Effect of macromolecular crowding in DNA folding

Author: Arnau Farré Melero
*Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.**

Advisor: Isabel Pastor del Campo

Abstract: The effect of macromolecular crowding has already been studied in protein folding, but there are almost no precedents about its impact on the DNA folding. In this work, we have studied the folding of a DNA hairpin in the presence of different concentrations of PEG as a crowder agent using optical tweezers. The obtained results suggest that macromolecular crowding impacts in the studied DNA hairpin folding/unfolding process by varying its folding free energy and its coexistence force, becoming a more stable structure. Nevertheless, further studies are required to collect more data and explore other conditions to infer consistent conclusions on this topic.

I. INTRODUCTION

The understanding of behaviour of the biomolecules has become a fundamental issue and has aroused the interest of many researchers during the last decades. Some studies have already shown that the folding of molecules like proteins or DNA might have a direct effect on the development of diseases like Alzheimer or Parkinson [1], but they are also showing promising results in fields like Pharmacology [2]. Over the last few decades, Biophysics has introduced a new field of interest thanks to the technological advances in instrumentation like optical tweezers: the single-molecule experiments [3]. Before, the study of biomolecules was limited to bulk experiments, which showed only their collective phenomena. Nowadays, one can isolate one single molecule and retrieve measures (e.g., *pulling* or *hopping*), which give information about the variability between molecules.

In particular, nucleic acids are maybe the most fundamental biomolecules in terms of life, as they carry the genetic information stored in a cell. Hence, the study of its structure and stability is critical. Most of the studies on DNA folding are carried out under aqueous solution, although its activity is mainly developed inside the cells. The cellular cytoplasm is a heterogeneous solution that includes biopolymers (proteins, lipids), inorganic compounds (mainly ions like K^+, Na^+, Cl^-) and other small solutes. The total volume occupied by all the molecules is not negligible, as it could reach up to 40% [4]. Therefore, it is reasonable to expect that the results might not match the real behaviour of the acid nucleics, as their conditions are far away from the reality. Particularly, the macromolecules that populate the cytosol can be considered as crowder agents, so as a whole they form what is called *crowding*. The macromolecules can interact with the biomolecules in a nonspecific way due to excluded volume effects, among others. Therefore, the thermodynamics and kinetics of the DNA can result affected [5].

The effect of the macromolecular crowding has already been widely studied with proteins [6], but there are almost no studies on nucleic acids. In general, describing a simple process such as the transition between the native-denatured (e.g., folded-unfolded) conformations of biopolymer chains surrounded by crowder agents is not easy due to the lack of knowledge about the nature of the effective interactions between them. That, added up to the fact that single-molecule experiments demand a high amount of data in order to fully describe the behaviour of a molecule, significantly increases the complexity of these type of experiments. Due to temporal limitations, we did not expect to build a consistent model about the relation between the DNA folding and the concentration of macromolecules, but to take a first look on the topic. Then, the main aim of this work was to bring a preliminary approach on the issue. We consequently simplified the problem to only one type of crowder agent and a small set of concentrations.

Here, we have studied the folding/unfolding process of a 20 base pair (bp) DNA hairpin with a tetraloop in the presence of different Polyethylene Glycol (PEG, M_W = 8000 kDa) concentrations. This crowder is ideal for our purpose as it is a neutral random coil polymer widely used to mimic the intracellular crowded media and it is not expected to electrostatically interact with the DNA chains [7]. We considered PEG concentrations of 5, 23, and 50 mg/mL, while a free-PEG environment was taken as a control group. Experiments were performed in Tris buffer (formed by 10mM Tris, 1M NaCL, 5mM EDTA and 0.01 NaN_3) using optical tweezers. The results obtained suggest that PEG crowding might have a thermodynamical effect on our hairpin, while its folding kinetics remain almost unchanged.

II. EXPERIMENTAL

A. Synthesis of molecular construct

Firstly, it was necessary to prepare a molecular construct containing the molecule of interest (see *Fig. 1(A)*. We chose a short, well-characterized 20 bp hairpin due to its simple two-state behaviour between the folded and

*Electronic address: afarreme8@alumnes.ub.edu

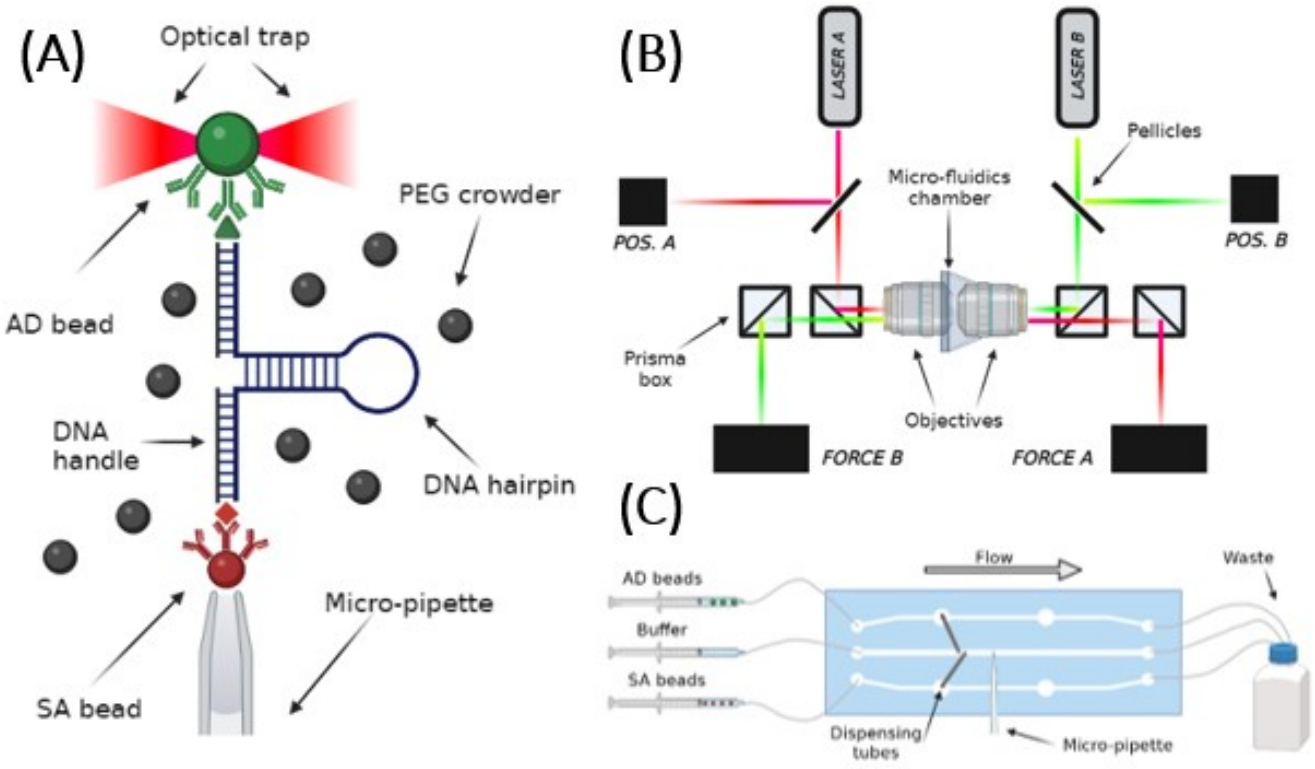

FIG. 1: Overview of the experimental setup of the experiments: (A) Sketch of a DNA hairpin molecular construction and experimental setup. (B) Scheme of the setup of the Minitweezers. (C) Scheme of a microfluidic chamber.

unfolded states, defined by the reaction:

$$F \underset{k_{UF}}{\overset{k_{FU}}{\rightleftharpoons}} U \tag{1}$$

The construct could be manipulated thanks to the *handles*, which are 29 bp double strand DNA sandwiching the hairpin that act as spacers. The 5' end was biotinated, which let it interact with streptavidin-coated beads (SA beads). Analogously, the 3' end was labelled with digoxigenin, which induced the antibody-antigen union with the anti-digoxigenin-coated beads (AD beads) [8]. Briefly, we carried out a ligation between two different oligonucleotides (supplied by Sigma-Aldrich), preceded by a tailing reaction that labeled the ends of the construct and the annealing reactions that formed the double-stranded structure of the *handles*.

B. Optical tweezers setup

In this work we used a *Minitweezers* setup, a double laser beam optical tweezers which proved to be very useful in single molecule experiments, as it creates a single optical trap being able to capture a bead. It is especially advantageous compared to the single-beam optical tweezers, as here the laser beams are exactly the same but with opposite propagation direction, cancelling the scattering forces on the optical trap. A scheme can be seen in *Fig. 1(B)*.

A few part of the laser light is directed to a position detector sensor (PSD) through a pellicle, and the rest is deflected and polarized through a prism box into one objective. Then, leaving the other objective, the polarized beam is deflected into a second prism box to the force detectors, which are able to measure its three components. The other beam follows the symmetric path, creating the optical trap at the center of the objectives [9]. Thus, one can move the computer-controlled optical trap thanks to a piezoelectric located at the end of the laser optical fibers.

The optical trap is formed at the center of a microfluidic chamber (shown in *Fig. 1 (C)*), whose three channels allow to distribute the beads separately through the lateral channels, which are connected to the central one by the dispensing tubes. The micro-pipette has air-suction capacity, allowing to fix the SA beads during the experiments.

C. Pulling experiments

In pulling experiments the trap is cyclically moved at a constant speed (here, 100 nm/s) between two fixed positions while the force is recorded. At low forces, the hair-

pin is in a folded state, with the stem in a double helix configuration, while at high forces the hairpin unfolds and the stem becomes a single-stranded DNA. Transitions between both states can be identified as force suddenly jumps in force-distance curves (FDCs, see *Fig. 2(A)*) .

In order to remove the possible bias of an experiment, we would ideally need at least 5 different molecules per considered medium. Besides, in order to retrieve enough data, each molecule should give around 100-120 trajectories. When we obtained a reasonable amount of data for a certain PEG concentration, we changed to a higher concentration and repeated the process.

D. The Bell-Evans model

To study the kinetics of our data, we used one of the simplest models in terms of single-molecule experiments, which is an approach proposed by Evans and based on the Bell model [10]. Here, Evans assumes that a two-state system presents only one kinetic barrier, $\Delta E(f)$, which can be expressed as:

$$\Delta E(f) = \Delta E(0) - f\Delta x_{TS} \tag{2}$$

Where Δx_{TS} is the distance between the pre-transition and transition states. Then, we can rewrite the kinetic rates as:

$$k_{FU} = k_m \exp \frac{f x_{F-TS}}{k_B T} \; ; \quad k_{UF} = k_m \exp \frac{\Delta G_{FU} - f x_{TS-U}}{k_B T} \tag{3}$$

By defining $f = f_c$ as the coexistence force (i.e., the value at which both states would coexist) and imposing $k_{FU}(f_c) = k_{UF}(f_c)$, we are left with:

$$\Delta G_{FU} = f_c x_{FU} \tag{4}$$

And we can also rewrite, as $x_{FU} = x_{F-TS} + x_{TS-U}$:

$$\log \frac{k_{FU}(f)}{k_{UF}(f)} = (f - f_c)\frac{x_{FU}}{k_B T} \tag{5}$$

We can derive the kinetic rates from the survival probability of the folded (unfolded) state during the unfolding (folding), $P_F(f)$ ($P_U(f)$). Hence, taking a set of folding (f_F) and unfolding (f_U) forces we can compute:

$$P_F(f) = 1 - \frac{N(f_U < f)}{N_U} \; ; \; P_U(f) = 1 - \frac{N(f_F > f)}{N_F} \tag{6}$$

Being N_U and N_F the number of considered unfolding and folding trajectories, respectively. Note that, as at the $f = f_c$ the survival probability for both states has to be the same for definition, we also find $P_F(f_c) = P_U(f_c)$. Finally, we get an expression for the kinetic rates introducing the loading rate r:

$$k_{FU}(f) = -r\frac{\frac{dP_F(f)}{df}}{P_F(f)} \; ; \; k_{UF}(f) = r\frac{\frac{dP_U(f)}{df}}{P_U(f)} \tag{7}$$

To deal with all the amount of data we implemented a *Python* program that was able to extract only the useful information (i.e., the time, position, force components and status of the protocol), cut the whole data series into folding or unfolding trajectories, apply a pre-processing (i.e., applying a median filter and aligning them) and extracting the corresponding sets of folding and unfolding forces.

III. RESULTS AND DISCUSSION

In *Fig. 2(A)* an example of the FDCs obtained is shown for the different PEG concentrations used. It is observed a slight increase on the force at which the hairpin folds as the PEG concentration increases. This increase can also be observed on the unfolding forces. Therefore, the average force seems to increase as the PEG concentration does. It is possible to determine the force at which the rip is observed for the first time, which is called first-rupture unfolding (folding) force for each unfolding (folding) trajectories. Using these forces we can compute the survival probability curves as it was explained in the previous section. These curves are represented in *Fig. 2(B)*, where it can be seen a displacement to the right (i.e., at higher forces) for higher concentrations of PEG. Thus, the addition of macromolecules in our media implies that higher forces are needed to both, fold and unfold, the DNA hairpin. However, they seem to keep its shape reasonably similar between them, which can suggest that the macrocrowding may be affecting the thermodynamics, but not the kinetics of the DNA.

The immediate consequence from this displacement is that these curves also cross at higher values. By taking as a first instance assumption that we are near to the equilibrium, we can consider the crossing points as the coexistence force. Hence, we can assume from our data that the f_c of the hairpin also suffers a change with the addition of PEG, tending to increase as the concentration of crowder does. This relation can be observed at *Fig. 2(C)*. Nevertheless, with this reduced set of points, we cannot infer which kind of relation (linear, exponential...) present the magnitudes.

Assuming that PEG does not bind to the hairpin we can use the Bell-Evans model to study the kinetic of our system. Using Eq.(7) we can obtain the kinetic rates (*Fig. 2(D)*). We find again a displacement to higher forces of both folding and unfolding transitions due to the increase of PEG concentration. Thus, the crossing points (f_c) occur at higher forces, which is consistent with the described before.

From Eq.(2) we could expect a linear distribution of the kinetic rates in the graph, as it is represented logarithmically. However, one can notice that these are certainly far from ideal linearity. In fact, this shows one of the limitations of the Bell-Evans model, which is only satisfied at the vicinity of the coexistence force.

Then, we can define $\Delta f = f - f_c$ and under Eq.(5)

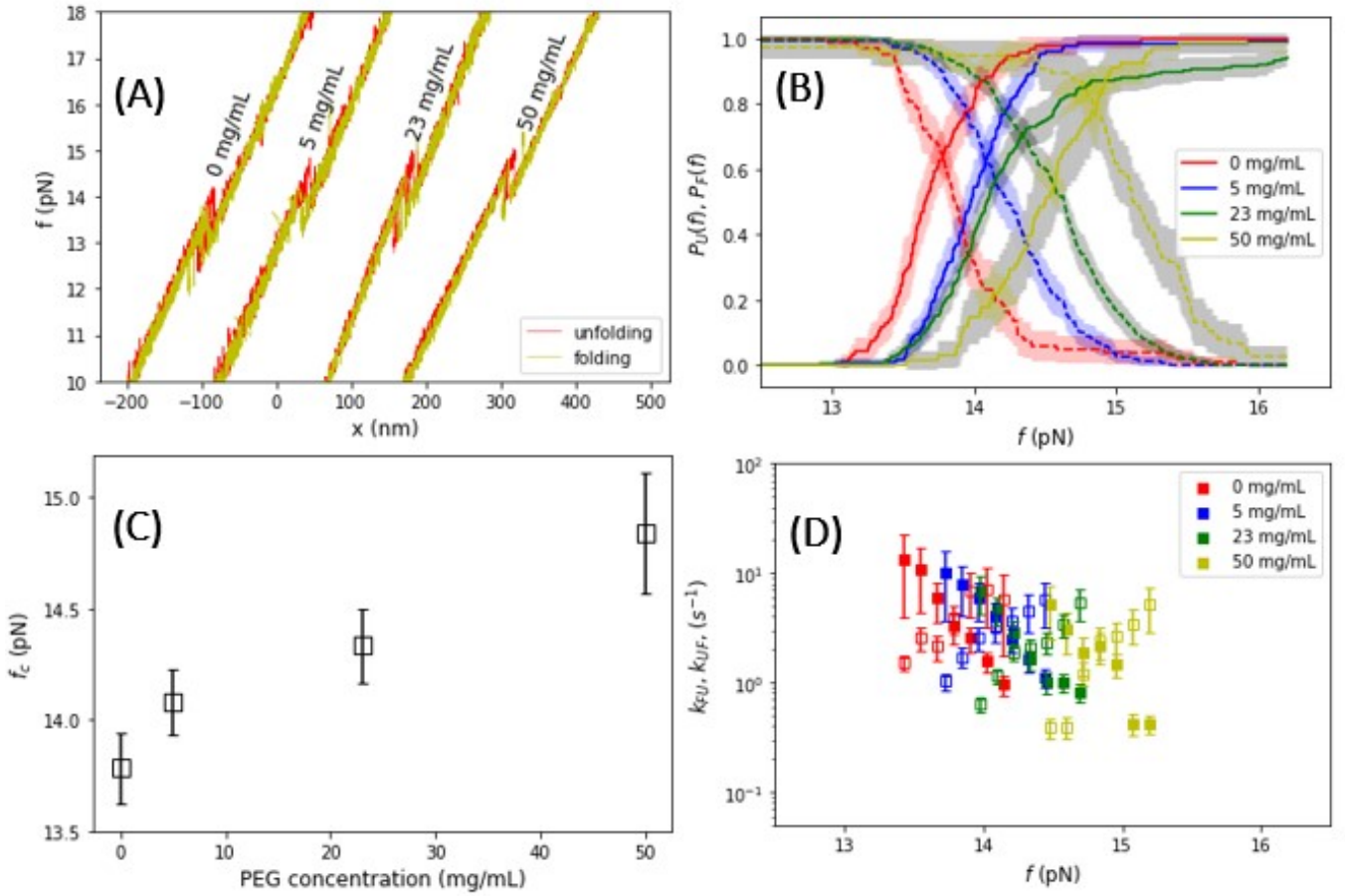

FIG. 2: (A) Example of one unfolding and folding force-distance curve (FDC) for each PEG concentrations studied. (B) Survival probabilities of the folded (dashed) and unfolded (solid) states for each concentration. The shadow shows the associated error computed by bootstrapping. (C) Experimental coexistence force values in function of the PEG concentration. (D) Kinetic rates of the unfolding (solid) and folding (empty) in function of the force.

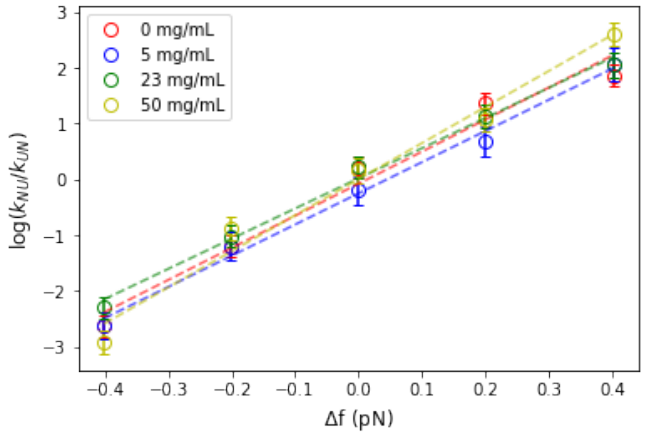

FIG. 3: Logarithmic ratio of the kinetic rates at the vicinity of the coexistance force.

we can represent again our data, which is presented in *Fig. 3*. Here, we are taking only close points (i.e., not far from 0.5 pN) to the coexistence force. Then, we observe linear behaviour of our data.

Finally, we can also study how the difference of free energy between the folded and unfolded states is affected. As a consequence of what we just described, we used the set of points in *Fig. 3* to perform a fit and extract the distance between both states, x_{FU}. In good approximation we can take this distance as a constant for all concentrations, as the slope shows a slight variation, obtaining $x_{FU} = 20 \pm 1\ nm$. Hence, using Eq.(4) we can compute the change experienced by ΔG_{FU}, which is represented in *Fig. 4*.

ΔG_{FU} shows a variation as it raises with the PEG concentrations. Therefore, the increase of the coexistence force is responsible for the increase of ΔG_{FU}. These results suggest that the DNA hairpin becomes more stable in a PEG crowded medium, which is in accordance with the fact we observed at *Fig. 2* that showed how we need higher forces to both fold and unfold the hairpin.

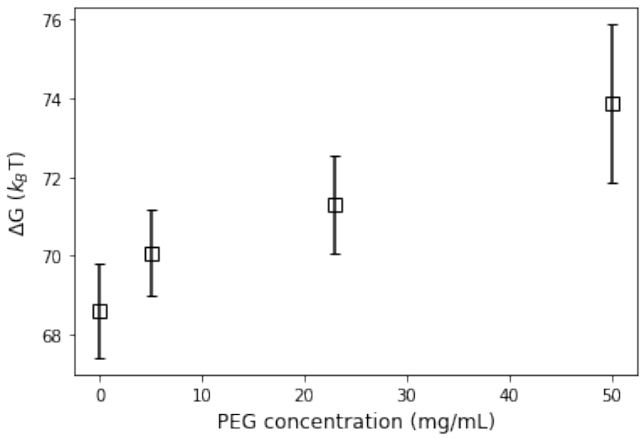

FIG. 4: Free-energy increment for the different considered PEG concentrations.

IV. CONCLUSIONS

To sum up, the results obtained suggest that the concentration of PEG might have an effect on the thermodynamics of the molecule, while the kinetics seems to remain almost invariable. In that sense, the increase of PEG concentration can lead an increase of the f_c of the hairpin and the ΔG_{FU}. This may suggest that our hairpin, under the crowded studied conditions, is a more stable structure than in aqueous solution.

However, it is clear that due to the time limitation, the amount of data collected during this work is very reduced compared to what a full biophysics study demands in order to extract solid conclusions. Consequently, we cannot extend our results to the macrocrowding effect on the DNA folding, generally speaking. Then, further studies need to be performed in order to confirm or not its impact. In the future, more data needs to be collected, either on the already proposed and on new concentration values of PEG. Furthermore, different crowder agents (e.g. PEGs with other M_W, dextrans of different sizes or *ficolls*) and different DNA hairpins with higher lengths should be explored.

In any case, this work has let me introduce into a real physics problem and live a first experience into a real laboratory. Besides, I feel I could integrate some acquired knowledge during the degree, including subjects such as Biophysics, Thermodynamics, Systems out of Equilibrium, Computational Science and other fundamental subjects (e.g., Basics of Laboratory).

Acknowledgments

I want to give my sincere thanks to Prof. Felix Ritort for allowing me to work in the *Small Biosystems Lab* and to Isabel, not only for tutoring this work, but also for making me feel really comfortable in the lab and encourage me to persist even in the most stressful moments. I have learnt many things from you during these months. Besides, I do not want to forget the other members of the Lab (Álex, Jaime, Marc and Paolo), who gave me their unconditional support whenever I needed it. Finally, I would like to express my gratitude to everyone who has accompanied me on this beautiful journey over the past few years.

[1] Krone, M.G. et al. "Effects of Familial Alzheimer's Disease Mutations on the Folding Nucleation of the Amyloid β-Protein" *J.Mol.Biol* 381(1): 221–228 (2008).
[2] Udomprasert, A., Kangsamaksin T. "DNA Origami Applications in Cancer Therapy." *Cancer science* 108(8): 1535–1543 (2017).
[3] Ritort, F. "Single-Molecule Experiments in Biological Physics: Methods and Applications." *J. Condens. Matter Phys.* 18(32): R531–R583 (2006).
[4] Alberts B. et al. "Molecular Biology of the Cell", (Garland Science, New York 2002, 4th ed.).
[5] Balcells C. et al. "Macromolecular Crowding upon In-vivo-like Enzyme-kinetics: Effect of Enzyme-obstacle Size Ratio." *Front. Chem.* 24(1): 3– (2015).
[6] Kuznetsova, I.M., Turoverov K., Uversky V. "What Macromolecular Crowding Can Do to a Protein." *Int. J. Mol. Sci.* 15(12): 23090–23140 (2014).
[7] Murphy, L.D., Zimmerman S.B. "Condensation and Cohesion of Lambda DNA in Cell Extracts and Other Media: Implications for the Structure and Function of DNA in Prokaryotes." *Biophys. Chem* 57 (1): 71–92 (1995).
[8] Camuñas-Soler J. "Force-Spectroscopy of Small Ligands binding to nucleic acids". Tesi doctoral. Dipòsit UB (2015).
[9] Alemany i Arias A. "Dynamic Force Spectroscopy and Folding Kinetics in Molecular Systems". Tesi doctoral. Dipòsit UB (2014).
[10] Bell G. I. "Models for the specific adhesion of cells to cells". Science 200: 618-627 (1978).