Determination of the elastic properties of short ssDNA molecules by mechanically folding and unfolding DNA hairpins

SUPPLEMENTARY INFORMATION

A. Alemany[1], F. Ritort[1,2]

[1] Small Biosystems Lab, Departament de Física Fonamental, Universitat de Barcelona, C/ Martí i Franquès 1, 08028 Barcelona, Spain

[2] CIBER-BBN de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, C/ Sinesio Delgado 4, 28029 Madrid, Spain

Correspondence to: Felix Ritort; e-mail: fritort@gmail.com

S1 Derivation of equation (5)

We define the change in force Δf as f_F-f_U, being f_F (f_U) the force in the folded (unfolded) branch just before (after) the unfolding or folding transition. By dividing Δf by k_{eff}^F we get:

$$\frac{f_F - f_U}{k_{eff}^F} = \frac{\Delta f}{k_{eff}^F} \tag{S1}$$

$$= \frac{\Delta f}{k_h} + \frac{\Delta f}{k_b} + \frac{\Delta f}{k_d} \tag{S2}$$

$$= \Delta x_h + \Delta x_b + \Delta x_d \tag{S3}$$

$$= (x_h(f_F) - x_h(f_U)) + (x_b(f_F) - x_b(f_U)) + (x_d(f_F) - x_d(f_U)) \tag{S4}$$

$$= (x_h(f_F) + x_b(f_F) + x_d(f_F)) - (x_h(f_U) + x_b(f_U) + x_d(f_U)) \tag{S5}$$

In equation (S2) we replaced k_{eff}^F by the effective rigidity of the three serially connected springs present in the folded branch, given by the handles k_h, the bead in the optical trap k_b, and the diameter of the hairpin k_d. In equation (S3) we used the definition $k_i = \Delta f / \Delta x_i$ (i=h,b,d).

At the unfolding and folding transition $\Delta\lambda$=0. Hence:

$$x_h(f_F) + x_b(f_F) + x_d(f_F) = x_h(f_U) + x_b(f_U) + x_{ssDNA}(f_U) \tag{S6}$$

Combining equations (S5) and (S6) we obtain:

$$\frac{f_F - f_U}{k_{eff}^F} = (x_h(f_U) + x_b(f_U) + x_{ssDNA}(f_U)) - (x_h(f_U) + x_b(f_U) + x_d(f_U)) \tag{S7}$$

$$= x_{ssDNA}(f_U) - x_d(f_U) \tag{S8}$$

from which it is straightforward to obtain equation (5) in the main paper.

S2 Derivation of the elastic behavior of the hairpin double helix

The contribution of the hairpin double helix diameter at force is modeled as a single bond of length d (hereafter taken equal to the helix diameter, $d=2$ nm) that is oriented due to the action of a force [1,2]. The energetic contribution that describes this effect is given by $E=-fd\cos\theta$, being $\theta\in[0,\pi]$ the relative angle between the bond and the force axes. The corresponding partition function is:

$$Z \propto \int_{-1}^{1} d(\cos\theta) e^{fd\cos\theta/k_BT} = \frac{k_BT}{fd}\sinh\left(\frac{fd}{k_BT}\right) \tag{S9}$$

where k_B is the Boltzmann constant and T is the absolute temperature taken equal to 298 K. Hence, the average distance x_d projected along the force axis is:

$$x_d(f) = \langle d\cos\theta \rangle = k_BT\frac{\partial \log Z}{\partial f} = d\left[\coth\left(\frac{fd}{k_BT}\right) - \frac{k_BT}{fd}\right] \tag{S10}$$

which is equal to the freely-jointed chain model describing the elastic behavior of a polymer with identical Kuhn and contour lengths.

S3 Derivation of equation (7)

From equation (2) in the main paper we can write:

$$\frac{1}{k_{eff}^{F}(f)} = \frac{1}{k_h(f)} + \frac{1}{k_b(f)} + \frac{1}{k_d(f)}, \quad \frac{1}{k_{eff}^{U}(f)} = \frac{1}{k_h(f)} + \frac{1}{k_b(f)} + \frac{1}{k_{ssDNA}(f)} \tag{S11}$$

and by subtracting the two equations one directly gets:

$$\frac{1}{k_{eff}^{U}(f)} - \frac{1}{k_{eff}^{F}(f)} = \left(\frac{1}{k_h(f)} + \frac{1}{k_b(f)} + \frac{1}{k_{ssDNA}(f)}\right) - \left(\frac{1}{k_h(f)} + \frac{1}{k_b(f)} + \frac{1}{k_d(f)}\right) \tag{S12}$$

$$= \frac{1}{k_{ssDNA}(f)} - \frac{1}{k_d(f)} \tag{S13}$$

It is remarkable that no assumption on the force dependence of k_b and k_h is required to obtain equation (S13).

S3 Simulation of pulling experiments and data analysis

In order to simulate pulling experiments we proceed as described in reference 3. For each DNA hairpin (L06, L12, L16 and L20) four different models were simulated:

Model 1: **ssDNA elastic properties based on the WLC model (ref. 4) and inclusion of the helix diameter.** The elastic behavior of ssDNA is modeled according to the WLC using the interpolation formula derived in reference 4. The persistence length was set to $P=1.3$ nm and the interphosphate distance to $d_b=0.59$ nm/base. The diameter of the double helix, present when the DNA hairpin is folded, is modeled as a single bond of length $d=2$ nm that aligns with the applied force according to equation (S10).

Model 2: **ssDNA elastic properties based on the WLC model (ref. 4) and exclusion of the helix diameter.** The elastic behavior of ssDNA is modeled according to the WLC using the interpolation formula derived in reference 4. The persistence length was set to $P=1.3$ nm and the interphosphate

distance to $d_b=0.59$ nm/base. The contribution of the DNA hairpin diameter to the molecular extension and stiffness in the folded state is neglected.

Model 3: **ssDNA elastic properties based on the WLC model (ref. 5) and inclusion of the helix diameter.** The elastic behavior of ssDNA is modeled according to the WLC using the interpolation formula derived in reference 5. The persistence length was set to $P=1.3$ nm and the interphosphate distance to $d_b=0.59$ nm/base. The diameter of the double helix, present when the DNA hairpin is folded, is modeled as a single bond of length $d=2$ nm that aligns with the applied force according to equation (S10).

Model 4: **ssDNA elastic properties based on the WLC model (ref. 5) and exclusion of the helix diameter.** The elastic behavior of ssDNA is modeled according to the WLC using the interpolation formula derived in reference 5. The persistence length was set to $P=1.3$ nm and the interphosphate distance to $d_b=0.59$ nm/base. The contribution of the DNA hairpin diameter to the molecular extension and stiffness in the folded state is neglected.

Each model is then simulated for each of the four hairpins (L06, L12, L16 and L20). In all cases, 100 unfolding and folding FDC were simulated. In Figure S1 we show FDC obtained from pulling experiments and from simulation according to model 1 for the four hairpins. Data for simulations performed according to models 2, 3 and 4 is not shown, but results are practically the same. The obtained FDC are analyzed *a posteriori* using the four different models according to the two methods presented in the main paper.

First, the jump in force upon unfolding and folding is extracted, and using equation (S7) the elastic response $x_{ssDNA}(f)$ is obtained. Simulated data for the four different molecules is merged together in a single $x_{ssDNA}(f)/L_c$-plot as described in the main paper, and results are fitted to the WLC using both interpolating formulas proposed in references 4 and 5, taking into account or neglecting the presence of the hairpin (models 1-4). In this way, each set of simulations is analyzed according to the four different models presented above. Results obtained for the four simulations analyzed within the assumptions of each model are summarized in Table S1. Gray cells correspond to the case where the simulation and subsequent analysis use the same model. It can be seen that values of P and d_b recovered from the data analysis are in agreement, within error bars, with values used to simulate pulling experiments.

Second, the force-dependent stiffnesses of the folded and unfolded branches are extracted from the simulated FDC. Using equation (S13) the stiffness of ssDNA is then obtained. Again, data is analyzed using the WLC with both interpolating formulas proposed in references 4 and 5 to model ssDNA, and taking into account or neglecting the presence of the hairpin. Results obtained for the four simulations are summarized in Table S2. It can be seen that values of P and d_b recovered from the data analysis are in agreement, within error bars, with values used to simulate pulling experiments.

These simulations are useful for two reasons. First, they demonstrate the consistency of our data analysis methods. Second, they confirm one of our conclusions: neglecting the presence of the double helix diameter when the hairpin is folded persistently leads to shorter values of d_b, whereas values for P then to be larger (Tables S1 and S2). Despite of these small differences, all results are compatible within error bars.

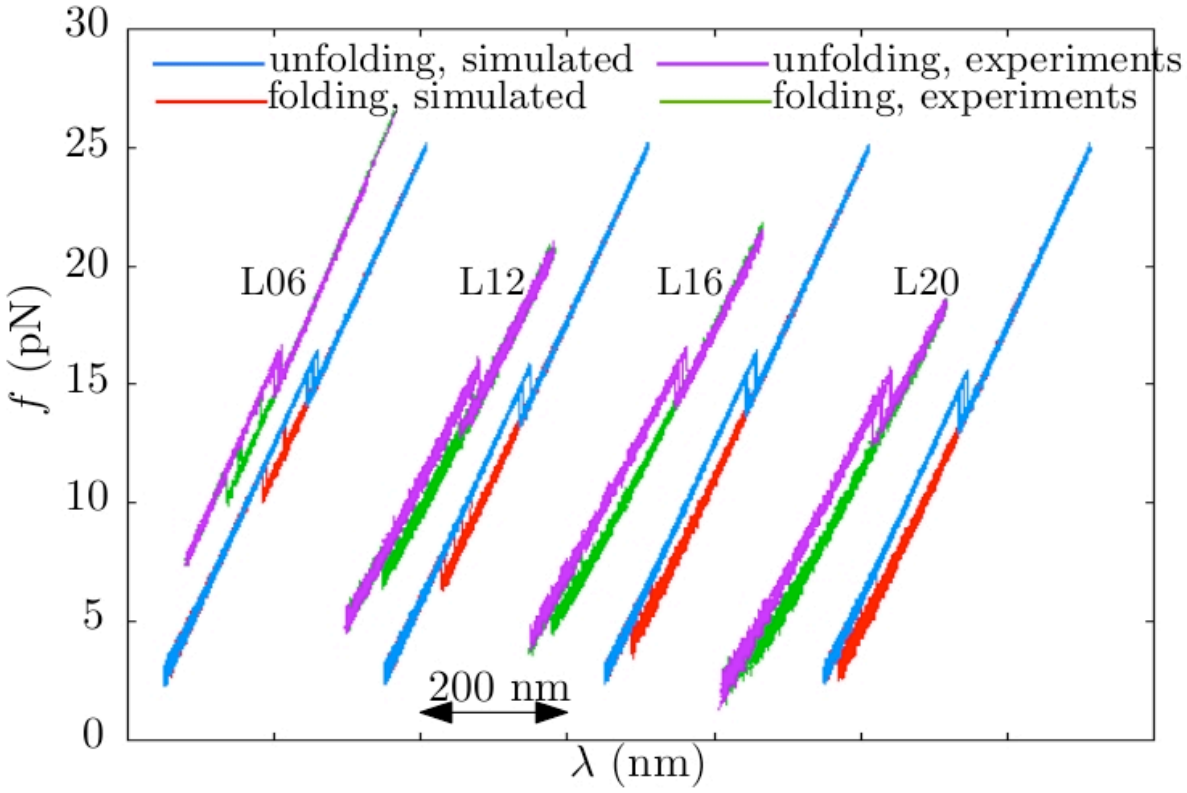

Figure S1. Comparison between FDC obtained from pulling experiments and from simulation according to model 1 for the four DNA hairpins L06, L12, L16 and L20.

Simulations with Model 1	P (nm)	d_b (nm/base)
Analysis with Model 1	1.3±0.1	0.60±0.02
Model 2	1.3±0.1	0.56±0.02
Model 3	1.1±0.1	0.60±0.02
Model 4	1.1±0.1	0.57±0.02
Simulations with Model 2	P (nm)	d_b (nm/base)
Analysis with Model 1	1.2±0.1	0.66±0.02
Model 2	1.3±0.1	0.60±0.02
Model 3	1.1±0.1	0.66±0.02
Model 4	1.1±0.1	0.61±0.02
Simulations with Model 3	P (nm)	d_b (nm/base)
Analysis with Model 1	1.4±0.1	0.60±0.02
Model 2	1.5±0.1	0.55±0.02
Model 3	1.2±0.1	0.60±0.02
Model 4	1.3±0.1	0.56±0.02
Simulations with Model 4	P (nm)	d_b (nm/base)
Analysis with Model 1	1.4±0.1	0.65±0.02
Model 2	1.4±0.1	0.60±0.02
Model 4	1.2±0.1	0.65±0.02
Model 5	1.3±0.1	0.61±0.02

Table S1. **Force-jump analysis of simulations performed according the the four different models.** Persistence length P and interphosphate distance d_b obtained by analyzing data simulated according to models 1-4 using the force-jump measurement and models 1-4. Gray cells correspond to the case where simulation and subsequent analysis use the same model. It can be observed that parameters of the simulation (P=1.3 nm, d_b =0.59 nm/base) are recovered within error bars.

Simulations with Model 1	P (nm)	d_b (nm/base)
Analysis with Model 1	1.3±0.1	0.59±0.02
Model 2	1.2±0.1	0.51±0.02
Model 3	0.8±0.1	0.50±0.02
Model 4	0.9±0.1	0.48±0.02
Simulations with Model 2	P (nm)	d_b (nm/base)
Analysis with Model 1	1.3±0.1	0.60±0.02
Model 2	1.4±0.1	0.60±0.02
Model 3	1.0±0.1	0.58±0.02
Model 4	1.0±0.1	0.54±0.02
Simulations with Model 3	P (nm)	d_b (nm/base)
Analysis with Model 1	1.5±0.1	0.56±0.02
Model 2	1.5±0.1	0.50±0.02
Model 3	1.4±0.1	0.61±0.02
Model 4	1.5±0.1	0.61±0.02
Simulations with Model 4	P (nm)	d_b (nm/base)
Analysis with Model 1	1.5±0.1	0.61±0.02
Model 2	1.7±0.1	0.60±0.02
Model 4	1.3±0.1	0.62±0.02
Model 5	1.4±0.1	0.61±0.02

Table S2. **Stiffness analysis of simulations performed according the the four different models.** Persistence length P and interphosphate distance d_b obtained by analyzing data simulated according to models 1-4 using the measurement of the stiffness and models 1-4. Gray cells correspond to the case where simulation and subsequent analysis use the same model. It can be observed that parameters of the simulation (P=1.3 nm, d_b=0.59 nm/base) are recovered within error bars.

References

1. Woodside, M. T.; Anthony, P. C.; Behnke-Parks, W. M.; Larizadeh, K.; Herschlag, D.;

Block, S. M. Science 2006, 314, 1001–1004.

2. Forns, N.; de Lorenzo, S.; Manosas, M.; Hayashi, K.; Huguet, J. M.; Ritort, F. Biophys. J. 2011, 100, 1765–1774.

3. Manosas, M.; Ritort, F. Biophys. J. 2005, 88,3224-3242.

4. Bustamante, C.; Marko, J. F.; Siggia, E. D.; Smith, S. Science 1994, 265, 1599–1601.

5. Bouchiat, C.; Wang, M.; Allemand, J.-F.; Strick, T.; Block, S.; Croquette, V. Biophys. J. 1999, 76, 409–413.