Simulated pulling experiments of an RNA hairpin.

Author: Joana Rosell Mirmi
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.

Advisors: Jaime Aspas Cáceres
Félix Ritort Farrán
(Dated: June 18, 2020)

Abstract: Different studies have shown how to approach the folding process of the RNA, with optical tweezers. On this basis, we've present a simple case of the simulation of a single RNA hairpin pulling experiment. During this work, the experimental device simulated is described and modeled, where also its elastic properties have been considered, as they are an important part of the study. Subsequently, information is extracted from the data and the results obtained on the simulation, related with the thermodynamics of kinetics the unfolding/folding of RNA molecule. The simulated model reproduces qualitatively the experimental results.

I. INTRODUCTION

The study of biological systems at the molecular level has become one of the most prolific areas of science. Since Watson and Crick discovered the double helix structure of DNA in 1953 [1], a lot of tools and methods have been developed to look deeply into the molecules that contain the genetic information of living organisms, viruses, etc [2].

Nowadays, biological and biomedical research has shown to be crucial for society. Everyday we find more evidence about the nature of disease, and it has been proved that some of them are related with some conformational changes on bio-molecules as proteins, others are caused by viruses, whose principal constituent are just genetic information in the form of DNA or RNA. So understanding the behaviour of this bio-molecules is necessary.

In this work we focused on the RNA, which is crucial for life, and it is a candidate for being the life-promoter molecule in Earth. Besides, RNA stores the information needed for synthesizing a protein, and also is the main piece of the ribosome, the molecular machine in charge of making proteins. On the other side, RNA behaviour does not only depends on its sequence. The secondary structure is also a key aspect, RNA can play the role of an enzyme by catalyzing molecular reactions, as in some regulatory aspects in translation: the process of making a protein from a messenger RNA (mRNA) molecule [3]. Techniques for RNA research are very diverse [4], [5]. The Small Biosystem Laboratory is specialized in single molecule techniques as magnetic and optical tweezers [5]. In these kind of experiments, we are able to manipulate one molecule at a time. This allow us to extract precise information about the studied molecule using non-equlibrium statistical mechanics applied to small systems [6]. From this approach, some crucial quantities, such as free energies and kinetic rates can be determined in a very precise way [7].

We focused on CD4-RNA, a simple RNA hairpin which forms a short double chain (*double stranded, dsRNA*) and ends with an unpaired bases forming a loop. When this molecule is stretched, the dsRNA part of the molecule unzipps, so the systems goes from its folded configuration to the unfolded one, in which the molecule forms as a simple polymer chain (*single stranded, ssRNA*).

II. DEVELOPING SECTIONS

A. Experimental setup and modeling the system

The mini-tweezers setup [5], [7], described in FIG. 1a and 1b, uses two highly focused counter propagating laser beams to trap a polystyrene bead. It is explained by considering the same intensity for all the rays, then, the moment that the laser loses when passing through the bead, is the same that the bead gains, by conservation of the moment. Hence for the difference in momentum, a force is applied to the bead, that can be directly measured, while the other end is fixed.

On the molecular RNA studied, two polystyrene beads are functionalized one with streptavidin (SA) and the other with antidigoxigenin (AD). The SA bead is fixed by air suction on the tip of a micro-pipette meanwhile the AD bead is captured in the optical trap. The molecular construct consists of the RNA molecule being studied linked to two identical handles, of identical double-stranded hybrid DNA/RNA. These handles are modified with biotin and digoxigenin at their ends, in order to be attached to the beads, thanks to this highly specific antibody-antigen interaction [8]. The AD beads are incubated with the molecular construct. Flowing in an adequate order the solutions containing the beads, and manipulating the optical trap, the configurations shown in Figure 1a and 1b are achieved.

In our experimental pulling protocol, the system starts from a known point, being this the RNA molecule folded. Then, we apply the protocol, in other words, we apply an external force through the light tweezers, hence we

achieve to move the trap up during the time of application, to the other stable state, the unfolded state. Finally, by reversing the motion, meaning applying the same force but in the other direction possible, so the bead moves down, we managed to reverse the protocol back to the initial state, RNA molecule is folded again.

Since the nature of the unfolding/folding process is stochastic, the unfolding force, at which the molecule opens, as well as the refolding force, at which the molecule closes again, will differ from experiment to experiment, so cannot be predicted, other than by the probability that at a given force, the molecule will be unfolded or else, it will remain folded. Be the same for the folding process.

In this study, two different elastic models are used to describe the molecular setup. Firstly, ssRNA and the double-stranded handles are considered as a polymers. Therefore, the first model we used to describe polymers is the freely-jointed chain (FJC). It assumes that a polymer is a random walk, without any interaction between monomers. For a given force, the extension of the polymer is given by:

$$x(f) = d\left[coth\left(\frac{fd}{k_BT}\right) - \frac{k_BT}{fd}\right] \tag{1}$$

where k_B is the Boltzmann constant, T is the temperature. For the folded state of the RNA molecule, it can be assumed to behave as a single bond of length $d = 2nm$.

Secondly, the worm-like chain model (WLC) is used for semi-flexible rigidity polymers, that is, polymers with a bendy segments that interact with each other. This model takes into account the persistence length [9]. The RNA molecule in its open state, as well as the behavior of the DNA handles will be described by this elastic model. The WLC model gives the elastic force of the polymer as a function of its molecular extension x,

$$f = \frac{k_BT}{P}\left[\frac{1}{4(1-x/L_c)^2} - \frac{1}{4} + \frac{x}{L_c}\right] \tag{2}$$

being the P the persistence length of the molecule modeled and L_c its molecular contour length, taken equal to $L_c = N \cdot d_b$, being N the total number of bases of the ssRNA, and the d_b is the inter-phosphate distance. This model is applied to the RNA in its unfolded state, and for the handles. The parameters used for the ssRNA molecule and for the handles are showed in the table I.
The bead captured in the optical trap satisfies:

$$|f| = k_b x_b \tag{3}$$

where the k_b is the stiffness of the optical trap [8]. From experimental data, $k_b = 0.073$ pN/nm.

As we mentioned before, the system only has two possible states. Its stochastic nature can be describe with a simple kinetic Bell-Evans model [12], [13]. The molecule goes from the folded state to the unfolded one, overcoming an energy barrier $B(\lambda)$, and the free energy difference between the two states is denoted as ΔG_{FU}. The kinetic rates corresponding to the Bell-Evans model are given by:

	Handles	**ssRNA**
P (nm)	20	0.75
d_b (nm/base)	0.27	0.655

TABLE I: **Summary table of the elastic parameters.** Shows the elastic parameters used for the handles and the RNA taken from [8], [10], [11].

$$k_{F\to U}(\lambda) = k_0 \exp\left(-\frac{B(\lambda)}{k_BT}\right) \tag{4}$$

$$k_{U\to F}(\lambda) = k_0 \exp\left(-\frac{B(\lambda) - \Delta G_{FU}}{k_BT}\right) \tag{5}$$

where k_0 is a pre-exponential factor, taken equal to $3 \cdot 10^4 s^{-1}$ [8], k_B is the Boltzmann constant and T the temperature.
From the point of view of the molecule, and taking into account the energy contributions from the experimental device, this energy potential can be seen as the free energy landscape, which maps all the possible configurations of a system with their corresponding free energies. To characterize the system, different reaction coordinates can be used, as the number of open base pairs, n, or the molecular extension. The FEL, at a fixed external force, f, is given by [8], [11]:

$$\Delta G_n(f) = \Delta G_n^0 + \Delta G_n^{st}(f) + \Delta G_n^d(f) \tag{6}$$

where the first contribution, ΔG_n^0 is on the free energy of formation at zero force of the configuration $n \in [0, N]$, being N the total of bases pairs that constitutes the molecule. The free energy of a base pair is given by the nearest-neighbor model (NN)[14], and the free energy of formation of the loop is taken from Mfold [15]; the second term, $\Delta G_n^{st}(f)$ refers to the work required to stretch the molecule when n base pairs have been opened to its equilibrium distance at force f so it has a dependence on the force applied; $\Delta G_0^d(f)$ is the contribution due to the orientation of the RNA molecule along the axis at which force is applied and only contributes for $n = 0$, when the RNA is folded.
The two states of the system correspond to two known positions on the FEL, so is possible to match the maximum of FEL with the energy barrier of the model, $B(\lambda)$. Also, it is possible to calculate the free energy difference between the two states, ΔG_{FU} as:

$$\Delta G_{FU}(\lambda) = \Delta G_{FU}^0 + \Delta W_{FU}^{st} + \Delta W_{FU}^h + \Delta W_{FU}^b \tag{7}$$

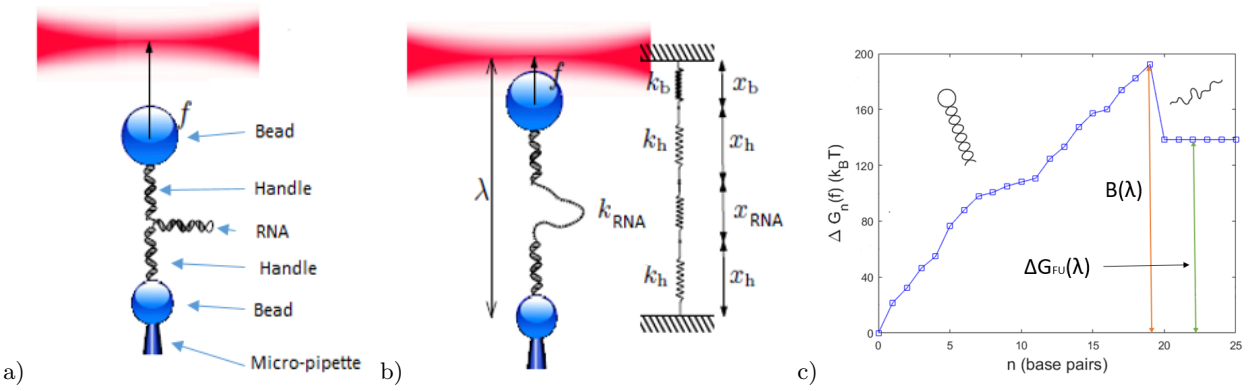

FIG. 1: **Scheme of the setup and the one-dimensional FEL of our system.** (a) The optical tweezers configuration. The beads are attached to the molecular system, one bead is captured by the optical trap, while the other one is immobilized in the tip of the micro-pipette by air suction. (b) The end-to-end distance, λ, with all the contributions simulated, and their respectively stiffnesses. (c) A graph of our one-dimensional free-energy landscape (FEL), for the two states system, folded (F) and unfolded (U), at force, $f = 10$ pN. The Bell-Evans model is used to approximate kinetics. It considers a single kinetic barrier at zero force, the height of which is $B(\lambda)$, and ΔG_{FU} is the difference of energies between the two states (F/U).

where ΔG^0_{FU} is the free energy of formation of the RNA at zero force; ΔW^{st}_{FU} is the work done to stretch the molecule so that it changes the state; ΔW^h_{FU} and ΔW^b_{FU} is the work done by the handles and the beads, respectively, when changing state.

B. Simulation of a pulling experiment

The code is designed to simulate a pulling protocol operated by the instrument. This protocol operates between two force values, going from the minimum value of the force to the maximum. We call this the *unfolding* pathway, or the forward part of the protocol. When the maximum force is reached, the optical trap returns to the initial point of the protocol, by the *folding* pathway, or the reverse part of the protocol. The experimental device repeats this protocol until the user stops it. The simulated working conditions are at room temperature, taken equal to 293 K, and the value of the Boltzmann constant multiplied by these temperatures is equal to 4.12 pN/nm with the relevant unit changes.

The control parameter in these studies is λ, defined earlier as the relative distance between the center of the optical trap and the bead immobilized in the tip of the micro-pipette. Then, for a given value of λ, it satisfies:

$$\lambda(f) = x^{\sigma}_{RNA}(f) + x_b(f) + 2 \cdot x_h(f) \qquad (8)$$

where, $x^{\sigma}_{RNA}(f)$ is the force-dependent distance of the RNA molecule, being σ the states folded and unfolded ($\sigma = F, U$), respectively; the $x_b(f)$ is the force-dependent distance between the bead captured by the optical trap and the center of the optical trap; and $x_h(f)$ is the force-dependent distance of the handles.

In this ensemble there are fluctuations in x_b and f given by [11][16] :

$$\langle \delta x_b^2 \rangle = \frac{k_B T}{k_x^{\sigma} + k_b} \qquad \langle \delta f^2 \rangle = \frac{k_B T k_b^2}{k_x^{\sigma} + k_b} \qquad (9)$$

where k_B is the Boltzmann constant, T is the temperature of the bath and k_b is the stiffness of the beads and k_x^{σ} is for the stiffness of the molecular setup (the handles and the RNA).

The key to the simulation is how to model properly the molecular transition from one state to another, since it is a stochastic process. As said, we consider our system with only two possible states, unfolded (U) and folded (F). In this case, the FEL of the system, in one dimension, can be schematized as 1c. The unfolding and folding rates of this process are given by equations (4) and (5).

Before we present the simulation, we have to take into account some considerations. The difference in size between the beads, the handles and the RNA molecule leads us to consider an instantaneous relaxation for the handles and bead to solve the dynamic equations [10]. To simulate the pulling experiment, we will solve for each given λ and for each state, folded and unfolded, equation (8), thus obtaining the distance-force relations, hence the path that will follow in the folded and unfolded regime.

The simulation starts in the folded state, with all bases pairs closed ($\sigma = F$). In the unfolded state, all bases pairs are open ($\sigma = U$). In what follows, the steps for the algorithm for an unfolding (folding) process are:

1. The optical trap is moved at a constant pulling speed v. Then, λ increases (decreases) following $\lambda_{i+1} = \lambda_i + v \cdot \Delta t$, where $\Delta t = 4 \cdot 10^{-3} s$ is the iteration time.

2. We solve equation (8) taking into account the elastic models described in section [II A] for $x^{\sigma}_{RNA}(f)$, $x_b(f)$ and x_h. We obtain the value of the force on the system, and to this value is added a Gaussian noise to beads position, x_b, and force, f, of zero mean and variance given by equation (9).

3. If the molecule is folded, it will unfold with probability $k_{F\to U}\Delta t$ Similarly, if it is unfolded, it will fold with probability $k_{U\to F}\Delta t$.

4. The system evolves until a maximum (minimum) value of the force is reached. When this condition is satisfied, the pulling speed changes its sign: $v_{i+1} = -v_i$

5. The code operates until a defined number of cycles is completed.

C. Results and discussion

We show the force-distance curves obtained by running our simulation for the following protocol parameters: $f_{\min}$ = 5 pN ; $f_{\max}$ = 25 pN, pulling velocity, v = 250 nm/s with 150 trajectories generated and v = 500 nm/s with 300 trajectories. The calculations for the kinetic rates and the survival probabilities have been carried out with the trajectories generated by the simulation at 500 nm/s.
As a first result, Figure 2a shows the force-distance curve (FDC) for data obtained from the simulation and compares it with the experimental data for this RNA molecule. The force-distance curves are drawn for a simulated experiment in magenta, for unfolding, and cyan for refolding, for a pulling velocity of v = 500 nm/s; and for experimental data in green for unfolding, and yellow for refolding. Experimental data was taken in a buffer containing 10 mM TRIS, 4mM $MgCl_2$ and 50 mM NaCl, at room temperature.

The force pattern of the FDC generated is shown to agree with the data collected from the experiments. The curves are shifted for clarity. As can be seen, the elastic behaviour of the experiment is well reproduced. The stochastic part of the simulation is illustrated in figure 2b, 2c and 2d. The rupture force of each cycle were collected and the histogram is shown in figure 2b. As we expected, higher pulling velocities have higher values of the rupture force, due to that in fact the systems dissipates more energy. From these histograms, we can calculate the survival probabilities as:

$$P_F(f) = 1 - \int_0^f \rho(f_U) df_U \tag{10}$$

$$P_U(f) = \int_0^f \rho(f_F) df_F \tag{11}$$

where, $\rho(f_U)$ and $\rho(f_F)$ are the unfolding and folding rupture force histograms, respectively.

The survival probabilities show that as long as the force increase, the probability of the folded state goes from 1 to 0, and the oppossite for the probability of the unfolded state. The molecule is expected to be unfolded over 15 pN, in agreement with the rupture force histograms. For the folding process the maximum in the force histograms are over 4 pN, over the same value from which the survival probability of the unfolded state increases. The last aspect studied is the validity of the kinetic model imposed. From the rupture force histograms and the survival probabilities, the kinetic rates, $k_{U\to F}$ and $k_{F\to U}$ can be obtain as in [8]:

$$k_{F\to U}(f) = -r\frac{1}{P_F(f)}\frac{dP_F(f)}{df} \text{ (unfolding)} \tag{12}$$

$$k_{U\to F}(f) = +r\frac{1}{P_U(f)}\frac{dP_U(f)}{df} \text{ (folding)} \tag{13}$$

where r is the loading rate, taken as $r = df/dt$.

The kinetic rates obtained can be checked by testing equations 4 and 5. As can be seen in figure 2d, when the kinetic rates are represented in logarithmic scale versus the force, it shows a linear behaviour. Finally, we are able to calculate the kinetic rates from the survivals probabilities, the histogram and the pulling speed [8], [11]. As expected, the probability of opening increases as the force applied is being also increased. And analogous, the probability of closing, once it has opened, will increase as the force decreases, which can be seen in figure 2c.

III. CONCLUSIONS

As conclusion, we can say that the application of the elastic models used to simulate the behavior of an RNA molecule are valid. The stochastic behaviour of the systems also is reproduced by the simulation, and from the analysis applied to the experimental data, one can see that the kinetic model holds. To summarize, we have developed a simulation that can be useful to complement the experimental work, as well as good way to get introduced in the understanding of these important and interesting systems.

Acknowledgments

I would like to thank A. Severino, P. Rissone and M. Rico for providing us of the experimental data to validate this simulation. I would like to give my sincere gratitude to Jaime Aspas Cáceres, for his guidance and teaching during this project. I would also like to thank Felix Ritort. A special thanks to my family and to you, Cristina, for your support, patience and love.

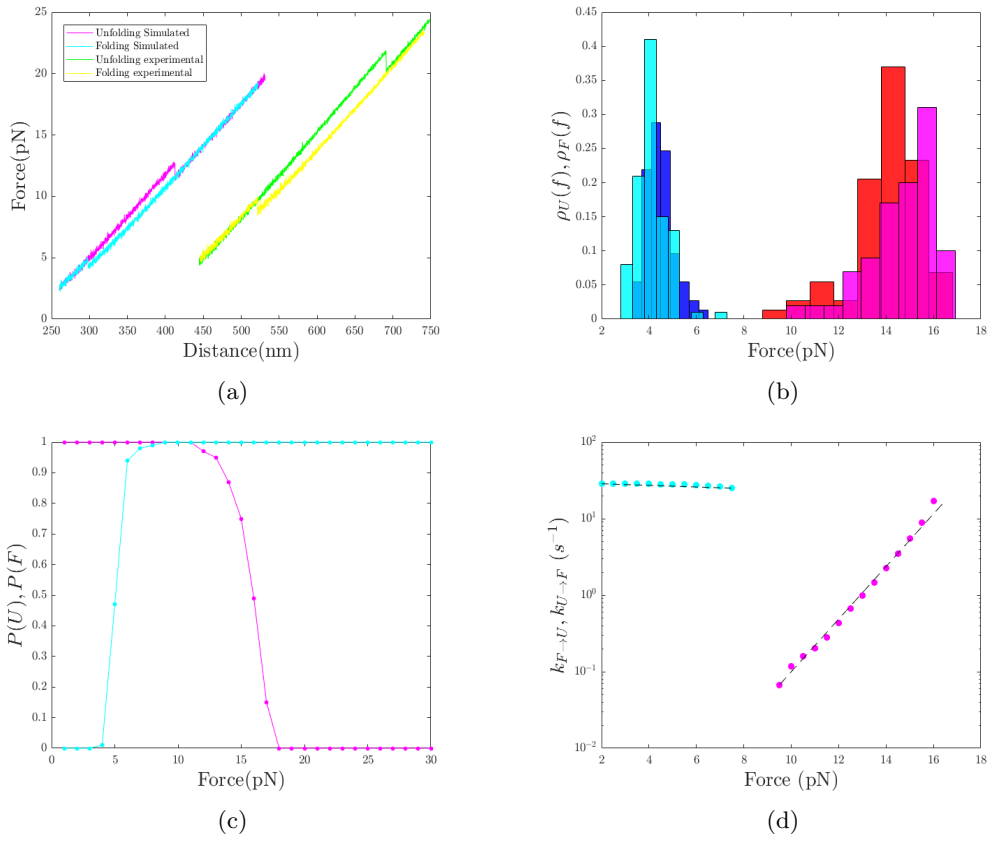

FIG. 2: **Pulling experiment and** (a) Force distance curves from simulation (magenta-cyan) and experimental (green-yellow). (b) Histograms for the rupture and the refolding forces for $v = 250$ nm/s (purple-orange) and $v = 500$ nm/s (magenta-cyan), calculated with the rice rule and with a Kernel density of probability function. (c) The survival probabilities are calculated for the unfolded process (magenta) and folded process (cyan) at v=500 nm/s. (d) The kinetic rates from equations 5 and 4 shown in logarithmic scale to il

[1] J. D. Watson, F. H. Crick, et al., Nature **171**, 737 (1953).
[2] F. Ritort, Journal of Physics: Condensed Matter **18**, R531 (2006).
[3] A. Galera-Prat, A. Gómez-Sicilia, A. F. Oberhauser, M. Cieplak, and M. Carrión-Vázquez, Current opinion in structural biology **20**, 63 (2010).
[4] K. C. Neuman and A. Nagy, Nature methods **5**, 491 (2008).
[5] F. Ritort, Revista Española de Física **32**, 39 (2018).
[6] A. N. Gupta, A. Vincent, K. Neupane, H. Yu, F. Wang, and M. T. Woodside, Nature Physics **7**, 631 (2011).
[7] J. M. Huguet, C. V. Bizarro, N. Forns, S. B. Smith, C. Bustamante, and F. Ritort, Proceedings of the National Academy of Sciences **107**, 15431 (2010).
[8] A. Alemany, Doctoral thesis (2014).
[9] J. Camunas-Soler, M. Ribezzi-Crivellari, and F. Ritort, Annual review of biophysics **45**, 65 (2016).
[10] A. Severino, A. M. Monge, P. Rissone, and F. Ritort, Journal of Statistical Mechanics: Theory and Experiment **2019**, 124001 (2019).
[11] A. Martínez, Doctoral thesis (2019).
[12] G. I. Bell, Science **200**, 618 (1978).
[13] E. Evans and K. Ritchie, Biophysical journal **72**, 1541 (1997).
[14] J. SantaLucia, Proceedings of the National Academy of Sciences **95**, 1460 (1998).
[15] M. Zuker, Nucleic acids research **31**, 3406 (2003).
[16] M. Ribezzi-Crivellari and F. Ritort, Biophysical journal **103**, 1919 (2012).