Jamming proteins with slipknots and their free energy landscape

Joanna I. Sułkowska\textsuperscript{1}, Piotr Sułkowski\textsuperscript{2,3,4} and José N. Onuchic\textsuperscript{1}
\begin{itemize}
\item \textsuperscript{1} Center for Theoretical Biological Physics, University of California San Diego, Gilman Drive 9500, La Jolla 92037,
\item \textsuperscript{2} Physikalisches Institute and Bethe Center for Theoretical Physics, Universität Bonn, Nussallee 12, 53115 Bonn, Germany
\item \textsuperscript{3} California Institute of Technology, Pasadena, CA 92215,
\item \textsuperscript{4} Institute for Nuclear Studies, Hoża 69, 00-681 Warsaw, Poland
\end{itemize}

Theoretical studies of stretching proteins with slipknots reveal a surprising growth of their unfolding times when the stretching force crosses an intermediate threshold. This behavior arises as a consequence of the existence of alternative unfolding routes that are dominant at different force ranges. Responsible for longer unfolding times at higher forces is the existence of an intermediate, metastable configuration where the slipknot is jammed. Simulations are performed with a coarse-grained model with further quantification using a refined description of the geometry of the slipknots. The simulation data is used to determine the free energy landscape (FEL) of the protein, which supports recent analytical predictions.

PACS numbers: 87.15.ap, 87.14.E-, 87.15.La, 82.37.Gk, 87.10.+e

The large increase in determining new protein structures has led to the discovery of several proteins with complicated topology. This new fact has raised the question if their energy landscape and the folding mechanism is similar to typical proteins. One class of such proteins includes knotted proteins which comprise around 1% of all structures deposited in the PDB database\textsuperscript{[1, 2]}. A related class of proteins contains more subtle geometric configurations called slipknots\textsuperscript{[3, 4]}. Recent theoretical studies using structure-based models (where native contacts are dominant) suggest that slipknot-like conformations act like intermediates during the folding of knotted proteins\textsuperscript{[5]}. This entire new mechanism is consistent with energy landscape theory (FEL) and the funnel concept\textsuperscript{[3, 8]}. It was shown that the slipknot formation reduces the topological barrier. Complementing regular folding studies, additional information about the landscape was obtained by mechanical manipulation of the knotted protein with atomic force microscopy\textsuperscript{[3]} both experimentally in\textsuperscript{[10, 11]} and theoretically in\textsuperscript{[12, 13, 14]}. For example,\textsuperscript{[12]} it has been shown that unfolding proceeds via a series of jumps between various metastable conformations, a mechanism opposite to the smooth unfolding in knotted homopolymers.

Most of our analysis is based on stretching simulations under constant force\textsuperscript{[16]}. The crucial signature for this process is the overall unfolding time from the beginning of the stretching until the protein fully unfolds. Normally one expects that the transition between the native and the unfolded basins to be limited by overcoming the free energy barrier, which gets effectively reduced upon an application of a stretching force. The rate by which this barrier is reduced depends on the distance between the unfolded basin and the top of the barrier measured along the stretching coordinate $x$. This idea was first developed in the phenomenological model of Bell\textsuperscript{[18]}, which states that the unfolding time $\tau$ decreases exponentially with applied stretching force $F$ as $\tau(F) = \tau_0 e^{-\frac{F}{k_B T}}$. A

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Dependence of the unfolding times $\tau$ on the stretching force $F$ for 1e2i (solid line, in red). In this Letter we describe this mechanism as a superposition of two unfolding pathways: I for small forces (dashed (lower) line, in blue), and II for intermediate and large forces (dashed-dotted (upper) line, green).}
\end{figure}
refined analysis performed in ref. [19] revealed that this
dependence is more complicated but still monotonically
decreasing.

The unfolding times for 1e2i measured in our simul-
tations are shown as the red curve in Fig. [1]. In contrast
to the above expectations, increasing the force in the range
3.3−5.5e/Å surprisingly results in a larger stability of the
protein. ε is the typical effective energy of tertiary na-
tive contacts that is consistent with the value ε/Å ≈ 71
pN derived in [15]. A solution for this paradox is accom-
plished by realizing that unfolding is dominated by two
distinct, alternative routes that are dominant at different
force regimes. A routing switch occurs when threshold is
crossed between weak and intermediate forces. At higher
forces, mechanical unfolding is dominated by a route that
involves a jammed slipknot. This jamming gives rise to
the unexpected dependence of unfolding time on applied
force. Characterizing this mechanism is the central goal
of this Letter.

Before discussing the stretching of 1e2i, we explain why
a slipknot formed by a uniformly elastic polymer should
smoothly unfold under stretching. To simplify the discus-
sion we approximate the threaded and knotting loops by
circles of size Rk and Rk*. These two loops shrink during
stretching and, when the threaded one eventually van-
ishes, the slipknot gets untied. If both loops have similar
sizes, the slipknot is very unstable and unties immedi-
ately. When the threaded loop is much larger than the
knotting one, Rk >> Rk, untightening can be explained
as follows. The elastic energy associated to local bend-
ing is proportional to the square of the curvature. If the
loop is approximated by a circle of radius R, then its local
curvature is constant and equals R−1. The total elastic
energy is $\int dsR^{-2} \approx R^{-1}$ [21]. From the assumption
Rk >> Rk we conclude that upon stretching it is ener-
getically favorable to decrease Rk rather than Rk. This
happens until both radii become equal and then, just as
above, the slipknot gets very unstable and untightens. In
this discussion we have not yet taken into account that
when a slipknot is stretched some parts of a chain slide
along each other. This effect could be incorporated by in-
cluding the friction generated by the sliding [22]. But in
the slipknot the sliding region associated with the knot-
ting loop is much longer than the region associated to

FIG. 2: A slipknot (left) consists of a threaded loop (k1 − k2,
in red) which is partially threaded through a knotting loop
(k2 − k3, in blue). An example of a protein configuration with
a tightened slipknot is shown in the right panel.

FIG. 3: The behavior of the slipknot during stretching (top)
is determined by the relative behavior of its two loops, en-
coded in the time dependence of k1, k2 and k3 (bottom). If the
threaded loop shrinks faster than the knotting loop, k1 merges
with k2 (bottom left) and the slipknot untightens (pathway I,
top left). If the knotting loop shrinks faster, k2 approaches k3
(bottom right, $\sim 14000\tau$) and the slipknot gets temporarily
tightened (pathway II, top right). This is a metastable state
which can eventually untie further stretching, with k1 finally
merging with k2 (bottom right, $\sim 19000\tau$). Kinetic stud-
ies were performed slightly above folding temperature using
overdamped Langevin dynamics with typical folding times of
10000$\tau$. 
the threaded loop. Thus this effect results in a faster tightening of the threaded rather than the knotting loop, facilitating even more the untightening of the slipknot.

The above argument should apply to slipknots in biomolecules because they are characterized by a persistence length that in principle is simply related to their elasticity [24]. For DNA this effect is described by worm-like-chain models (WLC) [24] and it has been confirmed experimentally. Although WLC models are too simple to describe the protein general behavior, they are useful in some limited applications. Thus at first sight one might expect that slipknots in proteins should smoothly untie upon stretching. Proteins, however, are much more complicated than DNA or uniformly elastic polymers. The presence of stabilizing native tertiary contacts leads to a jumping character during stretching [12]. In addition their bending energy is not uniform along the chain due to the heterogeneity of the amino-acid sequence. As a consequence it turns out that the intuition obtained through the above analysis of polymers or WLC models is misleading.

Our analysis of the evolution of the endpoints $k_1, k_2, k_3$ (Fig. 3 bottom) reveals that for various stretching forces unfolding proceeds along two distinct pathways (Fig. 3 top). In pathway I the slipknot smoothly unties, which is observed for relatively weak forces. At intermediate forces pathway II starts to dominate and the knotting loop can shrink tightly before the threaded one vanishes. In this regime the protein gets temporarily jammed (Fig. 3 right), leading to much longer unfolding times (catch pathway). The probability of choosing pathway I at different forces is shown in Fig. 4. This pathway competition explains the nontrivial total unfolding time dependence observed in Fig. 1.

The two different pathways I and II arise from completely different unfolding mechanisms. Pathway I starts and continues mostly from the C-terminal side, along 16α, 15β, 14α, 13β, 12(helices bundle), 11α (here the number denotes a consecutive secondary structure as counted from N-terminal, and α or β specifies whether this is a helix or a β-sheet; for more details about the structure of 1e2i see the PDB). This is followed by unfolding of helices 11α, 10α that allows breaking of the contacts inside the β-sheet created by the N-terminal, with unfolding proceeding also from the N-terminal. Pathway II also starts from the C-terminal but rapidly (as soon as helix 15 is unfolded) switches to the N-terminal. In this case, differently from pathway I, the β-sheet from the N-terminal unfolds even before 13β. These scenarios indicate that the pathway I should be dominant at weak forces since they are not sufficient to break the β-sheet during first steps of unfolding. The jammed pathway is typical only if stretching forces are sufficiently strong for unfolding to proceed from the two terminals of the protein.

A similar phenomenon was firstly proposed in ref. [25] and referred to as catch-bonds. Experimental evidence suggesting this mechanism was first observed for adhesion complexes [24, 27]. Using AFM, at large forces the ligand-receptor pair becomes entangled and therefore expands the unfolding time. A theoretical description of this mechanism was given in ref. [26].

The kinetic data can also be used to determine the associated free energy landscape (FEL) [7]. In an initial simplification we associate the barriers along the stretching coordinate as the the kinetic bottlenecks during the mechanical unfolding event. Generalizing Bell’s model, a recent description of two-state mechanical unfolding in the presence of a single transition barrier has been developed in [19], with the rate equation

$$\tau(F) = \tau_0 \left(1 - \frac{e^{-\frac{\frac{1}{2} x^2}{\Delta G}}}{\Delta G}\right)^{1-\nu} e^{-\frac{\nu F x^T}{\Delta G}} \left(1-\nu F x^T/\Delta G)^{1/\nu}\right),$$

where $\nu$ encodes the shape of the barrier. Here $x^T$ denotes the distance between the barrier and the unfolded basin (in a first approximation it can be regarded as $F$ independent) and lies on the reaction coordinate along the AFM pulling direction. It can be experimentally determined by measuring how the stretching force modulates the unfolding times $\tau$. The height of the barrier is denoted by $\Delta G$. Fig. 1 (unfolding times are given by solid red line) shows that this single barrier theory is not sufficient for the full range of forces. As described before, in the higher force regime, additional basins have to be included in the FEL. Models with several metastable basins have been called multi-state FEL models [31]. Evidence supporting the need of multi-states FEL was confirmed by AFM experiments in different systems [32, 33].

To construct a multi-state FEL that incorporates two unfolding pathways I and II we use a linear combination of eq. (1)-like expressions with different shapes and barrier heights. Each one of them essentially accounts for the distinct barrier along a relevant unfolding route. Fitting the stretching data to eq. (1) with a cusp-like
\(\nu = 1/2\) approximation (another possibility \(\nu = 2/3\) for the cubic potential in general leads to similar results \cite{19} ) determines accurately the location and the height of the potential barriers. Pathway II involves two barriers: first until the moment of creation of the intermediate which is followed the untieing event. They are characterized by \((x_1, \Delta G_1)\) and \((x_2, \Delta G_2)\) arising respectively from the lower and upper fits in Fig. 5 (left). The superposition of these two fits gives the overall mean unfolding time for pathway II (dotted-dashed curve in green in Fig. 1). For the ordinary slipknot unfolding (pathway I), the results \(x_I\) and \(G_I\) arise from the dashed blue curve in Fig. 1. This analysis leads to the results

\[
\begin{align*}
\Delta G_1 &= 8.0 k_B T, \\
\Delta G_2 &= 4.2 k_B T, \\
\Delta G_1 &= 4.7 k_B T.
\end{align*}
\]

We conclude that the free energy landscape consists of two “valleys”. The force-dependent probability of choosing one of the valleys during stretching depends on the details of the protein structure. It is determined from our simulations as shown in Fig. 1. Using these probability and the parameters above for \(x\) and \(\Delta G\), we can accurately represent the simulation data using a linear combination of equations of the form \(\Pi\). This agreement supports our analytical analysis and generalizes eq. \(\Pi\) for the full of range forces. In addition it demonstrates that structure-based models sufficiently capture the major geometrical properties of a slipknoted protein. A schematic representation of the free energy landscape for pathway II is shown in Fig. 5 (right).

Summarizing, we have analyzed the process of tightening of the slipknot in protein 1e2i and determined the corresponding free energy landscape. Its main feature is the presence of a metastable configuration with a tightened slipknot, which is observed for sufficiently large pulling forces. This phenomenon does not exist for uniformly elastic polymers. In this Letter we concentrated on protein 1e2i but similar behavior has also been observed for other proteins with slipknots, e.g. 1p6x. Our results provide testable predictions that can now be verified by AFM stretching experiments.

We appreciate useful comments of O. Dudko. The work of J.S. was supported by the Center for Theoretical Biological Physics sponsored by the NSF (Grant PHY-0822283) with additional support from NSF-MCB-0543906. P.S. acknowledges the support of Humboldt Fellowship, DOE grant DE-FG03-92ER40701FG-02, Marie-Curie IOF Fellowship, and Foundation for Polish Science.

\begin{thebibliography}{99}
\bibitem{1} M. L. Mansfield, *Nature Struct. Biol.* 1 213 (1994).
\bibitem{2} P. Virnau, L. A. Mirny and M. Kardar, *PLOS Comp. Biol.* 2 1074 (2006).
\bibitem{3} T. O. Yeates, Todd S. Norcross and N. P. King, *Curr. Opinion in Chem. Biol.* 11 596 (2007).
\bibitem{4} W. R. Taylor, *Comp. Biol. and Chem.* 31 151 (2007).
\bibitem{5} J. I. Sulkowska, P. Sulkowski and J. N. Onuchic, *PNAS* 106 3119 (2009).
\bibitem{6} D. Bolinger, J. I. Sulkowska, Hsiao-Ping Hsu, L. A. Mirny, M. Kardar, J. N. Onuchic and P. Virnau, submitted.
\bibitem{7} P.E. Leopold, M. Montal and J.N. Onuchic, *PNAS* 89 8721 (1992).
\bibitem{8} J. N. Onuchic and P. G. Wolynes, *Curr. Opin. Struct. Biol.* 14 70-75 (2004).
\bibitem{9} M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez and H. E. Gaub, *Science* 276 1109 (1997).
\bibitem{10} M. T. Alam, T. Yamada, U. Carlsson and A. Iki, *FEBS Lett.* 519 35 (2002).
\bibitem{11} T. Bornschloegl, D. M. Austrum, E. Mey, J. Dzubiella, M. T. Rief and K. Forest, *Biophys. J.* 96 1508 (2009).
\bibitem{12} J. I. Sulkowska, P. Sulkowski, P. Szymczak and M. Cieplak, *Phys. Rev. Lett.* 100 058106 (2008).
\bibitem{13} J. I. Sulkowska, P. Sulkowski, P. Szymczak and M. Cieplak, *PNAS* 105 19714 (2008).
\bibitem{14} J. Dzubiella, *Biophys. J.* 96 831 (2009).
\bibitem{15} J. I. Sulkowska and M. Cieplak, *J. Phys. Cond. Mat.* 19 283201 (2007).
\bibitem{16} N. D. Socci, J. N. Onuchic and P. G. Wolynes *Proc. Natl. Acad. Sci. USA* 96 2031-2035 (1999).
\bibitem{17} J. Vogt, R. Perozzo, A. Pautsch, A. Prota, P. Schelling, B. Pilger, G. Folkers, L. Scapozza and G. E. Schulz, *Proteins: Struct., Funct., Genet.* 41 545 (2000).
\bibitem{18} G. I. Bell, *Science* 200 618 (1978).
\bibitem{19} O. K. Dudko, G. Hummer and A. Szabo, *Phys. Rev. Lett.* 96 108101 (2006).
\bibitem{20} K. Koniaris and M. Muthukumar, *Phys. Rev. Lett.* 66 2211 (1991).
\bibitem{21} L. D. Landau and E. M. Lifshitz *Theory of Elasticity* 3rd edition
\bibitem{22} B. Audoly, N. Clauvelin and S. Neukirch, *Phys. Rev. Lett.* 99. 164301 (2007), R. Gallotti and O. Pierre-Louis, *Phys. Rev. E* 75 031801.
\bibitem{23} C. Bustamante, J. F. Marko, E. D. Siggia and S. Smith, *Science* 265 1599 (1994).
\bibitem{24} C. Bouchiat, M.D. Wang, J.-F. Allemand, T. Strick, S.M.
[25] M. Dembo, D.C. Torney, K. Saxman and D. Hammer, 
   Proc. R. Soc. Lond. B. Biol. Sci. 234 (1988).
[26] B. T. Marshall, M. Long, J. W. Piper, T. Yago, R. P. 
   McEver and Ch. Zhu, Science 319 630 (2003).
[27] W. Thomas, M. Forero, O. Yakovenko, L. Nilsson, P. 
   Vicini, E. Sokurenko and V. Vogel, Biophys. J. 90 753 
   (2006).
[28] D. Bartolo, I. Derenyi and A. Ajdari, Phys. Rev. E. 65 
   051910 (2002).
[29] E. Evans, A. Leung, V. Heinrich and Ch. Zhu, PNAS 
   101 11281 (2004).
[30] V. Barsegov and D. Thirumalai, PNAS 102, 1835 (2005).
[31] I. Schwaiger, M. Schleicher, A. A Noegel and M. Rief, 
   EMBO 6 46 (2005)
[32] P. M. Williams, S. B. Fowler, R. B. Best, J. L. Toca- 
   Herrera, K. A. Scott, A. Stward and J. Clarke Nature 
   422 446 (2003).
[33] C. Cecconi, E. A. Shank, C. Bustamante and S. Mar- 
   qusee, Science 23 2057 (2005).