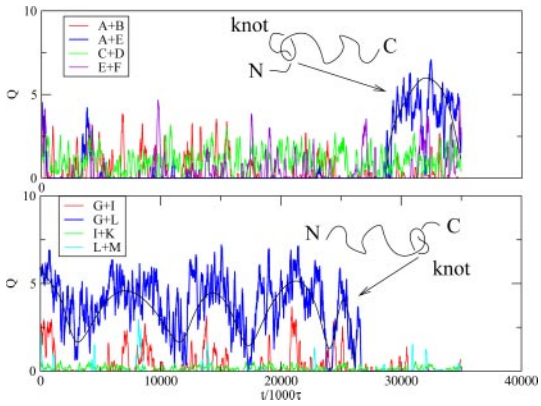
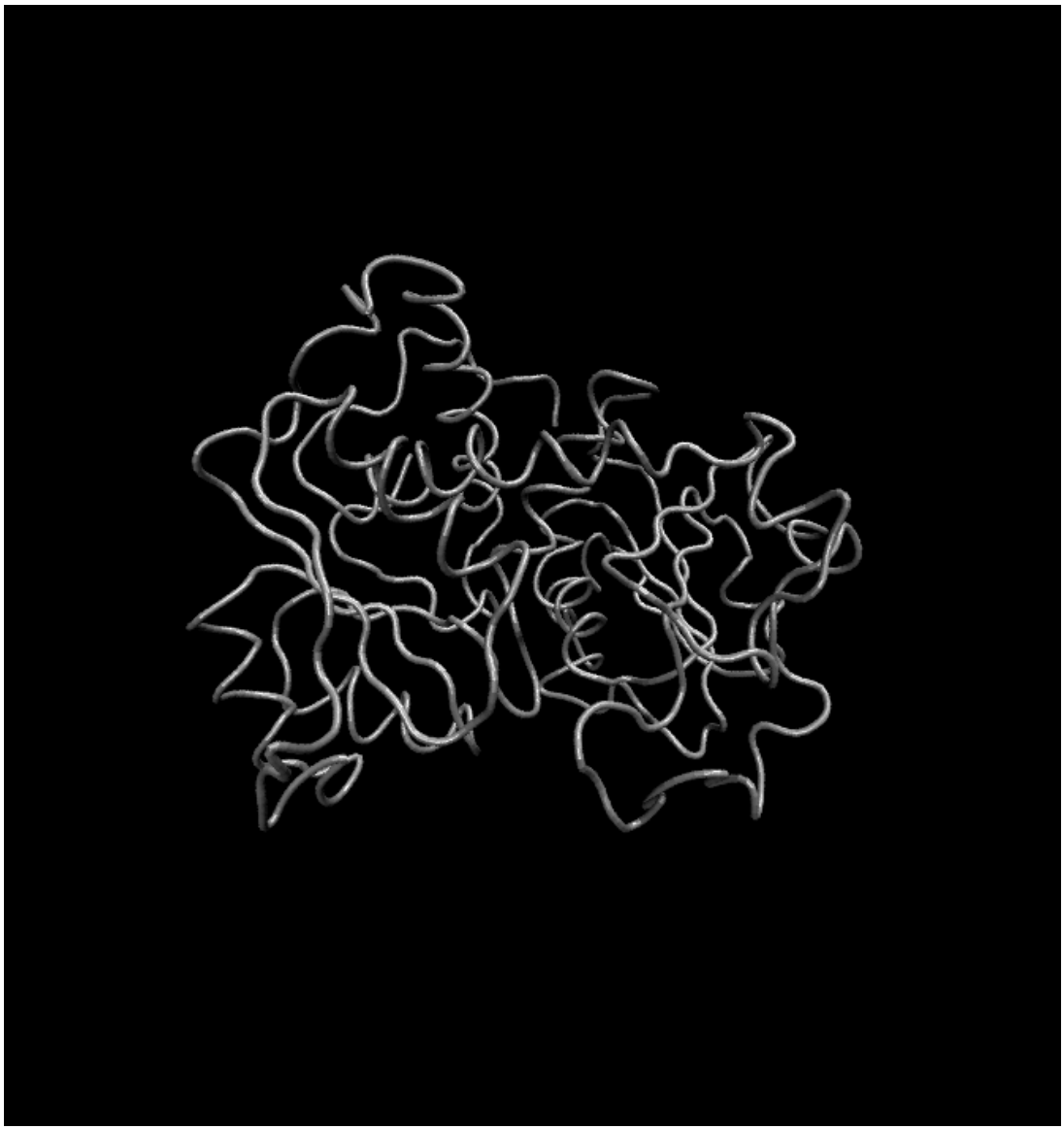
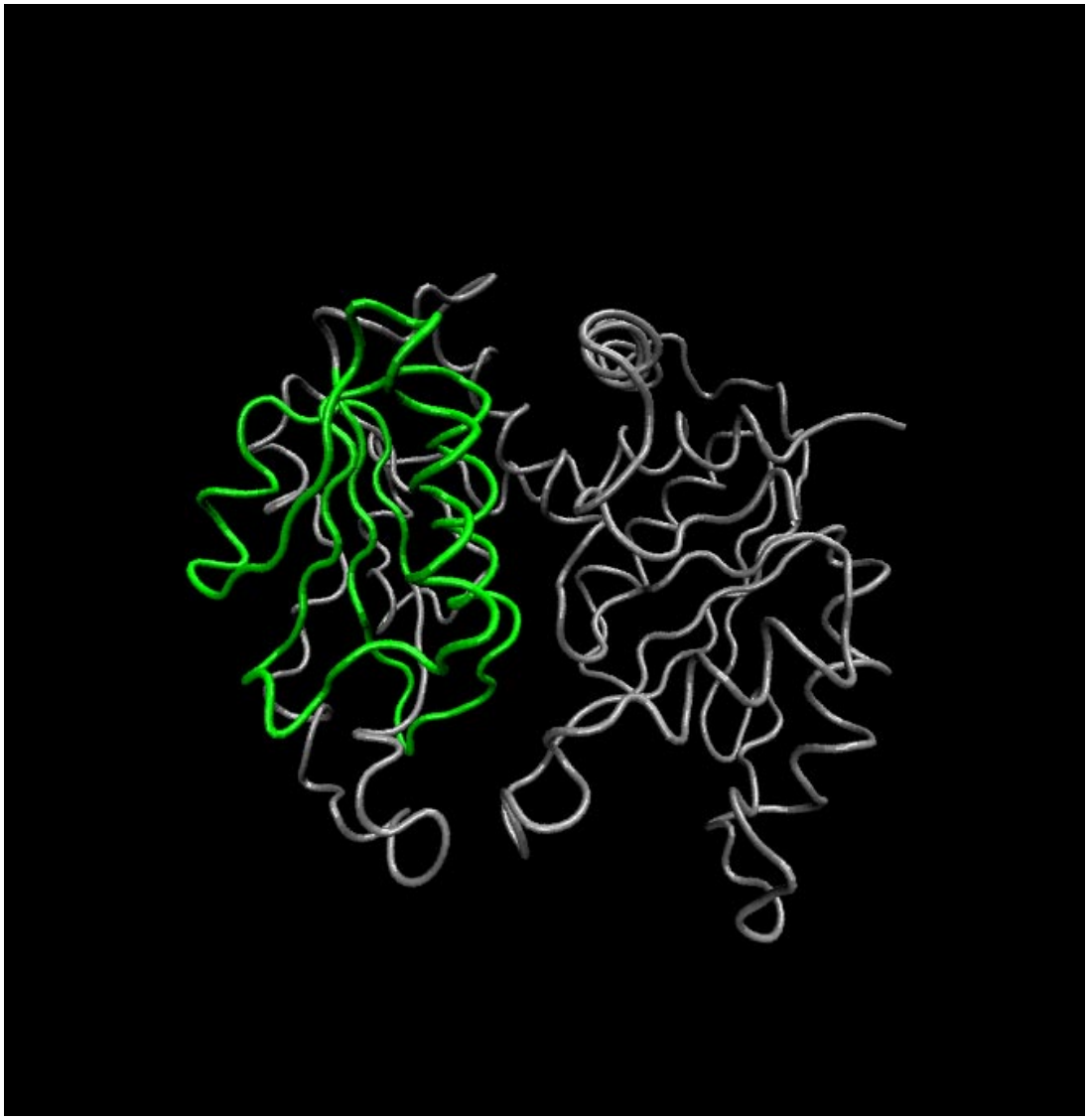


Fig. S3. Thermal untying of the protein accompanied by backtracking. *Lower* and *Upper* show, respectively, the number of contacts Q in domains b and a during unfolding. The black line approximates periodic breaking of contacts G+I and I+K in the first phase of unfolding, when the knot is still localized in domain b. The knot moves from domain b to a at $\approx 2{,}800\tau$, and eventually slides off the chain through the terminus N.



Movie S1. Stretching simulation of unknotted protein, 1c9y. Structure of protein which corresponds to knotted region in protein 1yh1 is marked in green, during stretching simulation.

Movie S1 (MPG)



Movie S2. Stretching simulation of the knotted protein. The native knotted region in 1yh1 is marked in green during all tightening processes.

Movie S2 (MPG)