A Newly Identified Class of Protein Misfolding in All-atom Folding Simulations Consistent with Limited Proteolysis Mass Spectrometry

Quyen V. Vu[1], Ian Sitarik[2], Yang Jiang[2], Divya Yadav[3], Piyoosh Sharma[3], Stephen D. Fried[3, 4*], Mai Suan Li[1, 5], Edward P. O'Brien[2, 6, 7*]

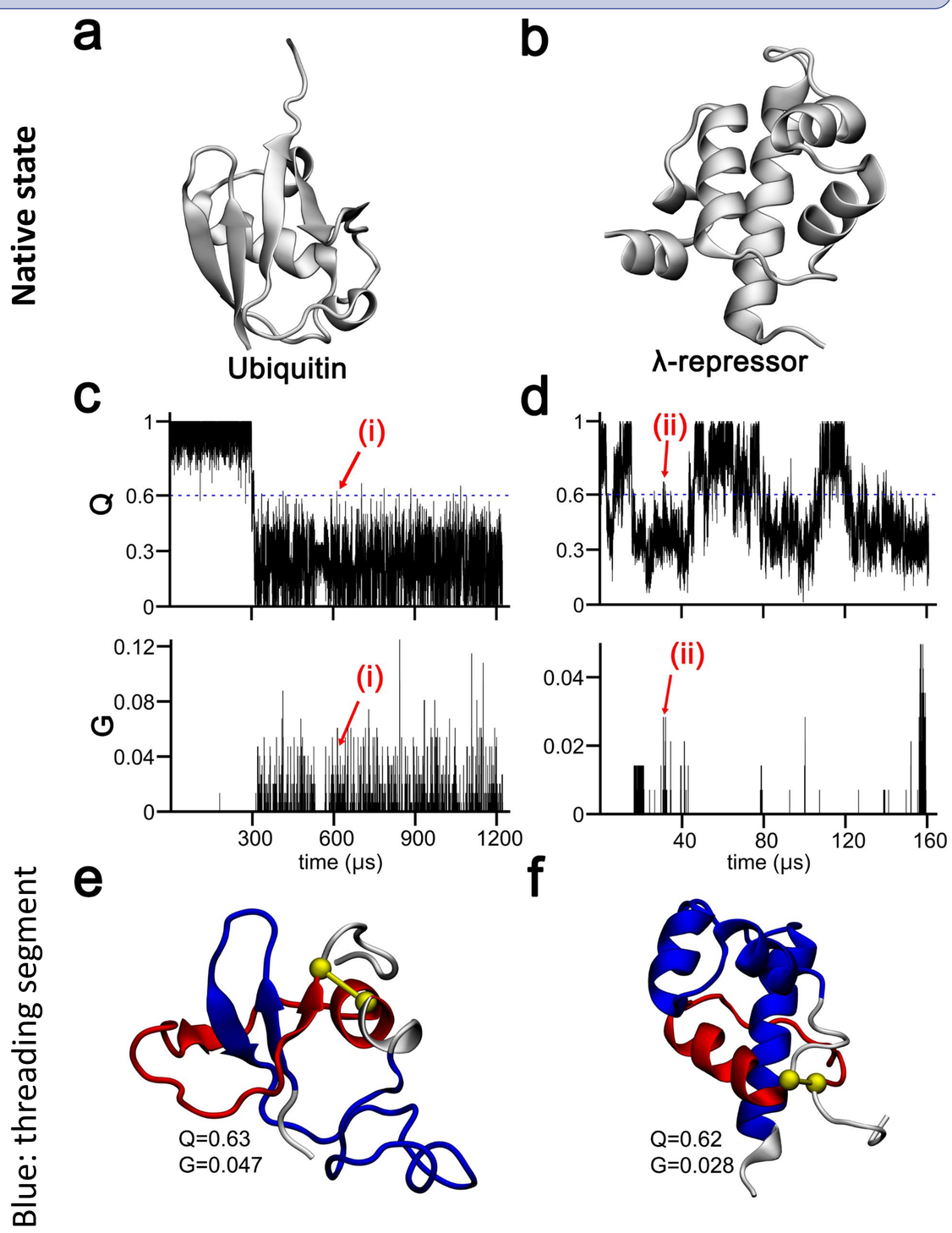
[1]Institute of Physics, Polish Academy of Sciences; Al. Lotnikow 32/46, 02-668 Warsaw, Poland. [2]Department of Chemistry, Penn State University; University Park, Pennsylvania, United States. [3]Department of Chemistry, Johns Hopkins University, Baltimore, MD 21218, United States. [4]T. C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, MD 21218, United States. [5]Institute for Computational Sciences and Technology; Quang Trung Software City, Tan Chanh Hiep Ward, District 12, Ho Chi Minh City, Vietnam. [6]Bioinformatics and Genomics Graduate Program, The Huck Institutes of the Life Sciences, Penn State University; University Park, Pennsylvania, United States. [7]Institute for Computational and Data Sciences, Penn State University; University Park, Pennsylvania, United States. *Corresponding authors: sdfried@jhu.edu, epo2@psu.edu

Email me at:
vuqv.phys@gmail.com

Abstract

Several mechanisms intrinsic to a protein's primary structure are known to cause monomeric protein misfolding. Coarse-grained simulations, in which multiple atoms are represented by a single interaction site, have predicted a novel mechanism of misfolding exists involving off-pathway, non-covalent lasso entanglements, which are distinct from protein knots and slip knots. These misfolded states can be long-lived kinetic traps, and in some cases are structurally similar to the native state according to those simulations. Here, we examine whether such misfolded states occur in long-time-scale, physics-based all-atom simulations of protein folding. We find they do indeed form, estimate they can persist for weeks, and some have characteristics similar to the native state. Digestion patterns from Limited Proteolysis Mass Spectrometry (LiP-MS) are consistent with the presence of changes in entanglement in these proteins. These results indicate monomeric proteins can exhibit subpopulations of misfolded, self-entangled states that can explain long-timescale changes in protein structure and function in vivo.

Entanglement exists in all-atom Folding Simulations



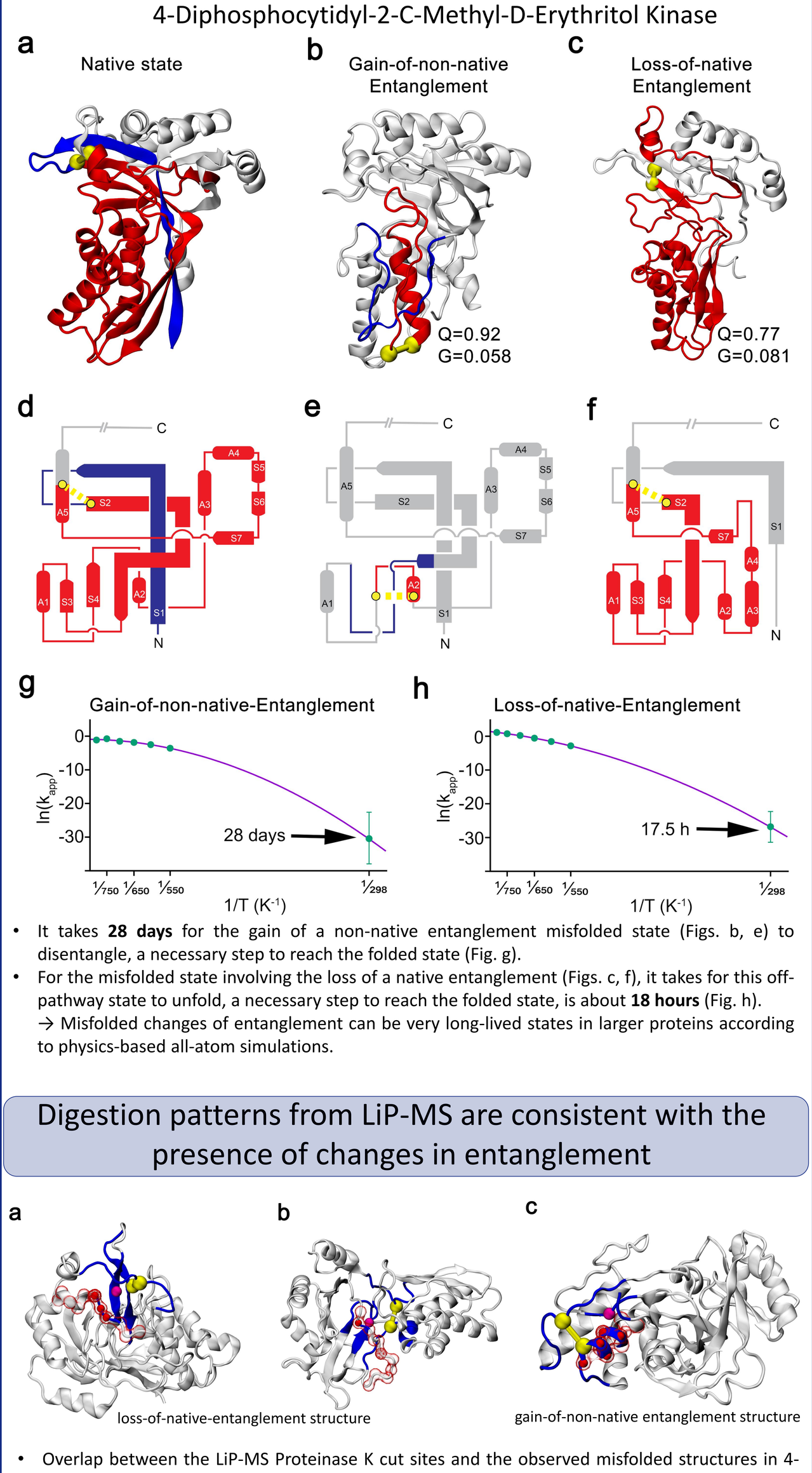
Results:

- We analysed whether non-native entanglements are populated during the folding trajectories of Ubiquitin and λ-repressor totalling 8 milliseconds and 643 microseconds of simulation time, respectively.
- We observed several non-native entangled states (denoted by high G values) during folding transitions → Protein self-entangled states are observed in all-atom models.
- We never observe in these trajectories direct transition from the misfolded state to the native state → These misfolded states are off pathway – they must unfold to reach the native state.
- These entangled states are clearly short-lived (lasting just a few nanoseconds according to Figs. c and d). Ubiquitin (76 residues) and λ-repressor (80 residues) are very small proteins

Entanglement is long-lived in a larger protein: Size effect

- To test this "size-effect" hypothesis we calculated the lifetime of two misfolded states of a larger protein, E. coli's 4-Diphosphocytidyl-2-C-Methyl-D-Erythritol Kinase (Fig. a; 283 residues; gene ispE)
- This protein was chosen from our dataset of proteins that exhibit entanglements in coarse-grained simulations because of:
  1. its large size (283 residues, Fig. a)
  2. it exhibits both classes of misfolding: a conformation with a gain of a non-native entanglement and another with a loss of a native entanglement (Figs. b, c)
- We back-mapped to all-atom model then, carried out in silico temperature jump experiments and Arrhenius analyses (Figs. g-h) to get the disentanglement/unfolding rate of these misfolded states.

4-Diphosphocytidyl-2-C-Methyl-D-Erythritol Kinase



- It takes **28 days** for the gain of a non-native entanglement misfolded state (Figs. b, e) to disentangle, a necessary step to reach the folded state (Fig. g).
- For the misfolded state involving the loss of a native entanglement (Figs. c, f), it takes for this off-pathway state to unfold, a necessary step to reach the folded state, is about **18 hours** (Fig. h).
  → Misfolded changes of entanglement can be very long-lived states in larger proteins according to physics-based all-atom simulations.

Digestion patterns from LiP-MS are consistent with the presence of changes in entanglement

- Overlap between the LiP-MS Proteinase K cut sites and the observed misfolded structures in 4-Diphosphocytidyl-2-C-Methyl-D-Erythritol Kinase.
- The two residues form the native contact to close the loop, the residue that crosses the plane of the loop on the threading segment, and the backbone segment making the loop are shown, respectively, as two yellow spheres, a magenta sphere, and a blue segment. The remaining portions of the protein structure are shown in gray.
- The residues overlapping between these entangled residues and the Proteinase K cut site are shown as red-solid spheres.

Conclusions

- Non-native self-entanglements occur during protein folding in all-atom models.
- Some of these states are long-lived, and have properties similar to the native state.
- This new class of protein misfolding opens up novel avenues of research including new targets for drug design, expanding our understanding of protein homeostasis in cells, and the impact of these states on protein function

Acknowledgments

National Institutes of Health

Checkout our Preprint: