

## Force fluctuations assist nanopore unzipping of DNA

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2010 J. Phys.: Condens. Matter 22 454122

(<http://iopscience.iop.org/0953-8984/22/45/454122>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 158.144.176.250

The article was downloaded on 30/10/2010 at 06:59

Please note that [terms and conditions apply](#).

# Force fluctuations assist nanopore unzipping of DNA

V Viasnoff, N Chiaruttini, J Muzard<sup>1</sup> and U Bockelmann

Nanobiophysics Lab, UMR CNRS Gulliver, ESPCI, 10 rue Vauquelin, F-75005 Paris, France

E-mail: [virgile.viasnoff@espci.fr](mailto:virgile.viasnoff@espci.fr)

Received 27 April 2010, in final form 8 July 2010

Published 29 October 2010

Online at [stacks.iop.org/JPhysCM/22/454122](http://stacks.iop.org/JPhysCM/22/454122)

## Abstract

We experimentally study the statistical distributions and the voltage dependence of the unzipping time of 45 base-pair-long double-stranded DNA through a nanopore. We then propose a quantitative theoretical description considering the nanopore unzipping process as a random walk of the opening fork through the DNA sequence energy landscape biased by a time-fluctuating force. To achieve quantitative agreement fluctuations need to be correlated over the millisecond range and have an amplitude of order  $k_B T/\text{bp}$ . Significantly slower or faster fluctuations are not appropriate, suggesting that the unzipping process is efficiently enhanced by noise in the kHz range. We further show that the unzipping time of short 15 base-pair hairpins does not always increase with the global stability of the double helix and we theoretically study the role of DNA elasticity on the conversion of the electrical bias into a mechanical unzipping force.

(Some figures in this article are in colour only in the electronic version)

## 1. Introduction

Proteinaceous and artificial nanopores are under increasing investigation as ways to efficiently detect and count single molecules. The transient blockade of the pore ionic current concomitant to the translocation of a molecule of interest has been studied to characterize the passage of single-stranded DNA (ssDNA) [1, 2], proteins [3–5], neutral polymers [6] or to study local chemical reactions [7]. Nanopores can also be used as force transducers. The electrophoretically driven translocation of molecules larger than the pore diameter can be obtained at high enough bias provided that the molecules deform to accommodate into the pore lumen. Double-stranded DNA (dsDNA) [8–10], loosely folded proteins [4, 11] or DNA-bound enzymes [12–15] can be forced to translocate by application of a suitable voltage bias. In the case of dsDNA the translocation process can be associated with unzipping of the base-pair sequence if the pore is smaller than 2 nm. Force spectroscopy [16, 17] or constant force unzipping [8, 9] have been performed on rather short dsDNA (10–20 bp). Unlike other micromanipulation techniques, such as optical or magnetic trapping, nanopore unzipping does not allow the

measurement of the unzipping force nor of the position of the DNA strand in the pore. Instead, a controlled biasing voltage is applied and the distributions of total DNA passage time are measured. The electrophoretic drive plays the role of the magnetic or optical force but cannot be directly measured. To a first approximation it can be related to the applied voltage  $V$  as  $F = q_{\text{eff}} * \frac{V}{d}$ , where  $d$  is the pore length and  $q_{\text{eff}}$  is an effective charge, the value of which ranges from 0.1 to  $0.4 e^-/\text{bp}$ . This approach was successfully used to explain unzipping of short DNA hairpins under constant or ramped voltages. Several studies tried to decipher the mechanism of DNA unzipping in nanopores as a function of the pore size [18], sequence length and base composition [8–10, 19, 20]. So far, the theoretical descriptions of nanopore unzipping consider the diffusion of the unzipping fork at constant force or constant loading rate over a single effective energy barrier for short hairpins [9, 21] or in a sequence-dependent energy landscape for larger sequences [22, 23, 17, 21, 24]. All-atom molecular dynamics simulations were performed to look at the unzipping process of short hairpins in artificial or proteinaceous pores [25–27]. However, to date, all models fail to quantitatively predict the translocation time of dsDNA in nanopores although the same approaches accurately describe DNA unzipping using optical or magnetic traps.

<sup>1</sup> Present address: School of Chemistry and Chemical Biology, University College Dublin, Republic of Ireland.

**Table 1.** Hairpin and duplex sequences used in this study. Duplex sequences given here are hybridized to their complementary part attached to a polyA moiety.

Hairpins	
hp1	5' <b>AGGCTGCGGCGTCGGATA</b> <i>GTTG</i> TATCCGACGCCGCAGCCT(A) <sub>30</sub> 3'
hp2	5' <b>AGGCTGCGATTTCGTCCTG</b> <i>GTTG</i> CAGGACGAATCGCAGCCT(A) <sub>30</sub> 3'
hp3	5' <b>AGGCTGCGATCTGTTCGT</b> <i>GTTG</i> ACGAACAGATCGCAGCCT(A) <sub>30</sub> 3'
hp4	5' <b>AGGCTGCGTTATCGGTTC</b> <i>GTTG</i> GAACCGATAACGCAGCCT(A) <sub>30</sub> 3'
hp5	5' <b>AGGCTGCGAAATGCTGTT</b> <i>GTTG</i> AACAGCATTTGCGCAGCCT(A) <sub>30</sub> 3'
Duplex	
Dupl	5' <b>AGGCTGCGGCATTTTGTCCGCGCCGGGCTTCGCTCACTGTTTCAGG</b> 3'

It was previously reported that the nanopore unzipping process depends on the pore geometry [18, 28], the base-specific pore/DNA interactions [29] and local base-pairing stability [20, 28]. In this paper we argue that considering the base-pair stabilities alone is not enough to account for nanopore unzipping experiments. Force fluctuations and/or local molecule elasticity may also play an important role. In the first part, we show that the voltage dependence of the translocation of a 45-mer DNA duplex can be quantitatively explained by the presence of correlated time fluctuations of order  $1k_B T/\text{bp}$  in the unzipping force. In the second part we show experimental evidence that 15-mer hairpins do not always unzip according to their global stability. We hypothesize that this behavior may partly originate from cooperative opening of several bases due to local differences in molecular elasticity. We provide an illustrative 1D model of force spreading along a DNA molecule unzipped in a shear mode.

## 2. Materials and methods

### 2.1. Hairpins and duplex

All DNA oligonucleotides were purchased PAGE purified from Eurogentec SA. The hairpins were slowly annealed after thermal denaturation at 95 °C. They were used at a final concentration of 2.5  $\mu\text{M}$ . For the duplex, the amount of single-stranded DNA in the solution was minimized by an agarose gel purification (without staining) of the hybridized strands. The sequences of the molecules are given in table 1. Common features of the sequences are highlighted in bold and loops are hyphenated. Hairpin sequences are given in full. The duplex sequence is complemented by the appropriate other DNA strand with a polyA tail of 30 bases at its 3' end. We used TE Buffer 1x, pH = 7.4 supplemented with 1 M KCl. All experiments were carried out at 24 °C.

### 2.2. Experimental set-up

We used a classical nanopore set-up already described in [20]. Black lipid membranes were formed on a hole of 30  $\mu\text{m}$  in diameter, punctured in a Teflon septum. The DPhPC (Avanti Lipids) bilayer was about 10  $\mu\text{m}$  in diameter with a lipid annulus around. This configuration allowed the membranes to resist voltages as large as 350 mV without rupture. Alpha hemolysin was purchased from Sigma.

Melting curves were obtained by measuring the change in absorbance at 260 nm of the DNA hairpins at 1  $\mu\text{M}$ . The

temperature was varied by 0.2 °C  $\text{min}^{-1}$  between 20 and 95 °C. The following buffer was used: LiCaco 20 mM, KCl 100 mM, pH 7.2. The dissociation coefficients were extracted using the standard protocol described in [30].

### 2.3. Numerical calculation of the first passage time

We theoretically describe the translocation of DNA through a nanopore by the biased diffusion of a fictitious point particle representing the unzipping fork in the energy landscape  $E_j$  created by the unzipping of the first  $j$ th bases. We coined  $w$  the mechanical and electrostatic energy difference  $w$  between two bases paired on the *cis* side of the membrane and two bases, unpaired, one on each side of the membrane. We assume that it does not depend on the sequence. The bases are numbered from 1 to  $N$  starting at the first base after the poly A overhang. We call  $E_j$  the free enthalpy needed to unzip the first  $j$ th bases. The unzipping energy landscape then is

$$E_j = \sum_{i=1}^j (\Delta H_i - T \Delta S_i) - j w \quad (1)$$

where  $\Delta H_i$  and  $\Delta S_i$  are the tabulated bulk melting values of each base pairs according to the nearest-neighbor model [31]. The energy  $w$  linearly biases the energy landscape estimated for DNA thermal melting in solution. Distributions of DNA unzipping time are calculated from the distributions of the first passage time of the unzipping fork at base  $N$ . In a previous work [24] we solved the diffusion dynamics in this landscape by thermally activated jumps from one to the next base pairs using a Gillespie method [32]. This method is not very efficient at calculating very long translocation times. In order to compute first passage times over 13 decades in time, we adopt a more continuous approach by solving the Fokker-Planck equation associated with the diffusion process. Let  $f(x, t)$  be the probability of finding the unzipping fork at position  $x$  at time  $t$ . The time evolution of  $f$  follows

$$\partial_t f = \partial_x \left\{ D \left[ \partial_x + \frac{1}{k_B T} \partial_x E_j \right] f \right\}. \quad (2)$$

The equation was solved by matrix multiplications representing discrete spatial steps equal to one base. The diffusion coefficient  $D$  is the only fitting parameter. It sets the unitary timescale  $t_o$  as  $\sqrt{D t_o} = 1$  base. The boundary conditions are reflecting at  $x = 1$  to prevent backward escape, and fully absorbing at  $x = N$ . The cumulative distributions of first passage time  $P_w(\tau)$  are calculated as  $1 - \int_0^N f(x, t) dx$ .

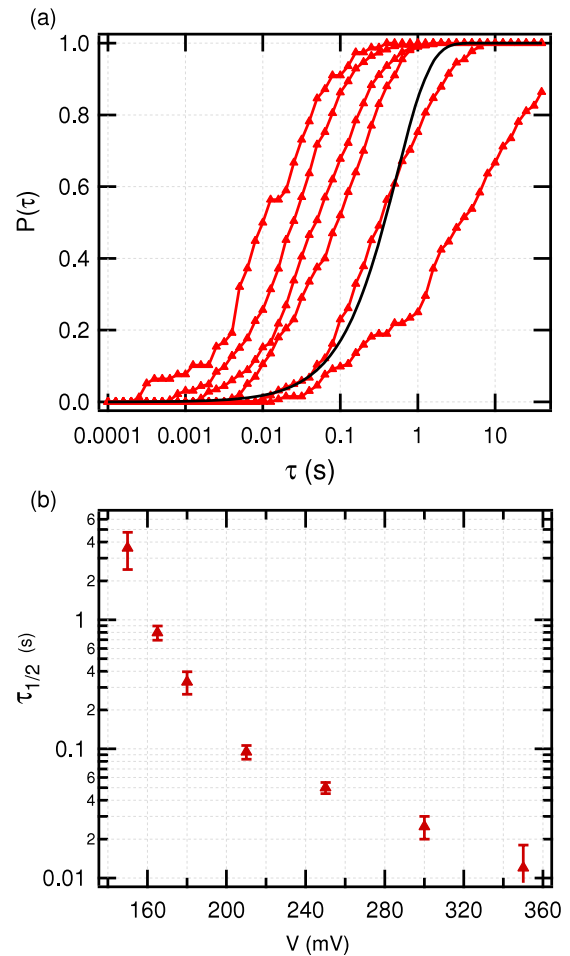
We use  $f(x, 0) = \delta(0)$  as the initial condition. In the range of common accessible times the Gillespie method or solving the Fokker–Planck equations led to similar results (data not shown).

### 3. Results and discussion

#### 3.1. Unzipping of long DNA duplexes: the role of force fluctuations

DNA unzipping under constant force using magnetic tweezers [33] have been implemented on DNA duplexes of several kilobase pairs. Above a certain force threshold (between 12 and 15 pN) the DNA mechanically melts. Around the transition a gradual opening of the molecules was observed with transient pauses at tightly bound regions resulting from a large local GC content. The unzipping time thus depends not only on the global duplex stability but on the local shape of the unzipping energy landscape. The same behavior could be expected for nanopore unzipping. In that case, however, the DNA duplexes are significantly shorter (typically less than 100 bp). For so few bases other parameters than base-pairing stability may also affect the unzipping dynamics. Among those are DNA/pore interactions, geometrical constraints and confinement, detailed process of single base unzipping and local distribution of mechanical stress. For very short sequences (10 bp or less) a small number of molecular events are expected to allow the translocation of the molecule. Indeed it has been reported that the translocation of short hairpins follow a Kramers description and can be reduced to a thermal passage over an energy barrier (Kramers event) corresponding to the opening of five bases [9]. This description was further refined allowing a finite escape time of the freed molecule [17, 21]. However, the value of the energy barrier could not be deduced directly from the sequence. For longer sequences the Kramers picture becomes inappropriate and it was demonstrated [24] that the unzipping process can be described by the sequential opening of the individual base pairs. One can expect that the detailed mechanisms of base-pair opening are averaged and that the DNA unzipping dynamics is mainly controlled by the large scale variations of base composition. As a result, the unzipping time depends mostly on the details of the energy landscape [24, 20]. In section 3.2 we study the voltage dependence of DNA unzipping of a 45-mer duplex coined *Dupl* and described in table 1. A previous study [20] showed that 45 bp is enough to consider the variations of base composition as the main factor controlling the DNA translocation dynamics. We then provide a random walk model under fluctuating force that quantitatively accounts for the voltage dependence of the unzipping dynamics under constant voltage.

We start by determining experimentally the voltage dependence of the 45 bp DNA duplex coined *Dupl* described in table 1. Figure 1(a) shows the translocation time cumulative distributions  $P(\tau)$  at biasing voltages ranging from 150 to 350 mV. The voltage dependence of the median time  $\tau_{1/2}$  defined as  $P(\tau_{1/2}) = 0.5$  is given in figure 1(b). Should the unzipping be a simple Kramers process over an



**Figure 1.** Experimental results. (a) Cumulative distributions of the translocation times  $P(\tau)$  for the 45-mer duplex *Dupl* at various voltages. From right to left,  $V = 150, 180, 210, 250$  and  $350$  mV. Translocation experiments were performed at  $24^\circ\text{C}$ , pH 7.4, 1 M KCl. The black solid lines correspond to an exponential fit of the distribution at 180 mV:  $P(\tau) = 1 - \exp(-\alpha\tau)$ . (b) Median time  $\tau_{1/2}$  of the distributions as a function of the applied voltage  $V$ . Notice that  $\tau_{1/2}$  does not vary exponentially with  $V$ . Doubling  $V$  results in a 2.5 decade decrease in  $\tau_{1/2}$ .

activation energy barrier  $E^*$ , one would expect, i—cumulative distributions  $P(\tau)$  of the form  $P(\tau) = 1 - \exp(-\tau/\tau_o)$  and ii—median times  $\tau_{1/2} = \ln(2) * \tau_o$  that scale as  $\tau_{1/2} \propto \exp(\frac{E^* - q_{\text{eff}}V}{k_B T})$ . Figure 1(c) shows that the predicted functional form does not match well our experimental distributions that are wider than exponential. Moreover, the log–lin plot of  $\tau_{1/2}$  versus  $V$  in figure 1 shows that  $\log(\tau_{1/2})$  does not vary linearly with  $V$ , especially for voltages below 200 mV. At low voltages  $\tau_{1/2}$  diverges more rapidly than exponentially with the voltage. Since none of the above features are experimentally observed, we conclude that the unzipping process cannot be described by two states (open and closed) separated by a fixed effective energy barrier. Spatial coordinates and the existence of several unzipped states should be taken into account.

We thus describe the unzipping process as a random walk of the opening fork. This dynamic description is similar to the approaches proposed to explain DNA unzipping using magnetic traps at constant force [33]. Unzipping of

DNA can be described theoretically as a random walk of a fictitious particle (representing the opening fork) in a free energy landscape. The landscape is calculated as the free enthalpy gain of the system as a function of  $i$ —the position of the DNA in the pore [23] and  $ii$ —the number of open base pairs [23, 24, 22]. The action of the transmembrane voltage is modeled by a bias of the energy landscape. The bias can be implemented in two ways. A ‘ratchet-like’ approach: the entry rate of a single base in the pore depends on the voltage, whereas the opening and closing of each base pair [23] is only thermally activated without any voltage dependence. Or a ‘mechanical’ approach: the opening fork always stands at the pore constriction, the threading of the single strand is instantaneous and the electrophoretic drive biases the opening and closing rates of the last closed base pair [34, 24]. In the following, we describe a 1D model that takes this ‘mechanical’ approach.

Let us assume that the opening fork always stands at the pore constriction, the threading of the single strand is instantaneous and the electrophoretic drive biases the opening and closing rates of the last closed base pair. The energy landscape is built as a function of the base number  $j$  as

$$E_j = \sum_{i=1}^j (\Delta H_i - T \Delta S_i) - j w \quad (3)$$

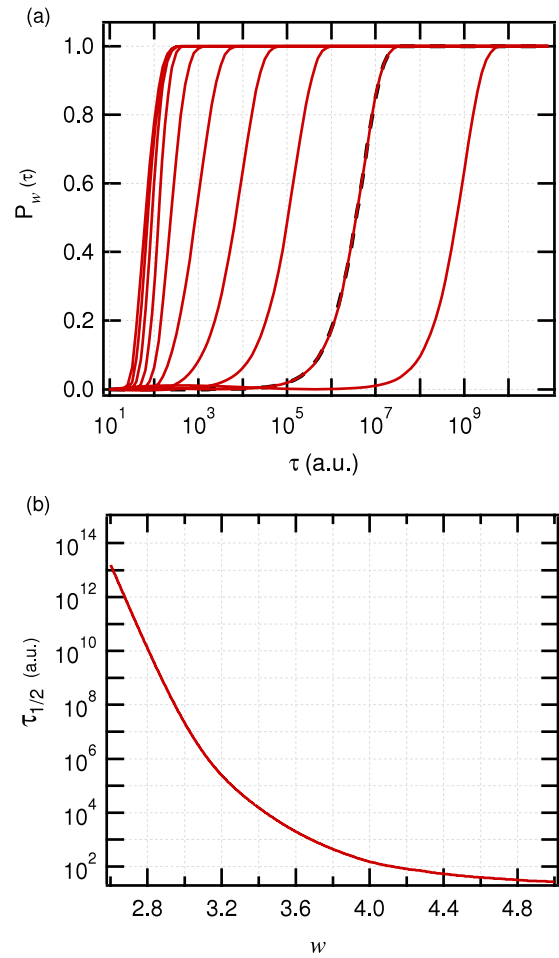
where  $\Delta H_i$  and  $\Delta S_i$  are the tabulated bulk melting values of each base pair according to the nearest-neighbor model [31]. The mechanical and electrostatic energy difference  $w$  between two bases paired on the *cis* side of the membrane and two bases, unpaired, one on each side of the membrane, is assumed to be sequence-independent. The energy landscape is thus tilted linearly by  $w$ . The translocation time of a particle equals its first passage time at base  $N$ . The latter is calculated numerically by solving the associated Fokker–Planck equation by discrete steps equal to one base, as described in section 2.

Figure 2(a) shows the distributions of translocation time calculated for the *Dupl* sequence at various biasing energies  $w$  ranging from 2.5 to 5  $k_B T$ . The dependence of the median times  $\tau_{1/2}$  with  $w$  is also represented in figure 2(b). The model makes both the following predictions:

- (i) For  $w$  smaller than  $\sim 3.3 k_B T$ , the distributions can be well fitted by a single exponential  $P_w(\tau) = 1 - \exp(-\tau/\tau_o)$  (see figure 2(a)). For values of  $w$  above 3.3  $k_B T$ , the distributions become steeper than exponential.
- (ii)  $\tau_{1/2}$  diverges very quickly at low biasing energy  $w$  and more slowly at low  $w$ . Doubling  $w$  results in an 11 orders of magnitude reduction of the translocation time.

Both effects can be explained by the shapes of the energy landscape  $E_j$  depicted in figure 3 for various  $w$ . At low  $w$ , the energy landscape has local minima that pin the unzipping fork. The pinning regions are highlighted in bold.

For  $w < 3k_B T$  the landscape displays a single minimum localized at  $n = 0$  corresponding to the fully closed molecule. The unzipping process is thermodynamically unfavorable and occurs only through a fluctuation over a very large effective energy barrier corresponding to the whole energy profile.



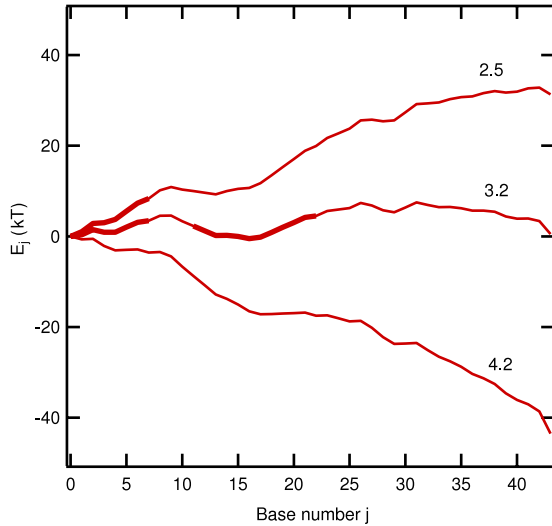
**Figure 2.** Predicted values of the distribution of first passage times assuming constant biasing energy  $w$ . (a) Cumulative distributions of first passage times  $P_w(\tau)$  for the *Dupl* sequence at various  $w = 2.89, 3.1, 3.25, 3.5, 3.7, 3.9, 4$  and  $4.2 k_B T$  (right to left). For small  $w$ , the distributions nicely fit to a simple exponential (dashed line). At larger  $w$ , the distributions become increasingly sharp. (b) Median first passage times  $\tau_{1/2}$  versus  $w$ . Notice the very strong dependence of  $\tau_{1/2}$  with  $w$  for  $w < 3.2k_B T$ . Doubling the bias energy results in an 11-decade decrease of the translocation time.

This explains the very large values of  $\tau_{1/2}$  at low  $w$  and the exponential form of  $P_w(\tau)$ .

For  $3 < w < 3.8k_B T$ , the unzipping process is thermodynamically favorable but should still occur by an activated jump over an energy barrier with a voltage-dependent width. Furthermore, in the present energy landscape a second minimum appears around base 15 that localizes the unzipping fork. The activation energy for DNA unzipping is then due to bases  $n > 15$ . The escape process is still thermally activated but the effective energy barrier is changed. It results that  $\tau_{1/2}$  has a non-monoexponential dependence with  $w$ .

For  $w > 3.8k_B T$ , there is no local minimum anymore. The unzipping fork is no longer pinned. The unzipping process is not dominated by a Kramers escape rate but amounts to a biased diffusion in a smooth landscape. In this ballistic limit  $\tau_{1/2}$  varies linearly with  $w$ . A detailed discussion of these regimes can be found in [24].





**Figure 3.** Energy landscape  $E_j$  for the *Dupl* sequence at  $w = 2.5k_B T$ ,  $3.2k_B T$  and  $4.2k_B T$ . Monitoring  $f(n, t)$ , the probability of finding the unzipping fork on base  $n$  at time  $t$ , we identified the pinning region of the unzipping process. These regions are highlighted in bold. At low  $w$  the unzipping fork is localized around base zero and translocation occurs by an activated jump over the energy barrier made by the whole sequence. At intermediate  $w$ , a second minima appears that pins the unzipping fork around base 15. The activation barrier is reduced. At large  $w$  the landscape does not present any minima and the translocation is no longer an activated process.

In order to relate the theoretical predictions to the experimental data we needed to evaluate the biasing energy applied to an individual base  $w$  as a function of  $V$ . In leading order  $w$  is expected to be proportional to the biasing voltage  $w = q_{\text{eff}} V$ . The effective charge  $q_{\text{eff}}$  should depend not only on the charge of the DNA but also on other transduction factors such as the local geometry and elasticity. It is left as a fitting parameter. According to the model, doubling of  $w$  results in a 11 orders of magnitude decrease of the predicted  $\tau_{1/2}$ . However, we experimentally observe a 2.5 orders of magnitude decrease upon doubling  $V$ . Hence the predicted localization of the unzipping fork is much too strong compared to our experimental observations. In addition, over the whole experimental voltage range, the measured distributions  $P_w(\tau)$  always span several time decades, while the model predicts a ballistic regime with very narrow distributions  $P_w(\tau)$  at large  $w$ . It is a clear indication that the ballistic regime is not observed in the available voltage range.

The model hence needs to be modified in order to fit our experimental results. In particular, the pinning transition needs to be more progressive. Since the nonlinear dependence of the translocation time with  $V$  is strong, one can expect that small temporal fluctuations of the biasing energy  $w$  may result in a large effect on the distribution of translocation times. Indeed  $w$  is prone to fluctuate, even at fixed applied voltage, since it should depend on the pore/DNA configuration, the instantaneous charge distribution and the instantaneous pore configuration. The translocation process might be assisted by thermal fluctuations in the biasing force. To investigate this idea, we implement such fluctuations by introducing a

time-dependent energy landscape calculated as

$$E_j(t) = \sum_{i=0}^j (\Delta H_i - T \Delta S_i) - j w(t) \quad (4)$$

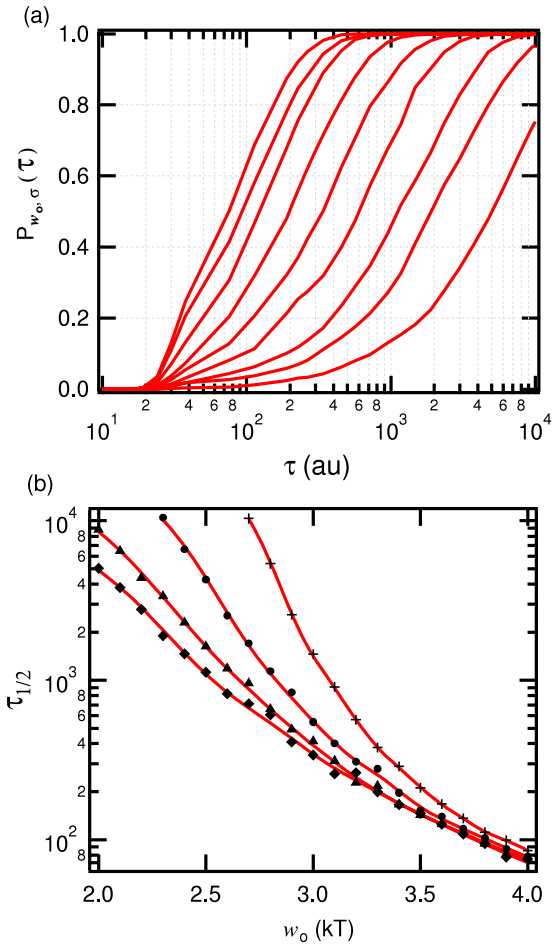
where  $w(t)$  is now a time-dependent stochastic variable with a Gaussian probability  $p(w) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\frac{(w-w_0)^2}{2\sigma^2})$  centered on  $w_0$  with  $w_0 = q_{\text{eff}} V$  and with a voltage-independent spread  $\sigma$  of order  $1k_B T$ . In order to simplify the resolution of the dynamic equation, we assume that  $w(t)$  is constant by part during a correlation time  $\tau_{\text{corr}}$ . We also assume that the fluctuations of  $w(t)$  happen on a much longer timescale than the thermal fluctuations acting on individual base pairs. Thermal noise and force fluctuations are thus additive. Strictly speaking, the Fokker–Planck equation is a noiseless equation derived from a Langevin equation with a random noise. The expression we use is valid only for white uncorrelated thermal noise. If the random force is ‘colored’ and correlated [35], the present expression of the Fokker–Planck equation is no longer valid. However, in the present case since  $w$  is constant by part, our treatment of the noise is correct. Under these assumptions the diffusion dynamics can still be described by the usual Fokker–Planck equation. For every  $\tau_{\text{corr}}$  the value of  $w$  is randomized to an uncorrelated value of probability  $p(w)$ . The unzipping probabilities are averaged over 500 noise realizations for each  $w_0$ . We thus have the following fitting parameters:  $q_{\text{eff}}$  that determines the force transduction efficiency,  $D$  that determines the unitary timescale of the simulation and  $\tau_{\text{corr}}$  that sets the correlation time of the fluctuations.  $\sigma$  is expected to be of order  $1k_B T$ .

Let us analyze different limits for the noise correlation time.

In the limit of  $\tau_{\text{corr}} \sim 1$  (a.u.), i.e. around the unitary time required to move by only one base, the noise effect is very small. The value of  $w$  is randomized at each step. Since the translocation process involves many single-base opening steps, it amounts to biasing the energy landscape by the value  $\bar{w} = w_0$ . This limit is equivalent to our previously described model at a fixed  $w$ .

The limit  $\tau_{\text{corr}} \rightarrow \infty$  depicts a situation where each single molecule translocates with a random but fixed  $w$ . The probability distributions of unzipping times should then be the average of the probabilities obtained by the previous model for  $w$  varying by a few  $k_B T$ . The typical translocation times for these distributions span the full time range covered by a  $1k_B T$  variance in  $w$ . This range is around 7–8 decades in time for small  $w_0$ . It would thus mean that the distributions should be extremely large with most molecules remaining stuck in the pore. Experimentally, though, we observe that almost all molecules end up translocating over two decades in time.

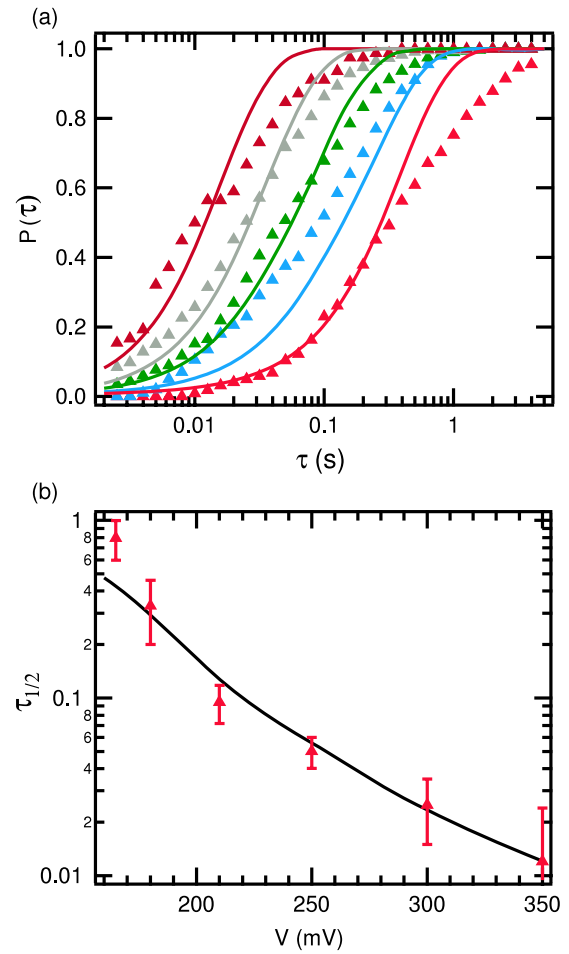
We thus investigate the regime where  $\tau_{\text{corr}}$  is of the same order of magnitude as the translocation time of DNA for rather large values of fixed  $w$ . Assuming  $\sigma = 1k_B T$ , we study the dependence of  $\tau_{1/2}$  versus  $w_0$  for various  $\tau_{\text{corr}}$ . As shown in figure 4 the effect of  $\tau_{\text{corr}}$  is rather strong. For  $\tau_{\text{corr}} = 30$  a.u. we find that  $\tau_{1/2}$  now decreases by only 2.5 orders of magnitude upon doubling  $w_0$ . The first passage time distributions are also more stretched than in the preceding



**Figure 4.** Effect of a Gaussian correlated noise of amplitude  $\sigma = 1 k_B T$  on the unzipping process. (a) Cumulative distributions of the first passage times  $P_{w_o, \sigma}(\tau)$  for *Dupl* sequence at various  $w_o$  and a correlation time  $\tau_{corr} = 30$  a.u.  $w_o = 2, 2.3, 2.5, 2.8, 3.0, 3.3, 3.6, 3.8$  and  $4k_B T$  (right to left). The distributions  $P_{w_o, \sigma}(\tau)$  are wider than in the previous model especially at large  $w_o$ . (b) Dependence of  $\tau_{1/2}$  with  $w_o$  for  $\sigma = 1 k_B T$  and  $\tau_{corr} = 10, 20, 30$  and  $40$  a.u. Notice the reduction of the divergence of  $\tau_{1/2}$  with increasing correlation time.

model, especially at large  $w_o$ . Introducing Gaussian correlated noise in the value of the biasing energy thus greatly reduces the dependence of  $\tau_{1/2}$  with  $w_o$ . It can be crudely understood by saying that, even if  $w_o$  lies in the highly pinned regime, the molecules translocate when the fluctuation is large enough to allow its unzipping during the correlation time. An analytical description of this effect can be found in the appendix.

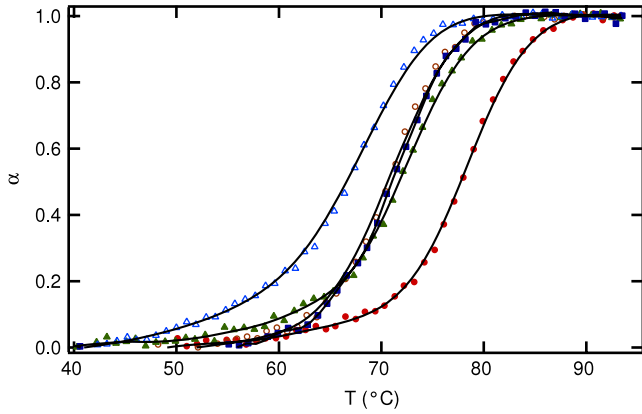
Our experimental data can now be fitted as shown in figure 5. We find  $q_{eff} = 0.22e$  (in agreement with previous publications [9, 23, 36]) and  $D$  is  $9 \text{ base}^2 \mu\text{s}^{-1}$  corresponding to a unitary time step of  $1.1 \mu\text{s}$  (in agreement with a typical DNA single-base unzipping and translocating rate [23, 20, 37]). The best fit is obtained for  $\sigma = 1.2k_B T$ . Our model then predicts that the correlation time of the fluctuations  $\tau_{corr} \sim 2$  ms. The fit of the unzipping time probabilities is given in figure 4(a) at various  $w_o$ . Although the predicted distributions do not perfectly fit the experimental curves, the agreement is very significantly improved by the introduction of time-correlated fluctuations in the bias energy.



**Figure 5.** (a) Fit of the experimental curves (symbols) by the predicted first passage time distributions for the following values of fitting parameters:  $\sigma = 1.2k_B T$ ,  $D = 9 \text{ base}^2 \mu\text{s}^{-1}$ ,  $q_{eff} = 0.22$  and  $\tau_{corr} = 2.1$  ms. (b) Fit of the dependence of the experimental median translocation time. The introduction of fluctuations in the biasing energy leads to a significantly improved agreement between experiment and theory.

This indicates that a stochastic resonance [35, 38] affects the unzipping process, in the sense that the presence of noise greatly facilitates the translocation. One might expect that the agreement between experiment and theory could be further improved by refining our description of the noise. A more rigorous treatment of the problem would be to solve a 2D master equation using the Gillespie method with both fork position and  $w$  as stochastic variables with appropriate transition rates. This approach is left for future work.

In addition to force fluctuations, one could also expect that the unzipping energy landscape in a nanopore differs from the landscape derived from the nearest-neighbor model measured in bulk. Entropic differences between the unzipped and zipped states are likely to be reduced due to the confined environment. This effect should tend to stabilize the paired state and minimize the effect of temperature. In contrast, the possible interactions through hydrogen bonds of the bases with the pore hydrophilic residues could lower the enthalpic contribution of the melting energy and favor DNA unzipping. The pore could then have a kind of ‘catalytic’ action on the



**Figure 6.** Thermal melting curves of the five short hairpins: the dissociation coefficient  $\alpha = \frac{[ssDNA]}{[dsDNA]}$  is plotted as a function of temperature for hp1 (●), hp2 (▲), hp3 (○), hp4 (■) and hp5 (△). The melting temperatures are ordered according to the predicted hairpin stabilities.

duplex unzipping process. In the absence of quantitative measurements of these effects, it is hard to conclude what is the main contribution of the pore on DNA unzipping.

### 3.2. Unzipping of short hairpins: putative role of cooperative unzipping

The model presented above assumes that the kinetics of DNA unzipping is governed by the duplex energy landscape. The influence of DNA stability on the distributions of translocation times was usually performed on small hairpins (around 10 bases). The sequence stabilities were modified either by lengthening the sequence [9, 19] or by introducing mismatches [8, 10, 19]. These modifications not only change the total base-pairing energy by several tens of per cent but also modify the number of base pairs to open. It was found that the translocation times of these molecules correlate either quantitatively [39] or qualitatively [17, 19] with the hairpin stability. These translocation process were described by a simple Kramers process. However, in a Kramers event the height of the energy barrier may not correlate with the stability difference between both equilibrium states. Furthermore, cooperative opening of bases as well as local force spreading along the DNA molecule could be expected to play an important role in the dynamics of unzipping of short molecules. In this section we experimentally show that short hairpins do not always translocate according to their global stabilities as usually described in the literature. We then derive a simple 1D athermal model that accounts for duplex rigidity in order to illustrate how shear forces spread along the DNA molecule during nanopore unzipping.

**3.2.1. Small hairpins' translocation is not always controlled by their global stabilities.** We designed 5 hairpins of 15 base pairs with the following characteristics: the first 8 bases (AGGCTGCG) and the loop sequence (GTTG) are identical for all molecules and an overhang of 30 polyA is added at their 3' moieties. Mfold Server was used to predict the thermodynamic

**Table 2.** Thermodynamical parameters of the DNA hairpins predicted by the Mfold Server at 1 M NaCl and 24 °C.  $\Delta H_o$  is in kcal mol<sup>-1</sup>,  $\Delta S_o$  in cal mol<sup>-1</sup> K<sup>-1</sup> and  $\tau_{1/2}$  in ms.

Name	$\Delta H_o$	$\Delta S_o$	$\Delta G_o$	A	T	G	C	$\tau_{1/2}$
Hp1	148.3	395.64	30.8	3	8	3	4	5
Hp2	149.1	402.8	29.5	5	6	4	3	18
Hp3	145.7	394.37	28.6	6	6	2	4	9
Hp4	145.7	394.7	28.5	6	6	2	4	39
Hp5	144.4	391.9	28	5	6	4	3	23

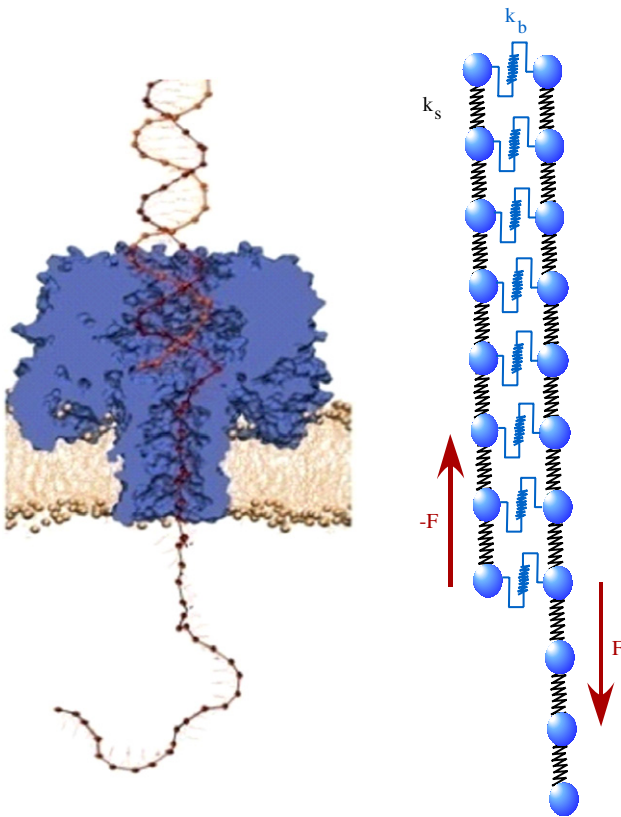
parameters of the hairpins. The hairpins are named according to their stability in descending order. The free enthalpy differences at 24 °C between hp1 and hp5 are 2.8 kcal mol<sup>-1</sup> =  $4.8k_B T$ /molecule. The values are summarized in table 2.

First, we checked that the thermal melting profiles of the hairpins vary according to the predicted stabilities. As shown in figure 6, the melting temperatures follow the free enthalpy order. The hairpins were then threaded at constant voltage (150 mV) through a single alpha-hemolysin pore in 100 mM *tris* buffer at pH = 7.4, 1 M KCl and 24 °C. The scatter plot of the normalized blocked currents  $i_b$  versus the translocation times  $\tau$  for each hairpin are shown on figure 7(a). Figure 7(b) shows the overlay of the cumulative translocation probabilities  $P(\tau)$  for the various molecules. The graphs clearly indicate that the blocked current is the same for all sequences, whereas the translocation times vary. The hairpins can be classified by increasing order of characteristic translocation times as: hp1 < hp3 < hp2 < hp5 < hp4. The corresponding values of  $\tau_{1/2}$  defined as  $P(\tau_{1/2}) = 0.5$  are given in table 2. Surprisingly, figure 7(b) indicates that hp1 translocates about ten times faster than hp4. The cumulative distributions  $P(\tau)$  cannot be perfectly fitted by a monoexponential. They display a long time tail larger than what is expected from an activated jump over a single energy barrier. The width of the distributions also increase as the characteristic time becomes longer. It results that the unzipping of these hairpins cannot be described by a simple two-state process separated by an activation energy barrier that scales with the global stability of the hairpins. The non-exponential distributions probably result from multiple intermediate states, the energies of which are not correlated with the global helix stability. In particular, DNA/pore interactions can act as catalysts as described in section 3.1. However, hp3 and hp4 have almost the same global stability and base content but translocate on different timescales. We do not have any final explanation for our observations: however, in the rest of the paper we would like to point out the possible role of the local elasticity of the DNA molecule on the force spreading along the bases and in turn on the unzipping cooperativity. We do not pretend that it explains our results but it just emphasizes the fact that local mechanical properties of the DNA helix can play an additional role to base-pair stability for unzipping of short DNA hairpins.

**3.2.2. Simple athermal 1D model.** It was previously reported that the unzipping time of a hairpin depends on the pore geometry and on the DNA position relative to the pore constriction [28]. Similarly, experiments and





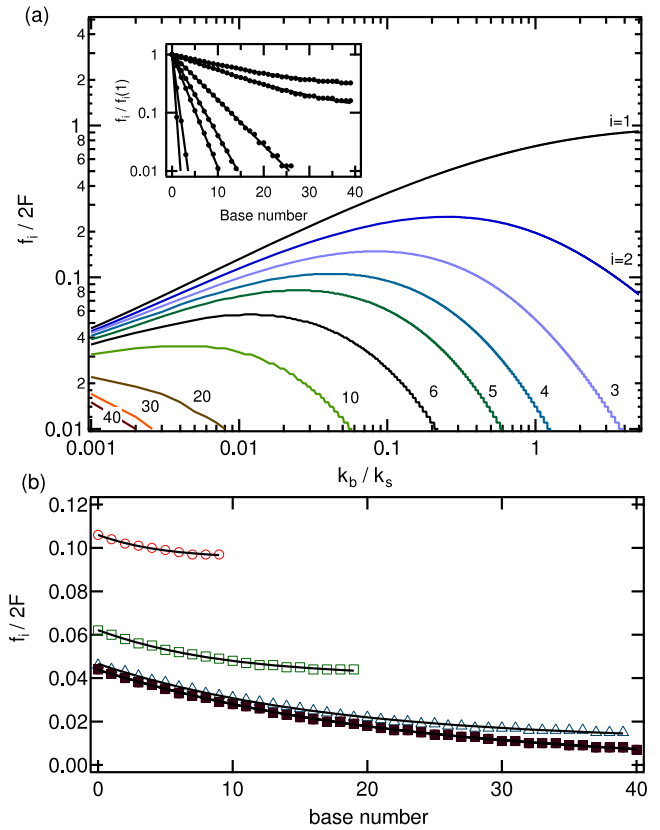


**Figure 8.** Left: putative position of the dsDNA in the pore. Right: 1D model of the DNA molecule under a shear stress. A point force  $f$  is applied on the first closed base pair. The elasticity of the DNA backbone is described by a spring constant  $k_s$  between each nucleotide. The elasticity of a base pair is modeled by the spring constant  $k_b$ . The force spreads exponentially along the DNA helix.

pairs will open simultaneously for  $F > N/2f_c$ . It is likely that the relevant values of  $r$  are below 0.1 if one compares the forces needed to unzip DNA (12 bp) compared to those needed to stretch it (60–120 pN). A good way to estimate  $r$  would be to run a simulation of DNA under shear and measure the spread of the deformation for a very long DNA duplex.

In the limit of  $N \lesssim \lambda$  finite size effects occur and the shear spreading is no longer exponential. Figure 9(b) shows the distribution of the force along the duplex for a fixed force  $F$ , a ratio  $r = 0.001$  and for different duplex lengths. For  $N = 200$  the force decreases exponentially along the molecule. As  $N$  decreases the force sustained by each base increases accordingly. Hence for a homogeneous duplex with  $r = 0.001$  and a critical force  $f_c$ , the critical shear  $F_c = 0.5 \frac{f_c}{0.1}$  for  $N = 10$  but increases to  $F_c = 0.5 \frac{f_c}{0.04}$  for  $N = 200$ . In this limit, and for a fixed external stress  $F$ , a short hairpin can open  $j$  base pairs at a time whereas a longer hairpin can stay closed. This result is not due to the global stability of the hairpin but to elastic spreading of the force along the molecule. This effect does not exist in optical trap DNA unzipping since the force is orthogonal to the backbone and  $F_c = f_c$  independently of the duplex length.

We conclude that molecules shorter than the typical length  $\lambda$  may open cooperatively at smaller forces than longer



**Figure 9.** A: normalized forces exerted on the base pair as a function of the ratio  $r = k_b/k_s$  for bases 1, 2, 3, 4, 5, 6, 10, 20 and 40.  $f_b(n)$  denotes the force sustained by base pair  $n$  and  $f$  is the externally applied force. At large  $r$ , the shear force is mostly supported by the first base pair. In the opposite limit, each base pair is equivalently sheared. Inset: force supported by each base pair for  $r = 5, 1, 0.1, 0.05, 0.015, 0.002$  and  $0.001$  from left to right. The shear force decays exponentially over a characteristic distance  $\lambda = \sqrt{\frac{k_s}{2k_b}}$ . B: finite size effects on the base pair load for fixed force  $F$  and given ratio  $r = 0.001$ . The various curves are plotted for 10 (O), 20 (□), 40 (Δ) and 200 (■) base pairs.

molecules and that more rigid molecules are expected to unzip at larger forces but in a more cooperative way. These conclusions do not involve any considerations on the base-pairing stability. For short hairpins the sequence-dependent local elasticity of the molecule is likely to play a role in the unzipping mechanism. More accurate evaluation of this effect could be tested by MD simulations.

#### 4. Conclusion

We showed that the nanopore unzipping of long DNA sequences can be understood as a diffusion of the unzipping fork along the sequence energy landscape, biased by a fluctuating force. Assuming Gaussian fluctuations with a standard deviation of  $1k_B T$  and a correlation time around 2 ms, we could reproduce the measured distributions of unzipping times and their voltage dependence. The comparison between experiment and theory indicates that thermal fluctuations of the biasing force are important for the translocation of molecules

through nanopores and that the associated unzipping is assisted by a stochastic resonance. We further showed examples of 15 base-pair hairpins that do not translocate according to the global stability of their helix, as usually reported in the literature. We discussed the influence of factors other than base-pair stability on the dynamics of translocation of short hairpins. We specifically focus on the problem of cooperativity due to local DNA elasticity.

## Acknowledgments

We would like to thank T Maggs, F Petrelis, A Aksimentiev and D Lacoste for fruitful discussions. This work was supported by ANR grant ANR-06-NANO-015.

## Appendix

### A.1. Analytical derivation of the cumulative first passage time distributions

In this appendix we derive a semi-analytical expression of the distribution of first passage time  $P_{w,\sigma}(\tau)$  assuming that the biasing energy is constant over a time  $\tau_{\text{corr}}$ . We study both limits:  $\tau \ll \tau_{\text{corr}}$  and  $\tau \gg \tau_{\text{corr}}$ . In the limit  $\tau \ll \tau_{\text{corr}}$ , the value of  $w$  is constant over  $0 < t < \tau$  but random for each molecule.  $P_{w,\sigma}(\tau)$  then is

$$P_{w,\sigma}(\tau) = \int_0^\infty p(w) P_w(\tau) dw = \langle P_w(\tau) \rangle \quad \text{for } \tau \ll \tau_{\text{corr}}.$$

In the limit  $\tau \gg \tau_{\text{corr}}$ , the value of  $w$  is constant by part over a time period of  $\tau_{\text{corr}}$ . The probability of translocating during a time interval of  $\tau_{\text{corr}}$  is thus  $\int_0^\infty p(w) P_w(\tau_{\text{corr}}) dw$ . It is constant over each interval  $\tau_{\text{corr}}$ . Let us call  $\tilde{P}_{w,\sigma}(\tau) = 1 - P_{w,\sigma}(\tau)$ . Then  $\tilde{P}$  is

$$\tilde{P}_{w,\sigma}(\tau + \tau_{\text{corr}}) = \left[ 1 - \int_0^\infty p(w) P_w(\tau_{\text{corr}}) dw \right] \times \frac{1}{\tau_{\text{corr}}} \tilde{P}_{w,\sigma}(\tau) \tau_{\text{corr}}.$$

where  $P_w(\tau)$  is the translocation time distribution calculated at constant  $w$ .

In the limit of large  $\tau \gg \tau_{\text{corr}}$ ,  $\tilde{P}_{w,\sigma}(\tau)$  satisfies the equation

$$\partial_\tau \tilde{P}_{w,\sigma}(\tau) = - \underbrace{\frac{1}{\tau_{\text{corr}}} \int_0^\infty p(w) P_w(\tau_{\text{corr}}) dw}_\nu \tilde{P}_{w,\sigma}(\tau).$$

Thus at long time the distribution of first passage times becomes exponential with a characteristic rate  $\nu$  that depends on the sequence and on the correlation time:

$$P_{w,\sigma}(\tau) = 1 - \exp(-\nu\tau) \quad \text{for } \tau \gg \tau_{\text{corr}}.$$

### A.2. Solution of the 1D spring model

The system of equations that determines the deformation of the molecule in this 1D model is the following.

For base  $i = 1$

$$k_s(u_2 - u_1) + k_b(v_1 - u_1) = f \quad (6)$$

$$k_s(v_2 - v_1) + k_b(u_1 - v_1) = -f \quad (7)$$

for base  $1 < i < N$

$$k_s(u_{i-1} - 2u_i + u_{i+1}) + k_b(v_i - u_i) = 0 \quad (8)$$

$$k_s(v_{i-1} - 2v_i + v_{i+1}) + k_b(u_i - v_i) = 0 \quad (9)$$

for base  $i = N$

$$-k_s(u_N - u_{N-1}) + k_b(v_N - u_N) = 0 \quad (10)$$

$$-k_s(v_N - v_{N-1}) + k_b(v_N - u_N) = 0. \quad (11)$$

This system is readily solved using the variable  $\Delta_i = u_i - v_i$  and  $\Sigma_i = u_i + v_i$ .

## References

- [1] Kasianowicz J J, Brandin E, Branton D and Deamer D W 1996 *Proc. Natl Acad. Sci. USA* **93** 13770–3
- [2] Akeson M, Branton D, Kasianowicz J J, Brandin E and Deamer D W 1999 *Biophys. J.* **77** 3227–33
- [3] Stefureac R, Long Y T, Kraatz H B, Howard P and Lee J S 2006 *Biochemistry* **45** 9172–9
- [4] Oukhaled G, Mathe J, Bianca A L, Bacri L, Betton J M, Lairez D, Pelta J and Auvray L 2007 *Phys. Rev. Lett.* **98** 158101
- [5] Makarov D E 2009 *Acc. Chem. Res.* **42** 281–9
- [6] Bezrukov S M and Kasianowicz J J 2001 *Biol. Membr.* **18** 453–7
- [7] Wu H C and Bayley H 2008 *J. Am. Chem. Soc.* **130** 6813–9
- [8] Sauer-Budge A F, Nyamwanda J A, Lubensky D K and Branton D 2003 *Phys. Rev. Lett.* **90** 238101
- [9] Mathe J, Visram H, Viasnoff V, Rabin Y and Meller A 2004 *Biophys. J.* **87** 3205–12
- [10] Sutherland T C, Dinsmore M J, Kraatz H B and Lee J S 2004 *Biochem. Cell Biol. Biochim.* **82** 407–12
- [11] Pastoriza-Gallego M, Gibrat G, Thiebot B, Betton J M and Pelta J 2009 *Biochim. Biophys. Acta-Biomemb.* **1788** 1377–86
- [12] Zhao Q, Sigalov G, Dimitrov V, Dorvel B, Mirsaidov U, Sligar S, Aksimentiev A and Timp G 2007 *Nano Lett.* **7** 1680–5
- [13] Hornblower B, Coombs A, Whitaker R D, Kolomeisky A, Picone S J, Meller A and Akeson M 2007 *Nat. Methods* **4** 315–7
- [14] Dorvel B, Sigalov G, Zhao Q, Comer J, Dimitrov V, Mirsaidov U, Aksimentiev A and Timp G 2009 *Nucleic Acids Res.* **37** 4170–9
- [15] Hall A R, van Dorp S, Lemay S G and Dekker C 2009 *Nano Lett.* **9** 4441–5
- [16] Tropini C and Marziali A 2007 *Biophys. J.* **92** 1632–7
- [17] Mathe J, Arinstein A, Rabin Y and Meller A 2006 *Europhys. Lett.* **73** 128–34
- [18] Comer J, Dimitrov V, Zhao Q, Timp G and Aksimentiev A 2009 *Biophys. J.* **96** 593–608
- [19] McNally B, Wanunu M and Meller A 2008 *Nano Lett.* **8** 3418–22
- [20] Viasnoff V, Chiaruttini N and Bockelmann U 2009 *Eur. Biophys. J.* **38** 263–9
- [21] Dudko O K, Mathe J, Szabo A, Meller A and Hummer G 2007 *Biophys. J.* **92** 4188–95

- [22] Bundschuh R and Gerland U 2005 *Phys. Rev. Lett.* **95** 208104–4
- [23] Lakatos G, Chou T, Bergersen B and Patey G N 2005 *Phys. Biol.* **2** 166–74
- [24] Bockelmann U and Viasnoff V 2008 *Biophys. J.* **94** 2716–24
- [25] Heng J B, Ho C, Kim T, Timp R, Aksimentiev A, Grinkova Y V, Sligar S, Schulten K and Timp G 2004 *Biophys. J.* **87** 2905–11
- [26] Heng J B, Aksimentiev A, Ho C, Marks P, Grinkova Y V, Sligar S, Schulten K and Timp G 2005 *Nano Lett.* **5** 1883–8
- [27] Heng J B, Aksimentiev A, Ho C, Marks P, Grinkova Y V, Sligar S, Schulten K and Timp G 2006 *Biophys. J.* **90** 1098–106
- [28] Muzard J, Martinho M, Mathe J, Bockelmann U and Viasnoff V 2010 *Biophys. J.* **98** 2170–8
- [29] Wanunu M, Sutin J, McNally B, Chow A and Meller A 2008 *Biophys. J.* **95** 4716–25
- [30] Mergny J L and Lacroix L 2003 *Oligonucleotides* **13** 515–37
- [31] Zuker M 2003 *Nucleic Acids Res.* **31** 3406–15
- [32] Gilespe D 1977 *J. Phys. Chem.* **81** 2340–61
- [33] Danilowicz C, Coljee V W, Bouzigues C, Lubensky D K, Nelson D R and Prentiss M 2003 *Proc. Natl Acad. Sci. USA* **100** 1694–9
- [34] Cocco S, Monasson R and Marko J F 2002 *Phys. Rev. E* **66** 051914
- [35] Petrelis F, Aumaitre S and Mallick K 2007 *Europhys. Lett.* **79** 40004–5
- [36] Nakane J, Wiggin M and Marziali A 2004 *Biophys. J.* **87** 615–21
- [37] Flomenbom O and Klafter J 2003 *Phys. Rev. E* **68** 041910
- [38] McDonnell M D and Abbott D 2009 *PLoS Comput. Biol.* **5** e1000348
- [39] Vercoutere W, Winters-Hilt S, Olsen H, Deamer D, Haussler D and Akeson M 2001 *Nat. Biotechnol.* **19** 248–52
- [40] Zhao Q, Comer J, Dimitrov V, Yemencioğlu S, Aksimentiev A and Timp G 2008 *Nucleic Acids Res.* **36** 1532–41
- [41] de Gennes 2001 *C. R. Acad. Sci., Paris IV* **2** 1505–8