Eur. Phys. J. E **10**, 153–161 (2003)
DOI: 10.1140/epje/e2003-00019-8

THE EUROPEAN PHYSICAL JOURNAL E

# Slow nucleic acid unzipping kinetics from sequence-defined barriers

S. Cocco[1,a], J.F. Marko[2,b], and R. Monasson[3,4,c]

[1] CNRS–Laboratoire de Dynamique des Fluides Complexes, 3 rue de l'Université, 67000 Strasbourg, France
[2] Department of Physics, The University of Illinois at Chicago, 845 West Taylor Street, Chicago, IL 60607-7059, USA
[3] CNRS–Laboratoire de Physique Théorique de l'ENS, 24 rue Lhomond, 75005 Paris, France
[4] CNRS–Laboratoire de Physique Théorique, 3 rue de l'Université, 67000 Strasbourg, France

Received 29 July 2002 / Received in final form 5 February 2003
Published online: 16 April 2003 – 

**Abstract.** Recent experiments on unzipping of RNA helix-loop structures by force have shown that ≈40-base molecules can undergo kinetic transitions between two well-defined "open" and "closed" states, on a timescale ≈1 sec [Liphardt *et al.*, Science 297, 733-737 (2001)]. Using a simple dynamical model, we show that these phenomena result from the slow kinetics of crossing large free energy barriers which separate the open and closed conformations. The dependence of barriers on sequence along the helix, and on the size of the loop(s) is analyzed. Some DNA and RNA sequences that could show dynamics on different time scales, or three(or more)-state unzipping, are proposed. Our dynamical model is also applied to the unzipping of long (kilo-basepair) DNA molecules at constant force.

**PACS.** 87.15.-v Biomolecules: structure and physical properties

## 1 Introduction

Helix-loops are the basic secondary-structure elements of folded single-stranded nucleic acids (ssNA). Recent physical studies of single helix-loop RNAs have revealed that despite their simple structures, they can display interesting dynamics [1]. When placed under moderate tensions ≈15 pN, telegraph-noise-like "switching" behavior can be observed. The characteristic time of this switching has been seen to be on the ≈1 sec timescale, surprisingly large given the small size (≈10 nm) of the molecules. The purpose of this paper is to present a simple theory capable of reproducing these slow switching kinetics.

In Section 2 we describe the model that we use for the energy of a helix-loop structure. We consider states of partial "unzipping" [2], using the current best quantitative descriptions of base-pairing interactions and single-strand nucleic acid elastic response. Section 3 discusses the free energy and equilibrium states associated with our energy model, and we show that there are large barriers between open and closed states for specific molecules. The appearance of a large free energy barrier between open and closed states for helix-loop structures is found and qualitatively explains experimentally observed discrete-state switching kinetics.

Section 4 combines this model with Eyring-Kramers transition-state theory to obtain a dynamical theory. We find that the free energy barrier requires a long timescale to be crossed. In Section 5 we show how with only the intrinsic attempt frequency of single-base-pair motion as a free parameter, our dynamical model describes the observed features of experiments on short (<30 bp) RNA helix-loop structures. In Section 6, we revisit the question of the free energy barrier, from a simple scaling point of view, and then in Section 7 we show how our theory can be generalized to describe branched helix-loop RNA structures. We also show how three-state-switching can occur for a specific helix-loop structure, for a particular choice of sequence.

In Section 8 we apply our dynamical model to the unzipping of much larger genomic DNA molecules at constant force. Bockelmann, Essevaz-Roulet and Heslot [3,4] have been able to observe sequence-related force variations during DNA unzipping at controlled end-separation. The large-scale inhomogeneity of genomic DNA sequences gives rise to a rough unzipping free energy landscape [3]. Lubensky and Nelson have noted that tension allows one to "tilt" the free energy landscape, and have predicted barrier-crossing-controlled kinetics for unzipping of large DNAs at constant force [5]. We use our model to make explicit calculations for the expected unzipping kinetics for the $\lambda$-DNA usually studied in such experiments. For this 48502-base-pair molecule, we find a huge variation in barrier heights. Unzipping occurs as a sequence of steps

[a] e-mail: cocco@ldfc.u-strasbg.fr
[b] e-mail: jmarko@uic.edu
[c] e-mail: monasson@lpt.ens.fr

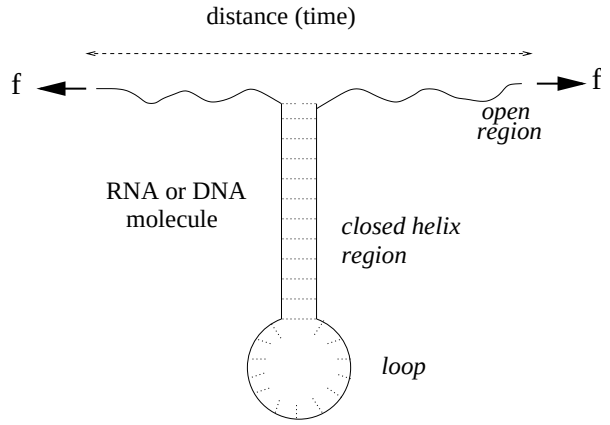


**Fig. 1.** Fixed force unzipping experiment. A constant force is applied to a helix-loop structure while the distance between molecule ends is measured.

between a series of deep free energy minima, with wide variations from realization to realization of the dynamics. Finally, in the conclusion we discuss a few more applications for our dynamical theory.

## 2 Model of base-pairing and ssNA elasticity

Our description uses the series of molecule configurations obtained by successively breaking base-pairs, starting from the molecule ends (Fig. 1) [2]. In addition to the base-pairing free energy of the double-stranded part of the molecule, we take into account the elastic response of the extended, unpaired, single-stranded part of the molecule [3,5–7]. The free energy change associated with opening the base pair $i$ at the boundary between the open and closed region is therefore

$$\Delta g(i,f) = g_0(i) - 2\, g_s(f,i). \qquad (1)$$

Here $g_0(i)$ is the free energy of opening base pair $i$, *i.e.* the free energy difference between paired and unpaired bases at zero force. $g_0(i)$ is sequence specific *via* the hydrogen bonds between the two bases of the pair and the stacking interactions between adjacent base pairs.

In the standard experimental conditions (pH 7, room temperature, 150 mM NaCl) $g_0$ ranges between 1 $k_BT$ and 6 $k_BT$ for Watson and Crick base pairs. The $\{g_0(i)\}$ have been obtained for each sequence of Figure 2 by the MFOLD server [8]. $g_s(f,i)$, is the free energy of the opened DNA base stretched by the constant unzipping force $f$; the factor 2 arises because for each base pair that opens two bases are stretched. This free energy thus includes only the contribution from the external force and does not depend on the nucleotide type. Furthermore it depends on the sequence because internal unpaired "bubbles" *e.g.* the final loop in Figure 1 are considered to open along with the base pair immediately preceding them; those base opening steps therefore pick up a multiple-base $g_s$ contribution.

The ssNA elastic behavior is complicated due to competing effects of electrostatic self avoidance and

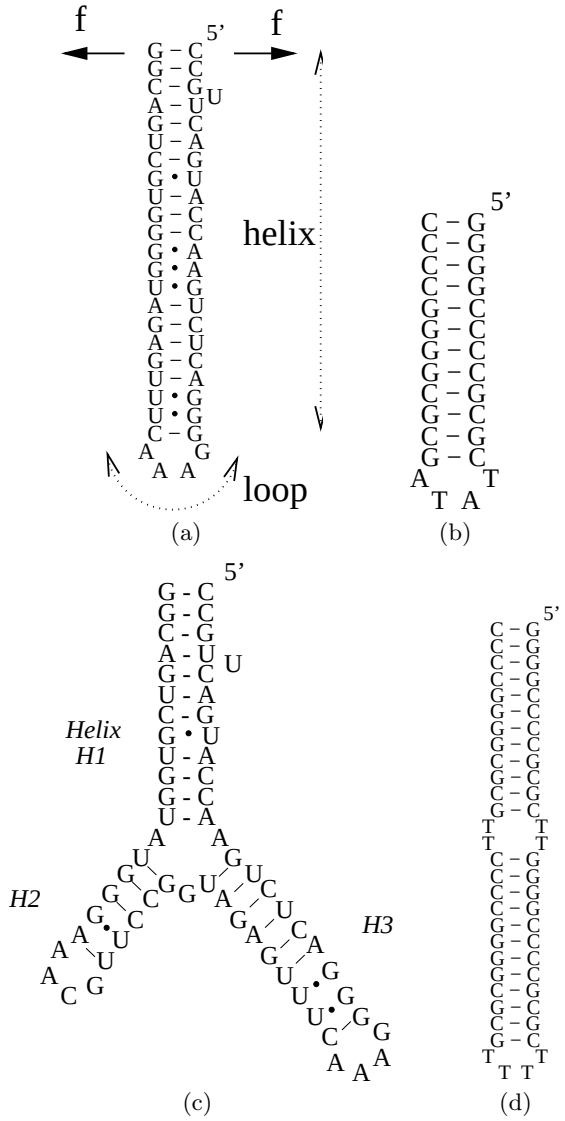


**Fig. 2.** Molecules studied: (a) P5ab RNA, a single helix-loop structure present on the P4-P6 domain of a self splicing group I intron of the *Tetrahymena thermophila*. The structure shown is predicted by Mfold [8] apart from the G-A weak pairs [14] indicated here with dots, and with the U-bulge translocated; (b) DNA hairpin consisting of a poly(GC) helix terminated with an ATAT loop [10]; (c) P5abc$\Delta$A RNA, a variant of P5ab with an additional helix giving a Y-branched structure at zero force; (d) Hypothetical RNA molecule obtained by ligation of two poly(GC) helices as in (b) and replacing A bases with Ts.

nucleotide-nucleotide interaction [9–13]. However, in the range of force of 10 pN characteristic for unzipping the experimental force-extension curve for ssDNA is well reproduced by a freely jointed chain like model (FJCL) with Kuhn length $d = 15$ Å and effective nucleotide length $l_{ss} = 5.6$ Å [11]. The corresponding free energy for forces up to 20 pN is:

$$g_{\rm s}^{\rm FJCL}(f) = k_B T l_{ss}/d\, \ln[\sinh(u)/u] \qquad (2)$$

with $u \equiv d\, f/[k_BT]$.

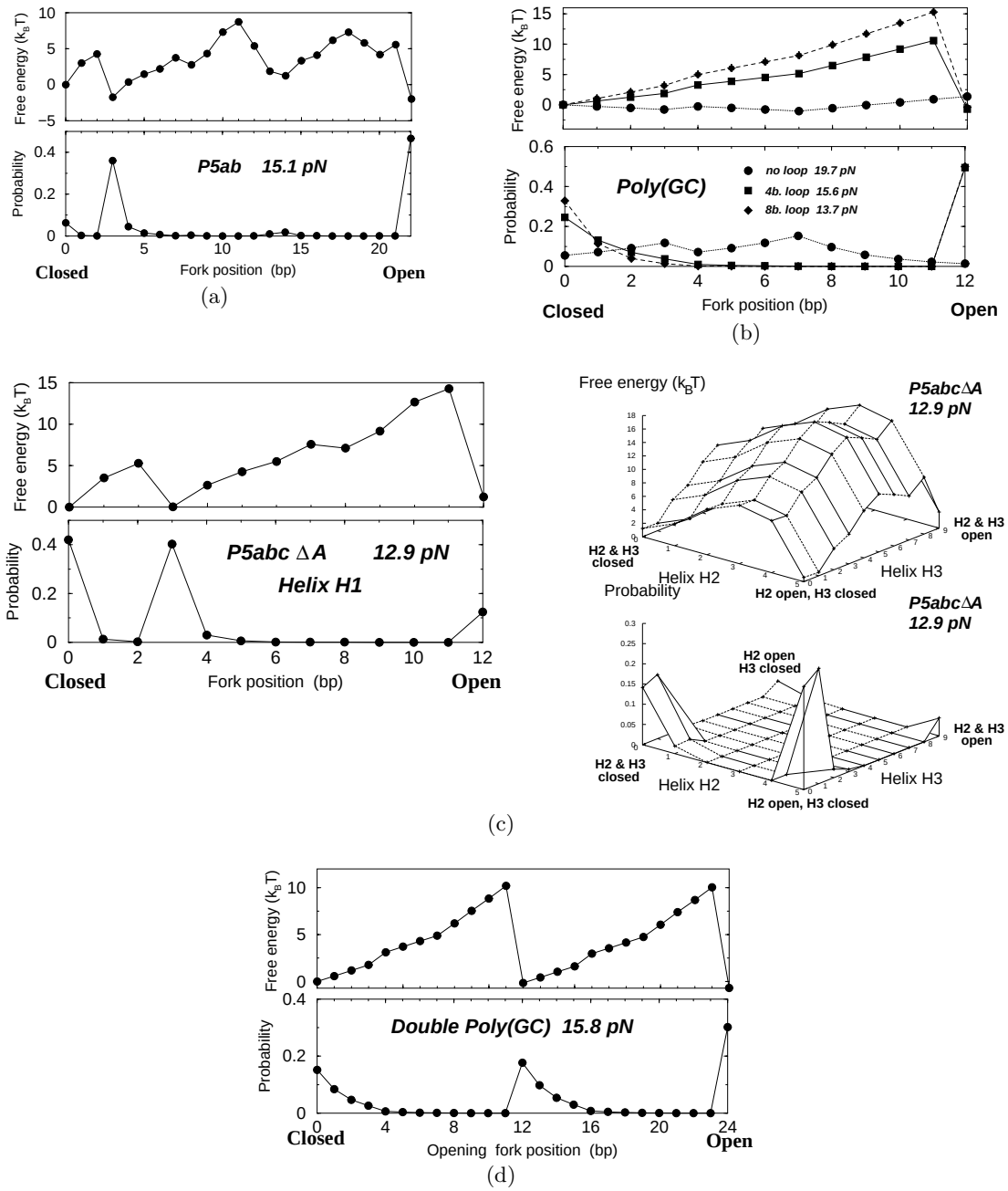


**Fig. 3.** Free energy landscapes (in $k_BT$, top) and probability distributions for the position of the opening fork (bottom) at the critical force for molecules of Figure 2. (a) P5ab; (b) Poly(GC) RNA with: a 8bp loop (diamond), a 4 bp loop (squares); no loop (circles). (c) P5abc$\Delta$A; (left: opening positions $n_1 = 0, \ldots, 12$ corresponding to helix H1, right: opening positions $(n_2, n_3)$ with $n_2 = 0, 5$ – H2 – and $n_3 = 0, 9$ – H3 –); by definition configurations $n_1 = 12$ and $(n_2 = 0, n_3 = 0)$ coincide. (d) Hypothetical double Poly(GC) (as in Fig. 2d) for force $f = 15.8$ pN.

The free energy to unzip the first $n$ base pairs is just the sum of the free energies of the individual base-opening steps,

$$G(n, f) = \sum_{i=1}^{n} \Delta g(i, f). \tag{3}$$

We emphasize that both the base-pairing and the elasticity contributions to (1) and (3) are free energies, *i.e.* are coarse-grained over atomic-scale fluctuations. This "zipper" model [2] is roughly equivalent to that described in the Supplementary Materials of [1].

## 3 Equilibrium behavior of a single helix-loop structure

Figure 3 shows $G(n, f)$ and the equilibrium probability $P(n, f) \propto \exp(-G(n, f)/k_BT)$ for a simple RNA hairpin,

called P5ab [14] (Fig. 2a), at the critical force $f^* \approx 15$ pN where two-state switching is observed [1]. At the critical force, the open and closed states ($n = 22$ and $n = 3$, note $n$ is just the number of broken base pairs) dominate; below or above this force, the free energy landscape is tilted either to favor the $n = 3$ or $n = 22$ state.

At the critical force, there are a few barriers due to drops in free energy resulting from openings of the U bulge ($n = 3$), the weak non Watson-Crick GA central pairs ($n = 13$), and the final loop ($n = 21$). The largest barrier has a height $\simeq$11 $k_BT$, and must be crossed to reach the closed state from the open one, and *vice versa*. The two-state behavior observed for P5ab follows from the partition of probability into two peaks separated by a rarely-accessed barrier region (see Supplementary Materials, Ref. [1]).

## 4 Dynamical model

To reach a quantitative understanding of switching, we introduce a dynamical model for the motion of the "fork" separating the base-paired and opened regions of the molecule shown in Figure 4. We propose the following expressions for the elementary rates of opening and closing base pair $n$ (*i.e.* to move the boundary between the open and closed portion of the molecule from $n$ to $n-1$ or to $n+1$):

$$r_o(n) = r\ \mathrm{e}^{-g_0(n)/k_BT}\ , \qquad r_c(f,n) = r\ \mathrm{e}^{-2g_s(f,n)/k_BT}. \tag{4}$$

Here $r$ is essentially the microscopic rate for a base pair to move together or apart in the absence of tension or base-pairing interactions, or roughly the inverse self-diffusion time for a few-nm-diameter object [17], $r = k_BT/(2\pi\eta\ell^3) \approx 5 \times 10^6$ s$^{-1}$, with $\ell = 5$ nm, $\eta = 0.001$ Pa sec, and $k_BT = 4 \times 10^{-21}$ J.

In (4) we have made the simplifying approximation that the opening rate $r_\mathrm{o}$ has no force dependence, and is simply proportional to the exponential of the base-pairing free energy of (1). Eyring–Kramers transition–state theory applied to breaking of a chemical bond considers indeed the potential energy to be "tilted" by a force-times-displacement contribution, as shown in Figure 5 [6,15, 16]. Because hydrogen bonds break for relatively small displacements ($\approx$0.1 nm) the reduction in the potential energy of the single-base-opening transition state will be roughly 15 pN $\times$ 0.1 nm = 0.3 $k_BT$. This can be neglected with respect to the base-pairing free energy of a few $k_BT$, which is a lower bound to the energy of the transition state associated with breaking a single base pair.

Detailed balance then determines the closing rate $r_\mathrm{c}$ to be proportional to the exponential of force times displacement, *i.e.* to the energy of a fluctuation that is able to pull the two bases back together in opposition to the applied force. The rates (4) lead to a master equation for the probability $\rho_n(t)$ for the boundary to be at site $n$ at time $t$:

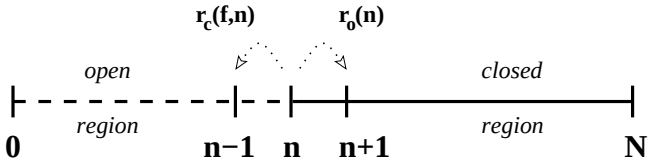


**Fig. 4.** The dynamical model for opening and closing of a single helix structure at constant force. The boundary between open and close portions of the molecule is located at base pair number $n$. Elementary closing ($n \to n-1$) and opening ($n \to n+1$) rates are $r_\mathrm{c}(f,n)$ and $r_\mathrm{o}(n)$ respectively.

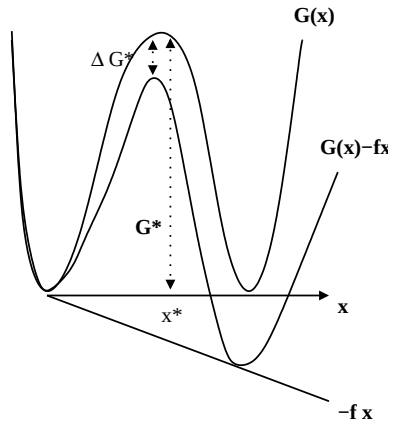


**Fig. 5.** Schematic representation of the free energy tilt in presence of an external force $f$ in a two state system with reaction coordinate $x$. The position of the barrier $x^*$ is called transition-state coordinate. Notice that the force shifts the height $G^*$ of the barrier in $x^*$ by the amount of $\Delta G^* \simeq -f\ x^*$ . The switching rate between the two state is simply $\propto\exp(-G^*/k_BT)$. This picture is general for two state systems and applies to various physical situations. For instance, the reaction coordinate $x$ is along the hydrogen bond direction when referring to discussion of equation (4) in Section 4, and equals the position of the opening fork, or number of unzipped base pairs, when referring to the opening dynamics in Sections 5, 6.

$$\begin{aligned}\rho_n(t+\delta t) &= \rho_n(t) - \delta t\ \rho_n(t) \times (r_\mathrm{o}(n) + r_\mathrm{c}(f,n)) \\ &+ \delta t\ \rho_{n-1}(t) \times r_\mathrm{o}(n-1) + \delta t\ \rho_{n+1}(t) \times r_\mathrm{c}(f,n+1)\end{aligned} \tag{5}$$

or equivalently, in the time continuum limit,

$$\frac{\mathrm{d}\,\rho_n(t)}{\mathrm{d}\,t} = -\sum_{m=0}^{N} T_{n,m}\ \rho_m(t). \tag{6}$$

This $(N+1) \times (N+1)$ matrix $T_{n,m}$ is tridiagonal, with nonzero entries $T_{m-1,m} = -r_\mathrm{c}(f,m)$, $T_{m+1,m} = -r_\mathrm{o}(m)$, and $T_{m,m} = r_\mathrm{c}(f,m) + r_\mathrm{o}(m)$.

## 5 Switching kinetics of a single helix-loop structure

We have solved (diagonalized) (6) for P5ab (Fig. 2a). The smallest eigenvalue is 0; the corresponding eigenvector

**Table 1.** Theoretical results for the molecules of Figure 2 compared to experimental values from [1,10] (in bold) when available. Columns indicate the free energy $\Delta G = G(N, f = 0)$ at zero force, the critical force $f^*$, the rate of opening-closing $k^*$, the switching time $t^*$ and the free energy $G^*$ of the highest barrier at criticality, the variation of the opening and closing rates upon force, and the ratio of the two largest non zero eigenvalues. This ratio is small when open and closed states are well defined, and close to one otherwise. Uncertainties in base-pairing free energy are at most $\delta g \approx 0.5\ k_BT$/bp, with consequent total uncertainties of $\approx N^{1/2}\delta g$ (or about 3 $k_BT$ for $N = 25$) for $\Delta G$ and of $\approx (N/2)^{1/2}\delta g \simeq 2\ k_BT$ for $G^*$, and $\approx 3\ k_BT/20$ nm $\approx 0.6$ pN for the critical force.

| Molecule (N) | $\Delta G$ $k_BT$ | $f^*$ (pN) | $\ln(k^*$ r) | $t^*$ (sec) | $\ln k_o(f)$ $(\sec^{-1})$ | $\ln k_c(f)$ $(\sec^{-1})$ | $G^*$ $(k_BT)$ | $\lambda_1/\lambda_2$ |
|---|---|---|---|---|---|---|---|---|
| P5ab | 57.2 | 15.1 | −13.8 | 0.25 | $-42.9 + 1.93$ f | $27.5 - 2.74$ f | 10.7 | 0.023 |
| (49) | **66.7** | **14.5** | | **0.25** | $\mathbf{-39 \pm 2.3 -}$ | $\mathbf{41 \pm 1.9 +}$ | | |
| | **±8.5** | **±1** | | **±0.1** | $\mathbf{(2.9 \pm 0.2)f}$ | $\mathbf{(2.8 \pm 0.1)f}$ | | |
| Repeated | | | | | | | | |
| AU (50) | 44.8 | 12.3 | −6.7 | 0.0002 | | | 0 | 0.99 |
| GC (50) | 135 | 27.4 | −10.3 | 0.008 | | | 0 | 0.99 |
| Poly(GC) | | | | | | | | |
| no loop (24) | 41.3 | 19.7 | −6.5 | 0.0002 | | | 2.4 | 0.42 |
| 4b loop (28) | 34.8 | 15.6 | −14.1 | 0.37 | $-46.7 + 2.02$ f | $1.7 - 1.01$ f | 11.3 | $2\ 10^{-4}$ |
| | $\mathbf{43 \pm 3}$ | **16** | | | | | | |
| 8b loop (32) | 33.2 | 13.7 | −18.0 | 16.7 | $-46.8 + 2.1$ f | $1.3 - 1.4$ f | 15.7 | $3\ 10^{-6}$ |
| double (56) | 71.3 | 15.8 | −14.4 | 0.48 | $-53.2 + 2.26$ f | $8.34 - 1.44$ f | 10.9 | 0.38 [a] |
| P5abc$\Delta$A | 70.6 | 12.9 | −17.1 | $10^b$ | $-43.8 + 2.06$ f | $9.4 - 2.05$ f | 14.3 | $9\ 10^{-4}$ |
| (64) | **71.9** | **12.7** | | | $\mathbf{-39 \pm 9.3 +}$ | $\mathbf{58 \pm 7.5 -}$ | | |
| | **±11.5** | **±0.3** | | | $\mathbf{(2.7 \pm 0.7)f}$ | $\mathbf{(4.2 \pm 0.5)f}$ | | |

[a] the ratio is close to one due to the presence of three, and not two, states giving rise to two large barrier crossing times (the remaining fluctuation times are much shorter: $\lambda_1/\lambda_3 \simeq 0.0001$);
[b] the predicted rate (0.1/sec) for P5abc$\Delta$A is very close to the lowest frequency (0.05 Hz) resolved experimentally [1], thus its value is known only roughly.

is the equilibrium Boltzmann distribution. At the critical force where the molecule is on average half-open, the smallest non zero eigenvalue is $\lambda_1 = 2.1 \times 10^{-6} r$. This is the rate of the slowest mode of fluctuation, the "switching" of the boundary of the open region between $n \approx 3$ and $n \approx 22$. The remaining 21 eigenvalues are all well separated from the leading ones. The second largest eigenvalue is $\lambda_2 = 0.9 \times 10^{-4} r$ (Tab. 1). Thus the theoretical dynamics of P5ab involve one slow opening-closing transition, combined with many other transitions occurring more than 50 times faster. The net rates of opening ($k_o$) and closing ($k_c$) can be computed from $\lambda_1 = (k_o + k_c)$, and the ratio of the open and closed equilibrium probabilities equal to $k_c/k_o$. To compare our theoretical results with the experiments of [1] we have fitted our result for $k^* \equiv k_o = k_c$ at the force $f^*$ where the open and closed states have equal probability to experimental data [18], giving $r = 3.6 \times 10^6\ \sec^{-1}$.

Figure 6b shows time series from Monte Carlo simulations of the dynamical model of equation (4). Slow two-state switching is seen on a ≈0.25 sec timescale, on top of which occur much faster small fluctuations. When we convolve these data with a 20 Hz low-pass filter (as used experimentally [1]) the result Figure 6c is essentially the same as the experiment Figure 6a. The variation of the rates with force given by theory are also in good agreement with experiment (Figure 7). The transition state coordinate can be inferred from the relative slopes of the logarithm of the opening and closing rates of Figure 7 around the critical force [1,16], see Figure 5. These slopes and thus the transition state coordinate are independent of the fitted parameter $r$. As in the two states system of Figure 5, the transition state is located at the top of the free energy barrier, in $n^* = 11 - 12$ (Fig. 3a).

## 6 Barriers from generic helix-loop structures

Figure 6d shows a simulated time series for a molecule of the same length as P5ab, but with repeated AU sequence and no terminal loop. This might be realized in an experiment on a double helix with covalently crosslinked end bases, or some other arrangement where there is not a terminating "loop" of bases. In this case all base pairs in the interior of the sequence have the same $g_0(n) \equiv g_0^{AT} = 1.7\ k_BT$ giving a smooth energy landscape. The molecular motion is essentially diffusion in a flat free energy potential. This situation shows none of the two-state character of the P5ab RNA. For the repeated AU sequence of $n = 25$ base pairs, the switching time is just the diffusion-like time (see Tab. 1)

$$t^* \approx 2N^2/(\pi^2 r_o^*) = 2 \times 10^{-4}\ \sec \qquad (7)$$

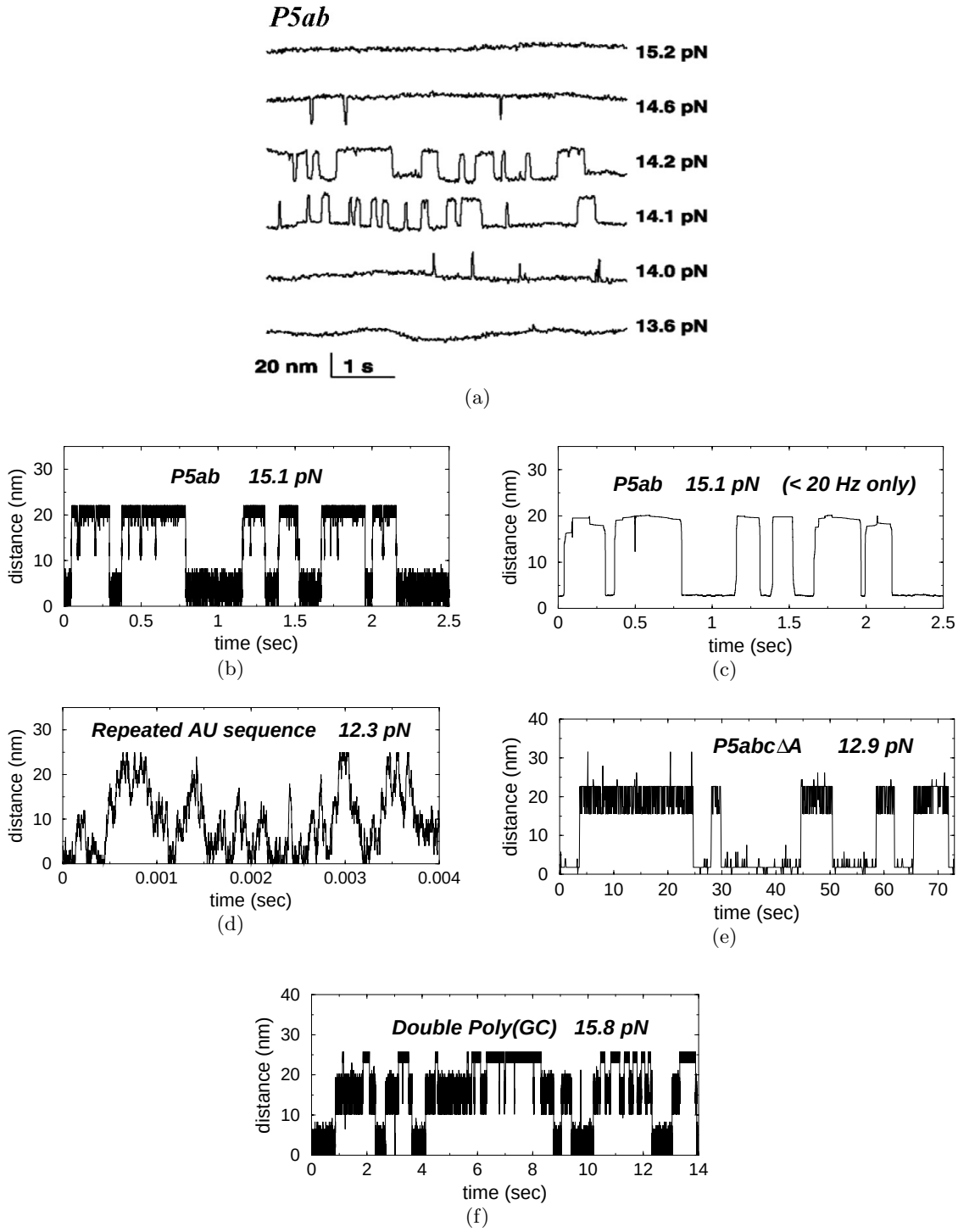


**Fig. 6.** Unzipping kinetics at the critical force. Distances between extremities of the molecule are shown for: (a) Experimental data for P5ab RNA showing two-state switching behavior [1]]; the relative time spent in each state depends on the applied force; (b) Theoretical time course for P5ab; there is a slow switching between the $n \simeq 3$ and $n \simeq 22$ configurations and fast transitions between configurations $n$ around these ones (Fig. 3a). (c) Convolution of (a) where oscillations faster than 20 Hz are averaged out; (d) Expected behavior for 50-base repeated AU sequence with no loop; experimentally the no loop condition can be realized by 25 AU base pairs terminated by a G-C helix, finally closed by a loop. At the critical force for the AU portion only this AU portion open and close because the critical force for the GC region is larger. The folding under the force strongly defavorize the formation of slipped bases. Note the 'ragged' time course, characteristic of diffusion on a time scale $\approx 10^3$ times shorter than in (a); (e) P5abc$\Delta$A; there is a slow switching between the closed molecule and the molecule with $H1$ opened and $H2$ opening or closing on a shorter time scale, the opening of $H3$ (distance between extremities of 31 nm) is a rare event at the force of 12.9 pN; (f) the hypothetical RNA molecule of Figure 2d.

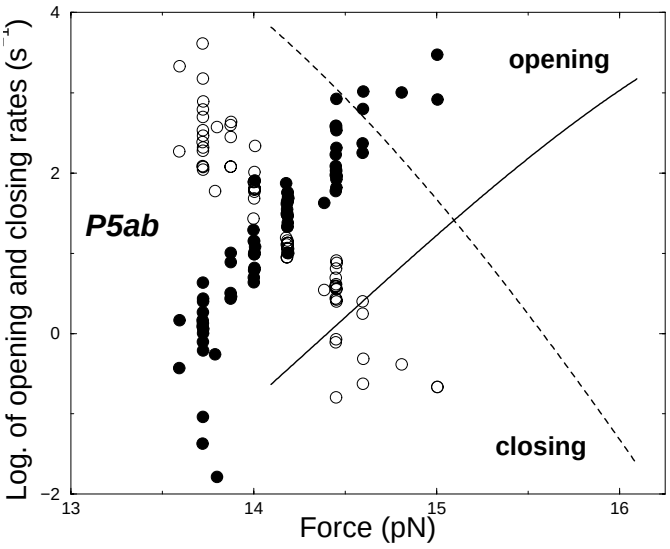


**Fig. 7.** Log. of the opening and closing rates for P5ab as measured in [1] as a function of force (full circles: opening, empty circles: closing), compared to theory (full line: opening, dashed line: closing). The theory fits the experimental data when shifting the critical force, at which the opening and closing rate coincide, from $f^* = 15$ pN to $f^* = 14.3$ pN to overlap the crossing point of theoretical curve with the experimental one. The slopes of $\ln k_o, \ln k_c$ do not depend on the fitted value of $r$. They give the relative positions $n_o = 8, n_c = 11$ of the transition state from the closed ($n = 3$) and open ($n = 24$) states respectively, with an absolute location in $n^* \simeq 12$.

where $r_o^* = r_c(f^*) = r\,\exp(-g_0^{AU})$. Thus, the 1 000-times longer switching time of P5ab comes from the 11 $k_BT$ barrier of Figure 2a.

The presence of a loop is sufficient to generate such a large barrier and, consequently, two-state switching. The simplest illustration is given by a homogenous DNA sequence ending with a loop. We have considered a 24-base Poly(GC) homogeneous-sequence helix terminated with a 4-base loop (Fig. 2b) [10]. Our theory predicts a two-state switching behavior on the same time scale as P5ab. The switching time is 50 times larger for a longer 8-base loop (Tab. 1).

The free energy barrier $G^*$ at criticality for a S-base-pair double-helix "stem" (uniform pairing free energy $g_0$) followed by a L-base loop (closing free energy $g_{\rm loop}(L)$ at zero force) (Fig. 1) can be simply estimated. The critical force $f^*$ is given by the condition that the free energy of the open molecule equals the free energy of the closed molecule, $G(0, f^*) = G(S, f^*)$, that is, $0 = S\,g_0 - (2S + L)\,g_{ss}(f^*) - g_{\rm loop}(L)$. The barrier height $G^*(S, L) \equiv G(S-1, f^*)$ then reads

$$G^*(S,L) = (S-1)(g_0 - 2\,g_{ss}(f^*)) = \frac{(S-1)\,(g_0\;L + 2\;g_{\rm loop}(L))}{L+2S}. \qquad (8)$$

Table 1 shows that, for a fixed helix length $S$, the critical force decreases with the length $L$ of the loop, while the free energy barrier $G^*$, and the switching time $t^*$ increase.

For short and random non-repeated sequences, (8) is only approximate when substituting $g_0$ with an average pairing free energy; it allows an estimate of how the switching time depends on $S$ and $L$. Note that the critical barrier essentially depends on the smaller of the two lengths $S, L$.

## 7 Branched-helix molecules and multiple-state-switching

Our approach can be extended to more complicated situations *e.g.* nucleic acids with branched structure. An example is the P5abc$\Delta$A RNA molecule [1] of Figure 2c, with free energy landscape as shown in Figure 3c. The opening of the first 12 bases (helix H1) follows as above, but going past the branch, description of the independent opening of the two helix regions requires a three-dimensional free energy landscape *i.e.* free energy as a function of the positions of the two unzipping boundaries [19]. A rich behavior emerges, shown in Figure 3e. At the critical force $f^* = 12.9$ pN, H1 switches on a long time scale $t^* = 10$ sec (Tab. 1), while the short lateral helix (H2) opens and closes with a much shorter characteristic time $t_2 \simeq 9$ msec. The lateral long helix (H3) opens very rarely at this force. Predictions for the rates and barrier heights are reported in Table 1.

It is possible to design molecules with multiple-state dynamics. Consider the molecule of Figure 2d, which has two well-bound Poly(GC) regions separated by an unpaired bubble, and terminated by a loop. For this system, three-state switching occurs (Figs. 2d, 3f), and should be observable in the frequency range accessible in experiments such as that of [1].

## 8 Opening dynamics of a long DNA

For long random sequence equation (8) is not applicable, because $g_0$ cannot be substitute by an average pairing free energy. For uncorrelated base pair distributions of the free energies $g_0(i)$, the free energy $G(n, f)$ (2) behaves as a random walk, resulting in a maximal barrier $G^* = O(\sqrt{N})$ [5]. In this section we focus on the fixed force "unzipping" dynamics of the 48 502 base pair phage $\lambda$ DNA molecule. This large molecule has a known and inhomogeneous sequence typical of genomic DNAs. To unzip this molecule, a rough free energy landscape must be traversed, with many local free energy minima separated by high barriers. The broad distribution of minima free energy and barrier heights must give rise to a series of jumps as the molecule unzips at constant force.

The base-pairing free energy for $\lambda$-DNA, shown in Figure 8, is calculated with the Mfold program [8], at room temperature and $Na^+ = 150$ mM. We consider unzipping starting from base 1, the direction studied experimentally [3,4]. The opening free energy at the critical force $f^* = 15.4$ pN, has a giant barrier of 3000 $k_BT$, due to the first half of the sequence being richer in GC content than the second one. This barrier, much larger than the

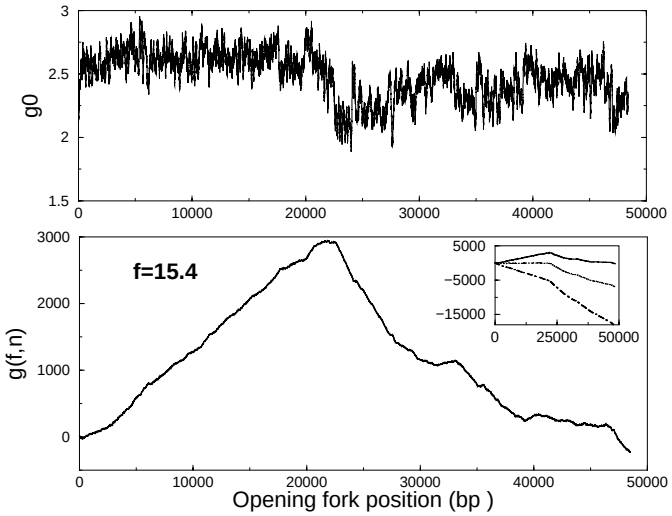


**Fig. 8.** Top: Pairing free energy along the sequence for the $\lambda$ DNA, calculated with the Mfold program for DNA, averaged over a window of 100 base pairs around each base. Bottom: Free energy landscape for the position of the opening fork at the critical force of 15.4 pN, and for forces of 16 pN (Inset, dotted line) and 17 pN (Inset, dot-dashed line). At the critical force there is a giant barrier of 3 000 $k_BT$ in the middle of the sequence between the GC rich first part and the AT rich second part.

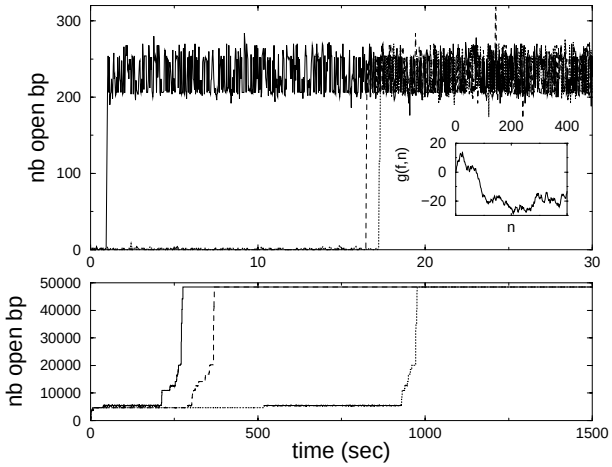


**Fig. 9.** Unzipping kinetics for $\lambda$ DNA. Time is in seconds, with the same value for $r = 3.6 \times 10^6$ sec$^{-1}$ as in previous figures. Top: at the critical force of 15.4 pN, opening the molecules requires first overcoming a barrier of the order of 10 $k_BT$ at the very beginning of the sequence (see the opening free energy at the corresponding force in Inset). The time at which the molecule overcomes this first barrier extends to a few tens of seconds, and is shown for three Monte Carlo runs (full, dashed and dotted lines). Only 300 base pairs open on a time scale of hours. The opening fork has indeed to pass a barrier of the order of 20 $k_BT$ at base pair 270. Bottom: opening kinetics at a force of 17 pN for three different Monte Carlo runs; after roughly one hour the molecule opens, after thermal activation over one large free energy barrier.

$\approx$200 $k_BT$ expected for a random sequence of 50 000 base pairs makes impossible to open the whole molecule and to observe switching behavior between the open and closed configurations. As force is increased, the free energy landscape is tilted (Fig. 5) and the barrier is decreased (Inset of Fig. 8).

Using the same microscopic rate $r = 3.6 \times 10^6$ sec$^{-1}$ as before, at the critical force the molecule opens only a few hundred base pairs on the time scale of hours (Fig. 9). For a force of 17 pN the molecule opens on the time scale of an hour and never closes again. In order to reduce the barriers to a level where the transition occurs in a reasonable experimental time of an hour requires forces well above the point where the fully open and fully closed states are in thermal equilibrium (15.4 pN).

Thus, unzipping at constant force proceeds by successive steps, each corresponding to the crossing of a free energy barrier. The location of these steps gives a kind of "fingerprint" which is specific to the DNA sequence. However, the kinetics of overcoming each barrier is a stochastic and slow process. There is thus a wide run-to-run variation of time needed for unzipping. This can be seen in Figure 9: complete unzipping is seen to take from 200 to 800 s. to occur, in different simulation runs. Experimentally, this might be best studied by observing a large number of DNAs at once as they are unzipped using constant force.

## 9 Conclusion

A few improvements might be added to this model. First, opening may occur through the cooperative nucleation of a few-base-pair bubble [6]. Second, mismatches might take place during closing [20], although they are highly limited by the presence of the 15 pN force. Finally it would be interesting to be able to describe unzipping events involving breaking of tertiary structures; these are thought to be present in the molecule P5abc$\Delta$A in presence of $Mg^{2+}$ [1,21].

A general result of our work is that a slow switching character should be quite generic for small biological molecules with helix loop structure. We note that our approach could also be used to analyze the opening-closing dynamics of nucleic-acid-detecting DNA "beacons" [22], both on their own, and in the presence of their targets. In this case the "unzipping" forces are applied by the hybridization interactions instead of by a large force transducer. Since such experiments amount to molecule recognition processes it is not impossible that slow barrier-crossing transitions of the sort discussed here occur *in vivo*.

### *Note added in proof*

While this article was in press, Danilowicz *et al.* reported constant-force experimental results for unzipping of lambda-DNA [Proc. Natl. Acad. Sci. USA **100**, 1694 (2003)]. Although the direction of unzipping was the opposite of that considered here, the observed kinetics are in quite good accord with our results (Sect. 8).

We thank P. Cluzel for helpful discussions and M. Zuker for advice on use of MFOLD. Work at UIC was supported by NSF Grants DMR-9734178 and DMR-0203963, by the Petroleum Research Foundation of the American Chemical Society, by the Research Corporation, by a Biomedical Engineering Research Grant from the Whitaker Foundation, and by a Focused Giving Grant from Johnson & Johnson. S.C. was partly funded by an A. della Riccia grant. R.M. was supported in part by the MRSEC Program of the NSF (DMR-9808595).

## References

1. J. Liphardt, B. Onoa, S.B. Smith, I.Jr. Tinoco, C. Bustamante, Science **297**, 733 (2001)
2. D. Poland, H.A. Scheraga, *Theory of Helix-Coil Transitions* (Academic Press, 1970)
3. U. Bockelmann, B. Essevaz-Roulet, F. Heslot, Phys. Rev. E **58**, 2386 (1997)
4. U. Bockelmann, P. Thomen, B. Essevaz-Roulet, V. Viasnoff, F. Heslot, Biophys. J. **82**, 1537 (2002)
5. D.K. Lubensky, D.R. Nelson, Phys. Rev. Lett. **85**, 1572 (2000); Phys. Rev. E **65**, 031917 (2002)
6. S. Cocco, R. Monasson, J.F. Marko, Proc. Natl. Acad. Sci. USA **98**, 8608 (2001); Phys. Rev. E **65**, 041907 (2002)
7. U. Gerland, R. Bundschuh, T. Hwa, Biophys. J. **81**, 1324 (2001)
8. M. Zuker, Curr. Opin. Struct. Biol **10**, 303 (2000); To reproduce the experimental ionic conditions (10 mM Tris, 250 mM Na+, 10 mM Mg2+ or EDTA), a correction on the free energy equal to –0.193 $k_BT \ln([Na^+]+3.3[Mg]^{0.5})$ (suggested by M. Zuker) has been applied. Pairing free energies for Non Watson-Crick GA base pairs have been obtained from their GU counterparts, with $\simeq 1$ $k_BT$ corrections measured in M. Meroueh, C.S. Chow, Nucl. Acids Res. **27**, 1118 (1999). Calculations for DNA sequence have been performed on DNA MFOLD server, at room temperature and 0.15 $Na^+$, using free energy parameters from J.Jr. Santa Lucia, Proc. Nat. Aca. Sci. USA **95**, 1460-1465 (1998)
9. Y. Zhang, H. Zhou, Z.C. Ou-Yang, Biophys. J. **81** 1133, (2001)
10. C. Bustamante, S.B. Smith, J. Liphardt, D. Smith, Curr. Opin. Struct. Biol. **10** (2000)
11. S.B. Smith, Y. Cui, C. Bustamante, Science **271**, 795 (1996)
12. M.N. Dessinges, B. Maier, Y. Zhang, M. Peliti, D. Bensimon, V. Croquette, Phys. Rev. Lett. **89**, 248102 (2002)
13. S. Cocco, J. Marko, R. Monasson, A. Sarkar, J. Yan, Eur. Phys. J. E **10**, 249 (2003)
14. J.H. Cate *et al.*, Science **273**, 1678 (1996)
15. H.A. Kramers, Physica (Utrecht) **7**, 284 (1940)
16. E. Evans, K. Ritchie, Biophys. J. **72**, 1541 (1997)
17. M. Doi, S.F. Edwards, *The Theory of Polymer Dynamics* (Clarendon, Oxford, 1986)
18. Possible effects of the feedback mechanism on the force used in [1] (private communication) are not considered here, but could be included in our model
19. P5abc$\Delta$A includes three helices with $N_1 = 12$ (H1), $N_2 = 5$ (H2), and $N_3 = 9$ (H3) base pairs. A single (respectively double) boundary between open and closed regions arises if less (resp. more) than 12 base pairs are open. The number of configurations of the molecule is thus $N = N_1 + (N_2 + 1)(N_3 + 1) = 72$. From configurations where H2 and H3 are partially opened, four transition steps are possible
20. A. Ansari, S.V. Kuznetsov, Y. Shen, Proc. Natl. Acad. Sci. USA **98**, 7771 (2001)
21. X. Fang, T. Pan, T.R. Sosnick, Biochemistry **38**, 16840 (1999)
22. G. Bonnet, S. Tyagi, A. Libchaber, F.R. Kramer, Proc. Natl. Acad. Sci. USA **96**, 6171 (1999)