Fluctuation Relations applied to characterize heterogeneous molecular ensembles

Alvaro Martinez-Monge,[1] Anna Alemany,[1] and Felix Ritort[1]

[1]*Small Biosystems Lab, Departament de Física Fonamental, Universitat de Barcelona, C/ Martí i Franquès 1, 08028 Barcelona, Spain*
e-mail address: fritort@gmail.com

The identification and characterization of heterogeneous populations of biomolecular systems (e.g. cells, proteins, nucleic acids) has aroused a great interest in molecular biology. For instance, the state of the art in cancer research are the targeted and personalized therapies. Such therapies require to identify and characterize with extremely high accuracy the healthy cells and the cancer cells. Thus, it is clear the importance of understanding and characterization of heterogeneous molecular systems.

Thanks to the recent advances in single-molecule experiments[1] we can manipulate with high precision and sensitivity individual molecules, providing us a formidable tool to study and characterize an ensemble of different molecules at the nanoscopic scale.

An essential step in this project is to answer the following question:

Is it possible to discriminate the folding free energies of different structures of a mutational ensemble of DNA hairpins?

To answer this question, we have taken a standard and well-known canonical DNA hairpin, such as CD4[2], which consists on a stem of 20 base-pairs ending in a tetraloop (see figure 1-A), and we randomized 4 base-pairs (blue circles in figure 1-A) so that several populations of hairpins are obtained. Any of the four nitrogenous bases presents in DNA (adenine, guanine, thymine and cytosine) may be found in each mutated location, so we have $4^8 = 2^{16}$ different populations.

This purpose cannot be done in bulk measurements where only an average behavior of the system is measured. Thus, the context of single-molecule experiments is the perfect scenario to characterize a sample composed by mutated DNA hairpins.

To this end, we carry out pulling experiments using a miniaturized dual-beam laser optical tweezers apparatus[3] that allows us to stretch an individual molecular construct (i.e. the DNA hairpin). This device is able to measure the applied force and the end to end distance (distance λ, see figure 1-B). The thermodynamical properties of the molecule under study can be inferred from the resulting force-distance curve. For example, we can extract the free energy difference between the folded and unfolded state of the DNA hairpins using Fluctuation Theorems.

In particular, Crooks Fluctuation Theorem[4] allows us to extract free energy differences of formation of native molecular structures from bidirectional pulling experiments (that is, combining work measurements along unfolding and refolding paths):

$$\frac{P_F(W)}{P_R(-W)} = e^{\beta(W-\Delta G)} \tag{1}$$

Where W is the work exerted on the system that can be calculated using[5]:

$$W = \int_{\lambda_i}^{\lambda_f} f d\lambda \tag{2}$$

$\Delta G = G(\lambda_f) - G(\lambda_i)$ is the free energy between the folded and unfolded state and $P_F(W)$, $P_R(-W)$ are the work distributions along the forward (F, unfolding) and the reverse (R, refolding) processes, respectively and $\beta = 1/k_B T$.

In consequence, this method allows us to characterize the energy spectrum of the mutant ensemble (see figure 1-C). In consequence, we can say that it is possible to discriminate between the elements of an ensemble using the free energy of formation at zero force.

Besides, we have find out that the global work distributions of the mutational ensemble satisfy the Crooks Fluctuation Relation with an effective temperature higher than the actual temperature (see figure 1-D). In this work we have developed a theoretical model that, combined with the previous result, make possible to quantify and characterize the degree of heterogeneity of a mutant population of nucleic acids.

This work is the previous step that makes clear the fact that we can study evolution at a molecular scale using single-molecule techniques.

Since 1859, when Charles Darwin founded the grounds of evolutionary biology in his work "The Origin of Species"[6], the study of evolution has aroused a great interest within all scientific disciplines. Darwin explained for the first time how species evolve in time through mutations and selective amplifications of the fittest. In other words, the main features of evolution are variation and selection.

On the other hand, we know that evolution does not only takes place in macroscopic organisms. In fact, evolution can be regarded as a process where environmental information is encoded into the genetic code of organisms[7]. Therefore, variability (i.e. heterogeneity) and selection are also present at molecular scale.

Thus, combining our experimental and theoretical framework with directed evolution methods we can make feasible to study molecular evolution.

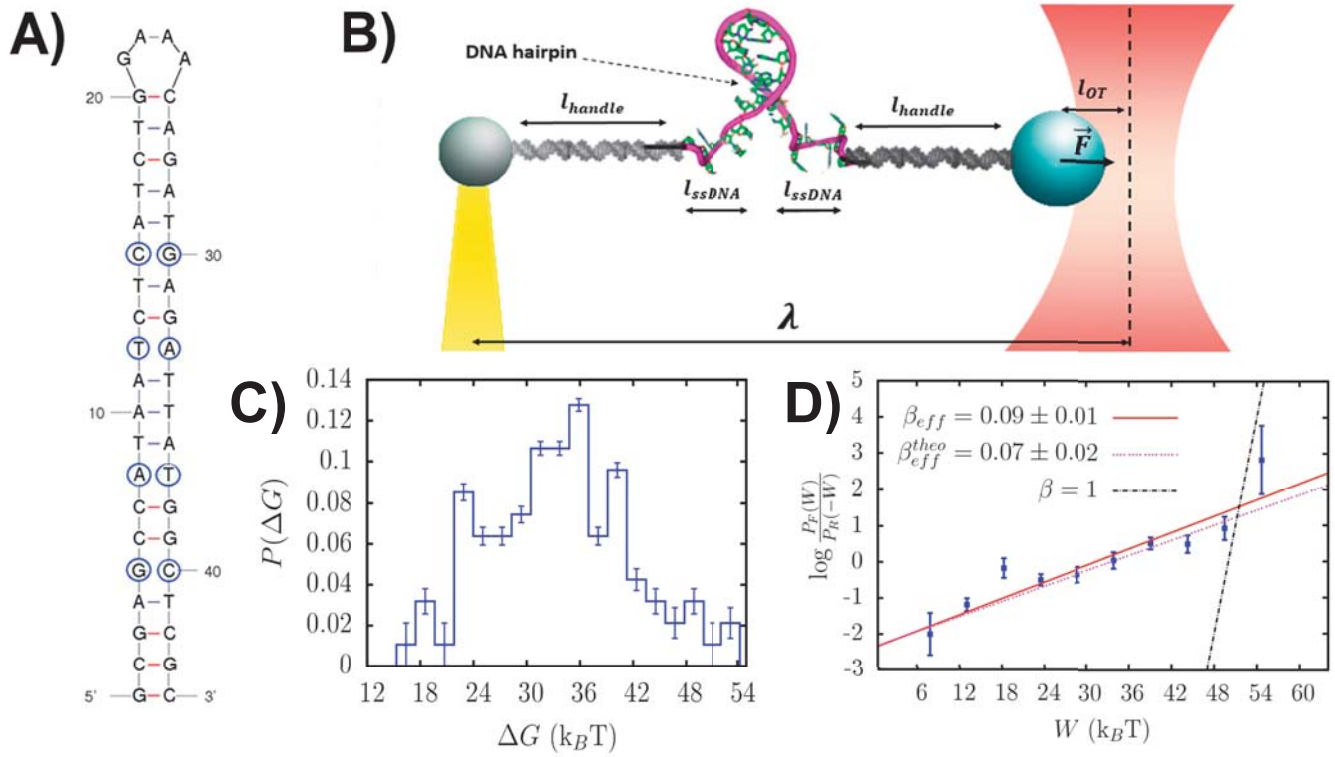

FIG. 1. **DNA hairpins, experimental setup and folding free energy distribution A)** Sequence and structure of cd4. Locations randomized are indicated in circles. **B)** Experimental setup. **C)** Free energy distribution at zero force in k_BT units. **D)** Logarithm of the ratio of the global work distributions, linear fit (solid straight line) and comparison to theoretical model (dotted line). Comparison with actual temperature (dashed line).

[1] Ritort, Felix. Journal of Physics: Condensed Matter **18.32**: R531 (2006)

[2] Manosas, Maria, et al. Physical review letters **96.21**: 218301 (2006)

[3] Huguet, Josep M., et al. Proceedings of the National Academy of Sciences **107.35**: 15431-15436 (2010)

[4] Crooks, Gavin E. Physical Review E **60.3**: 2721. (1999)

[5] Mossa, Alessandro, et al. The Journal of chemical physics **130.23**: 234116. (2009)

[6] Darwin, Charles. *The origin of species.* (GoodBook Classics, London, 1951)

[7] Smith, John Maynard. Philosophy of Science: 177-194. (2000)