Pergamon

Biophysics, Vol. 42, No. 6, pp. 1245–1252, 1997

0006–3509/97 $24.00 + .00

PII: S0006-3509(98)00040-4

ANALYSIS OF THE DYNAMICS OF THE TRANSITIONS IN *LEPTOMONAS COLLOSOMA* SPLICED LEADER RNA AND THE POSSIBLE ROLE OF KNOTS ON SPLICING*

Yu. B. MAGARSHAK

Courant Institute of Mathematical Sciences, New York, U.S.A.

(*Received* 24 *July* 1995)

The existence at the 5′-end of two competing secondary structures of spliced leader RNA of *Leptomonas collosoma* was shown in [1]. On changing from one stable secondary structure to the other and back, RNA passes through a series of states in which the chain ties into knots. In the case of reversible transitions, the knots both in the forward and backward directions are figurative, although the resulting secondary structures are quite different and the forward and backward pathways do not coincide. It is assumed that under certain conditions the chain may tie into the following knot of series 5_2, which most probably does not untie. This knot tightens around the GG nucleotides, with possible migration of a hole, as a result of which the chain may tear. Independent experiments show that during splicing the RNA chain tears precisely at this site. On the basis of these results the possible role of knots in enzymatic activity is discussed. © 1998 Elsevier Science Ltd. All rights reserved.

Triple interactions between nucleotides were discovered much later than the double helix determined by complementary, i.e. binary, interactions between nucleotides. These interactions are primarily responsible for the processes of migration of branchings or bifurcations [2]. For a bifurcation to migrate, the sequence of nucleotides in each of the chains forming it must possess definite symmetry. In the course of the jump of the bifurcation to one nucleotide to any side, two Watson and Crick bonds tear and two are born. The question of whether migration of the bifurcation, in which only one bond tears and one is created, is possible, is perfectly natural. A model of this kind in which in the loop the bond jumps from the nucleotide following the 3′ end of the loop to the nucleotide present at its 5′-end (or vice versa from the nucleotide preceding the 5′-end of the loop to the nucleotide present at its 3′-end) was first predicted theoretically in [3, 4]. This was followed by an analysis of the possible structures formed by dynamic loops [5, 6]. Thus, if two loops migrate over regions with common nucleotides, they may completely or partially annihilate, unite, repel, etc. However, the secondary structures determined by the jump of one single bond need not necessarily generate movement of the loop. What really counts is the migration of the hole between two chains interacting with the third chain (see Fig. 1). Analysis and classification of such possible structures is not the aim of the present paper and requires special examination. We merely note that migration of the hole only terminologically resembles

* *Biofizika*, **42**, No. 6, 1216–1223, 1997.

for example, on relaxation. The difference in the pathways in the forward and backward directions does not contradict the principles of statistical mechanics. In particular, it naturally does not contradict the principle of detailed balance since on relaxation the system is in general not in equilibrium.

At first sight, the hypothesis on topological control over enzymatic activity, at least in RNA, although it looks fine, is nevertheless fantastic. It is all the more surprising that apparently there is convincing experimental evidence that such control exists.

As shown in [1], the 5′ half of spliced leader RNA of *Leptomonas collosoma*:

```
           10          20          30          40          50
5'>AACUAAAACA AUUUUUGAAG AACAGUUUCU GUACUUCAUU GGUAUGUAGA GACUUCU >3'
```

containing the splicing region, may switch between two alternative structures, namely between form 1 [1]

```
         γ           β                      α
      GACA  a  GAAGU  uuuuaacaa,  AAUCA  a                   5′
      ||||     |||||              ||.||
  uuu CUGU  a  CUUCA              UUGGU  auguagagacuuc       3′
         δ           ε                      ζ
```

and form 2:

```
             3            4
  ucauu  GGUAU  gu  AGAGACU  uc                              3′
         .|||.      |||.|||
         UCAUG      UCUUUGA  caagaaguuuuuaacaaaaucaa,        5′
             2            1
```

According to experiments, these two structures differ only slightly in stability and pass into each other in a time less than 1 s.†

We analysed the possible transitions between configurations on the assumption that there are no intermediate states occurring either in form 1, or inform 2 obtained in the experiment [1]. All the following results essentially depend on this assumption, which satisfies the "Ockham's razor" principle.

Note that there is no pair of nucleotides in form 1 which would remain undisrupted in form 2, and conversely. Therefore, finding the low-energy transition between these forms is extremely problematical. However, we found a low-energy pathway of such a transition on the assumption that the hole migrates between certain states (see the Appendix). Below we analyse to which topological transformations of the chain in three-dimensional space these transformations of secondary structure correspond.

† The standard notation of secondary structures is insufficient to describe the dynamics of their transition. We therefore use a special system of notation.

1. The nucleotide sequences are written in odd lines. The nucleotides are denoted by letters. Paired nucleotides are denoted by capital letters and free nucleotides by lower case letters.

2. Odd lines denote interactions between nucleotides. Vertical dashes correspond to Watson–Crick bonds. Weak UG-bonds are denoted by a point. The migration of holes and also Holiday bifurcations are denoted by asterisks. The arrows may indicate the direction of migration of the holes.

3. If the sequence is not placed on one line, at the start and end of the line the corresponding numbers of the 3′- and 5′-ends are indicated so that the complete sequence is readily reconstructed.

4. The gap between nucleotides arises if necessary to depict the secondary structure. It has no special meaning, so that the nucleotides written in the line with a gap between them or without a gap in any event are adjacent.

Let us begin with form 2. First, as a result of fluctuation the long 5′-end may approach the helix 2–3 and interact with it. As a result of the interaction with portion 3 the helix α–ζ forms and the secondary structure becomes "pseudoknotted". In three-dimensional space the 5′-end so twists into a helix that a knot forms. Now it should be borne in mind that any helical structure is chiral. About every ten nucleotides form one turn. If the initial helix of the pin makes about one turn, the knot into which the molecule is tied is the figurative knot 4_1. If the helix of the pin makes about one and half revolutions, then the following knot of the series 5_2 forms. In this case the pin had 12 bonds and the most probable knot into which the molecule is tied is the knot 4_1. Then the curvature and elasticity of the chain promote destruction of helix 1–4 and facilitate the self-formation of the helix γ–δ, competing with it with the loop UUU in the middle. At this stage a long loop also forms between the portions 1 and α. Then the elastic forces of the chain move the liberated 3′-end through the loop. As a result the knot unties. Note that the figurative knot 4_2 is the second knot in the trefoil series and there is obviously no problem with untying the knot when the 3′-end simply slips through the loop. The resulting structure without knots allows the helices β–ε to form and the system arrives at the form 1 [1].

The reverse pathway does not coincide with the forward one. This occurs for the same reason that it is incomparably less difficult to withdraw a thread from a needle than to put it back into the eye of the needle. In this case the probability of the arbitrary entry of the 3′-end into the loop is even more difficult than into the eye of a needle because the eye (loop) heavily crashes into the helix β–ε after the chain has emerged from it, which, of course, never happens with an ordinary needle.

The reverse pathway from form 1 to form 2 begins with fixation of the 3′-end. It embraces the loop γ–δ and interacts with the portion γ by the mechanism of hole migration. Part of the thick helix 1–4 forms. Note the specific character of this transition, since when each bond ruptures two new ones form. This transition is obviously sufficiently advantageous energy-wise to compensate for the elasticity of the bent 3′-end and to fix it. Simultaneously, the helix β–ε is partially or completely destroyed. The liberated region 2 interacts with the portion GGU in positions 41–43. Then, as a result of hole migration in the GGU nucleotides, the 5′-end is freed. However, later the situation essentially depends on the knot into which the molecule is tied. In the experiment a pathway exists from form 1 to form 2. This can only mean that this knot is a figurative knot and unties, as occurs on the pathway from form 2 to form 1. If the molecule is tied into the following knot of series 5_2, self-untying of the chain from it is unlikely. As a result of hole migration in the GGU triplet in positions 41–43 the 5′-end is freed, as in the preceding case, although this does not lead to untying of the knot. All the above transformations seem intricate in a verbal description. But they are exceptionally effective if they are made, for example, with a cord as a model of the polynucleotide chain.

Note that in the knot 5_2 on the GG bond the molecule is very heavily bent and on this pair of nucleotides as may be seen from the Appendix, the knot tightens. From independent experiments, it is known that splicing occurs precisely on this bond. This is a strong indirect argument in favour of the idea that this reconstruction is correct.

DISCUSSION OF THE RESULTS

Some additional features of the forms 1 and 2 facilitate the formation of the knots described above. For a figurative knot to form, the molecule must bend into a ring, made easier by the loop GU in form 1. Simple analysis shows that this loop bends the molecule precisely in the requisite direction to form the given knot. On the other hand, when knot 5_2 forms, the free end must

connect with the helix, which makes one and half revolutions. This interaction is promoted by a pair of loops in form 2 consisting of only one nucleotide A present precisely one opposite the other.

Then the sequence UUUUU is in the middle of the loop from which the 3′-end slips out. This may have a special meaning. First, uracil is a pyrimidine and consequently is smaller than each of the purine molecules. Second, of the two pyrimidines it has a lower energy of formation of complementary bonds. Both these factors help to untie the knot. We would expect the region of the loop from which the 3′-end slips out and also the region of GG pairs in which the hole migrates and the knot ties, to be conservative in the course of evolution.

Of course, the dynamics of the knots under the control of the primary sequence may be incomparably more complex than that developed in the present work. From this point of view the pin is something like a hydrogen atom in quantum mechanics, i.e. the model that is the simplest to analyse. The fact that this mechanism was detected in [1] may be considered not only a scientific achievement, but also a great success.

Of course, it is extremely tempting to speculate, on the basis of these results, on the topological regulation of enzymatic activity. In fact, the results of this paper appear to give grounds for the assumption that regulation using the dynamics of knots, at least in some processes *in vivo*, actually takes place.

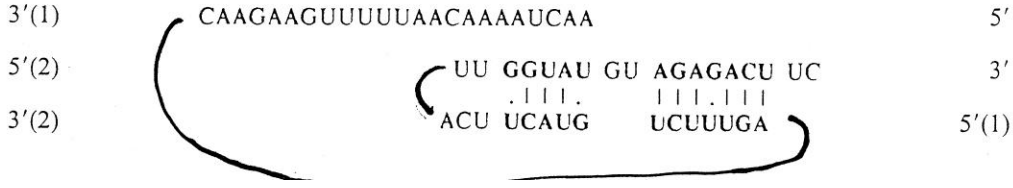
Recently, the theory of secondary DNA structures was proposed on the basis of topological invariants [8]. In this theory the apparatus of the theory of knots was used, although in this case knots proper were not considered. The results of the present paper encourage the creation of a more general theory, including a description both of the secondary structures and the knots, which they form in three-dimensional space.

It seems to us that the sequence of transitions described in broad terms is the only possible chain of low-energy transitions. Theoretically, there is only a solitary trivial alternative to it, namely, that transitions are not in general low-energy, but simply times of the order 0.1 s, sufficient for all the bonds in forms 1 and 2 [1] to tear and unite anew. Experiments undoubtedly will soon show which of these two processes — trivial or with topological transitions — actually occurs. However, in the trivial case there is no direct link of forms 1 and 2 with splicing. We again recall that in the model considered the knot tightens precisely on the pair of GG nucleotides on which splicing occurs. If this is a mere coincidence, it is striking, since the probability of it happening is exceptionally low. If this is not mere coincidence, then it appears that there is no alternative to the chains of transformations presented in the present paper.

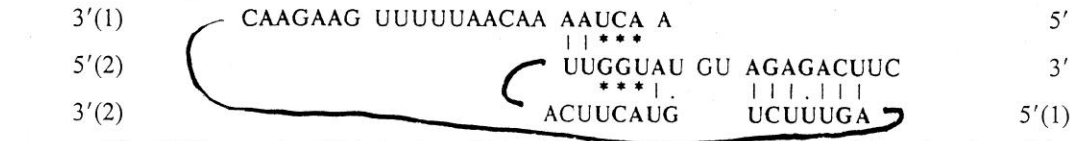
Of course, the situation *in vivo* may be much more complex than the proposed model. Possibly, some details will be amended or refined in due course. However, the main thing appears to us to be beyond doubt, namely, the key role of RNA knots and their transformations in the course of splicing.

APPENDIX

The proposed sequence of low-energy transitions of the 5′-half of *Leptomonas collosoma* spliced Leader RNA from form 2 to form 1.

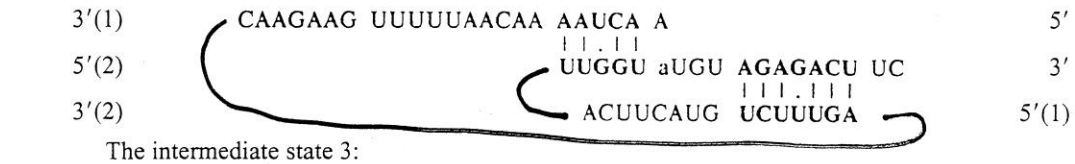


The intermediate state 1:

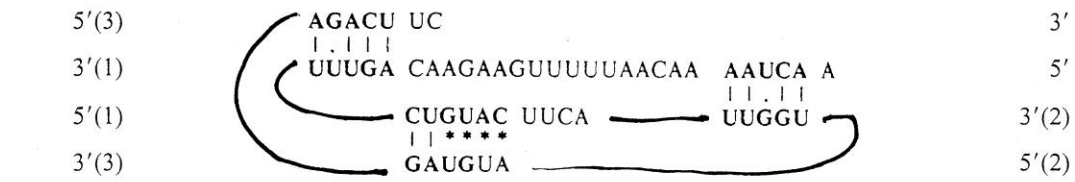


The GGU zone, in which *in vivo* splicing takes place, is in the middle of the migration of the bifurcation. However, on passing in this direction, rupture of the chain cannot occur since the molecule is still not tied into a knot.

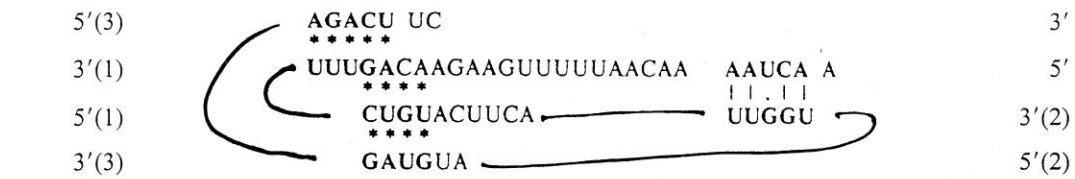
The intermediate state 2:



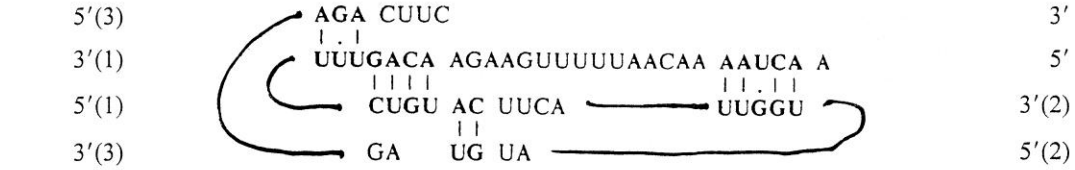
The intermediate state 3:



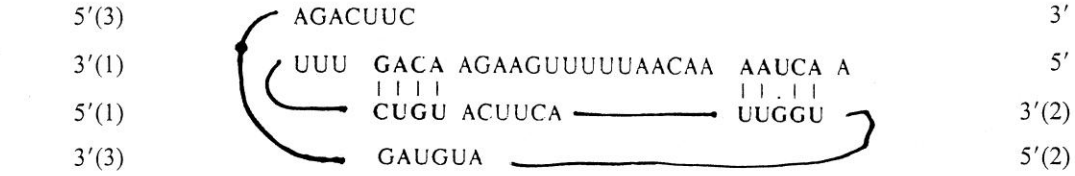
The intermediate state 4:



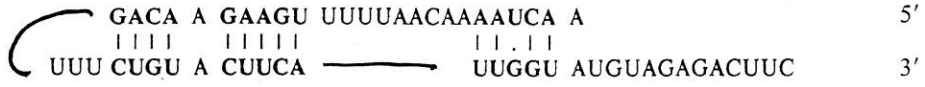
The intermediate state 5:



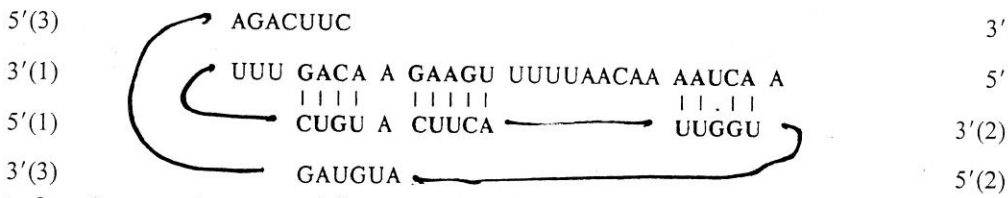
The intermediate state 6:



And, finally, form 2:

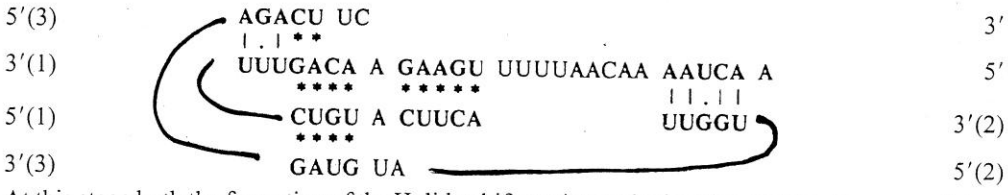


The transitions in the reverse direction, i.e. starting from form 1



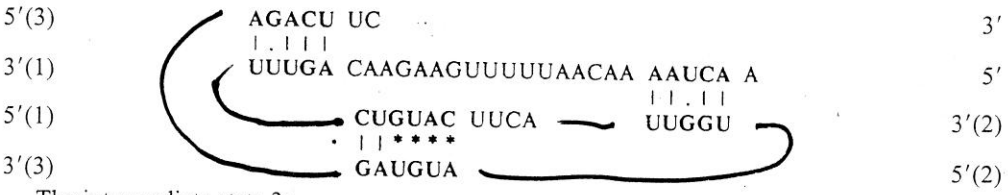
to form 2 apparently occur as follows.

The first intermediate state:

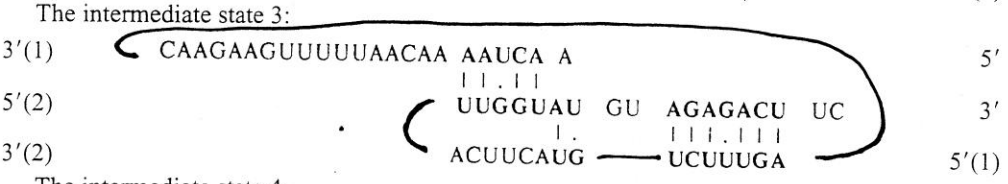


At this stage both the formation of the Holiday bifurcation and migration of the hole are possible.

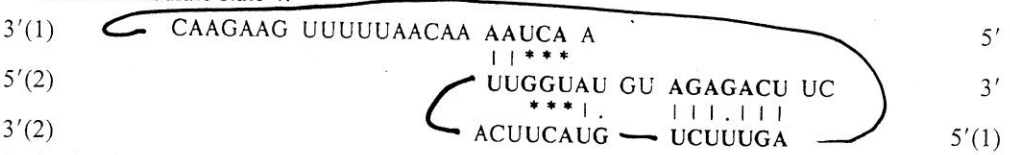
The second intermediate state:



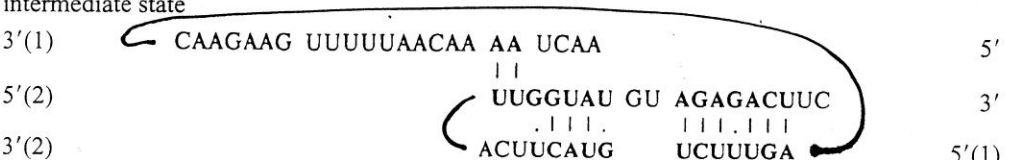
The intermediate state 3:



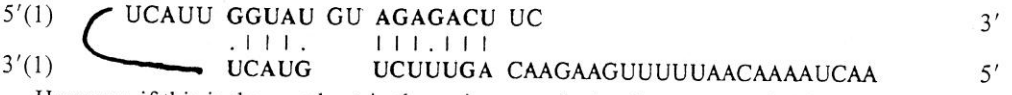
The intermediate state 4:



In the GGU zone migration of the hole occurs. Then the dynamics depends on the knot into which the molecule ties. If this is knot 4_2, then the molecule unties from it and through the intermediate state



arrives at the final form 2:



However, if this is the next knot in the series, namely 5_2, then most probably the chain cannot arrive at form 2. The 5′-end is liberated, although untying of the knot is unlikely. The hypothesis that rupture of the chain on the bent GG bond (splicing) results from the liberation of the 5′-end

appears reasonable. However, it is possible that a role may also be played in this rupture by additional factors going beyond the transformations of the knots to which this paper is devoted.

REFERENCES

1. K. A. LeCuyer and D. M. Crothers, *Biochemistry*, **32**, 5301 (1993).
2. I. G. Panautin and P. Hsieh, *Proc. Nat. Acad. Sci.*, **91**, 2021 (1994).
3. A. A. Mironov and A. E. Kister, *J. Biomol. Str. Dyn.*, **4**, 1 (1986).
4. A. A. Mironov and A. E. Kister, *Molec. Biol.*, **23**, 61 (1989).
5. Y. Magarshak and C. Benham, *J. Biomol. Str. Dyn.*, **10**, 465 (1992).
6. Y. Magarshak, *Biosystems*, **30**, 21, Elsevier Science Publishers (1993).
7. L. H. Kauffman, *World Scientific* (1991).
8. L. H. Kauffman and Y. B. Magarshak, *J. Knot Theory and its Ramifications*, **3**, 233 (1994).