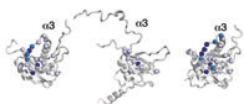


# RNA Biology



 Taylor & Francis

ISSN: 1547-6286 (Print) 1555-8584 (Online) Journal homepage: <http://www.tandfonline.com/loi/krnb20>

## The elusive quest for RNA knots

Aaron S. Burton, Marco Di Stefano, Niles Lehman, Henri Orland & Cristian Micheletti

To cite this article: Aaron S. Burton, Marco Di Stefano, Niles Lehman, Henri Orland & Cristian Micheletti (2016) The elusive quest for RNA knots, *RNA Biology*, 13:2, 134-139, DOI: [10.1080/15476286.2015.1132069](https://doi.org/10.1080/15476286.2015.1132069)

To link to this article: <http://dx.doi.org/10.1080/15476286.2015.1132069>



© 2016 The Author(s). Published with  
license by Taylor & Francis Group, LLC  
This article not subject to US copyright law



[View supplementary material](#)



Accepted author version posted online: 01  
Feb 2016.  
Published online: 01 Feb 2016.



[Submit your article to this journal](#)



Article views: 489



[View related articles](#)



[View Crossmark data](#)

CrossMark



Citing articles: 2 [View citing articles](#)

Full Terms & Conditions of access and use can be found at  
<http://www.tandfonline.com/action/journalInformation?journalCode=krnb20>

POINT OF VIEW

 OPEN ACCESS

## The elusive quest for RNA knots

Aaron S. Burton<sup>a,\*</sup>, Marco Di Stefano<sup>b,\*,\*\*</sup>, Niles Lehman<sup>c</sup>, Henri Orland<sup>d,e</sup>, and Cristian Micheletti<sup>b</sup>

<sup>a</sup>NASA Johnson Space Center, Houston, TX, 77058 USA; <sup>b</sup>SISSA, Via Bonomea 265, I-34014 Trieste, Italy; <sup>c</sup>Department of Chemistry, Portland State University, Portland OR, 97207 USA; <sup>d</sup>Institut de Physique Théorique, Commissariat à l'énergie atomique CEA, IPHt CNRS, UMR3681, F-91191 Gif-sur-Yvette France; <sup>e</sup>Beijing Computational Science Research Center, Haidian District Beijing, 100084, China

### ABSTRACT

Physical entanglement, and particularly knots arise spontaneously in equilibrated polymers that are sufficiently long and densely packed. Biopolymers are no exceptions: knots have long been known to occur in proteins as well as in encapsidated viral DNA. The rapidly growing number of RNA structures has recently made it possible to investigate the incidence of physical knots in this type of biomolecule, too. Strikingly, no knots have been found to date in the known RNA structures. In this Point of View Article we discuss the absence of knots in currently available RNAs and consider the reasons why knots in RNA have not yet been found, despite the expectation that they should exist in Nature. We conclude by singling out a number of RNA sequences that, based on the properties of their predicted secondary structures, are good candidates for knotted RNAs.

### ARTICLE HISTORY

Received 2 November 2015  
Revised 13 November 2015  
Accepted 17 November 2015

### KEYWORDS

Physical knots; RNA knots;  
RNA structure

The systematic characterization and classification of knots has engaged scientists since the introduction of Tait's knots table, whose first entries are shown in Fig. 1A. It is readily seen that the different types of entanglement are permanently trapped in the closed curves. In fact, any two curves cannot be continuously deformed into one another unless they are cut open, rearranged and resealed.

In our everyday life, the knots that typically surround us differ from these well-defined mathematical ones because they are tied in open chains (Fig. 1B). These knots are clearly not permanently trapped and hence are termed “physical knots”. Many of the physical knots that we encounter have been purposely introduced in, for example, the yarn of sewn clothes or in ropes secured to anchoring points. Not infrequently, however, we have to deal with their unwanted spontaneous occurrence, as in the earphone cords pulled out of our pocket.

These spontaneous forms of entanglement arise by statistical necessity in microscopic filaments too. As a matter of fact, equilibrated polymers cannot escape the formation of knots, if they are sufficiently long,<sup>1</sup> especially if they are densely packed.<sup>2,3</sup> It is therefore a remarkable fact that the statistical incidence of topological entanglement in biomolecules such as proteins and DNA filaments, which are both long and compact, is substantially smaller than expected for general models of equilibrated polymers with equivalent length and packing conditions.<sup>4–6</sup>

The observed limited incidence of knots probably results from evolutionary pressure to limit the likely detrimental implications on the biological viability and functionality of these

molecules. For instance, if a protein was as prone as a generic flexible polymer to the stochastic formation of knots, then its folding route would be riddled with kinetic traps and dead ends. It is accordingly plausible, as remarked by Levinthal in more general contexts,<sup>7,8</sup> that protein sequences have evolved to encode not only the thermodynamically-stable native fold, but also the sequence of steps leading to the native structure formation so to minimize the incidence of misfolded states, including knotted ones which would be very challenging to backtrack.

Arguably, the most vivid illustration of the seemingly general aversion to entanglement of biomolecular systems is provided by DNA molecules in higher eukaryotes, where they are hierarchically organized in chromatin fibers and form very long chromosomes. It has long been a puzzle how the chromosomes could dynamically reconfigure through the decondensed interphase configurations and the rod-like mitotic ones without being hindered by the *cis*- and *trans*-chromosome entanglement expected to arise from their dense packing in the nucleus.<sup>9</sup> The complementary advancements in imaging techniques<sup>10,11</sup> and computer modeling<sup>5,12</sup> have nowadays clarified that the incidence of intra- and inter-chromosome entanglement is kept at a minimum by the fact that the cell cycle lasts only a tiny fraction of the time needed to equilibrate decondensed chromosomes. Consequently, interphase chromosomes are quenched in a state that retains memory of the disentangled mitotic state and are thus efficiently primed for the subsequent recondensation step. These considerations support

**CONTACT** Cristian Micheletti  [michelet@siswa.it](mailto:michelet@siswa.it)

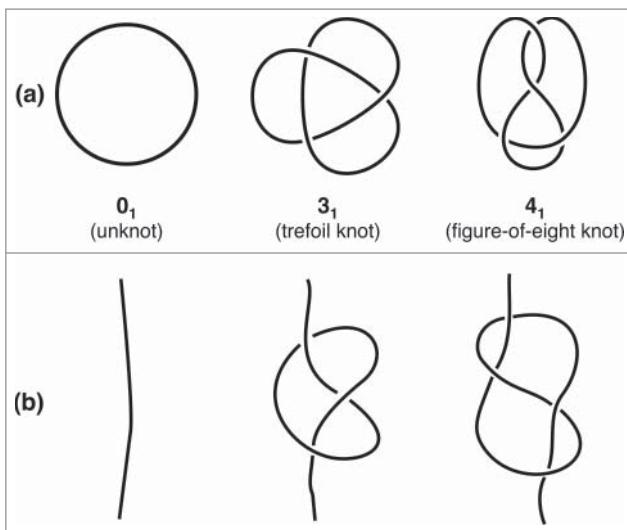
\*These authors equally contributed to this work.

\*\*present address: Structural Genomics group, Center Nacional d'Anàlisi Genomica - Center de Regulació Genomica (CNAG-CRG), Baldíri i Reixac 4, 08028 Barcelona, Spain

 Supplemental data for this article can be accessed on the publisher's website.

Published with license by Taylor & Francis Group, LLC. This article not subject to US copyright law.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.



**Figure 1. The simplest types of knots.** The simplest mathematical knots are shown in panel (a) and are labeled by the number of crossings in the simplest two-dimensional projection followed by a conventional indexing subscript. The corresponding physical knots, which are obtained by opening the closed chains, are shown in panel (b).

the long-held notion that biomolecules should typically be free of significant forms of self- and mutual entanglement.<sup>13</sup>

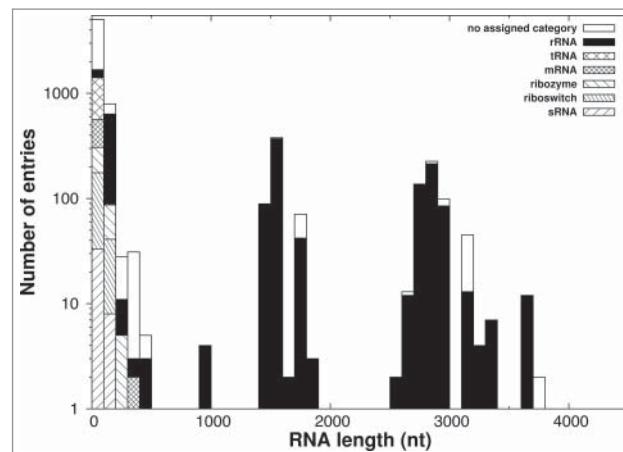
Yet, although expectedly disfavoured, knots do occur both in proteins and in DNA.

For instance, it is known that about 1 – 2% of the proteins in the protein data bank (PDB) contain physical knots.<sup>14–16</sup> Furthermore, electrophoretic experiments have shown that viral dsDNA, once circularised either in bulk or inside viral capsids, is substantially knotted.<sup>17–19</sup> In particular, the incidence of knots in P4 DNA, which is 10kb-long, is at least 95%.<sup>18</sup> The incidence of knots in longer DNA packed inside viral capsids is expected to be even higher, though at the moment this has not been ascertained yet due to limitations of current topological profiling techniques.

Motivated by these considerations, last year some of us undertook a survey of the occurrence of knots in the RNA structures deposited in the PDB as of June 2014.<sup>20</sup> This was motivated by the fact that, since the early study of ref. 21, no systematic profiling of RNAs topology had apparently been carried out, despite the rapid increase of the database of publicly available structures. At that time, the PDB contained several thousand RNA entries, spanning a total of 5,466 chains. Assuming a knot incidence comparable to that found in proteins, ~1%, tens of knotted entries could have been expected. Strikingly, no genuinely-knotted RNA chain was identified in the survey.

For this Point of View article, we considered the reasons why knots in RNA have not yet been found, despite the expectation that they should exist in Nature.

To initiate this query, we first updated our 2014 survey by profiling the RNA structures deposited in the PDB as of August 2015 (Fig. 2 and Supplemental Material). Again, despite the appreciable increase in the number of distinct chains over one year, from 5,466 to 7,013, no genuinely-knotted RNA structure was found. To be precise, the two surveys did return three knotted entries, which are listed in Table 1, but, as already noted in



**Figure 2. RNA chain length and types.** The distribution of lengths (number of phosphates) of all 7,013 RNA molecules with more than five phosphates deposited in the PDB as of August 2015 is shown. The color code describes the contribution of each categorized subdivision to the total amount of chain.

ref. 20 they are most likely artifactual results of cryo-EM reconstructions.

Though an equal footing comparison with proteins is difficult for the different typical length (number of monomers) of the two types of molecules and the PDB entry redundancy, the total absence of knots in the > 7,000 available RNAs structures is notable.

Assuming that this dataset provides an unbiased representation for naturally-occurring RNAs, it is appealing to speculate on the biological rationale behind the apparent aversion of RNA to knots. A first general consideration regards RNA folding kinetics, particularly its possible co-transcriptional component.<sup>22,23</sup> It appears plausible that newly transcribed nucleotides in a nascent RNA chain could anneal with other nearby nucleotides establishing ephemeral local secondary structures. While these temporary motifs may be traded off for longer range ones at later folding stages, their presence would make it more difficult to develop knots by increasing the effective thickness and local rigidity of the chain. The possible out-of-equilibrium accretion of the RNA chain growing sequentially in layers around the already-transcribed, compact, portion would clearly also push toward entanglement-free structures, too. It is worth noting that the sequential modular growth of RNA chains could be also relevant from an evolutionary point of view. In fact, it has been recently suggested that RNA ribosomal complexes have reached their current large and articulated conformations by the sequential addition of small modular structures.<sup>24</sup> Similarly to the kinetic ones, such evolutionary mechanisms should limit the incidence of self-entanglement too.

It also appears plausible that RNA knots, and other forms of entanglement, are not favored thermodynamically. In fact, independent of kinetic considerations, naturally occurring RNAs appear to be primed for acquiring native structures with rather low geometrical complexity. This observation is prompted by the analysis of secondary-structure predictions of algorithms that are exclusively based on the minimization of the (model) free energy of RNA sequences. These phenomenological approaches clearly do not account for any kinetic effects

**Table 1.** Knotted RNAs in the survey data set. The knots are likely artifactual results of cryo-em reconstruction. The entries are the same as those shown in Table 1 of ref. 20, though some of the PDB codes differ because the PDB introduced a new archiving system of large structures (with no change of atomic coordinates). The overall knot type was established by using the minimally-invasive closing procedure of ref. 46. The # sign in the knot label of the last entry denotes the composition (concatenation) of various prime knots.

Putative knot	Molecule (Organism)	Length (nt)	PDB ID	Suspicious clasp (essential crossings)
4 <sub>1</sub> 16 <sub>124834</sub> 3 <sub>1</sub> #3 <sub>1</sub> #4 <sub>1</sub> #3 <sub>1</sub>	23S rRNA (E. coli) 26S rRNA (T. lanuginosus) 23S rRNA (E. coli)	2,740 3,169 2,904	4V4V:B0 4V7H:B5 1C2W:B	G1478-C1480 and C1558-C1561 C618-A622 and A1401-U1405 A710-G714 and U2775-A2780 G263-C264 and G363-C364 C1454-G1455 and C1526-G1527 U1476-A1477 and G1514-A1515 U1856-G1857 and C1887-G1888

and yet their returned structures are geometrically much simpler than randomly reshuffled variants of the sequences. For instance, even when using structure prediction algorithms based on planar graph representations one observes that viral RNA sequences have a significantly smaller graph diameter, a proxy for three-dimensional size, than their reshuffled versions.<sup>25,26</sup> This point is reinforced by the analysis of the minimum energy structures returned by McGenus.<sup>27,28</sup> Specifically, it is found that the predicted non-planar graphs of the reshuffled variants have a topological complexity, as measured by the graph genus, that is more than an order of magnitude larger than for the wild-type counterparts.<sup>20</sup>

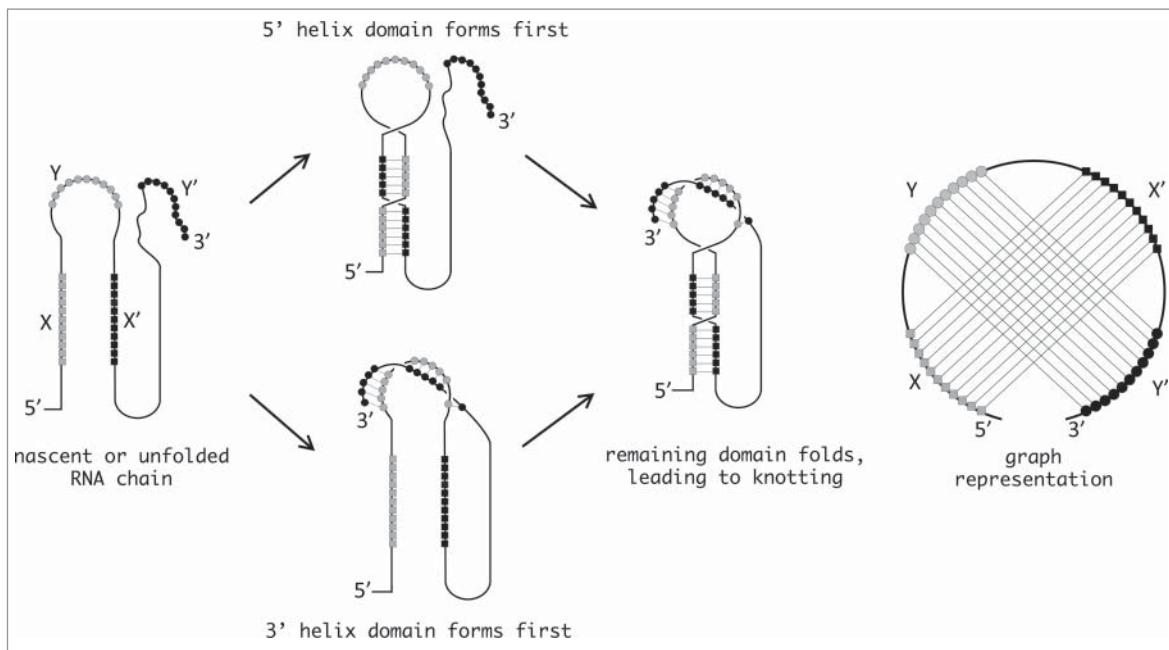
Although RNA knots may be elusive today, it is conceivable that they were important in earlier evolutionary stages, even during the origins of life. The RNA world concept posits a period of time when RNA, or something chemically similar, was responsible for all living processes including heredity and metabolism. RNA-RNA interactions have often been proposed to be critical to maintaining the physical integrity of networks of distinct sequences.<sup>29-31</sup> Typically these interactions have focused solely on secondary structure interactions, but many environments on an early Earth could have experienced rather extreme temperature fluctuations that could disrupt weak base-pairing. More severe entanglements such as knots – but without covalent bonding – would have been a means to preserve spatial sympatry of RNA network individuals in harsh environments. The observation that many catalytic RNA species have structures that are pseudoknotted seems to hint at closer inter- and intra-molecule associations as one goes back in time.

It is, again, important to recall that these conclusions and considerations are based on the set of currently known RNA structures which, though not negligible, is inevitably still limited. There is also almost certainly a selection effect among the solved RNA structures for molecules that adopt a single stable conformation, meaning that if RNA adopts a knotted conformation as one of multiple conformations or as a misfolding event, then it is less likely to be detected. In this regard, a possible parallel can be drawn with proteins, where the discovery of knotted structures was inextricably tied with the growth of the PDB size. In fact, while the occurrence of a knot in transcarbamylase was reported as early as 1977,<sup>32</sup> its genuine character was disputed for its proximity to one of the termini. It was only after the protein database had grown enough that later surveys of Mansfield and Taylor established beyond doubt that genuine, deeply-tied knots were present in naturally occurring proteins.<sup>33,34</sup> It is therefore appealing to speculate that a similar pathway may follow for RNA too. With the number of established structures rapidly increasing, it is possible that genuine knotted RNAs can be found.

The plausibility of knotting in biological RNAs is illustrated strongly by two considerations based on the seminal work of N. Seeman and coworkers. In a remarkable set of experiments,<sup>35</sup> they designed synthetic ~100-nt long RNA sequences that could spontaneously fold in circularised knotted structures and be unknotted by the Top3 bacterial DNA. More recently, the same assay has been used to demonstrate that the human Top3 $\beta$  enzyme also has RNA topoisomerase activity.<sup>36</sup> Two important facts emerge from these experiments. First, they vividly demonstrate that even short RNA molecules can be knotted. Secondly, that the molecular machinery of the cell is already predisposed to deal with RNA knots and simplify them. From the principle of biological parsimony, such simplifying action would hardly be expected if complex RNA topologies were not present *in vivo*, particularly when one considers the exquisite specificity of enzymes that act on nucleic acids for RNA or DNA, but almost never both.

In view of these considerations, the occurrence of knots in naturally occurring RNA appears at least possible, and could be very likely. They would arise from a balance between evolutionary pressures to avoid detrimental entanglement and the functional advantage that knotted RNAs could have in specific, and arguably rare cases. For instance, physical entanglement might allow RNAs to modulate gene expression or to enhance their resilience to *in vivo* degradation. The latter aspect might be relevant for the recently discovered circRNAs<sup>37-39</sup> which thanks to their circularised form, could trap long-lived knots, as suggested by Frank-Kamenetskii.<sup>40</sup> The knotted forms of any RNA would be expected to be only a fraction of total folding space, such that any given genotype could have multiple phenotypes. These could arise from the delicate interplay of folding kinetics and thermodynamics or from the interaction with partner molecules. Perhaps only 5–10% of a "knottable" RNA is actually knotted *in vivo*. This could help explain why they have eluded detection so far.

To guide the search for knotted RNA, we carried out specifically for this Point of View article a bioinformatics survey aimed at identifying candidate RNA sequences that could be susceptible to knotting. As the Seeman laboratory elegantly demonstrated,<sup>35</sup> nucleic acid sequences designed to contain two self-complementary 11-nucleotide sequences (X, Y, X', Y' in which X and X' pair and Y and Y' pair) can spontaneously form knots. We have diagrammed this scenario in an example RNA, and show pathways through which knotted structures could form, at least in some sub-population of folded structures (Fig. 3). We note that in addition to forming physical knots, these molecules are also pseudoknotted, in the sense that the strands involved in the two helices are interleaved. Thus, pseudoknotted RNA sequences with relatively long base-pairing



**Figure 3. Possible pathways leading to RNA knotting.** The sketch illustrates possible pathways through which two sets of interleaved self-complementary sequences (marked with squares and circles, respectively) can fold and produce a knotted RNA structure. The latter is tied in a figure-of-eight knot and the diagram representing the two helices features the characteristic pseudoknot crossings.

stretches appear to be the most likely candidates to form knots. The minimal number of basepairs needed to form a knot can be estimated from a theoretical standpoint, based on the requirement of a minimum of three strand-crossings for the simplest knot. Because  $\sim 10$  basepairs are necessary for two strand crossings (a full helical turn), pseudoknotted structures containing multiple helices of at least ten basepairs would be prime candidates for knotting. A review of the RNA pseudoknot database reveals two candidates that approach or meet these criteria (Table 2). Specifically, the plasmid Colib-P9 mRNA (with helices of 18 and 10 basepairs), and *Homo sapiens* CCR5 PRF (with helices of 25 and 13 basepairs) and telomerase RNA (helices of 22 and 9 basepairs) appear very likely to have at least a slip-knotted structure (in which a free-end has not been completely threaded through a loop). Incidentally we note that such structures have also been observed in proteins.<sup>41–43</sup> These candidate sequences are found within larger RNA mole-

cules, which might seem to preclude the threading events necessary for knotting. In proteins, however, it has been demonstrated *in vitro* that addition of large structured domains to the N- and C-termini of a knotted protein does not prevent it from adopting its native, knotted conformation so it should not be assumed *a priori* that longer RNAs cannot form knots.<sup>44</sup> In addition, we have also identified a number of pseudoknotted RNAs with one helix of 10 or more basepairs and a second helix of seven or more basepairs. Although the lowest energy conformations of these molecules would likely be only pseudoknotted, the potential would seemingly exist for knotting to occur to at least a small extent as an alternative or misfolded conformation (Table 2). As a matter of fact, their secondary structures, as predicted by McGenus,<sup>27,28</sup> are not *a priori* incompatible with proper knots.

The table of RNA sequences we have identified, while not exhausting the repertoire of possible knotted candidates, should nevertheless provide a starting point for a targeted search for knotted RNAs. Previous biochemical analyses raised the possibility that two naturally-occurring RNA intron sequences were capable of adopting knotted conformations, but the presence of a physical knot could not be confirmed irrefutably with the available data.<sup>45</sup> A significant challenge for the unambiguous confirmation of true knots in RNA is the difficulty associated with solving RNA structures; this challenge could be overcome by the development of an appropriate high-throughput assay to detect knotted RNAs, especially when present as isoforms coexisting with other stable unknotted conformations. The lack of such systematic assays is, at present, the main obstacle toward advancing our understanding of the incidence of physical knots in RNAs.

In conclusion, the lack of physical knots in presently available RNA structures is particularly intriguing and also unexpected by comparison with other biopolymers, such as proteins and

**Table 2. Candidate knot-forming RNA molecules from Pseudoknot database (<http://ekevanbatenburg.nl/PKBASE/PKBGETSUM.HTML#TOP>; updated November 28, 2014, accessed October 11, 2015).**

Molecule	Helix 1 length (nt)	Helix 2 length (nt)
CCR5_PRF ( <i>H. sapiens</i> )	13	25
Colib-P9_repZ (Plasmid Colib-P9)	18	10
hTER ( <i>H. sapiens</i> )	22	9
pMU407_repBA (plasmid pMU407.1 [InCL/M])	19	8
RSV (Rous sarcoma virus)	14	8
Ec_alpha ( <i>E. coli</i> )	12	8
SESV (Seal louse virus)	11	8
Sc_18S-PKE21-7/8 ( <i>S. cerevisiae</i> )	10	8
Mm_Edr ( <i>M. musculus</i> )	10	9
Bovine coronavirus	8	10
MHV (mouse hepatitis virus)	8	10
Qbeta_PKdX (bacteriophage Q- $\beta$ )	8	9

encapsidated viral DNA where knots are known to occur. It is possible that the sequence of naturally-occurring RNAs, some of which need to be efficiently translocated through biological pores, have evolved to harness folding kinetics and thermodynamics so as to minimize the incidence of various forms of self-entanglement, including knots, in their native structures. At the same time, the fact that knotted RNAs with as few as  $\sim 100$  nucleotides have been successfully designed and synthesized,<sup>35</sup> suggests that there is no fundamental reason why relatively short knotted RNAs should not exist in nature, possibly in circularised form where knots could be essential to enhance their mechanical stability. According to this standpoint, the current situation would parallel the historic route taken for proteins, where the occurrence of knots was once deemed implausible, if not impossible, due to the inevitable initial limitations in size and representation bias of the set of available protein structures. To aid the ongoing efforts to detect knotted RNAs we have presented a list of sequences whose predicted native structures are expectedly susceptible to the formation of knots, possibly as isoform competing with other unknotted ones.<sup>40</sup> We hope that the list of selected candidates will serve as a useful starting point toward the systematic search of RNA knots.

## Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

## Acknowledgments

We thank Sandro Bottaro and Giovanni Bussi for useful discussions.

## Funding

We acknowledge financial support from the Italian Ministry of Education, grant PRIN 2010HXAW77

## References

- Sumners DW, Whittington SG. Knots in self-avoiding walks. *J Phys A Math Gen* 1988; 21:1689-94; <http://dx.doi.org/10.1088/0305-4470/21/7/030>
- Micheletti C, Marenduzzo D, Orlandini E. Polymers with spatial or topological constraints: theoretical and computational results. *Physics Reports* 2011; 504:1-73; <http://dx.doi.org/10.1016/j.physrep.2011.03.003>
- Meluzzi D, Smith DE, Arya G. Biophysics of Knotting. *Annu Rev Biophys* 2010; 39:349-66; PMID:20192771; <http://dx.doi.org/10.1146/annurev.biophys.093008.131412>
- Lua RC, Grosberg AY. Statistics of Knots, Geometry of Conformations, and Evolution of Proteins. *PLoS Comput Biol* 2006; 2:e45; PMID:16710448; <http://dx.doi.org/10.1371/journal.pcbi.0020045>
- Grosberg AY. Crumpled globule model of DNA packing in chromosomes: from predictions to open questions. In: Mondaini RP, editor. BIOMAT 2011 Singapore: WORLD SCIENTIFIC; 2011. pages 17-28
- Mirny LA. The fractal globule as a model of chromatin architecture in the cell. *Chromosome Res* 2011; 19:37-51; PMID:21274616; <http://dx.doi.org/10.1007/s10577-010-9177-0>
- Levinthal C. How to fold graciously. *Mossbauer Spectroscopy Biol Syst* 1969; 67:22-4
- Dill KA, Chan HS. From Levinthal to pathways to funnels. *Nat Struct Mol Biol* 1997; 4:10-9; PMID:8989315; <http://dx.doi.org/10.1038/nsb0197-10>
- Sikorav JL, Jannink G. Kinetics of chromosome condensation in the presence of topoisomerases: a phantom chain model. *Biophysical J* 1994; 66:827-37; PMID:8011915; [http://dx.doi.org/10.1016/S0006-3495\(94\)80859-8](http://dx.doi.org/10.1016/S0006-3495(94)80859-8)
- Meaburn KJ, Misteli T. Cell biology: Chromosome territories. *Nature* 2007; 445:379-81; <http://dx.doi.org/10.1038/445379a>; PMID:17251970
- Cremer T, Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genetics* 2001; PMID:11283701; 2:292-301; <http://dx.doi.org/10.1038/35066075>
- Rosa A, Everaers R. Structure and Dynamics of Interphase Chromosomes. *PLoS Comput Biol* 2008; 4:e1000153; PMID:18725929; <http://dx.doi.org/10.1371/journal.pcbi.1000153>
- Virnau P, Mallam A, Jackson S. Structures and folding pathways of topologically knotted proteins. *J Phys Condens Matter* 2010; 23:033101; PMID:21406854; <http://dx.doi.org/10.1088/0953-8984/23/3/033101>
- Potestio R, Micheletti C, Orland H. Knotted vs. unknotted proteins: evidence of knot-promoting loops. *PLoS Comput Biol* 2010; 6:e1000864; PMID:20686683; <http://dx.doi.org/10.1371/journal.pcbi.1000864>
- Virnau P, Mirny LA, Kardar M. Intricate Knots in Proteins: Function and Evolution. *PLoS Comput Biol* 2006; 2:e122; PMID:16978047; <http://dx.doi.org/10.1371/journal.pcbi.0020122>
- Taylor WR. Protein knots and fold complexity: Some new twists. *Computational Biol Chem* 2007; 31:151-62; PMID:17500039; <http://dx.doi.org/10.1016/j.compbiochem.2007.03.002>
- Rybenkov VV, Cozzarelli NR, Vologodskii AV. Probability of DNA knotting and the effective diameter of the DNA double helix. *PNAS* 1993; 90:5307-11; PMID:8506378; <http://dx.doi.org/10.1073/pnas.90.11.5307>
- Arsuaga J, Vazquez M, McGuirk P, Trigueros S, Sumners DW, Roca J. DNA knots reveal a chiral organization of DNA in phage capsids. *PNAS* 2005; 102:9165-9; PMID:15958528; <http://dx.doi.org/10.1073/pnas.0409323102>
- Shaw SY, Wang JC. Knotting of a DNA chain during ring closure. *Science* 1993; 260:533-6; PMID:8475384; <http://dx.doi.org/10.1126/science.8475384>
- Micheletti C, Di Stefano M, Orland H. Absence of knots in known RNA structures. *Proc Natl Acad Sci USA* 2015; 112:2052-7; PMID:25646433; <http://dx.doi.org/10.1073/pnas.1418445112>
- VanLoock MS, Harris BA, Harvey SC. To knot or not to knot? Examination of 16S ribosomal RNA models. *J Biomol Structure Dynamics* 1998; 16:709-13; PMID:10052626; <http://dx.doi.org/10.1080/07391102.1998.10508282>
- Boyle J, Robillard GT, Kim S-H. Sequential folding of transfer RNA: A nuclear magnetic resonance study of successively longer tRNA fragments with a common 5' end. *J Mol Biol* 1980; 139:601-25; PMID:6997498; [http://dx.doi.org/10.1016/0022-2836\(80\)90051-0](http://dx.doi.org/10.1016/0022-2836(80)90051-0)
- Nussinov R, Tinoco I Jr. Sequential folding of a messenger RNA molecule. *J Mol Biol* 1981; 151:519-33; [http://dx.doi.org/10.1016/0022-2836\(81\)90008-5](http://dx.doi.org/10.1016/0022-2836(81)90008-5); PMID:6279861
- Petrov AS, Bernier CR, Hsiao C, Norris AM, Kovacs NA, Waterbury CC, Stepanov VG, Harvey SC, Fox GE, Wartell RM, et al. Evolution of the ribosome at atomic resolution. *Proc Natl Acad Sci USA* 2014; 111:10251-6; PMID:24982194; <http://dx.doi.org/10.1073/pnas.1407205111>
- Yoffe AM, Prinsen P, Gopal A, Knobler CM, Gelbart WM, Ben-Shaul A. Predicting the sizes of large RNA molecules. *Proc Natl Acad Sci USA* 2008; 105:16153-8; PMID:18845685; <http://dx.doi.org/10.1073/pnas.0808089105>
- Tubiana L, Božić AL, Micheletti C, Podgornik R. Synonymous Mutations Reduce Genome Compactness in Icosahedral ssRNA Viruses. *Biophysj* 2015; 108:194-202; PMID:25564866; <http://dx.doi.org/10.1016/j.bpj.2014.10.070>
- Bon M, Orland H. TT2NE: a novel algorithm to predict RNA secondary structures with pseudoknots. *Nucleic Acids Res* 2011; 39:e93-3; PMID:21593129; <http://dx.doi.org/10.1093/nar/gkr240>
- Bon M, Micheletti C, Orland H. McGenus: a Monte Carlo algorithm to predict RNA secondary structures with pseudoknots. *Nucleic Acids Res* 2012; 41:1895-900; PMID:23248008; <http://dx.doi.org/10.1093/nar/gks1204>

29. Strigles JC, Martin MB, Schmidt FJ. Frequency of RNA–RNA interaction in a model of the RNA World. *RNA* 2006; 12:353-9; PMID:16495233; <http://dx.doi.org/10.1261/rna.2204506>
30. Briones C, Stich M, Manrubia SC. The dawn of the RNA World: toward functional complexity through ligation of random RNA oligomers. *RNA* 2009; 15:743-9; PMID:19318464; <http://dx.doi.org/10.1261/rna.1488609>
31. Vaidya N, Manapat ML, Chen IA, Xulvi-Brunet R, Hayden EJ, Lehman N. Spontaneous network formation among cooperative RNA replicators. *Nature* 2012; 491:72-77; PMID:23075853; <http://dx.doi.org/10.1038/nature11549>
32. Richardson JS. Beta-sheet topology and the relatedness of proteins. *Nature* 1977; 268:495-500; PMID:329147; <http://dx.doi.org/10.1038/268495a0>
33. Mansfield ML. Are there knots in proteins? *Nat Struct Mol Biol* 1994; 1:213-4; PMID:7656045; <http://dx.doi.org/10.1038/nsb0494-213>
34. Taylor WR. A deeply knotted protein structure and how it might fold. *Nature* 2000; 406:916-9; <http://dx.doi.org/10.1038/35022623>; PMID:10972297
35. Wang H, Di Gate RJ, Seeman NC. An RNA topoisomerase. *PNAS* 1996; 93:9477-82; PMID:8790355; <http://dx.doi.org/10.1073/pnas.93.18.9477>
36. Xu D, Shen W, Guo R, Xue Y, Peng W, Sima J, Yang J, Sharov A, Srikanthan S, Yang J, et al. Top3 $\beta$  is an RNA topoisomerase that works with fragile X syndrome protein to promote synapse formation. *Nat Neurosci* 2013; 16:1238-47; PMID:23912945; <http://dx.doi.org/10.1038/nn.3479>
37. Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, Marzluff WF, Sharpless NE. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* 2013; 19:141-57; PMID:23249747; <http://dx.doi.org/10.1261/rna.035667.112>
38. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregeren LH, Munschauer M, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 2013; 495:333-8; PMID:23446348; <http://dx.doi.org/10.1038/nature11928>
39. Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO. Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types. *PLoS One* 2012; 7:e30733; PMID:22319583; <http://dx.doi.org/10.1371/journal.pone.0030733>
40. Frank-Kamenetskii MD. RNA topology. Artificial DNA: PNA & XNA 2014; 4:35-6; PMID:23603781; <http://dx.doi.org/10.4161/adna.24680>
41. King NP, Yeates EO, Yeates TO. Identification of Rare Slipknots in Proteins and Their Implications for Stability and Folding. *Journal of Mol Biol* 2007; 373:153-66; PMID:17764691; <http://dx.doi.org/10.1016/j.jmb.2007.07.042>
42. Sulkowska JI, Rawdon EJ, Millett KC, Onuchic JN, Stasiak A. Conservation of complex knotting and slipknotting patterns in proteins. *Proc Natl Acad Sci USA* 2012; 109:E1715-1723; PMID:22685208; <http://dx.doi.org/10.1073/pnas.1205918109>
43. Jamroz M, Niemyska W, Rawdon EJ, Stasiak A, Millett KC, Sułkowski P, Sulkowska JI. KnotProt: a database of proteins with knots and slipknots. *Nucleic Acids Res* 2015; 43:D306-314; PMID:25361973; <http://dx.doi.org/10.1093/nar/gku1059>
44. Mallam AL, Onuoha SC, Grossmann JG, Jackson SE. Knotted fusion proteins reveal unexpected possibilities in protein folding. *Mol Cell* 2008; 30:642-8; PMID:18538662; <http://dx.doi.org/10.1016/j.molcel.2008.03.019>
45. Burton AS. Characterization of Novel Functions and Topologies in RNA. PhD Thesis, Portland State University, 2011;
46. Tubiana I, Orlandini E, Micheletti C. Probing the entanglement and locating knots in ring polymers: a comparative study of different arc closure schemes. *Prog Theor Phys Supplement* 2011; 191:192-204; <http://dx.doi.org/10.1143/PTPS.191.192>