Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 350 (2005) 95–107
PHYSICA A
www.elsevier.com/locate/physa

Driven translocation dynamics of polynucleotides through a nanopore: off-lattice Monte-Carlo simulations

C.-M. Chen

Physics Department, National Taiwan Normal University, Taipei, Taiwan

Available online 8 December 2004

Abstract

The driven translocation dynamics of a polynucleotide chain through a nanopore is studied using off-lattice Monte-Carlo simulations, which plays an important role in the nanopore sequencing of polynucleotides. We report a detailed study on the dependence of translocation dynamics on the chain length and the local geometry near the nanopore. In particular, we find that the length dependence of the infection time of the chain could exhibit very different behaviors for different geometries.

PACS: 84.35.+i; 87.19.La; 87.17.Aa; 05.45.Xt; 87.18.Sn

1. Introduction

Studying the dynamics of driven polymer translocation through a nanopore has attracted considerable attention from experimental [1–10] or theoretical [11–20] perspectives in recent years. Understanding this process is very helpful for us to comprehend many phenomena in cell biology [21,22], such as the movement of RNA molecules and transcription factors across nuclear pores, translocation of proteins from the cis to trans side of a membrane through channels, and injection of DNA from a virus head into the host cell. Such an understanding on forced permeation of

E-mail address: cchen@phy.ntnu.edu.tw (C.-M. Chen).

0378-4371/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2004.11.026

polymers is also important in chemistry to separate and purify synthetic and biological polymers. Various potential applications of the polymer transport process in biotechnology has been proposed, including gene therapy, controlled drug delivery, and DNA sequencing.

Particularly, an inexpensive and accurate device for rapid DNA decryption can be designed using the single-molecule method of DNA sequencing by monitoring the variation of the ionic current due to an applied electric field which drives single-stranded polynucleotides through a nanopore in a thin film [1,5,19]. During such a polymer translocation process, the entropy of the polymer is considerably reduced and a driven electric field is required to lower the free energy of the system. In the mean time, the ionic current through the nanopore induced by the applied electric field fluctuates with polymer sequence as a result of the size blockade effect of monomers in the nanopore. Recently, several elegant experiments have demonstrated the capability of this method to distinguish long stretches of the same nucleotides, such as 30 adenines followed by 70 cytosines. These nanopores can be prepared by using ion channels imbedded on a lipid bilayer or nanopores sculpted in a Si_3N_4 membrane. The geometry of a lipid bilayer is approximately planar, and the polymer translocation dynamics through ion channels in a lipid bilayer has been extensively studied theoretically or computationally. However, ion channels in a lipid bilayer are labile and the channel diameters they provide cannot easily be adjusted. Recently developed ion-beam sculpting technique [5,9,10] can fabricate robust, solid-state nanopores with any desired diameter on a thin film, which could provide a new platform for the understanding of polymer transport and will make it possible to create robust single-molecule-sensing devices to characterize polynucleotides. In such a case, a curved geometry near the nanopore is typical and special designs near the pore is possible by modern nano-technologies. In this work, the effects of a curved geometry near the nanopore on the translocation dynamics of polymers will be investigated. Specifically, the translocation process involves with nanopore searching of the chain ends, effective transfection, and sliding. Our goal is to study dependence of translocation dynamics on the geometry near the nanopore and interactions between the nanopore and polymer molecule by Monte-Carlo (MC) simulations. In the present model a schematic illustration of the polymer moving across the membrane is shown in Fig. 1. Our model and computer simulations for a curved geometry near the pore is presented in Section 2. In Section 3 we apply this model to study the polymer translocation dynamics in the presence of a potential barrier near the pore. Our conclusions are summarized in Section 4.

2. Model for polymer translocation

Consider a polymer chain consisting of N monomers of size a, which is driven through a nanopore of length L and diameter D by an applied electric field, as shown in Fig. 1. The geometry of the thin film near the pore is curved and there is a barrier surrounding the pore (assuming the upper surface of the film has a shape of a Mexican hat). Such a barrier is also present for ion channels in a lipid bilayer. In the

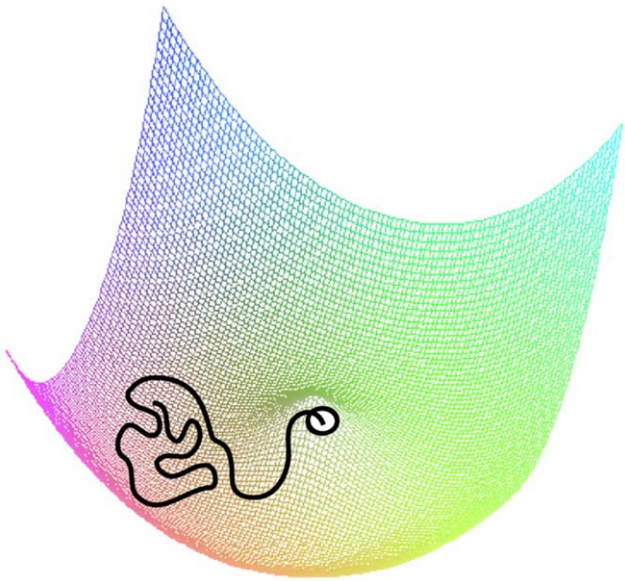

Fig. 1. A chain of DNA molecule near a nanopore on a curved surface is shown. The DNA chain is driven through the pore due to an applied electric field.

presence of a perpendicular electric field of amplitude E, we assume that the potential near the pore on the upper surface of the membrane can be expressed as $V(r) = \varepsilon(r - r_0)^2$, where r is the distance away from the pore and the valley of the potential occurs at r_0. The value of ε depends on the curvature of the thin film near the pore and the strength of electric field. Different shape of the upper surface can be assumed, but it would not change the general characteristics of the translocation dynamics. In our simulations, the polymer chain is represented by a bead–spring model, and its motion is simulated by the Metropolis MC algorithm in a continuous space at a constant temperature T. Each bead represents a monomer of size a and the bond length between two consecutive monomers is allowed to fluctuate around its equilibrium value b_0 with a spring constant e_s. At each instant, a nucleotide is picked up at random and attempts to move in any direction, and the move is accepted with probability $p = \min[1, \exp(-\Delta U/kT)]$, where ΔU is the energy change of the chain and kT is thermal energy. In our model, the energy of polynucleotides can be expressed as $U = U_{\text{bend}} + U_{\text{electric}} + U_{\text{bond}} + U_{\text{H-bond}}$, where $U_{\text{bend}} = \sum_{i=1}^{N-2} e(1 - \cos\theta_i)$ is the bending energy of the chain with rigidity e and bending angles $\{\theta_i\}$ [23–25], U_{electric} is the electric potential energy due to a constant electric field in the z-direction, $U_{\text{bond}} = \sum_{i=1}^{N-1} e_s(b_i - b_0)^2$ is the sum of bond energy between consecutive monomers $\{i, i+1\}$ (of bond length b_i), and $U_{\text{H-bond}}$ is the hydrogen bonding energy of (A, T) and (G, C) pairs. Here we consider the special case of negligible hydrogen bonding between bases, which can be realized by adjusting pH value, raising temperature, or adding ureas. This model was also used to study nanopore

sequencing of polynucleotides assisted by a rotational electric field in our previous work [19].

An important issue in our model is to relate a MC step to a duration in real time. To estimate this MC step, we consider the above sliding process of a polymer chain through a hole. For such a driven transport in a viscous fluid of viscosity η, the sliding velocity of a monomer is approximately $v_s = qE/(6\pi\eta a)$, where q is the net electric charge per monomer and a is the monomer size. In our simulations, the driving electric force is of order 10^{-12} J/m, the water viscosity η is about 10^{-3}N s/m^2, and $a \approx 10^{-9}$ m. Therefore, we know that the sliding velocity is about 0.1 m/s. Since the sliding time for one monomer to cross a thin film of thickness 10^{-8} m in our simulation is about 100 MC steps, we estimate that 1 MC step is equivalent to 10^{-9} s. To check the validity of our estimation, we compare our simulation results with experimental data in Refs. [1,6]. For a chain of 100 monomers driven by an electric field $E \approx 10^7$ V/m, the average blockade time in our simulations is of order 1000 MC steps (or 1 μs). Experimentally, the blockade time of a polymer chain of 100

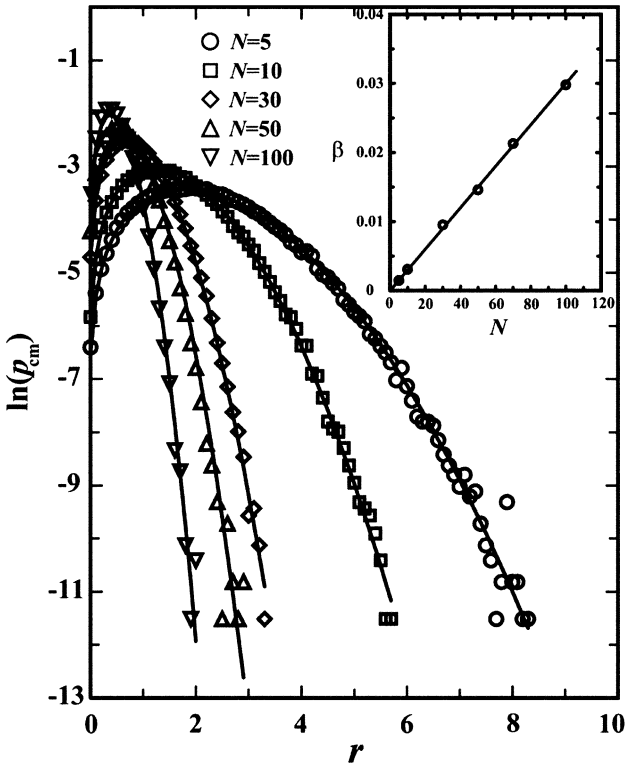

Fig. 2. The probability distribution of the center of mass of a DNA chain at a distance r away from a nanopore for various chain lengths ($N = 5$, 10, 30, 50, and 100). The solid lines are fitted curves using the Maxwell–Boltzmann distribution to the data from our simulations. The inset shows the dependence of the coefficient β on chain length N.

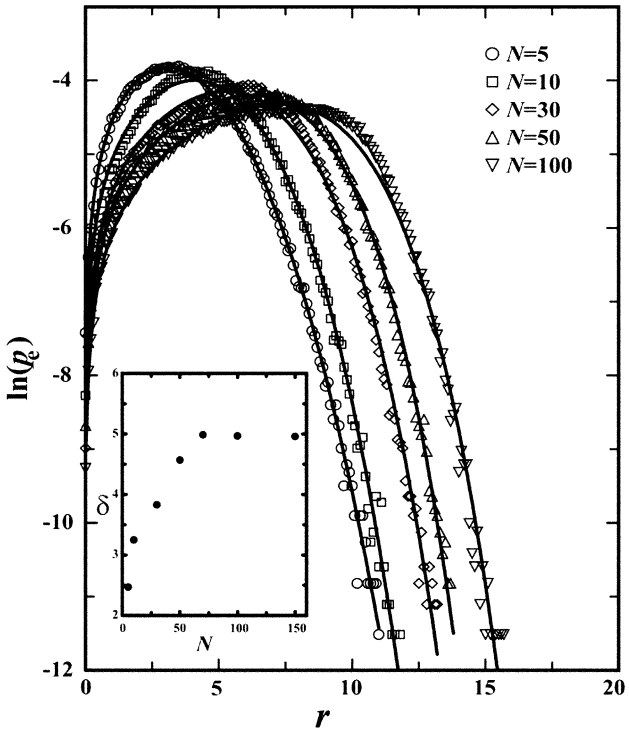

Fig. 3. The probability distribution of the end monomers of a DNA chain at a distance r away from a nanopore for various chain lengths ($N = 5$, 10, 30, 50, and 100). The solid lines are fitted curves to the data from our simulations. The inset shows the dependence of the coefficient δ on chain length N.

monomers is about 100 μs in Ref. [1] using an electric field $E \approx 10^6$ V/m. Since the sliding velocity is approximately proportional to E^2, as shown in Ref. [6], our estimated time of 1 MC step gives translocation times of the order observed in the experiments.

3. Results of simulations

To study the kinetics of a polymer chain passing through a nanopore, we have simulated its translocation process 100 times for each set of parameters. Each chain is randomly deposited above the membrane. Due to the applied electric field, the polymer chain will land on the membrane surface and search for the nanopore, from which we define a nanopore searching time of the chain ends as t_{search}. After the nanopore was found by one chain end (the nanopore can only accommodate a single-strand DNA chain), the infection of the polymer chain may or may not succeed since thermal energy could move the chain away from the pore. After n_{inf}

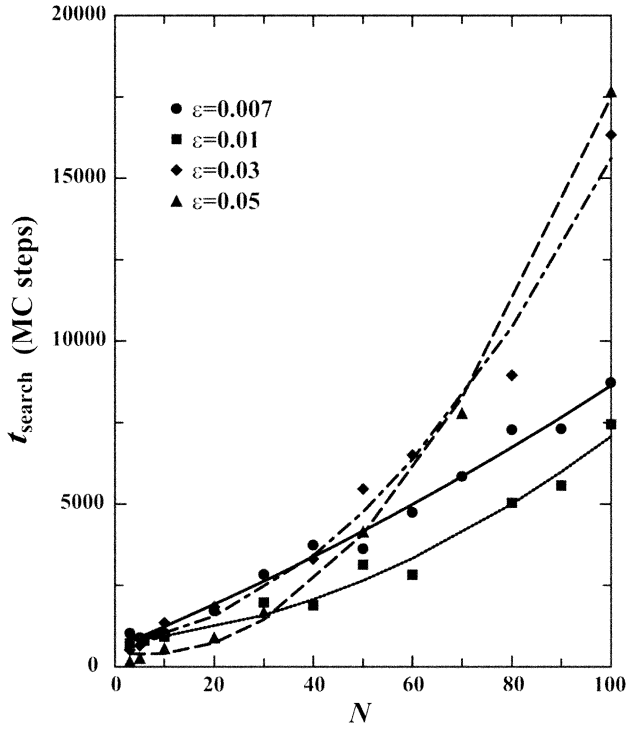

Fig. 4. The dependence of the searching time of the chain, t_{search}, on chain length N for various values of ε. In this case, $r_0 = 0$.

times of trial, the chain finally penetrates the nanopore with a complete sliding process. Thus we define the infection time as t_{inf} and the sliding time as t_{slide}. The total translocation time of the polymer chain is defined as $t_{all} = t_{inf} + t_{slide}$. In our simulations, the radius (a) of each monomer is set to be 1 unit length (approximately 0.5 nm) and the equilibrium bond length (b_0) between two consecutive monomers is 2.5 unit length. Furthermore, we choose the pore size $D = 4$, thermal energy $kT = 1$, the constant electric field amplitude $E = 0.5$, the spring constant of the bond $e_s = 30$, and the bending rigidity $e = 0.2$. Here thermal energy and electric charge of each monomer have been set to unity, and the corresponding electric field is of order 10^7 V/m.

We first consider the translocation dynamics of a polymer chain through a nanopore in a parabolic potential, i.e., $r_0 = 0$. In this case, as soon as the polymer chain lands on the upper surface, it drifts toward the nanopore guided by the potential. The probability distribution of the center of mass of the polymer chain at various distances (r) away from the nanopore can be well approximated by a Maxwell–Boltzmann distribution function (in two-dimension), i.e., $p_{cm} \propto r\exp(-\beta r^2)$. As shown in Fig. 2, data points are simulation results for polymer

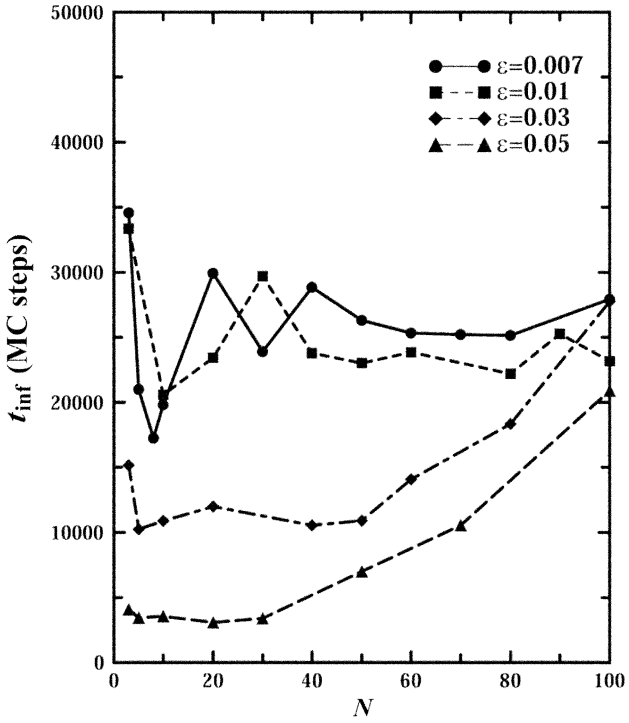

Fig. 5. The dependence of the infection time of the chain, t_{inf}, on chain length N for various values of ε. In this case, $r_0 = 0$.

chains of $N = 5$, 10, 30, 50, and 100 for the case of $\varepsilon = 0.03$, and the solid lines are fitted curves of p_{cm} using the Maxwell–Boltzmann distribution. The coefficient β is mainly determined by the electric potential energy of the chain and thermal energy. At a constant temperature, since the electric potential energy of the chain is proportional to chain length, β increases linearly with chain length N as shown in the inset. For long chains, translational entropy is negligible and the center of mass of the chain is localized near the nanopore. In Fig. 3, we show the probability distribution of polymer chain ends as a function of r for $N = 5$, 10, 30, 50, and 100 for the case of $\varepsilon = 0.03$. For long chains, their chain ends are not popularly distributed near the nanopore due to the exclusion from other monomers in the chain and the distribution of polymer chain ends is not localized. From our simulations we found that the end distribution function p_e is proportional to $r\exp(-\beta' r^{\delta})$, where $\beta' \neq \beta$ and $\delta > 2$ is a function of N. This indicates an effective potential experienced by chain ends that is proportional to r^{δ}, instead of the original parabolic potential. For long chains of $N > 75$, the exponent $\delta = 5$ reaches a constant value. This excluded volume effect considerably increases the nanopore searching time of chain ends.

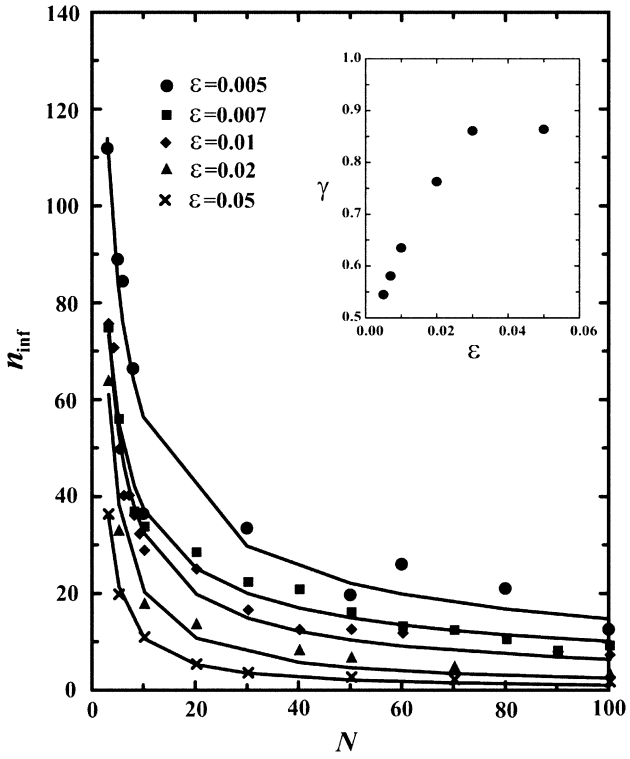

Fig. 6. The dependence of the number of attempted trials of infection, n_{inf}, on chain length N for various values of ε. The solid lines are fitted curves using $N^{-\gamma}$ to the data from our simulations. In this case, $r_0 = 0$. The inset shows the dependence of the coefficient γ on chain length N.

The searching time t_{search} is defined to be the duration for chain ends to find the nanopore after the chain lands on the surface. For a flat surface, the mobility of the chain determines the searching time. For a curve surface, it is dominated by how often the pore is blocked by other monomers. In Fig. 4, t_{search} is plotted as a function of N for $\varepsilon = 0.007$, 0.01, 0.03, and 0.05. The searching time increases with N for $\varepsilon \geqslant 0$ due to both the chain mobility and the blocking effect. For short chains, the searching time decreases with ε since the chain mobility increases with ε. However, for long chains the searching time increases with ε since the blocking effect is more prominent at large ε. In our simulations, the chain dependence of the searching time at various ε can be described by a quadratic function, i.e., $t_{search} = a_0 + a_1 N + a_2 N^2$. The coefficient a_2 is 0.19 for $\varepsilon = 0.007$ and 1.96 for $\varepsilon = 0.05$. Our results indicate that the N^2 term in the searching time is mainly due to the surface curvature.

The infection time t_{inf} is shown in Fig. 5. For nearly flat surfaces ($\varepsilon \leqslant 0.01$), the infection time is not sensitive to the chain length. Although the searching time increases with the chain length, this effect on the infection time is cancelled by the

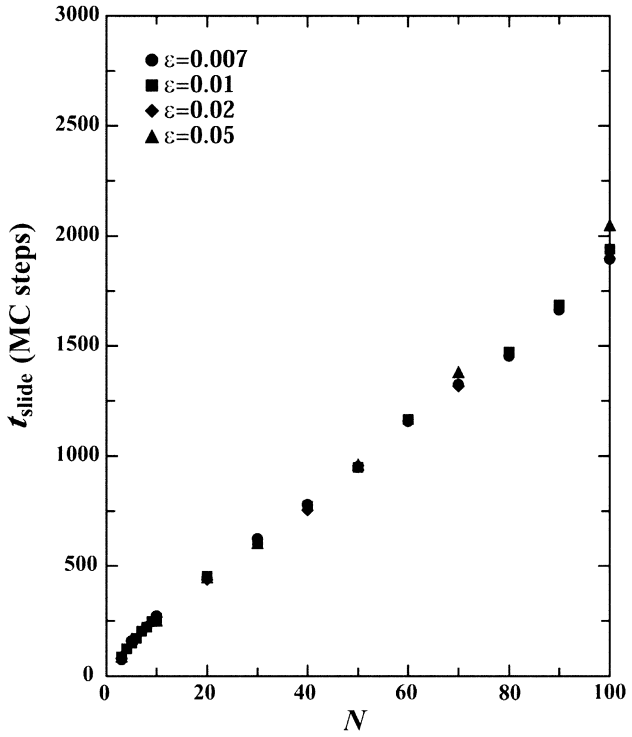

Fig. 7. The dependence of the sliding time of the chain, t_{slide}, on chain length N for various values of ε. In this case, $r_0 = 0$.

number of attempted infections n_{inf}. For short chains, it is easy for the chain to diffuse away from the nanopore before the chain can successfully slide through the pore. The high mobility of short chains leads to a large number of attempted infections. In Fig. 6, the chain length dependence of n_{inf} is plotted for $\varepsilon = 0.005$, 0.007, 0.01, 0.02 and 0.05. From our simulations, it is found that $n_{inf} \propto N^{-\gamma}$. The exponent γ, as shown in the inset of Fig. 6, increases linearly with ε and then saturates at about 0.9 for $\varepsilon \geqslant 0.03$. Therefore, for small values of ε, the infection time of the chain does not grow with its chain length. Since the infection time is typically much longer than the sliding time, this result indicates that the value of ε should be less than 0.03 in a nanopore sequencing experiment in order to reduce the sequencing time. For $\varepsilon \geqslant 0.03$, the infection time increases with the chain length. The chain length dependence of the infection time for curve surfaces can also be approximated by a quadratic function. This property can be used to demix a DNA mixture by driving DNA chains through a series of nanopores, particularly for separating long DNA molecules. It is known that the standard methods, such as the gel electrophoresis, for separating DNA by length deteriorates seriously for long DNA molecules since the mobility of long DNA chains is independent on their chain length. Therefore, one

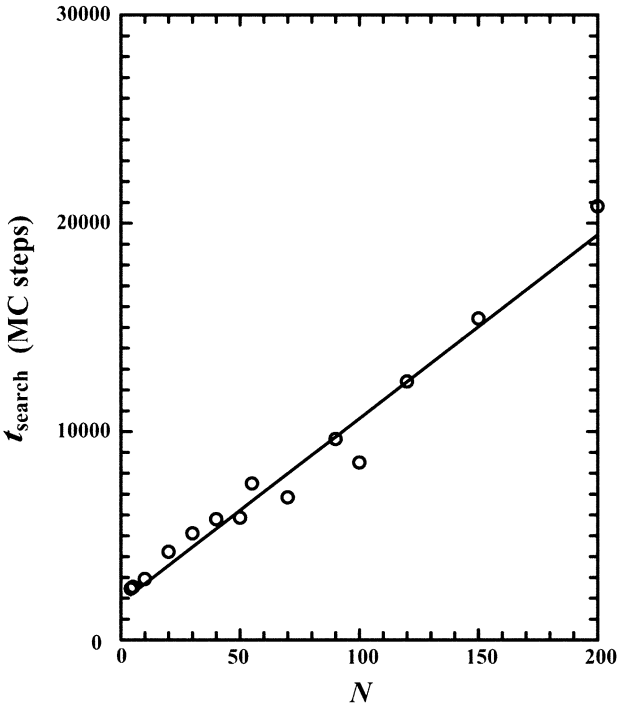

Fig. 8. The dependence of the searching time of the chain, t_{search}, on chain length N. In this case, $r_0 = 5$ and $\varepsilon = 0.05$.

can design a demixing device consisting of a series of parallel thin films on which nanopores are fabricated.

The sliding time of the chain is defined to be the time required for the first monomer to enter the upper surface and the last one to leave the lower surface. Therefore, we can express the sliding time as $t_{slide} = (Na + L)/v_s$, where $L = 8$ is the pore thickness. The chain length dependence of t_{slide} is plotted in Fig. 7 for $\varepsilon = 0.007$, 0.01, 0.02, and 0.05. The data collapse of t_{slide} in Fig. 7 shows no major effect of surface curvature on the sliding time of the chain. The simulation curve of the sliding time can be expressed as $t_{slide} = 63.5 + 18.2N$. The slope of t_{slide} $(1.82 \times 10^{-8}\,s)$ agrees with its theoretical value $a/v_s \approx 2.5 \times 10^{-8}\,s$. The value of $a/L = 2.5/8$ in the t_{slide} curve is also consistent with the value from the fitted curve 18.2/63.5. It is found in our simulations that the linear dependence of t_{slide} on chain length indicates a constant velocity of the chain for $N = 5 \sim 100$. We note that a nonlinear increase in the sliding velocity (decrease in the sliding time) for decreasing length for short chains ($N = 3$–5) is observed in our simulations. The threshold length value closely matches the pore thickness $L = 8$ which is in agreement with the experimental observations [6]. This difference in the sliding velocity may be due to the polymer segment outside the pore. For chains of length longer than the pore thickness,

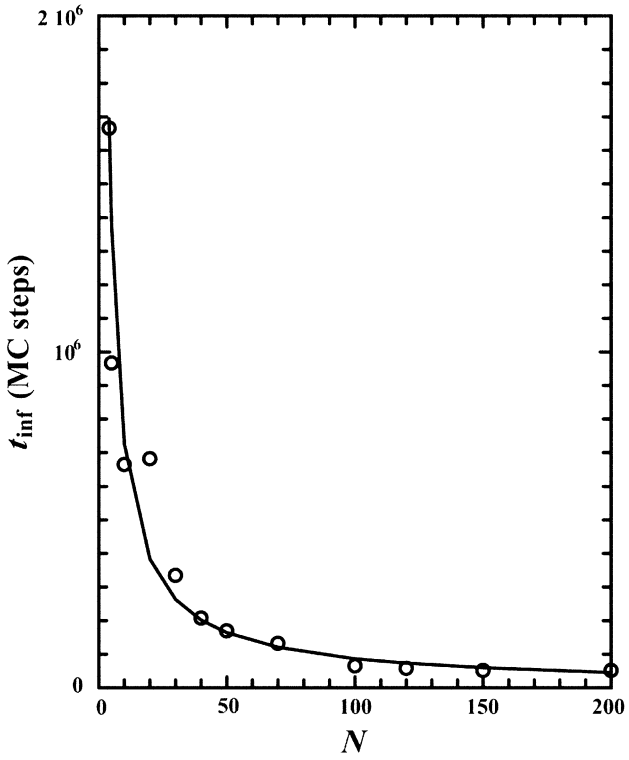

Fig. 9. The dependence of the infection time of the chain, t_{inf}, on chain length N. In this case, $r_0 = 5$ and $\varepsilon = 0.05$.

fluctuations and vibrations of the monomers outside the pore are transmitted to those monomers inside the pore making it relatively slower.

Furthermore, we study the translocation dynamics of DNA chains in a Mexican-hat potential, i.e., $r_0 \neq 0$. As an example, we choose $r_0 = 5$ and $\varepsilon = 0.05$. As shown in Fig. 8, the searching time t_{search} increases linearly with the chain length since the nanopore is not blocked by DNA. For chains of length longer than 100, the average searching time of a DNA chain in a Mexican-hat potential is much smaller than that in a parabolic potential. The infection time of the chain t_{inf} as a function of N is plotted in Fig. 9. Interestingly, we find the infection time decreases as the chain length increases ($t_{\text{inf}} \propto N^{-0.92}$). This is very different from the $r_0 = 0$ case where we find $t_{\text{inf}} \propto N^2$. This behavior has been experimentally demonstrated in a microfabricated entropic trap array [26] which has been designed to separate long DNA molecules. In Fig. 10, the sliding time of the chain increases linearly with chain length. A nonlinear decrease of t_{slide} occurs only for very short chains. The sliding velocity of the chain in the $r_0 = 5$ case is found to be twice smaller than t_{slide} in the $r_0 = 0$ case due to the drag from monomers outside the pore.

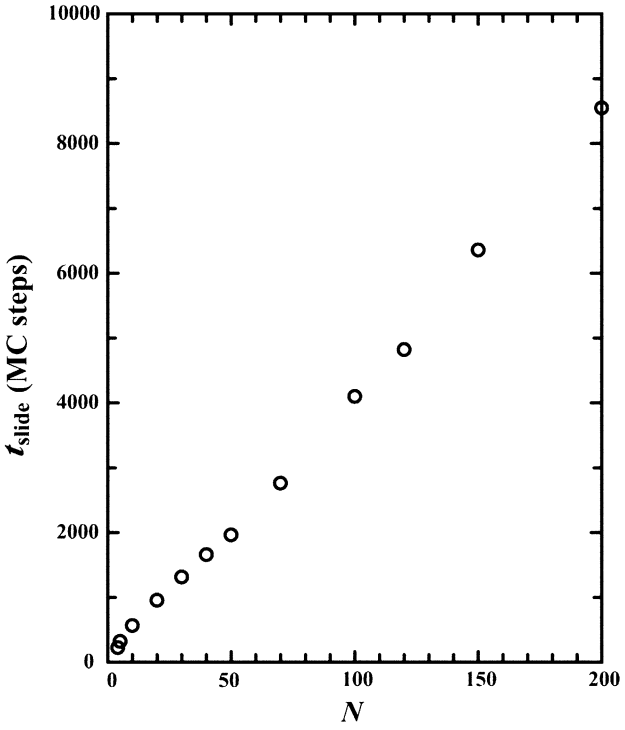

Fig. 10. The dependence of the sliding time of the chain, t_{slide}, on chain length N. In this case, $r_0 = 5$ and $\varepsilon = 0.05$.

4. Conclusions

In conclusion, we have applied off-lattice Monte-Carlo simulations to study the translocation dynamics of a single-strand DNA chain through a nanopore on a curved thin film driven by an applied electric field. Two different types of curved surfaces are considered including a bowl shape and a Mexican-hat shape surfaces. The translocation dynamics of the DNA chain is characterized by the searching time, the infection time, and the sliding time of the chain. For a nearly flat surface or a surface of Mexican-hat shape, the searching time increases linearly with the chain length. For a curved surface of bowl shape, the searching time increases quadratically with chain length due to the blockade of other monomers. More interestingly, we find that the infection time of the chain has very different dependence on the chain length for these two surface geometries. Based on our simulations, a series of parallel thin films fabricated with nanopores can be designed to demix DNA fragments of different chain length, particularly for long DNA molecules. The sliding time of DNA chains increases linearly with chain length in all simulations, except for a nonlinear decrease for very short chains. This behavior is consistent with experimental observations.

Acknowledgements

The author acknowledges the support from NSC of Taiwan under Grant No. 92-2112-M-003-005 and from National Taiwan Normal University under Grant No. ORD92-3.

References

[1] J.J. Kasianowicz, E. Brandin, D. Branton, D.W. Deamer, Proc. Natl. Acad. Sci. USA 93 (1996) 13770.
[2] M. Akeson, D. Branton, J.J. Kasianowich, E. Brandin, D.W. Deamer, Biophys. J. 77 (1999) 3227.
[3] A. Meller, L. Nivon, E. Brandin, J. Golovchenko, D. Branton, Proc. Natl. Acad. Sci. USA 97 (2000) 1079.
[4] S.E. Henrickson, M. Misakian, B. Robertson, J.J. Kasianowich, Phys. Rev. Lett. 85 (2000) 3057.
[5] J. Li, et al., Nature 412 (2001) 166.
[6] A. Meller, L. Nivon, D. Branton, Phys. Rev. Lett. 86 (2001) 3435.
[7] L. Movileanu, H. Bayley, Proc. Natl. Acad. Sci. USA 98 (2001) 10137.
[8] S. Howorka, L. Movileanu, O. Braha, H. Bayley, Proc. Natl. Acad. Sci. USA 98 (2001) 12996.
[9] A.J. Storm, J.H. Chen, X.S. Ling, H.W. Zandbergen, C. Dekker, Nature Matr. 2 (2003) 537.
[10] J. Li, M. Gershow, D. Stein, E. Brandin, J.A. Golovchenko, Nature Matr. 2 (2003) 611.
[11] W. Sung, P.J. Park, Phys. Rev. Lett. 77 (1996) 783.
[12] E.A. Di Marzio, A.J. Mandell, J. Chem. Phys. 107 (1997) 5510.
[13] P.-G. de Gennes, Adv. Polym. Sci. 138 (1999) 91.
[14] P.G. de Gennes, Proc. Natl. Acad. Sci. USA 96 (1999) 7262.
[15] D.K. Lubensky, D.R. Nelson, Biophys. J. 77 (1999) 1824.
[16] M. Muthukumar, Phys. Rev. Lett. 86 (2001) 3188.
[17] K.L. Sebastian, A.K.R. Paul, Phys. Rev. E 62 (2000) 927.
[18] T. Ambjörnsson, S.P. Apell, Z. Konkoli, E.A. Di Marzio, J.J. Kasianowicz, J. Chem. Phys. 117 (2002) 4063.
[19] C.-M. Chen, E.-H. Peng, Appl. Phys. Lett. 82 (2003) 1308.
[20] E. Slonkina, A.B. Kolomeisky, J. Chem. Phys. 118 (2003) 7112.
[21] H. Lodish, et al., Molecular Cell Biology, third ed., Scientific American Books, New York, 1995.
[22] B. Alberts, et al., Molecular Biology of the Cell, third ed., Garland Publishing, New York, 1994.
[23] C.-M. Chen, Phys. Rev. E 63 (2001) 010901.
[24] C.-M. Chen, Y.-A. Fwu, Phys. Rev. E 63 (2001) 011506.
[25] C.-M. Chen, P.G. Higgs, J. Chem. Phys. 108 (1998) 4305.
[26] J. Han, H.G. Craighead, Science 288 (2000) 1026.