INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **35** (2002) L349–L356 PII: S0305-4470(02)35164-3

LETTER TO THE EDITOR

DNA sequence from the unzipping force? One mutation problem

Somendra M Bhattacharjee[1] **and D Marenduzzo**[2]

[1] Institute of Physics, Bhubaneswar 751 005, India
[2] International School for Advanced Studies (SISSA) and INFM, Via Beirut 2-4, 34014, Trieste, Italy

E-mail: somen@iopb.res.in and maren@sissa.it

Received 20 March 2002
Published 21 June 2002
Online at stacks.iop.org/JPhysA/35/L349

Abstract

The possibility of detecting mutations in DNA from force measurements (as a first step towards sequence analysis) is discussed theoretically based on exact calculations. The force signal is associated with the domain wall separating the zipped from the unzipped regions. We propose a comparison method ('differential force microscope') to detect mutations. Two lattice models are treated as specific examples.

PACS numbers: 87.14Gg, 87.80Fe, 05.20Gg

The possibility of a force-induced unzipping transition [1–6] has opened up new ways of exploring the properties of biomolecules. Since the critical force for a double stranded DNA depends on its sequence, an inverse problem can be posed: can the DNA sequence be detected from the force required to unzip it? A still simpler question would be: can the mutations in DNA be detected by force measurements? A positive answer to either of these questions would lead to the possibility of testing (and detecting) mutations and of sequence determination in a non-destructive way.

Based on a few simple models used earlier for the DNA denaturation and unzipping transitions [3–5], we show here the signature of mutations on the f versus r curve (force versus relative distance of the end points of the two strands). The inverse problem is then to get the position of the mutations from such an experimentally realizable curve. Our emphasis in this letter is on the one base pair mutation problem also known as point mutation. This is not just of academic interest. The replication of DNA is a high-fidelity process, thanks to the inbuilt proof-reading and repair mechanisms, so that mistakes are very rare though even one could play havoc.

Our proposal (solution to the inverse problem), based on exact calculations, is to obtain the force difference ('differential force') between identically stretched native and mutant DNA.

This could be done by using, e.g., two atomic force microscope tips: we name this apparatus as a 'differential force microscope'. The position of the mutation can be obtained from a calibration curve involving the extremum position (or its value) of the differential force curve. In our models, we can find the nature of the mutation from the sign of the differential force.

Let us model a double stranded (ds) DNA by two N-monomer polymers interacting at the same contour length (or monomer index) j of the strands through a contact attractive potential $-\epsilon_j$ ($\epsilon_j > 0$), which might depend on the contour length. The interaction energy is $H = -\sum_{j=1}^{N} \delta_{\boldsymbol{r}_j,\boldsymbol{0}}\epsilon_j$, where $\boldsymbol{r}_j$ denotes the relative distance of the jth base pair and δ denotes the Kronecker delta. Other features such as the self- and mutual-avoidance, base stacking energy, helicity, etc are ignored in this study (but see below) in order to focus on the base pairing energy. Such a model exhibits a denaturation transition at a model-dependent temperature $T = T_m$ [7] from a low T double stranded configuration to a high T phase of two unpaired single strands.

A few definitions: starting with a sequence $\{\epsilon_j | j = 1, \ldots, N\}$ of bases of DNA (to be called the native DNA), we define mutant DNA as one with almost the same base sequence as the native molecule *except* a few base pairs. Homogeneous DNA or homo-DNA is DNA with identical base pairs (i.e. $\epsilon_j = \epsilon$ for all j) while heterogeneous DNA is one with heterogeneity in the sequence (j-dependent ϵ_j). Note that the model we have defined does not consider base pair stacking interaction and therefore does not distinguish, e.g., between an AT and a TA (or CG and GC) base pair. Consequently, the mutations we are referring to are only those involving AT (or TA) ↔ CG (or GC).

Two specific examples are considered here because of their exact tractability (analytical and numerical): (Mi) a two-dimensional, $d = 1 + 1$, directed DNA model (with the strands directed along the lattice vector $(1, 1)$) with base-pair interaction and mutual avoidance (forbidden crossing of the strands), and (Mii) two Gaussian polymers with the base-pair interaction in d-dimensions. In both cases, only the relative chain, involving the separation $\boldsymbol{r}$ of the bases at the same contour length, need be considered. These models in addition to the denaturation transition also show, for $T < T_m$, an unzipping transition in the presence of a force at the free end ($j = N$) [3–5]. In this paper, we consider the conjugate fixed distance ensemble and henceforth restrict ourselves to $T < T_m$. Consequently, we put $\beta = (k_B T)^{-1} = 1$ unless explicitly shown, where k_B is the Boltzmann constant.

For any ds DNA having their first monomers ($j = 1$) joined and their last monomers ($j = N$) at a relative lattice distance $\boldsymbol{r}$, the force $\boldsymbol{f}_N(\boldsymbol{r})$ required to maintain this relative distance $\boldsymbol{r}$ is $\boldsymbol{f}_N(\boldsymbol{r}) \equiv \nabla\mathcal{F}(\boldsymbol{r})$ where $\mathcal{F}(\boldsymbol{r})$ represents the free energy of the system in the fixed-$\boldsymbol{r}$ ensemble[3]. By definition, $\int_0^\infty \boldsymbol{f}_N(\boldsymbol{r}) \cdot \mathrm{d}\boldsymbol{r} = \mathcal{F}(\infty) - \mathcal{F}(0)$ which gives the free energy of binding or the work required to unzip DNA completely.

The two ensembles, fixed-force and fixed-stretch, are expected to give identical results in the $N \to \infty$ limit, though for finite N inequivalence might be expected [9] (this is indeed the picture valid for the homo-DNA), but a more serious situation arises if, e.g. for the scalar case, $\partial f/\partial x < 0$ in the fixed-stretch ensemble because, in the fixed-force ensemble, $\frac{\partial\langle x\rangle}{\partial\beta f} \equiv \langle x^2\rangle - \langle x\rangle^2 \geqslant 0$, where $\langle\cdot\rangle$ denotes thermal averaging. This is the case for heterogeneous DNA as shown in figure 1. The regions of 'wrong' sign (reminiscent of pre-shocks in Burgers turbulence [8] or of 'slip' in [2]) go away only after quenched averaging but do survive in the thermodynamic limit for each individual realization. *The absence of self-averaging in the system is encouraging, because it implies that individual features, typical of a single*

[3] For simplicity of notation, we are treating the position $\boldsymbol{r}$ as a continuous variable. For our discrete lattice models (Mi) and (Mii), the derivatives and integrals are to be replaced by proper finite differences and sums. It is also to be noted that in the continuum the force for the Gaussian interacting polymers satisfies a Burgers' type equation: see, e.g., [8].

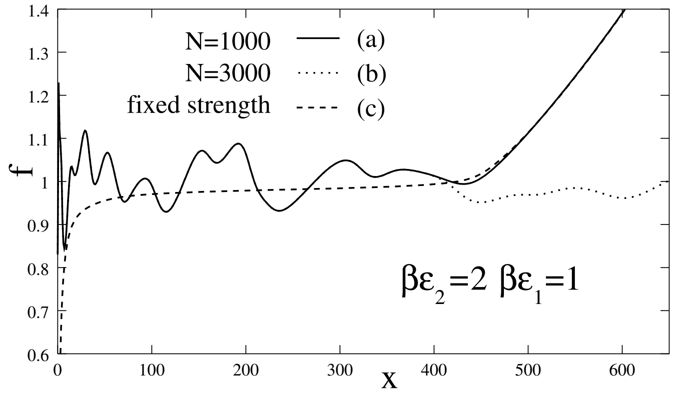


Figure 1. The force versus stretching distance curves for heterogeneous DNA (model (Mi)). The sequences are chosen randomly but both share the same sequence (1000 ϵ) from the open, pulled end. For the fixed stretch ensemble, curves (*a*) and (*b*), the pattern is identical over a region of x. Curve (*c*) is the fixed force ensemble phase coexistence curve with finite-size effect. The length of the unzipped part in units of base pairs is approximately x/y_0 (see equation (6)).

realization of the sequence, are maintained in the characteristic curves. We draw attention to the overlap of the f versus x curves over a range of x in figure 1 for different lengths with identical sequence near the open end. This overlap is a sign that the modulation is a characteristic of the sequence, and we have found this to be true quite generally.

Let us now consider the one mutation problem where the kth pairing energy of the native DNA has been changed from ϵ_k to ϵ_k'. For this one site change in H, the partition function $Z_{N|k}(\boldsymbol{r}, N)$ is given by

$$Z_{N|k}(\boldsymbol{r}, N) = Z_N(\boldsymbol{r}, N) + c_k Z_N(\boldsymbol{r}, N|\mathbf{0}, k). \tag{1}$$

where $c_k \equiv (\exp(\beta(\epsilon_k' - \epsilon_k)) - 1)$ and $Z_N(\boldsymbol{r}, N)$ is the native DNA partition function with last monomers at a relative distance $\boldsymbol{r}$, while $Z_N(\boldsymbol{r}, N|\mathbf{0}, k)$ is with the additional constraint of the kth pair being zipped. Note that equation (1) applies to both heterogeneous and homogeneous cases (and also to self avoiding strands).

We define a zipping probability $P(\boldsymbol{r}, N|\mathbf{0}, k) \equiv \frac{Z_N(\boldsymbol{r}, N|\mathbf{0}, k)}{Z_N(\boldsymbol{r}, N)}$, which is the conditional probability that site k is zipped when the free ends are at a distance $\boldsymbol{r}$. The force difference (to be called the differential force) to keep the free ends of both the mutant and the native DNA at the same distance $\boldsymbol{r}$ can be written as

$$\delta \boldsymbol{f}_{N|k}(\boldsymbol{r}) = -\frac{c_k}{1 + c_k P(\boldsymbol{r}, N|\mathbf{0}, k)} \nabla P(\boldsymbol{r}, N|\mathbf{0}, k) \tag{2}$$

$$= \frac{c_k P(\boldsymbol{r}, N|\mathbf{0}, k)}{1 + c_k P(\boldsymbol{r}, N|\mathbf{0}, k)} (\boldsymbol{f}_N(\boldsymbol{r}|\mathbf{0}, k) - \boldsymbol{f}_N(\boldsymbol{r})). \tag{3}$$

Here $\boldsymbol{f}_N(\boldsymbol{r}|\mathbf{0}, k)$ is a generalized force, the one necessary to keep the free ends of the native DNA at a relative distance $\boldsymbol{r}$ when the kth monomers are zipped. Except for c_k, all other quantities in equation (3) refer to the native DNA. This fact allows the inverse problem to be tackled as we show explicitly for a few cases. Since for DNA we have generally two possible choices for ϵ (ϵ_1 and ϵ_2), the sign of c_k determines the sign of δf in equation (3). Thus, in our simple model, the nature of the mutation can be identified from the sign of the differential force curve. Equations (1) and (3) can be generalized to more than one mutation

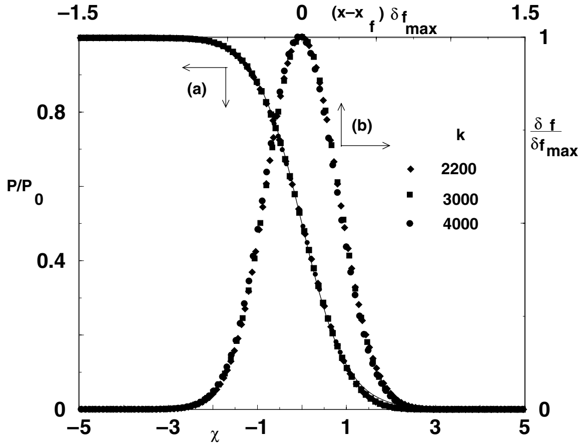


Figure 2. (*a*) The collapse of $P(x, N|0, k)/P_0$ versus χ where $x_{\rm d}(k)$, $w_{\rm d}(k)$ are obtained by fitting equation (4) (solid curve). For clarity only three cases of k are shown. (*b*) The collapse plot of $\delta f/\delta f_{\max}(k)$ versus $(x - x_f(k))\delta f_{\max}(k)$. A similar collapse is found even with $x_{\rm d}(k)$ and $w_{\rm d}(k)$ of (*a*). These are for model (Mii), homo-DNA, $N = 5000$, $d = 1$ and $\beta\epsilon = 2\beta\epsilon' = 1.5$. Arrows point towards the relevant axes.

and to models with other local energy parameters, though they become algebraically more involved. These more complicated cases will be discussed elsewhere. We consider the simplest case here.

If $\boldsymbol{r} \equiv (x, 0, \ldots, 0)$ is the direction of stretching, the quantity $P(\boldsymbol{r}, N|\boldsymbol{0}, k)$ has a kink-like behaviour (as in figure 2). In the fixed stretch ensemble, the chain separates into an unzipped and a zipped region separated by a domain wall, which to a very good approximation can be fitted by a tanh function:

$$\begin{aligned} P(\boldsymbol{r}, N|\boldsymbol{0}, k) &\equiv P(\boldsymbol{0}, N|\boldsymbol{0}, k) \exp\left\{-\beta \int {\rm d}\boldsymbol{r}\cdot(\boldsymbol{f}_N(\boldsymbol{r}|\boldsymbol{0}, k) - \boldsymbol{f}_N(\boldsymbol{r}))\right\} \\ &\approx P_0(1 + \tanh\chi)/2 \qquad \chi \equiv [x - x_{\rm d}(k)]/w_{\rm d}(k) \end{aligned} \tag{4}$$

where $x_{\rm d}(k)$ and $w_{\rm d}(k)$ are the position and the width of the wall (kink) respectively, and P_0 is a constant. Equation (4) suggests that $\delta f(x) \equiv w_{\rm d}(k)^{-1}\tilde{\rm f}(\chi)$, where $\tilde{\rm f}(\chi)$ is a scaling function.

It is possible to extend the above analysis to the general case where more than one mutation is present. For example, the partition function for the case with two mutations at positions k_1 and k_2, in an obvious notation, is

$$Z_{N|k_1,k_2}(\boldsymbol{r}) = Z_N(\boldsymbol{r}) + \sum_{i=1,2} c_{k_i} Z_N(\boldsymbol{r}N|\boldsymbol{0}k_i) + c_{k_1}c_{k_2} Z_N(\boldsymbol{r}N|\boldsymbol{0}k_1|\boldsymbol{0}k_2). \tag{5}$$

The last term in the above equation, representing correlation of the mutations, gives an additional contribution to the differential force over and above the individual contributions of the mutations. This additional contribution is negligible if the two mutations are far away or, more quantitatively, not in the same domain wall.

We now use these general results (equations (1)–(5)) for the particular cases (Mi) and (Mii). In the two-dimensional model (Mi), the partition function for two directed chains having their last monomers at a relative lattice distance x (along $(1, -1)$ and in units of the elementary square diagonal), and their first monomers joined, can be written in terms of N monomer-to-monomer transfer matrices W_j $(j = 1, \ldots, N)$ with matrix elements

$\langle x'|W_j|x\rangle \equiv ((\exp(\beta\epsilon_j) - 1)\delta_{x',0} + 1)\ (2\delta_{x',x} + \delta_{x',x+1} + \delta_{x',x-1})\ |x\rangle$, $|x'\rangle$ being the position vectors (with the constraint $x, x' \geqslant 0$).

For homo-DNA with contact energy ϵ, the largest eigenvalue of W determines the free energy and the thermodynamic properties in the limit $N \to \infty$. For $T < T_m \equiv \frac{\epsilon}{k_B \log\frac{4}{3}}$, the melting temperature, $\beta f_c = \cosh^{-1}\left(\frac{1}{2z_0} - 1\right)$ gives the critical force, with $z_0 = \sqrt{X} - X$, $X \equiv (1 - \mathrm{e}^{-\beta\epsilon})$. Indeed, in the fixed-stretch ensemble, any finite ($x \ll N$) stretch puts the chain on the phase coexistence curve.

Exploiting the equivalence of the ensembles valid for homo-DNA , we find $\beta f_N(x) \equiv \hat{F}(x/N)$, where

$$\hat{F}(y) = 2\tanh^{-1}(\max\{y, y_0\}) \qquad \left(y_0 \equiv (1 - 4z_0)^{\frac{1}{2}}\right) \tag{6}$$

is a piecewise continuous non-analytic function. For finite N, there is no singularity but the approximation of equation (6) still works quite well. At $x \sim y_0 N$ the force curve increases sharply (see also figure 1, curves (*a*) and (*c*)).

We now come to the explicit results for the one mutation case where one ϵ is replaced by $\epsilon' < \epsilon$. For homo-DNA, by starting from equation (6), using equations (3) and (4) one can find analytical approximations to the shapes of the previously introduced $P(x, N|0, k)$ and δf in terms of piecewise continuous functions. We find e.g. $P(x, N|0, k) = P(0, N|0, k)\exp(g(x))$, where (if $Ny_0 < \bar{k} \equiv N - k$)

$$g(x) = 0 \qquad \text{when} \quad x < \bar{k}y_0 \tag{7}$$

$$= g_{\bar{k}}(x) \qquad \text{if} \quad \bar{k}y_0 < x < Ny_0 \tag{8}$$

$$= g_{\bar{k}}(x) - g_N(x) \qquad \text{if} \quad Ny_0 < x < \bar{k} \tag{9}$$

and $g_k(x) = \log\left[\left(\frac{1+\frac{x}{k}}{1+y_0}\right)^{x+k}\left(\frac{1-y_0}{1-\frac{x}{k}}\right)^{x-k}\right]$. Equation (3) now simplifies because $f_N(x|0, k) = f_{\bar{k}}(x)$. The characteristics of the differential force curve δf versus x, such as the extremum value, its position and the width, $\delta f_{\max}(k)$, $x_f(k)$ and $w_f(k)$ respectively, can be determined from equations (7)–(9) as

$$\delta f_{\max}(k) \sim [w_f(k)]^{-1} \sim \bar{k}^{-1/2} \qquad x_f(k) = \bar{k}\tilde{x}_f(\bar{k}) \tag{10}$$

with $\tilde{x}_f(0) = 1$ and $\lim_{\bar{k}\to\infty}\tilde{x}_f(\bar{k}) = y_0$, where y_0 is defined in equation (6). The area of the peak, which yields the difference (with respect to the native case) in the work necessary to completely unzip the molecule, is constant as expected. The scaling form introduced after equation (4) suggests that the differential force is significant only in the domain wall region and the width of the domain wall $w_{\mathrm{d}}(k) \sim w_f(k)$ as we see in the numerical results.

For model (Mii), with one-dimensional Gaussian polymers ($T_m = \infty$), $P(x, N|0, k)$ has been calculated exactly by a transfer matrix method and is shown in figure 2(*a*). The validity of equation (10) ($\delta f_{\max}(k) \sim w_f(k)^{-1}$) is apparent from the data-collapse of the various δf curves in figure 2(*b*) where $\frac{\delta f}{\delta f_{\max}(k)}$ is plotted against $(x - x_f(k))\delta f_{\max}(k)$. The peak force difference $\delta f_{\max}(k)$ as a function of the mutant position k is in accord with the $\bar{k}^{-1/2}$ law of equation (10). The results for model (Mi) are shown in figure 3. For $d > 1$ ($\boldsymbol{r} = (x, 0, \ldots, 0)$ as above) the situation is similar to, e.g., the data-collapse of figure 2 and equation (10) remains valid.

Coming to the case of one mutation on heterogeneous DNA consisting of two energies ϵ_1 and $\epsilon_2 > \epsilon_1$ chosen with equal probability, the shapes of the zipping probability curves are found to be similar to the homo-DNA case, in fact indistinguishable on the plot of figure 2, with $x_{\mathrm{d}}(k)$ and $w_{\mathrm{d}}(k)$ sequence dependent. This indicates the validity of the domain wall

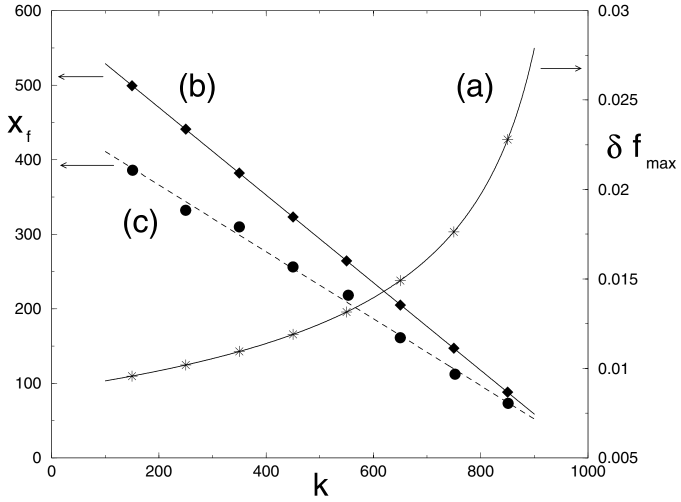
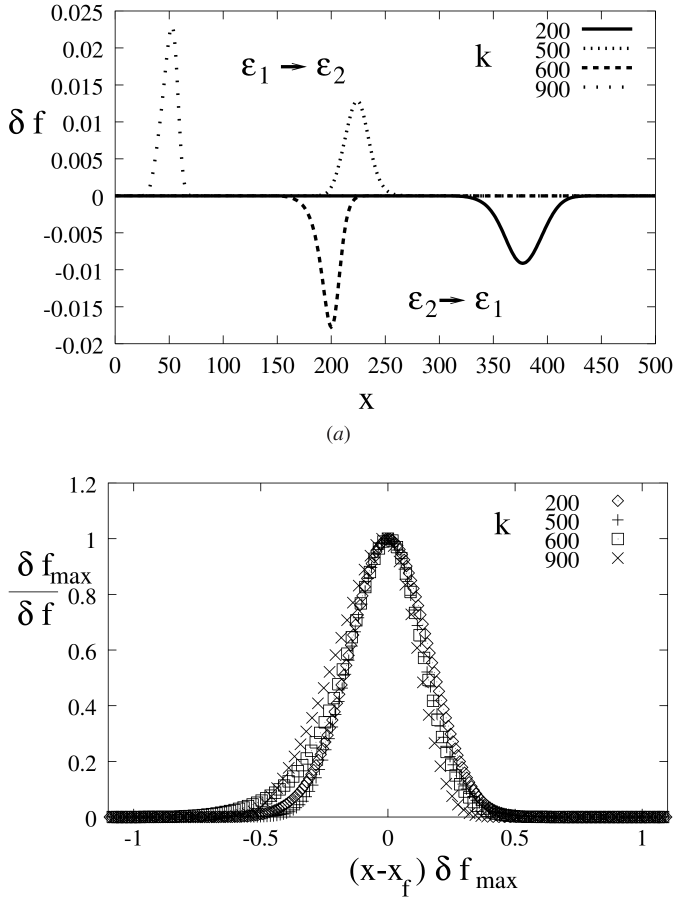


Figure 3. The 'calibration' curves $\delta f_{\max}(k)$ and $x_f(k)$ for (*a*) and (*b*) homogeneous DNA and (*c*) heterogeneous DNA (only $x_f(k)$ is shown). The curves fitting the data according to equation (10) are shown. Parameters are as in figure 1.

interpretation even for heterogeneous DNA. The mutation involves a change of the energy at site k (i.e. $\epsilon_1 \leftrightarrow \epsilon_2$). The signals δf for various mutations are shown in figure 4. As already mentioned, in our models the sign of δf tells us the nature of the mutation. These individual curves can again be collapsed on to a single one as for homo-DNA, though the nature of the collapse is not as good, mainly because the area under the curve is no longer strictly a constant. This reflects the importance of local sequences around the mutation point. Figure 3 (curves (*b*) and (*c*)) shows the k-dependence of $x_f(k)$. Unlike for homo-DNA, $\delta f_{\max}(k)$ does not seem to have the simple form of curve (*a*) in figure 3. Although the linearity is maintained for $x_f(k)$ as for the homogeneous case, there are regions of non-monotonicity at small scales which hamper the inversion.

Figure 3 gives a basis for a calibration curve. This could be $x_f(k)$ or $\delta f_{\max}$ versus k (or both) for homogeneous DNA, though for heterogeneous DNA we find the x_f versus k curve to be more reliable. Given the value of $x_f(k)$, one can look up in figure 3 for the corresponding k. The accuracy of the method relies on the ability to resolve close-by mutation points, i.e. mutation points in the same domain wall. There are differences in the full profiles of the δf curves for mutations at k, $k+1$, $k+2$, though translating that information back to the identification of the position is yet to be achieved. A better resolution is in any case obtained by changing the point at which the strands are pulled.

We now argue on how our calculation can compare with a potential experiment. In [1, 2], the typical force arising in the unzipping is between 10 and 15 pN, while the resolution is set below 0.2 pN, so in percentage it is <1.3–2%. Dynamical effects (important in [10]) are almost negligible at the lowest stretching velocity used in [1, 2] (20 nm s^{-1}) and are less important as the velocity is lowered. Our values for the typical force at $T = 1$ are at the border of this present day resolution (see figure 3, where it appears that the resolution is around 1% for $k \sim 500$, i.e. in the middle of the chains). In principle, they can be improved further by lowering T, though it is not clear to what extent this will work because the experimental temperature range is rather limited. In [11], the authors suggest that there will be experimental difficulties which would hamper the acquisition of the base-by-base sequence of DNA by



(*a*)

(*b*)

Figure 4. (*a*) The δf versus x curve for heterogeneous DNA. The sign of δf gives the nature of the mutation. (*b*) The collapse plot of the curves of (*a*) as in figure 2. The plots are for model (Mi), $N = 1000$ and $\epsilon_2 = 2\epsilon_1 = 1$ (with $T = 1$).

means of force measurements (but would however allow us to get information over groups of ~10 bases). This difficulty, though absent in our exact analysis of the models, might also set a lower limit on the error on the position of the mutation. Summarizing, we prove that mutations are detectable in the theoretical models. Numbers coming from our models suggest that this measurement might be a benchmark for present day real technology.

We finally propose an algorithm for sequencing DNA from the unzipping force in our models. This is defined so that the energy of the *j*th base pair is ϵ_2 if the average force at stretch $x = N - j$ is above the force signal of homo-DNA with an attractive energy $\epsilon \equiv \frac{\epsilon_1+\epsilon_2}{2}$ (T is low enough so that $y_0 \sim 1$). Our algorithm differs from the discussion in [2] because T and extra constraints (see below) play crucial roles. If we define the 'score' of the algorithm as the fraction of base pairs correctly predicted, from our data we observe that for any finite-size sample the score is 100% for $T < T_0(N) \sim N^{-\psi}$, with $0 \leqslant \psi \leqslant 1$ for $N \to \infty$. However, N_0 monomers near the open end can be sequenced at $T \cong T_0(N_0)$, no matter how big the total N is. Once this is done, we restart this time keeping the corresponding bases at position N_0 from the open end at a distance x with the constraint that the monomers at N are at a distance

$x' > N_0 - x$, to prevent rejoining in the already unzipped N_0 monomers. In this way, we would sequence another N_0 monomers and so on. We have verified this for models (Mi) and (Mii).

In conclusion, we studied the $\boldsymbol{f}$ versus $\boldsymbol{r}$ characteristic curve in the fixed stretch ensemble for simple models of DNA focusing on the base pairing energy only. We have seen that for homo-DNA, the force difference between a native and the corresponding one mutation case when pulled to the same distance contains enough signature to locate the position of the mutation. This could be the basis of a differential force microscope to detect mutations. For heterogeneous DNA, the mutation point cannot always be localized as accurately as for homo-DNA. Accuracy could be achieved by taking cognizance of the full features of the δf curve. We have shown that the differential force curve can be understood as due to the domain wall separating the zipped and the unzipped phases as the strands are pulled apart. Moreover, we found (figure 1) that the modulations in the force curve are connected to the local sequence. This holds the promise of extension of our proposal to cases beyond point mutation.

Acknowledgments

SMB thanks ICTP for hospitality. We thank A Maritan and F Seno for useful discussions.

References

[1] Essevaz-Roulet B, Bockelmann U and Heslot F 1997 *Proc. Natl. Acad. Sci. USA* **94** 11935
Strunz T *et al* 1999 *Proc. Natl Acad. Sci. USA* **96** 11277
Rief M, Clausen-Schaumann H and Gaub H E 1999 *Nature Struct. Biol.* **6** 346

[2] Bockelmann U, Essevaz-Roulet B and Heslot F 1997 *Phys. Rev. Lett.* **79** 4489
Bockelmann U, Essevaz-Roulet B and Heslot F 1998 *Phys. Rev.* E **58** 2386

[3] Bhattacharjee S M 2000 *J. Phys. A: Math. Gen.* **33** L423
Bhattacharjee S M 2000 *J. Phys. A: Math. Gen.* 9003(E)
Bhattacharjee S M 1999 *Preprint* cond-mat/9912297
Bhattacharjee S M 2002 *Ind. J. Phys.* A **76A** 69 (cond-mat/0010132)

[4] Lubensky D K and Nelson D R 2002 *Phys. Rev.* E **65** 031917
Lubensky D K and Nelson D R 2000 *Phys. Rev. Lett.* **85** 1572
Sebastian K L 2000 *Phys. Rev.* E **62** 1128
Zhou H 2000 *Preprint* cond-mat/0007015

[5] Marenduzzo D, Bhattacharjee S M, Maritan A, Orlandini E and Seno F 2002 *Phys. Rev. Lett.* **88** 028102
Marenduzzo D, Trovato A and Maritan A 2001 *Phys. Rev.* E **64** 031901

[6] Gerland U, Bundschuh R and Hwa T 2001 *Biophys. J.* **81** 1324

[7] Poland D and Scheraga H A 1966 *J. Chem. Phys.* **45** 1464
Peyrard M and Bishop A R 1989 *Phys. Rev. Lett.* **62** 2755

[8] Mukherji S 1994 *Phys. Rev.* E **50** R2407

[9] Neumann R M 1985 *Phys. Rev.* A **31** R3516
Guyer R A and Johnson J A Y 1985 *Phys. Rev.* A **32** 3661
Titantah J T *et al* 1999 *Phys. Rev.* E **60** 7010

[10] Peyrard M 1998 *Europhys. Lett.* **44** 271

[11] Thompson R E and Siggia E D 1995 *Europhys. Lett.* **31** 335