

# Ensemble View of RNAs and Proteins: Loops, Knots, Territories, and Evolution

Alexander Y. Grosberg<sup>1,\*</sup>

<sup>1</sup>Department of Physics and Center for Soft Matter Research, New York University, New York, New York

There is an old joke about a lazy doctor who, instead of looking at each patient individually, requested information about average body temperature of all patients in the hospital. The joke pretends to be funny because an ensemble view of patients is absurd. But in some other cases, including in the life sciences, an ensemble view is not only not absurd, but the right instrument, and Liu and Hyeon (1) in this issue of the *Biophysical Journal* offer a nice example: not only do they look at ensembles, they compare two different ones—ensemble of proteins and ensemble of RNAs. Their findings are remarkably thought-provoking.

Their main technical tool is the study of contact probability—the probability that two monomers of a specific polymer chain, separated by a contour distance of  $s$  monomers, are found in contact in three dimensions (Fig. 1). In polymer physics, a similarly defined quantity  $P(s)$  is widely used to distinguish between different phases of a polymer chain. In particular, for various fractal conformations,  $P(s)$  decays as a power law  $P(s) \sim s^{-\gamma}$ , including Gaussian (random walk) coils with  $\gamma = 3/2$  (in three dimensions), swollen self-avoiding coils ( $\gamma \approx 1.765$ ), and various crumpled

(2), fractal (3), and loopy (4) globules ( $\gamma$  close to unity). In all these cases,  $P(s)$  reveals no characteristic scale; by contrast,  $P(s)$  for an equilibrium globule has a distinct crossover at  $s \sim N^{2/3}$ , separating different fractal behaviors with a power-law drop  $\gamma \approx 3/2$  at  $s < N^{2/3}$ , and saturation to a constant ( $\gamma \approx 0$ ) at  $s > N^{2/3}$ . Furthermore, in the recently suggested steady but nonequilibrium (energy-consuming) loop extrusion model of chromatin (5–7),  $\gamma \approx 3/4 < 1$  at small  $s$ , crossing over to crumpled globule-type  $\gamma$  close to unity at larger  $s$  (and possibly further switching to an even faster decay).

Liu and Hyeon (1) found that, compared to an ensemble of globular proteins, an ensemble of RNAs is much more reminiscent of a crumpled than an equilibrium globule, with  $\gamma$  close to unity. Why is this important?

Slow decay of  $P(s)$  for a certain polymer ensemble is usually indicative of several other properties. Firstly, these ensembles are characterized by dramatically lowered incidence of conformations with knots. Indeed, it was recently shown (8) that all known RNAs are unknotted—unlike proteins, some of which do have knots, including sometimes pretty complex and deeply buried ones (9). Secondly, these conformations are typically territorial, in the sense that different parts of the chain tend to avoid overlap, with each of them carving out a piece of territory for itself. This concept of territoriality became known in the

context of interphase chromatin (see references in Halverson et al. (10)), but there is evidence that RNAs also exhibit territoriality, as in the case of influenza virus (11,12).

Although these three properties (i.e., slow fractal decay of  $P(s)$ , reduced abundance of knots, and territoriality) seem to go hand in hand, it is fair to emphasize that our understanding of such conformations is still incomplete. They were first implicated in the context of rapid collapse of a homopolymer (2), but more recent studies suggest that the process of rapid collapse may be more intricate than initially thought (13). Related conformations are observed in dense systems of unconcatenated rings, but the exact nature of these conformations is still under debate (4,10); trees, especially annealed ones (14), share many properties with unconcatenated rings, but whether they belong to the same or to a different universality class is not known. In the light of the work by Liu and Hyeon (1), these theoretical questions begin to impinge on the RNA world.

At the end of this comment, we should return to the beginning, to the question of ensemble versus individual approach. When we talk about, say,  $P(s)$  behavior for various states of a polymer, we deal with an ensemble in the standard statistical mechanics sense—with the ensemble of microstates of a fluctuating object, a polymer chain in this case. Neither the

Submitted April 5, 2016, and accepted for publication April 18, 2016.

\*Correspondence: ayg1@nyu.edu

Editor: Rohit Pappu.

<http://dx.doi.org/10.1016/j.bpj.2016.04.021>

© 2016 Biophysical Society.

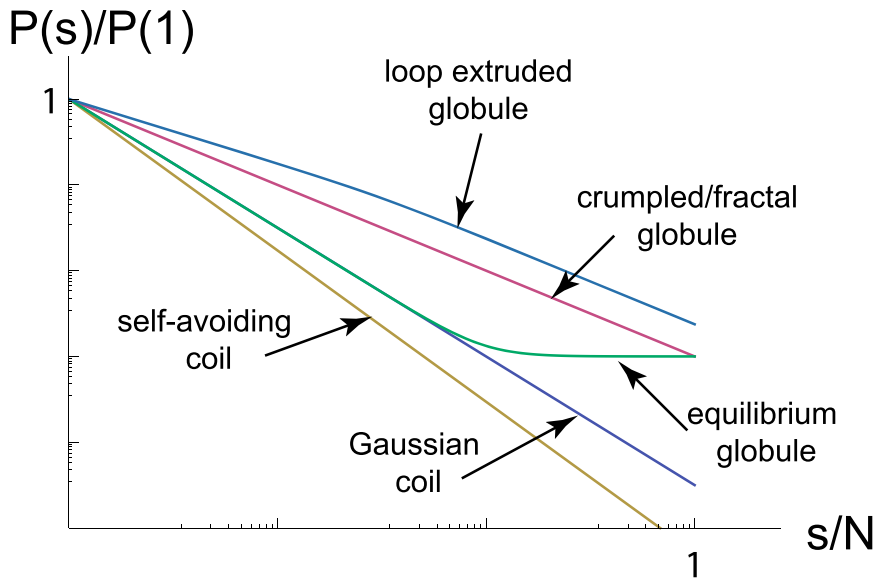


FIGURE 1 Ideal-world contact probability for various states of a polymer chain. Logarithmic scale. To see this figure in color, go online.

ensemble of protein conformations from the Protein Data Bank (PDB) nor similar RNA ensembles are like that. First, for both proteins and RNAs, there is an existing ensemble selected by evolution. Second, there is a subensemble of those molecules that are in the PDB, but the selection of proteins and RNAs in the PDB may be statistically biased. It is possible, for instance, that territoriality of a certain macromolecule may be correlated (positively or negatively) with its ability to crystallize. We are presently not aware (at least this writer does not know) of any specific statistical bias of the PDB with respect to  $P(s)$  or any related conformational property, but we must remain alert that such bias may exist, even if hidden from the eye.

As an example, consider the question of knot abundance. As mentioned above, we do know of a certain number of knotted proteins (9), but no knotted RNAs (8). As a matter of principle, there are two possibilities as far as protein knots are concerned. First, it is possible that normal proteins should

have no knots, and only when a knot is needed for some purpose (e.g., protection against destruction by a protease), then evolution manages to make a knot; in this scenario, the logical question is—what is the functional purpose of every particular knot found in a protein? But there is also a second possibility, that evolution is largely indifferent toward the presence or absence of knots, and then one should expect the knot abundance in protein ensemble to be about the same as in a suitably chosen random one; in this scenario, the main logical question is shifted—what is the right statistical ensemble? Now we should welcome RNA to the club of similar controversies: do RNAs lack knots because knots are bad for something and were eliminated by evolution, or is it that the ensemble of RNA conformations, chosen by evolution for some other reason, is simply not friendly to knots?

In this sense, Liu and Hyeon's findings (1) illuminate from an unexpected direction the entire landscape of protein and RNA evolution.

## REFERENCES

1. Liu, L., and C. Hyeon. 2016. Contact statistics highlight distinct organizing principles of proteins and RNA. *Biophys. J.* 110:2320–2327.
2. Grosberg, Y., A. S. K. Nechaev, and E. I. Shakhnovich. 1988. The role of topological constraints in the kinetics of collapse of macromolecules. *J. Phys. France.* 49:2095–2100.
3. Lieberman-Aiden, E., N. L. van Berkum, ..., J. Dekker. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science.* 326:289–293.
4. Ge, T., S. Panyukov, and M. Rubinstein. 2016. Self-similar conformations and dynamics in entangled melts and solutions of nonconcatenated ring polymers. *Macromolecules.* 49:708–722.
5. Sanborn, A. L., S. S. P. Rao, ..., E. Lieberman-Aiden. 2015. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. USA.* 112:E6456–E6465.
6. Goloborodko, A., J. F. Marko, and L. Mirny. 2016. Two regimes of chromosome compaction via active loop extrusion. *Biophys. J.* 110:2162–2168.
7. Grosberg, A. Y. 2016. Extruding loops to make loopy globule? *Biophys. J.* 110: 2133–2135.
8. Micheletti, C., M. Di Stefano, and H. Orland. 2015. Absence of knots in known RNA structures. *Proc. Natl. Acad. Sci. USA.* 112:2052–2057.
9. Jamroz, M., W. Niemyska, ..., J. I. Sułkowska. 2015. KnotProt: a database of proteins with knots and slipknots. *Nucleic Acids Res.* 43:D306–D314.
10. Halverson, J. D., J. Smrek, ..., A. Y. Grosberg. 2014. From a melt of rings to chromosome territories: the role of topological constraints in genome folding. *Rep. Prog. Phys.* 77:022601.
11. Chou, Y. Y., R. Vafabakhsh, ..., P. Palese. 2012. One influenza virus particle packages eight unique viral RNAs as shown by FISH analysis. *Proc. Natl. Acad. Sci. USA.* 109:9101–9106.
12. Noda, T., and Y. Kawaoka. 2012. Packaging of influenza virus genome: robustness of selection. *Proc. Natl. Acad. Sci. USA.* 109:8797–8798.
13. Bunin, G., and M. Kardar. 2015. Coalescence model for crumpled globules formed in polymer collapse. *Phys. Rev. Lett.* 115:088303.
14. Rosa, A., and R. Everaers. 2014. Ring polymers in the melt state: the physics of crumpling. *Phys. Rev. Lett.* 112:118302.