Eur. Phys. J. E **16**, 235–246 (2005)
DOI: 10.1140/epje/e2005-00026-9

THE EUROPEAN PHYSICAL JOURNAL E

Model for DNA hairpin denaturation

S. Cuesta-López[1,2,3], M. Peyrard[3,a], and D.J. Graham[3,4]

[1] Dept. Física de la Materia Condensada, Universidad de Zaragoza, c/Pedro Cerbuna s/n, 50009 Zaragoza, Spain
[2] Instituto de Biocomputación y Física de Sistemas Complejos, Universidad de Zaragoza, Zaragoza, Spain
[3] Laboratoire de Physique, École Normale Supérieure de Lyon, 46 allée d'Italie 69364 Lyon Cedex 07, France
[4] Department of Physics, Cornell University, Ithaca, NY 14853, USA

Received 26 July 2004
Published online 22 February 2005 –

Abstract. We investigate the thermal denaturation of DNA hairpins using molecular dynamics simulations of a simple model describing the molecule at a scale of a nucleotide. The model allows us to analyze the different interacting features that determine how an hairpin opens, such as the role of the loop and the properties intrinsic to the stem.

PACS. 87.15.Aa Nonlinear dynamics and nonlinear dynamical systems – 87.15.He Dynamics and conformational changes – 05.45.-a Nonlinear dynamics and nonlinear dynamical systems

1 Introduction

Understanding the physics of DNA is a challenge, and simple models of the molecule can help because they allow us to determine what are the essential features that are necessary to generate its remarkable properties. Due to its large amplitude motions, which are involved in many biological processes, DNA is a laboratory for nonlinear physics [1] and the models can also be used to analyze some of its biological properties [2]. However, such studies are only useful if the models can be validated by comparison with experiments, which requires a well-controlled experimental situation where precise results can be obtained. DNA melting, i.e. its thermal denaturation by separation of the two strands, is a purely physical process that can be accurately tracked by UV absorbance, and this is why it has been used to calibrate the parameters of a simple nonlinear model for DNA denaturation [3]. More recently, a new class of experiments has started to deliver very accurate results on DNA: in particular, we refer to studies of DNA hairpins that possess, at their ends, a fluorophore and a quencher [4]. These molecules are made of a single strand of DNA which carries sequences of bases that are complementary to each other in each of its two terminal regions. As a result, when the base pairs of these two sequences are formed, the molecule takes the shape of a hairpin (see a schematic picture in Fig. 1) made of a stem, which is a short segment of a DNA double helix, and a loop, which is a single strand carrying bases that are not paired. Such molecules exist in two states. Above a temperature T_m, called the "melting temperature", the double helix of the stem is denaturated and the molecule behaves like a standard polymer chain, which generally has its two ends far from each other. As a result, the fluorophore and the quencher, each bound to one end of the strand, are far from each other and the molecule is fluorescent. At temperatures significantly lower than T_m, the stem is formed. The two ends of the hairpin are close to each other so that the quencher prevents the fluorescence of the chromophore. Therefore, the fluctuations of the fluorescence, recorded with confocal microscopy, can be used to investigate the structural dynamics of these peculiar DNA samples [4].

The interest of these "molecular beacons" is double: first, as discussed above, they can be used to provide data on DNA denaturation and self-assembly; and second, they can be used as highly sensitive probes to detect some short sequences with a high specificity [5]. While the experiments can provide accurate measurements on the melting temperature T_m as a function of some properties of the beacon, such as the length of the loop or its rigidity (rigidity is determined by the bases in the loop), the understanding of the experimental results would be significantly improved if we had a precise idea of all the physical processes involved in the fluctuations of the hairpin.

Our aim in this study is to introduce a simple model that allows us to identify the basic phenomena involved in the statistical and dynamical properties of DNA hairpins. We use molecular dynamics simulations to investigate its properties. As the typical time scales of the fluctuations of DNA hairpins vary from ns to μs an all-atom simulation is extremely computationally demanding. To our knowledge, the only simulation of a hairpin which was able to follow a trajectory long enough to observe opening-closing events was a massively parallel stochastic simulation using 40 000 processors [6]. Very simple models that nonetheless preserve the basic physics of the system are an

[a] e-mail: michel.peyrard@ens-lyon.fr

essential complement to these studies, to allow statistical studies and provide a fundamental understanding of the basic phenomena governing hairpin fluctuations. We consider a generalization of a simple model for DNA denaturation [7–9], which turned out to allow some quantitative analysis of DNA melting curves [3] and biological processes [2]. For the study of DNA hairpins the simple model has to be significantly modified. Although it cannot pretend to give a full quantitative picture of the fluctuations of the hairpin, such a model can be useful to understand the processes that lead to the experimental observations because, although they are performed at the scale of a single molecule, experiments that detect the quenching of fluorescence are not sufficient to precise the microscopic structure of the molecule. Comparison with the experiments can also validate the model and tell us what are the appropriate values of its parameters, opening the way to other applications of this model in different contexts.

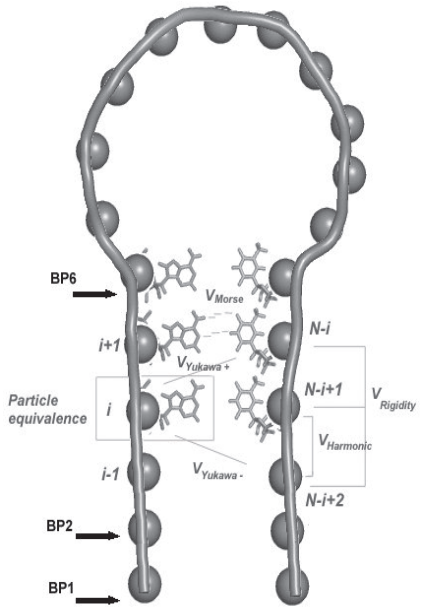


Fig. 1. Schematic view of the model and its interaction potentials. The model has been drawn using Pymol by DeLano Scientific [10].

2 Description of the model

Our model, schematically shown in Figure 1, is based on the PBD model [7,8], which has been used by several authors to study denaturation because it is very simple but nevertheless gives results that can be quantitatively compared with experiments [2,3]. However, in order to model the denaturation of a hairpin, the PBD model must be significantly extended.

The original PBD model does not care about the geometry of the DNA strands because it only describes the base pair opening through the relative distance between the two bases of a pair. The molecular structure only enters through a stacking interaction between adjacent base pairs, which depends on the stretching of the two interacting base pairs, not on the actual position of the bases; the actual positions are irrelevant for the analysis of the denaturation.

From the PBD model we keep the basic unit at the level of a base, or more precisely a nucleotide, which comprises a base, a sugar and a phosphate group, represented as an entity with a mass $m = 300$ a.m.u. (atomic mass unit). To study a hairpin we need to specify the actual geometry of the strand. Therefore, the strand is described as a chain of particles, the nucleotides, which are coupled by a stiff harmonic potential allowing small fluctuations around the average interparticle distance r_0, equal to 6 Å,

$$V_h = K_s \left[|\mathbf{r}_{i+1} - \mathbf{r}_i| - r_0\right]^2 \tag{1}$$

where $\mathbf{r}_i$ is the spatial position of particle i, and K_s is the coupling constant set to 0.18 eV Å^{-2}. In order to speed up the calculations, the dynamics is assumed to be two-dimensional and each nucleotide therefore has two degrees of freedom in a plane. This could appear as a severe restriction because it does not allow for the description of the helicoidal structure of the stem. For the denaturation of long DNA molecules, the helicoidal shape is indeed important [11] because a local opening, which eliminates the torsion in some region of the helix, transfers this torsion to other parts, thereby affecting the opening rates in these parts. For a stem of a few bases such an effect is expected to be much weaker for two reasons: (i) the local opening can at most extend over a few bases, so that the torsion to be transfered on other sites is small; (ii) the free end of the stem allows it to rotate, relaxing the extra torsion. Therefore it is neglected in our model, which allows us preserve the simplicity of the model which, as explained in the introduction, does not pretend to be fully quantitative.

Short DNA structures have been found to show highly sequence-dependent conformational dynamics where the flexibility associated with the base stacking plays an important role in the melting transition [4]. This is a point that we would like to understand in this study and therefore this aspect must be included in the model. It can be done by introducing a "rigidity potential" which couples three consecutive nucleotides and is minimum when the angle $\hat{\theta}_i$ between the two vectors joining nucleotides i and $i-1$ and nucleotides i and $i+1$ is equal to π. This potential has been chosen as

$$V_r = K_r \left[1 + \cos\hat{\theta}_i\right]. \tag{2}$$

The rigidity constant K_r plays a very important role because it determines the shape of the loop and strongly influences the deformations of the stem. However, at the mesoscopic scale of the present model, the actual value of K_r is very difficult to determine because what appears as a local bending of the strand occurs in the actual DNA structure through a complicated process. The strand can be bent by changing some angles between the covalent bonds, which has a high energetic cost, but also by a sequence of rotations around the bonds, which requires

much less energy. The relative contributions of these motions to the actual bending of the DNA strand depends on the motion that is performed. In order to understand the physical role played by the rigidity of the strand, we have varied K_r in a very broad range, from 10^{-6} eV to 0.35 eV. As discussed below, the comparison with the experiments leads to an estimate of the appropriate value of K_r, which could not be made a priori.

The two potential energies, (1) and (2), are sufficient to describe the properties of the strand in the loop, which can be viewed as a particular polymer. But for the stem, in which the base are assembled in pairs, we need to take into account additional energy terms. The first one is the potential that links two nucleotides in a pair. From the PBD model, it can be taken as a Morse potential V_M, which has the proper qualitative shape: a strong rise when the two nucleotides are pushed together closer than their equilibrium distance, and a rise followed by a plateau (corresponding to a vanishing force) when the nucleotides are pulled apart from each other,

$$V_M = D_M \left[e^{-\alpha[|\mathbf{r}_i - \mathbf{r}_{N-i+1}| - d_o]} - 1 \right]^2, \qquad (3)$$

where D_M represents the depth of the Morse potential, having a value of 0.22 eV and d_o is the equilibrium bond distance equal to 10 Å. For α, we have chosen a value of 4.45 Å^{-1} in agreement with the usual value for an AT base pair taken in previous studies [3,8,9]. It should be noticed that, like all the potentials in this model, V_M is an effective potential that describes a set of microscopic interactions. It includes the hydrogen bonds between the paired bases, but also the repulsion between the heavily charged phosphate groups in the base/sugar complex, as well as the solvent effects. As already pointed out for K_r, such effective potentials are difficult to calculate a priori but the value of D_M is however much better defined than K_r because the order of magnitude of the energy required to break a DNA base pair is known (and has for instance been evaluated for the PBD model); and moreover, it can be estimated from the comparison with experiments in order to get a correct thermal denaturation temperature for the stem. The Morse potential has the same qualitative shape as the Lennard-Jones potential, more often used in molecular dynamics simulations. Its choice for the PBD model was motivated by the possibility that it offered to carry some calculations analytically, but the studies of the PBD model show that the difference between the Morse and the Lennard-Jones potential is not significant for the statistical results that we are seeking in such a study.

Finally, an additional potential term is necessary to stabilize the geometry of the stem against shear distortion. It can be written as a Yukawa repulsive potential between a nucleotide and its two nearest neighbors of the complementary strand in the stem (see Fig. 1):

$$V_{Yuk} = K_{Yuk} \left[\frac{e^{-C_{Yuk}|\mathbf{r}_i - \mathbf{r}_{N-i}|}}{|\mathbf{r}_i - \mathbf{r}_{N-i}|} \right] + K_{Yuk} \left[\frac{e^{-C_{Yuk}|\mathbf{r}_i - \mathbf{r}_{N-i+2}|}}{|\mathbf{r}_i - \mathbf{r}_{N-i+2}|} \right]. \qquad (4)$$

Such potentials are appropriate to describe the screened-charge coulombic repulsion between the phosphate groups in the nucleotide entity. The values of C_{Yuk} in equation (4) has been set to 0.4 Å^{-1} which is the inverse of a typical Debye length for an ionic solution, while $K_{Yuk} = 50.0$ eV Å have been chosen in order to keep the V_{Yuk} small with respect to the other energies in the system, in particular the rigidity potential (see Fig. 7).

The calculations have been done with a stem having 6 base pairs, labeled $BP1 \ldots BP6$; $BP1$ is the base pair at the free end of the hairpin. $BP6$ is next to the loop: its nucleotides are attached to the two ends of the part of strand that makes the loop, and moreover each of its 2 nucleotides is linked by a Yukawa potential term to the nucleotide at the end of the loop on the opposite side of the hairpin. The model is comparable to the molecular beacons investigated experimentally [4]. To derive equations of motions from these potentials, one must also specify the boundary conditions. The two ends of the strand are free, i.e. the end nucleotides only have interactions with one other nucleotide along the strand, and, whenever base pair $BP1$ is formed, with the nucleotide of the opposite strand through the Morse potential V_M.

Finally, let us conclude this part by saying that although nature is much more complicated than this simplified vision, we have tried to design the model to take into account the main biophysical properties of DNA. In real molecules, loops have been shown to be quite flexible while the stem remains more rigid due to the pairing between complementary bases in a helical conformation [12]. Note that this feature is well represented in our model by the interactions present in the stem. One difficulty in such a modeling approach is the proper choice of the parameters. Fortunately, not all parameters are of equal importance. The precise value of the harmonic coupling constant K_s and the parameters of the Yukawa potential only have a small effect on the thermodynamics of the hairpin because the conformational changes that play a role when the base pairs open are only weakly coupled to the stretching of the strands or to the shear motion of the stem. These two potentials are only required to determine the geometry of the molecule. On the other hand, the depth D_M of the Morse potential has a very strong influence on the denaturation temperature because it measures the energy that has to be spent to open a base pair. This provides a convenient way to evaluate a proper value of D_M because the denaturation temperature is known. Similarly, in agreement with the experiments, the rigidity coefficient K_r is very important. This study will show that this coefficient influences denaturation through different effects. There is one role of K_r that is immediately clear: it controls the cooperativity of the denaturation because when one base moves out of the stacks, it tends to pull out its neighbors due to the rigidity of the strand.

3 Discussion of the simulation method

DNA hairpins fluctuate mainly between two families of states: low enthalpy ones, due to the base pairing in the

stem in the *closed state*; and high entropy states, the *open states*, which have many different conformations. As observed experimentally [4], the closing transition, i.e. the formation of the hairpin, is a long process that evolves through a path of partially folded or *misfolded states* before reaching the final, properly closed state. Closure occurs only after one particular collision of the two arms of the hairpin followed by the nucleation and the propagation of a base-paired stretch [13]. Therefore the closing of the hairpin, which typically requires a few μs, is beyond the time scales of molecular dynamics simulations–even with a simple model–unless exceptional computing facilities are used [6] or if the model is biased to avoid mismatches. Our model has such a bias because the bases of the stem, linked by the Morse potential never lose the memory of the corresponding base in the pair. The force between them may become vanishingly small when the hairpin is open, but only the correct closing is allowed. It would be a limitation to study the dynamics of the fluctuations that open and close the hairpin and compare it with experiments. But, as this dynamics occurs on time scales that are beyond simulation, we focused our investigation on the opening temperature of the hairpins, for which a proper description of the dynamics of the closing is not required. We did occasionally observe closing, but the observation only corresponds to the last part of the actual closing, when the two legs of the hairpin approach each other with a correct matching of the base pairs, followed by a fast zipping of the stem. On the contrary, the melting transition of the hairpin is accessible to simulations with a simple model because it only requires an energy fluctuation large enough to unzip all the base pairs of the stem. Melting occurs in a time scale that is is compatible with simulations. Although the model is not sufficient to allow a complete analysis of the fluctuations of DNA hairpins, its investigation is however important because it allows us to understand one of the two phenomena that determine the fluctuations, the opening event. As shown below, MD simulations clarify for instance the role of the loop in the opening and they exhibit some features that had not been inferred from the experiments.

A study of the thermally-induced denaturation transition requires a precise control of temperature. It can be achieved by coupling the system to a Nosé-Hoover chain of thermostats, which gives canonically distributed positions and momenta [14]. This method, which is a modification of the original Nosé dynamics [15,16], ensures a proper exploration of the phase space for small or stiff systems. Moreover we also noticed in our simulations that it leads to a faster relaxation to equilibrium than a single-thermostat calculation.

The process followed in our study is simple: after an initial heating and equilibration at a temperature well below the opening-transition temperature, we heat up the system step by step, choosing temperature intervals of 5 K. Each temperature variation is performed as a linear ramp lasting 50 ps, followed by 50 ps of equilibration at the new temperature before the MD trajectory is recorded and analyzed during 50 ns. We have checked that this time is sufficient to observe any relevant evolution in the model.

During the simulation carried out at each stabilized temperature, properties like energy and distance between Morse pairs are recorded and analyzed. This allows us to calculate some thermodynamic quantities like the specific heat, and to determine whether the hairpin is denaturated or not.

This determination raises however some technical difficulties because, in a small system such as the hairpin, a true thermodynamic transition between a closed and an open state does not exist. Instead of a sharp denaturation temperature, one expects a smooth evolution between a state in which the fluctuating hairpin is closed most of the time, and a state in which it is open most of the time. This is a typical problem of single-molecule studies, whether they are experimental or numerical. Besides these true opening and closing events, the dynamics of the model can pass through intermediate stages which can easily be wrongly interpreted as opening and closing, such as

- large amplitude fluctuations of a closed state which bring the system near full opening without actually breaking the base pairs may be wrongly analysed as an opening event. Such very large fluctuations are possible with the nonlinear interaction potentials of the nucleotides;
- in an open state, if the loop has a very small stiffness, it tends to take a rather compact globular shape due to entropic effects. This keeps its ends, hence the hairpin legs, close to each other altough they are not paired. This could be wrongly interpreted as a closed state.

In order to avoid this kind of artefacts, we have monitored the pairing of the bases from the distance between the complementary bases of the stem and we assign a state as denaturated when *all* the Morse pairs have been broken *for a period of time greater than 1* ps.

4 Influence of the rigidity on the denaturation temperature

DNA hairpin denaturation temperature has been shown to be highly dependent of the loop's characteristics, like its sequence and its length [4]. The experimentally measured melting temperature T_m of designed structures has been found to decrease as the length of the loop increases. On the other hand, a poly(A) hairpin denaturates at lower temperature than a poly(T) one. This suggests a rigidity dependence because poly(A) can be expected to be more rigid than poly(T). A purine base like A is made of two organic rings instead of one for a pyrimidine base like T. As a result, the stacking interaction between As is larger, reducing the flexibility of the strands. The spread of T_m with loop length is stronger in the case of adenine, where T_m drops from 54.1 C down to 29.8 C for loops from 12 to 30 bases than for thymine, where it varies in the range 58.1 C to 42.2 C.

Fig. 2. Typical snapshots of a fluctuating hairpin, showing different states of a moderately flexible (left) and highly flexible loop (right).

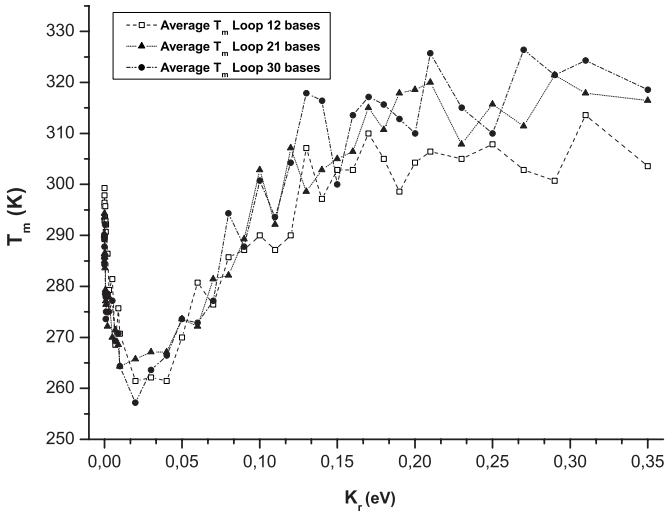


Fig. 3. Evolution of the denaturation temperature T_m in function of the rigidity parameter K_r for hairpins with different loop length.

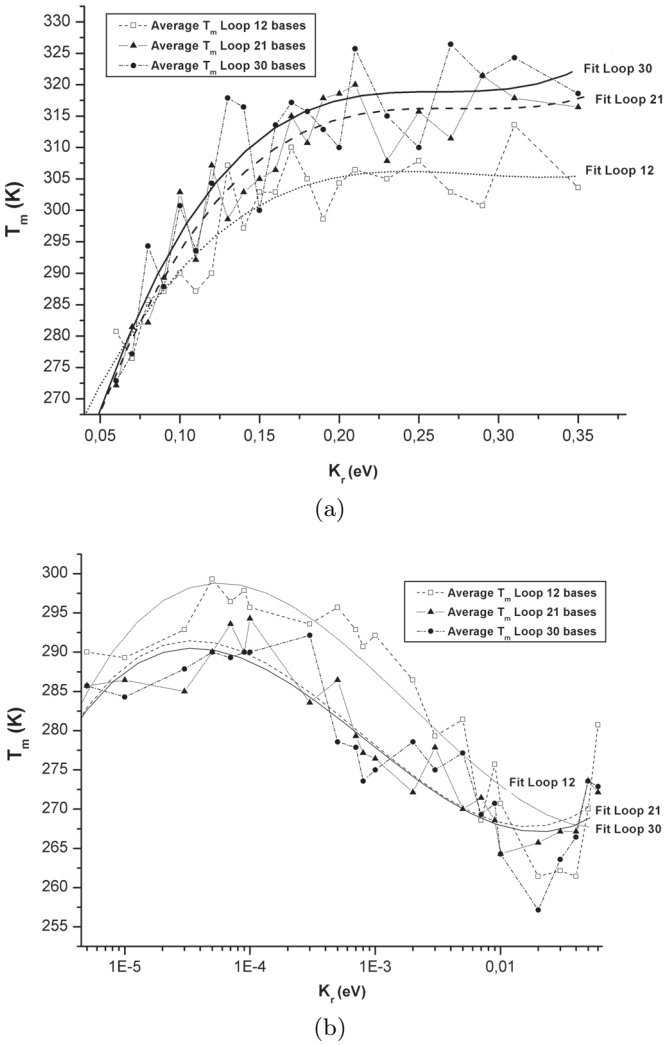


Fig. 4. Detailed view of the rigidity zones explored in our study. (a) shows polynomial fits for the evolution of T_m calculated in the case of high rigidity for different loop sizes. (b) magnifies the flexible zone corresponding to low values of K_r.

In experiments T_m results from the equilibrium between opening and closing, so that the role of the properties of the loop in the opening events is difficult to assess, although kinetic studies can give some data [13]. Numerical simulations of the model can provide detailed results on the opening and its dynamics.

A first series of studies has been made by assuming a homogeneous rigidity along the whole strand. We have estimated the denaturation temperature T_m for different loop lengths and for a wide range of values for the rigidity potential. Figure 2 shows typical aspects of the loop, depending on its rigidity. Figure 3 summarizes the results, which are presented in greater details in Figures 4a and 4b. Due to the difficulty of determining T_m as discussed in the previous section, each point on the figure results from an average of 7 calculations (requiring 25h CPU per point with a 2.0 GHz workstation). Even with such an averaging, large fluctuations subsist in the plot of T_m versus K_r but the general evolution is nevertheless clearly visible.

As shown in the graphs, the variation of the denaturation temperature versus the rigidity of the stem and the length of the loop is complex. When the rigidity increases, T_m exhibits a minimum at intermediate rigidities, which suggests a competition between different processes. At very low rigidity (see Fig. 4b) T_m depends only weakly on K_r. Then, as K_r increases by almost two orders of magnitude (from 10^{-4} to 10^{-2} eV), the denaturation temperature shows a slow decrease of about 20 K until it reaches the minimum located around $K_r = 0.05$ eV. After that, a very sharp rise appears as T_m increases by more than 50 K in a short range of rigidity from 0.05 to 0.15 eV. This rise can be related to the energy needed for a cooperative breaking of Morse pairs in the stem. But, instead of keeping on rising, T_m stabilizes for high rigidity values, which suggests that other phenomena come into play at higher rigidities.

It should also be noted that the influence of the loop size on T_m is itself complex. According to the polynomial fits shown in Figure 4b, for low values of K_r a bigger loop makes the opening transition easier, while for fairly rigid structures the open state appears to be more accessible for short loops (Fig. 4a).

Besides the determination of T_m, simulations also allow us to examine how the denaturation starts and evolves. Three different situations have been observed. High values of K_r, lead to an opening of the base pairs starting from the loop side and propagating quickly along the stem until it reaches the free end. On the contrary, when K_r is small, i.e. when the loop is very flexible, large amplitude motions can very often be seen on the stem in the form of *denaturation bubbles*. These modes, in cooperation with an opening of the base pair at the free end of the stem, lead to a regime where the base pairs that are close to the loop are the last ones to break. Finally, for an intermediate range of stiffness, the denaturation is the result of an opening that starts from both sides, sometimes with a clear predominance of one of the two.

All these data suggest that there is an interplay of several physical phenomena that determines how DNA hairpins denaturate. In the following sections we will try to clarify and quantify their influence through several numerical experiments that are designed to separate the different phenomena.

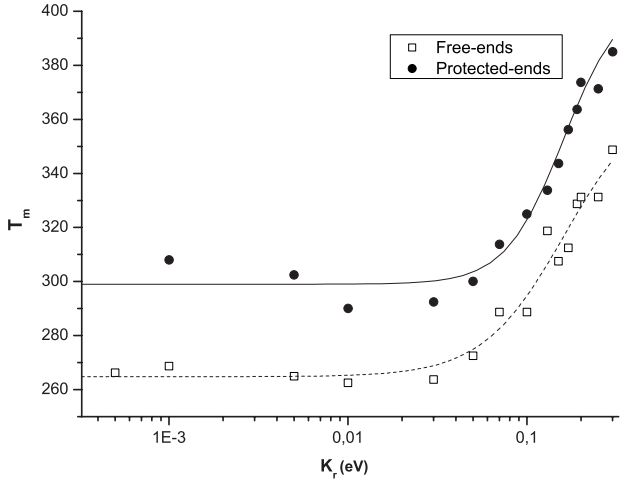
5 Analysis of the results

5.1 Properties intrinsic to the stem

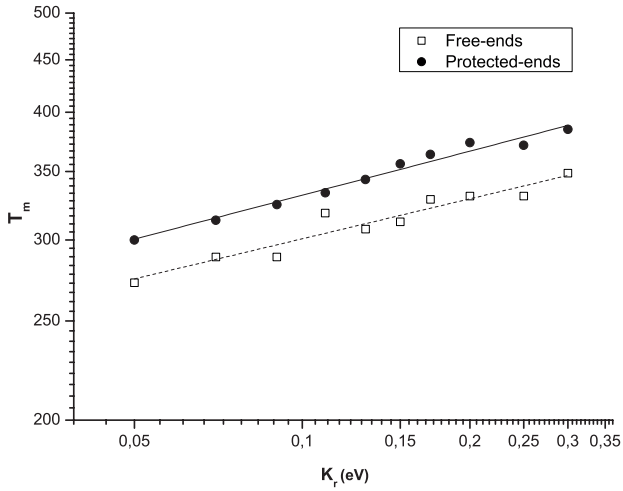
Let us begin with the influence of the characteristics inherent to the stem, which is obviously the part of the molecule which plays the predominant role in denaturation.

In order to precise the role of the rigidity of the chain we have carried out two different calculations to determine the behavior of an isolated stem. First we examined the 6-base-pair stem of the hairpin alone, i.e. with both ends free. At very low rigidity K_r, the stem transition temperature T_c is almost independent of K_r. On the contrary, when K_r reaches 0.05 eV, T_c rises quickly with K_r.

However, the simulation of the isolated stem does not correctly reflect properties that are *intrinsic* to the stem because the denaturation is strongly affected by boundary effects. In order to avoid these perturbations, for a second series of calculations, we designed a simulation of a stem with *protected ends*. A stem of 20 base pairs thermalized at temperature T, is surrounded by two regions (of 20 base pairs) maintained at a temperature $T' < T$ (in most of the calculations $T' = T - 20$ K) by separate thermostats. A buffer zone of 5 base pairs –simulated without thermostat, i.e. as a microcanonical domain– is introduced between two regions maintained at different temperatures to allow for a temperature gradient. The domain in which we study the denaturation of the stem is a section of 15 base pairs in the center of the region maintained at temperature T (Fig. 6). In such a configuration the denaturation of the domain of interest is not forced by the fluctuations of free ends, but reflects more accurately the intrinsic properties of the stem. Figure 5 shows that its denaturation occurs about 20 K above the denaturation of a stem with free ends. The variation of T_m as a function of the stiffness is however very similar whether the stem has free boundaries or not.



(a)



(b)

Fig. 5. Denaturation temperature of the stem versus rigidity. Squares: six-base-pair stem with free ends, circles: stem with ends protected by a domain maintained at a lower temperature. (a) log-lin scale (b) log-log scale, showing that, for $K_r > 0.05$ the denaturation temperature rises approximately as K_r^{α} with $\alpha = 0.129$ or $\alpha = 0.142$ in the case of a stem with free or protected ends, respectively.

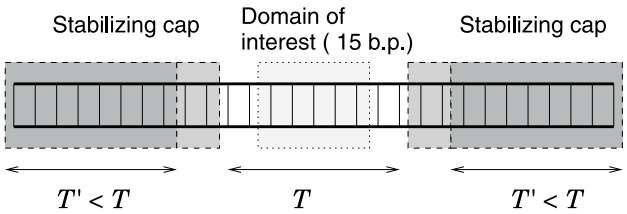


Fig. 6. Schematic picture of the configuration used to simulate the denaturation of the stem with protected ends. The shaded rectangles at both ends correspond to regions that are maintained at a temperature slightly lower than the domain of interest. The two smaller rectangles shaded in light grey are buffer zones, simulated without a thermostat. The central rectangle is the domain in which we study the stem denaturation.

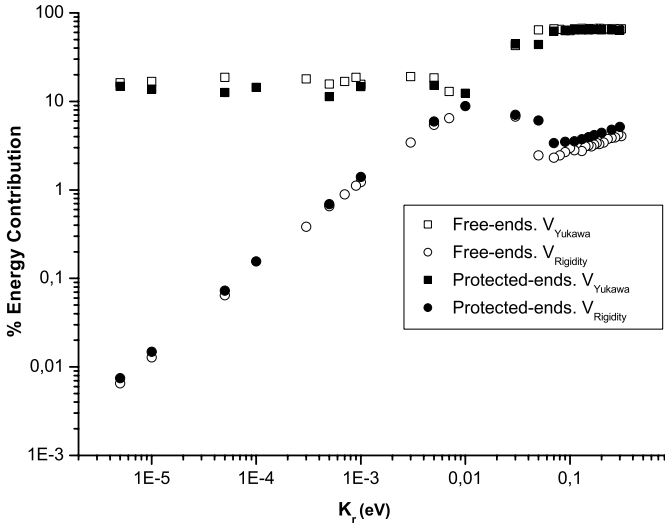


Fig. 7. Percentage of contribution to the global potential energy estimated both for the Yukawa and Rigidity potentials as a function of the stiffness parameter K_r. The Yukawa contribution to the potential energy dominates for $K_r < 0.01$ eV, and becomes negligible for $K_r > 0.01$ eV.

The increase of T_m versus K_r can be understood in terms of an increasing cooperativity of the transition: if several bases have to open at once, the barrier to overcome is higher and this drives the melting temperature up. A similar phenomenon has been noted earlier for the PBD model, where an analytical calculation of T_c is possible and shows that T_c grows as the square root of the coupling constant [17]. Here, for $K_r > 0.05$, the denaturation temperature rises approximately as K_r^{α} with $\alpha = 0.129$ or $\alpha = 0.142$ in the case of a stem with free or protected ends, respectively.

The plateau observed for the very low values of K_r arises from another contribution to the rigidity of the stem, the Yukawa potentials introduced to stabilize the stem against shear distortions. They are weak, but become nevertheless dominant when K_r becomes too small as shown in Figure 7.

The investigations of the stem alone show that the large increase of the melting temperature T_m of the hairpin, observed for K_r in the range $K_r = 0.05$–0.15 eV can be assigned to a similar increase of the transition temperature T_c of the stem.

5.2 Role of the loop

The loop is another structural component that has a strong influence on the dynamics and thermodynamics of the hairpin. The simulations distinguish three major effects, depending on the characteristics of the loop:

(i) the loop protects one end of the stem and limits its fluctuations,
(ii) a rigid loop may exert a mechanical effect that tends to stretch the base pairs of the stem,

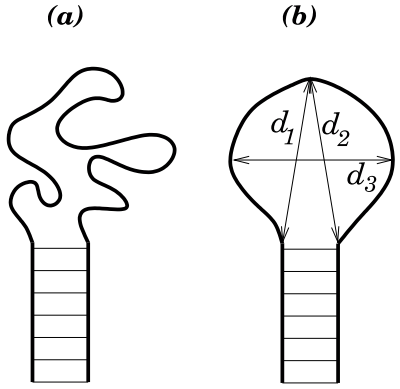


Fig. 8. Schematic picture of the shape of the loop: (a) globular loop. (b) Distances d_1, d_2, d_3 recorded to evaluate the fluctuations of the shape of the loop.

(iii) a long fluctuating loop applies a random force to one end of the stem, which tends to denaturate it.

Let us discuss separately these three possible effects of the loop.

5.2.1 Protective effect and shape fluctuations of the loop

For very low values of the rigidity, the loop takes a highly distorted globular shape, which has a higher entropy than an extended configuration (Fig. 2).

This compact conformation, linked to the last base pair of the stem, tends to maintain the two bases close to each other, isolating this end of the stem from the fluctuations of the solvent. As a result, this prevents an opening of the stem from one of its ends, raising the denaturation temperature.

In order to quantitatively follow the evolution of the shape of the loop and its fluctuations, we recorded the time evolution of the three characteristic distances, d_1, d_2, d_3 defined in Figure 8: d_1 and d_2 are the lengths of the segments that connect the particle situated in the middle of the loop with the last pair of bases closer to the loop, while d_3 is a diametral distance between the bases situated at 1/4 and 3/4 of the length of the loop. Note that these distances grow as the loop evolves from a compact globular shape to the almost circular shape that one would expect for a loop behaving like a bent rigid rod. Then, when this regular shape is reached, d_1, d_2, d_3 can be expected to stay approximately constant, except for small fluctuations.

Figure 9 shows the evolution of d_1, d_2, d_3 versus K_r, recorded in a series of simulations for different rigidities. Each simulation is performed at a temperature that is slightly below the denaturation temperature that corresponds to this particular rigidity, so the results reflect the behavior of the loop when the hairpin is close to opening. Therefore, these calculations give information on the processes that lead to the denaturation. In addition to the values of d_1, d_2, d_3 giving the shape of the loop, we have computed their *Root Mean Square Deviation (RMSD)* (or standard deviation) from their mean values, which measures the fluctuations of this shape. This value is defined as

usual by $RMSD(d_x) = \sqrt{[\sum_{i=1}^{N} (d_{xi} - \langle d_x \rangle)^2]/N}$, where d_{xi} is the value of the distance x at some time t_i, $\langle d_x \rangle$ its average along the calculation, and N is the total number of data points taken into account.

The monotonic increase of d_1, d_2, d_3 when the rigidity increases reflects the expected evolution toward an almost circular loop. It should be noted that the range of K_r for which the distances d_1, d_2, d_3 stay very small corresponds to the range in which the melting temperature of the hairpin decreases when K_r increases (Fig. 3), which is consistent with a protective role of a highly flexible globular loop. The variation of their RMSD is more interesting because, as K_r increases, it falls from a high value on the order of $d_3/2$ to a very small value, showing a fairly sharp change of behavior from a *distorted globular state* to a loop that only has small fluctuations. This evolution is sharper for longer loops. This suggests a collective behavior and, as expected, the transition occurs for a higher value of K_r when the length of the loop increases.

5.2.2 Mechanical stretching

When K_r is large enough, the loop adopts an *annular* shape (Fig. 2), as shown by the values of d_1, d_2, d_3, and behaves like an elastic rod that is kept bent by the stem, which keeps its two ends at a distance smaller than the persistence length of the strand that makes the loop. This generates a *mechanical force* that tends to separate the hairpin legs. This force is inversely proportional to the size of the loop. For large values of K_r this force is the main cause of denaturation. This is demonstrated by observing the dynamics of the denaturation for different lengths of the loop, at fixed K_r. Figure 10 shows a gray plot of the time evolution of the vibrational amplitude of the bases pairs at the transition temperature T_m, for two rigid molecules ($K_r = 0.15$ eV) with a loop composed of 12 and 30 bases, respectively.

Note how the denaturation starts in both cases from a high amplitude vibration located at the base pair closest to the loop (the sixth one) and propagates along the stem to unzip all the complementary pairs. This opening is slower in the case of a longer loop as expected because the elastic force generated by a bent rod decreases as its length increases.

5.2.3 Forces induced by the Brownian motion of the loop

Up to now, the effects of the loop that we have discussed are essentially static in the sense that they depend on its equilibrium shape, whether it is globular or similar to a bent rod. But the loop also has an influence on the stem through its dynamics. A hairpin loop can be seen as polymer with both ends attached to a pair of complementary bases part of the stem. This filament immersed in a fluid undergoes random fluctuations. As a result it applies a Brownian force on the stem, which favors the opening of the base pairs.

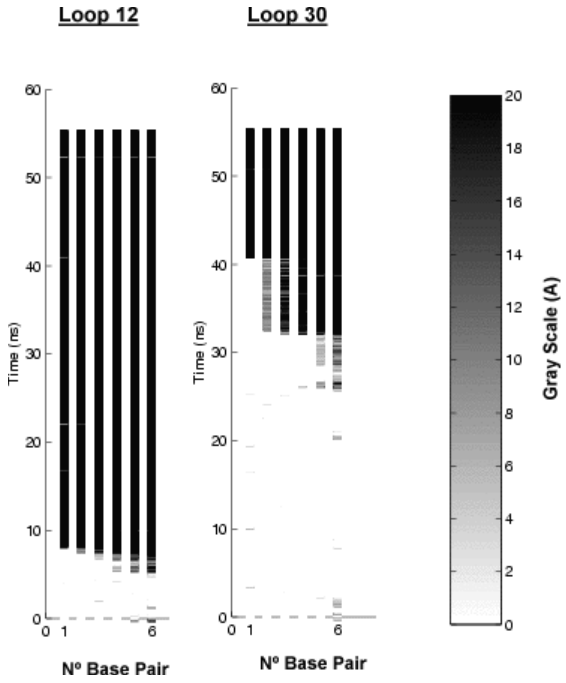


Fig. 10. Denaturation of two hairpins with different loop lengths and the same rigidity. The evolution of the Morse pairs' vibrating amplitude is presented in a gray plot as a function of the simulation time and the pair's position in the stem. Base-pair 6 is the base pair adjacent to the loop. The rigidity parameter is $K_r = 0.15$ eV and the temperature is $T = 295$ K for the 12-unit loop and $T = 300$ K for the 30-unit loop.

In order to evaluate this effect, we have determined the melting temperatures of modified hairpins in which the loop has been cut in the middle, leaving two dangling strands. This prevents the mechanical effect discussed in the previous section for the high rigidity case. Figure 11 shows that, when the loop is cut, the melting temperature does not depend on the length of the dangling strands. This indicates that the Brownian forces applied to the stem by the fluctuations of the loop play a negligible role in the denaturation of hairpins.

A direct measurement of the Brownian force exerted by a dangling strand coupled to a thermal bath can be made by simulating such a strand and recording the forces that it exerts on its attachment points. The results are shown in Figure 12 for a 6-base and a 15-base strand. The forces increase linearly with temperature, as one would expect from the Brownian dynamics of the strands, but it is not affected by the value of the rigidity constant K_r throughout a very broad range of rigidity. These numerical experiments indicate that the experimental observation of a variation of T_m when the rigidity of the loop changes cannot be related to the Brownian forces exerted by the loop on the stem.

This result is consistent with the observation of Figure 11 that T_m does not depend on the length of the loop when it has been cut into two dangling strands.

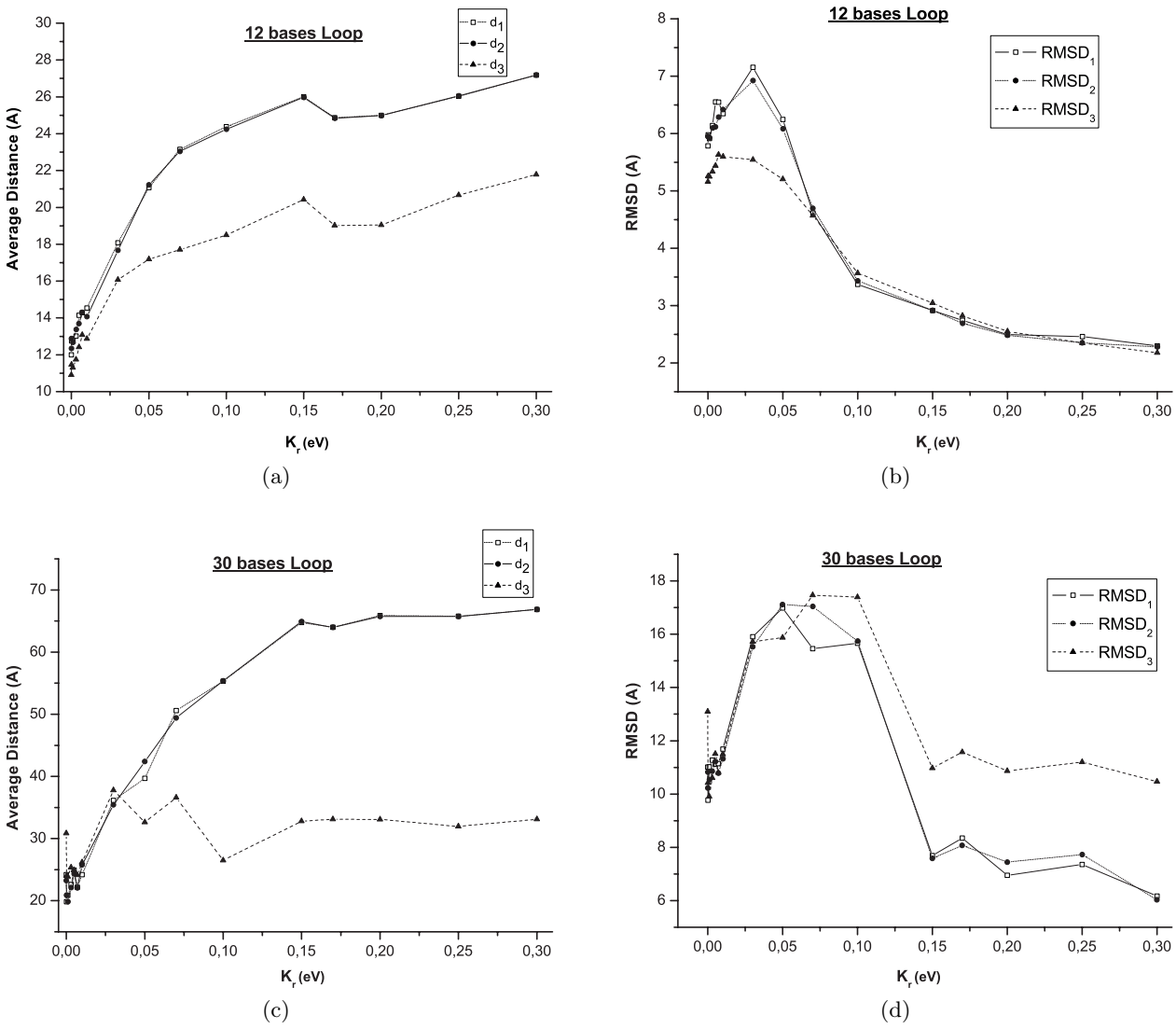


Fig. 9. Evolution of the loop shape ((a) and (c)) and of its fluctuations ((b) and (d)) versus the rigidity parameter K_r, determined from the measurement of the three characteristic distances d_1, d_2, d_3 and their root mean square deviation (RMSD). (a) and (b) represent the study of these quantities for a loop of 12 bases, while (c) and (d) repeat the same analysis for a 30-base loop. The small fluctuations around smooth evolutions are not significant because the results are sensitive to the exact value of the temperature $T < T_m$ that has been used in the simulations.

5.3 Inhomogeneous rigidity model

The studies of previous sections have been performed with a rigidity of the strand that is homogeneous along the whole hairpin, whether we consider the loop or the stem. This reduces the number of free parameters but, in order to compare with the experiments [4], we must vary the rigidity of the loop while keeping the properties of the stem unchanged. This result suggests that we complete our investigations by studying an inhomogeneous model, providing different degrees of stiffness for the loop and the stem.

Therefore, we have carried out a study assigning a constant value of $K_r^{'}$ to the bases taking part in the stem, and exploring the influence of a loop having different rigidities. For the stem, the value $K_r^{'} = 0.15$ eV has been chosen because it leads to dynamical properties of the stem that are in qualitative agreement with the observations. Excitations below the melting temperature are present in the stem but they are not large enough to strongly distort the double helix structure. Lower values of $K_r^{'}$ lead on the contrary to very large distortions of the double helix, even for $T < T_m$.

Figure 13 shows the variation of the melting temperature for various loop lengths. First one notices that the melting temperature is monotonically decreasing as the rigidity of the loop increases, in contrast to the results presented above. This indicates that the rise observed at large rigidities was entirely due to the increase of the stem denaturation temperature with increasing strand rigidity.

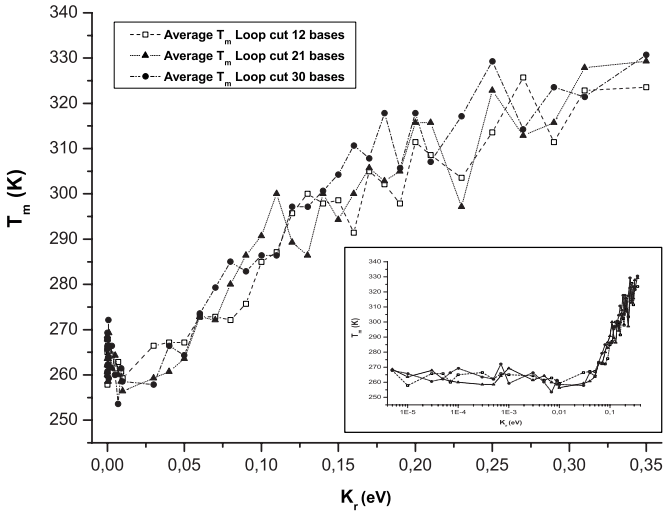


Fig. 11. Melting temperatures of modified hairpins, in which the loop has been cut in the middle, versus the rigidity of the strands, for 3 different loop lengths. The inset shows the variation in logarithmic scale for the low values of K_r.

Second, the polynomial fits show that the effect of the loop length is different for low and high rigidities. For $K_r > 0.05$ eV, the simulation shows that increasing the length of the loop raises the melting temperature T_m while the reverse is true for low values of K_r. Combined with the experimental results, this observation gives some useful informations on the melting process of DNA hairpins, as discussed in the next section.

6 Discussion and conclusion

In this section let us summarize what we have found, how it compares with experiments, and what it tells us about DNA hairpin denaturation.

We are now in a position to explain the full curves of Figures 3 and 4 in the light of the studies performed to investigate the different phenomena that enter into the melting of the hairpin. For the very low rigidity range shown in Figure 4b, the loop takes a globular shape that protects the stem from denaturation, and, as we have seen. For $K_r < 0.03$ eV, the denaturation of the stem does not depend on the rigidity, which explains why T_m stays approximately constant for $K_r < 10^{-3}$ eV. In the range $10^{-3} < K_r < 0.05$ eV, the loop gradually loses the globular shape that protected one end of the stem, while the stem transition temperature is still very weakly sensitive to the rigidity, as shown in Figure 5a. This explains the decrease of T_m observed in Figure 4a. Beyond $K_r = 0.05$ eV, an increase in rigidity starts to play two opposite roles: a more rigid loop tends to mechanically open the hairpin, while a more rigid stem becomes harder to denaturate because the denaturation is more cooperative. In the range $0.05 < K_r < 0.15$ eV, the effect of the stem dominates because the loop has not yet fully taken its rigid shape as shown in Figure 9. Beyond $K_r = 0.15$ eV, the effect of an increasing rigidity in the loop compensates the effect of the increasing stem rigidity, and T_m stays approximately constant.

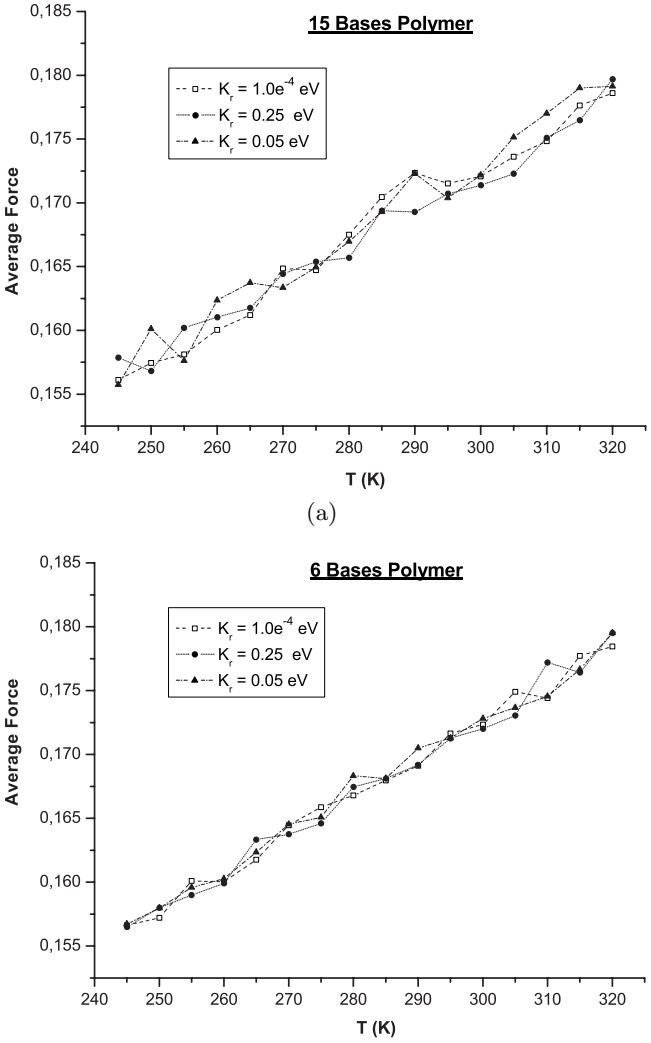


Fig. 12. Estimated force for polymers different in length, with variable rigidity potentials as a function of the temperature: (a) 15-base strand, corresponding to a 30-base loop; (b) 6-base strand, corresponding to a 12-base loop.

Let us now consider the simulations with a given stem and variable loop rigidities. This situation corresponds to the experiments of [4] which show that a longer loop decreases the melting temperature, whatever the rigidity of the loop. This indicates that the loop does not have a mechanical effect on the opening because, if it were the case, a longer loop would be easier to bend, reducing the mechanical force on the stem, so that the melting temperature would on the contrary increase together with the length of the loop. Therefore, the experiments indicate that the rigidity of the loop must be in the low K_r range of our studies, i.e. $K_r < 0.05$ eV. In order to precise the value of K_r we have initiated a series of simulations of the hairpin using an all-atom model of DNA [18]. By comparing the statistics of the fluctuations of the strands

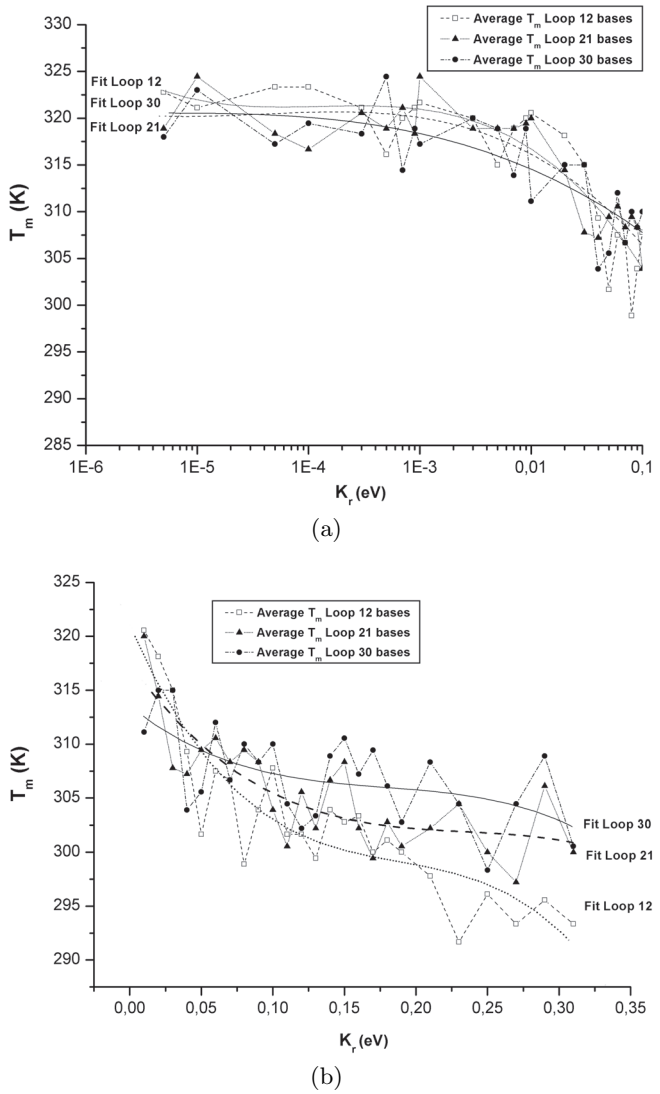


Fig. 13. Evolution of the denaturation temperature T_m for a hairpin with a fixed rigidity in the stem ($K_r' = 0.15$ eV) and variable loop rigidities, for different loop lengths: (a) linear scale for K_r; (b) logarithmic scale for K_r to show the variation at very low loop rigidity. The response to variations in the loop rigidity has been calculated both for cases of high and low rigidity values of K_r according to different loop sizes.

(defined for instance by the segments joining neighboring phosphates) in the full model and in the reduced model with different values of K_r, one can evaluate the range of K_r that gives the best agreement. Preliminary results based on Gaussian fits lead to values of K_r in the range $0.001 < K_r < 0.005$ eV, which are consistent with the conclusions of the above discussion.

Moreover, studies combining the measurement of T_m with observations of the fluctuations of fluorescence [4] concluded that the opening is not influenced by the loop rigidity. This is in good agreement with our results: as soon as we chose K_r small enough to have T_m that decreases for increasing loop length, we find that the rigidity of the loop only has a small influence on the opening of the hairpin (see Fig. 13a).

This suggests that the model is consistent with the melting of actual hairpins, but it also tells us something more about the role of the loop. In this range of K_r, the loop has a highly fluctuating, mostly globular shape. We have seen that it can have a *stabilizing* effect on the stem. This is a scenario that was not considered in [4].

At a first glance, our conclusions for the appropriate value of K_r may seem inconsistent because when we examined the case of a fixed stem rigidity, we noted that the value $K_r' = 0.15$ eV was appropriate for the stem, while the discussion leads us to $K_r < 0.05$ eV for the loop. In fact, this is perfectly consistent with the structure of DNA, as well as the conclusions of earlier models. As noted earlier, most of the flexibility of the strands comes from rotations around covalent bonds; these bonds allow large changes of the angle between the segments that make up the stands, without having to change the angles between two bonds connected to a given atom, which would have a high energetic cost. In other words, the flexibility of the strands comes from variations of dihedral angles rather than bond angles. But when the bases are paired, as is the case in the stem, the rotational freedom of the segments that make up the stem is drastically reduced. The dihedral angles are no longer free, and, in our simple model that does not explicitly include the dihedral angles, this translates into a higher rigidity parameter K_r in the stem than in the loop.

It should be noted that the statistical physics of DNA thermal denaturation with a simple model [19] leads to the same conclusion because a sharp transition, in agreement with experiments, can only be obtained if the base stacking interaction is nonlinear, and such that the effective interaction between open bases is significantly reduced with respect to the interaction between bases belonging to closed pairs. In the present model the coupling between bases along the strand is achieved by the strand rigidity. Thus, a lower rigidity for the bases belonging to the loop could be expected not only from DNA structure but also from the thermodynamics of its denaturation. In fact, to be more realistic our model should have a rigidity parameter K_r that depends on the state of the bases, so that when a base pair in the stem opens up, K_r switches to a smaller value. But this would introduce an additional complexity, and more parameters. This is not essential as long as we consider the opening of very short stems. Nonlinear base stacking, essentially leads to entropy effects, which only become significant for the denaturation of long DNA helices.

This study to analyze the fluctuations of DNA hairpins is not complete because we only investigated the *opening*. In order to analyze the experimental studies of hairpin fluctuations, the *closing* step must also be considered. Investigations in this direction are in progress but they must rely on a completely different approach because the time scales are very different, and the role of mismatches, which are partial closings of the stem, is crucial. The dynamical model that we have proposed is certainly

oversimplified but the basic phenomena that we have exhibited (such as the role of the loop shape and its fluctuations) do not rely on specificities of the model. A more accurate model could certainly give more quantitative results, but we think that the ideas would be preserved. Moreover there are now studies that show that simple models can give quantitative results for the structure of DNA hairpins. Recently, a model describing single-stranded DNA as a continuous, unshearable, unstretchabe and flexible thin rod has been shown to be very fruitful [20]. Although this model is more complex than our model and limited to the static structure, it demonstrates that simplifying DNA models may be a fruitful research path that is not restricted to qualitative studies.

Part of this work have been supported by the European cooperation program LOCNET (HPRN-CT-1999-00163) and by the project BFM 2002-00113 DGES (Spain). S. Cuesta is supported by the Spanish Ministry of Science and Education (FPU-AP2002-3492) and his stay in Lyon has been supported by *"Programa Europa XXI"* DGA(CONSI+D)-CAI CB-3/03. The stay of D.J. Graham has been supported by the National Science Foundation (USA).

References

1. M. Peyrard, Nonlinearity **17**, R1 (2004)
2. Chu H. Choi, G. Kalosakas, K.O. Rasmussen, M. Hiromura, A.R. Bishop, A. Usheva, Nucleic Acids Research **32**, 1584 (2004)
3. A. Campa, A. Giansanti, Phys. Rev. E **58**(3), 3585 (1998)
4. N.L. Goddard, G. Bonnet, O. Krichevsky, A. Libchaber, Phys. Rev. Lett. **85**, 2400 (2000)
5. G. Bonnet, S. Tyagi, A. Libchaber, F. Russel Kramer, PNAS **96**, 6171 (1999)
6. E.J. Sorin, Y. Min Rhee, B.J. Nakatani, V.S. Pande, Biophys. J **85**, 790 (2003)
7. M. Peyrard, A.R. Bishop, Phys. Rev. Lett. **62**, 2755 (1989)
8. T. Dauxois, M. Peyrard, A.R. Bishop, Phys. Rev. E **47**, 684 (1993)
9. N. Theodorakopoulos, T. Dauxois, M. Peyrard, Phys. Rev. Lett. **85**, 6 (2000)
10. W.L. DeLano, *The Pymol Molecular Graphics System* (Delano Scientific, San Carlos CA (USA), 2002)
11. M. Barbi, S. Lepri, M. Peyrard, N. Theodorakopoulos, Phys. Rev. E **68**, 061909 (2003)
12. J.K. James, I. Jr Tinoco, Nucleic Acid Res. **21**, 3287 (1993)
13. G. Bonnet, O. Krichevsky, A. Libchaber, Proc. Natl. Acad. Sci. USA **95**, 8602 (1998)
14. G.J. Martyna, M.L. Klein, M. Tuckerman, J. Chem. Phys. **97**, 2635 (1992)
15. S. Nosé, J. Chem. Phys. **81**, 511 (1984)
16. S. Nosé, Molec. Phys. **52**, 255 (1984)
17. T. Dauxois, N. Theodorakopoulos, M. Peyrard, J. Stat. Phys. **107**, 869 (2002)
18. S. Cuesta-López, Y.-H. Sannejouand, M. Peyrard, unpublished, 2004
19. T. Dauxois, M. Peyrard, A.R. Bishop, Phys. Rev. E **47**, R44 (1993)
20. C. Pakleza, J.A.H. Cognet, Nucleic Acids Research **31**, 1075 (2003)