

IBM Research Report

DNA Transistor

Stan Polonsky, Gustavo Stolovitzky, Stephen Rossnagel

IBM Research Division

Thomas J. Watson Research Center

P.O. Box 218

Yorktown Heights, NY 10598



Research Division

Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

DNA Transistor

Stas Polonsky, Gustavo Stolovitzky, Steven Rossnagel
IBM T.J. Watson Research Center

April 12, 2007

Abstract

We present a concept of a nano-electro-mechanical device capable of controlling the position of DNA inside a nanopore with single nucleotide accuracy. The principle of the device operation utilizes the interaction of discrete charges along the backbone of a DNA molecule with the electric field inside the nanopore. We propose an immediate application of our device as a replacement of capillary electrophoresis in DNA sequencing by separation.

I. Introduction

The ability to move along information-carrying polymers of nucleic acids with single nucleotide precision is vital for machinery of life. DNA polymerase, for example, has to translocate along the DNA molecule with single nucleotide accuracy in order to replicate it correctly. The same accuracy is needed for RNA polymerase to transcribe RNA from a DNA template. Another example is the ribosome, an organelle in the cell that needs a precision position control along the messenger RNA to translate it into protein.

As nanotechnology masters its ability to manipulate single molecules, the question whether man-made devices can position themselves along information-carrying polymers with single monomer precision becomes of practical interest. In this report we propose a nano-electro-mechanical device capable of trapping a single DNA molecule and translocating it with single nucleotide accuracy. We call this device a “DNA Transistor”. The report is organized as follows. In Section II we explain the operation principle of the DNA transistor – electrostatic control over ionized phosphate groups of DNA backbone. We analyze the device using a simplified “toy” model of DNA. The main goal of this section is to develop intuition and derive crude analytical estimates to characterize the DNA trapping capability of the device. In Section III we describe a practical implementation of DNA. In Section IV we estimate the performance of the device. In Section V discuss possible uses of the device in DNA sequencing, as well as implementation issues, and problems that may arise in the actual implementation of the device.

II. Principle of operation

DNA is a strong acid in aqueous solution. The negative charge distribution along DNA molecule is non-uniform and almost point-like (for details see [1]). The maxima of this distribution are located on phosphate groups of the molecule’s helical backbone (see

Figure 1), specifically, two oxygen atoms that do not form ester bonds with neighboring carbon atoms. The charge localized at these atoms is close to that of a single electron. The separation length of charges along the axis of the DNA molecule is $d = 0.34 \text{ nm}$, the radius of the helix is $r_h = 1 \text{ nm}$, the angle between neighbor charges is 36° , or, equivalently, approximately 10 charges per helix turn.

The discreteness of the charge distribution along the molecule is a key to our device operation. As we will argue, the interaction of discrete charges with an external electric potential well can be used to localize the polymer with single monomer resolution. A time-oscillating potential well can enable the translocation of the polymer by one monomer per oscillation. We will illustrate these ideas using the following simplified model. Consider an idealization of a charged linear polymer as an equidistant string of point charges q separated by a distance d . The string traverses through a symmetric trapezoidal electric potential well $V(z)$ (see Figure 2) of depth V_0 . The geometry of the potential is characterized by the sidewall thickness S and the well length W . For simplicity we assume d to be the unit distance, restoring d when necessary. For a given displacement δ of the polymer, the position of i -th charge is $i + \delta$. The energy of the string,

$$E(\delta) = q \sum_{i=-\infty}^{\infty} V(i + \delta), \quad (1)$$

is a periodic function of δ with a unit period. We refer to the difference between maximum and minimum values of $E(\delta)$ as trapping energy $\Delta E_{TR} = E_{\max} - E_{\min}$. If ΔE_{TR} is sufficiently larger than thermal energy (say at 300K) $E_T = kT/2 \approx 13 \text{ meV}$, the translocation of the string between two successive energy minima is unlikely, i.e. the potential V serves as a trap. We are interested in understanding under what conditions ΔE_{TR} reaches its maximum value. Besides the obvious proportionality to V_0 , the most important fact is that ΔE_{TR} is sensitive to the geometry of V . Figure 3 visualizes the mechanism of this sensitivity. Thick diagonal lines show positions of charges as a function of δ . Grayed rectangles mark two sidewalls – the only area with non-zero electric field $\mathcal{E} = |dV/dz| = V_0/S$. Consider the case of half-integer S and integer W (see Figure 3a). As the string moves towards the right starting at $\delta = 0$, the two charges (numbers 1, 2) inside the left sidewall contribute to force to the right (negative energy slope). A single charge inside the right sidewall (number 5) contributes to force to the left (positive energy slope). For δ between 0 and 1/2, there is a net force to the right of magnitude $q\mathcal{E}$, corresponding to potential of force equal to $-q\mathcal{E}\delta$. When δ reaches the value 1/2 the situation changes. Charge 2 leaves the left sidewall leaving charge 1 alone. At the same time charge 4 enters the right sidewall joining charge 5. The imbalance of charges inside sidewalls reverses: the number of charges inside the right sidewall exceeds that inside the left sidewall. The further increase of δ increases the total energy until $\delta = 1$ when the string translocates by one period. In summary, the energy minima E_{\min} are located at half-integer values and energy maxima E_{\max} – at integer values. For the trapping energy we have $\Delta E_{TR} = q\mathcal{E}/2$. Inserting back the unit of length d , we have that

$$\Delta E_{TR} = q\mathcal{E}d/2. \quad (2)$$

Similar considerations allow to analyze ΔE_{TR} for any other geometry of V . For example, if W is slightly less than integer (see Figure 3b), the numbers of charges in the left and right sidewalls are equal for certain displacements δ (shown as dashed regions). For such δ the total force applied to the string is zero, and this interval does not contribute to energy change. The net result is a decrease of ΔE_{TR} .

Even less energy change occurs in the example of Figure 3c where the sidewall dimension is an integer. The numbers of charges inside sidewalls are constant for any δ . There is no charge imbalance, so $\Delta E_{TR} = 0$.

Now we can formulate *the requirement for optimal geometry V that maximizes ΔE_{TR} : the sidewall charge imbalance should be nonzero for any displacement δ* . Any potential with half-integer sidewalls S and integer well W satisfies this requirement. Thus Figure 3a corresponds to the optimal geometry, and equation (2) gives the upper estimate on trapping energy.

In our simple model, the trapping energy depends only on how close to integer or half-integer values the trap dimensions are, and not on their absolute values. Assuming a potential with constant field \mathcal{E} , this statement is illustrated on Figure 4.

The trapping energy does not depend of the sign of V_0 , i.e. whether we have potential well or potential barrier. When V_0 changes sign, local energy minima swap their positions with energy maxima. The absolute change of energy is irrelevant since it does not have an effect on the trapping properties of the potential. Changing the sign of V_0 effectively “moves” the local minimum by half period. This sets the accuracy of string position control by the trap to $d/2$. This also suggests a simple mechanism for controlled translocation of the string through the trap (see Figure 5). Consider a trap oscillating with a frequency f in an external “dragging” electric field \mathcal{E}_{dr} . Every time the potential crosses zero, the string translocates by $d/2$ in the direction of the symmetry breaking external field. The translocation speed

$$\mathcal{V} = d f \quad (3)$$

is in principle independent on the value \mathcal{E}_{dr} . We call this effect *digital electrophoresis*. In practice, the above equality will not be exact due to Brownian fluctuations that may make the DNA molecule to choose an unintended minimum. We further discuss this issue in Section IV.

The potential trap should not necessarily be symmetrical. From a fabrication point of view, it might be easier to control the dimensions of only one sidewall. It is easy to show that ΔE_{TR} decreases by no more than a factor of 1/2 for any slope of the second sidewall.

ΔE_{TR} is a result of force imbalance inside the sidewalls, so it is critically important that at least one sidewall is well defined. In other words, the potential should have well articulated “angles”.

The above discussion considers the distance d between charges as constant, i.e. “stiff” string. Below we analyze the effect of string elasticity on ΔE_{TR} . Now our system is characterized by the vector $\vec{\delta}$, its component δ_i specifies the position of i th charge as $i + \delta_i$.

We model the energy as

$$E(\vec{\delta}) = q \sum_{i=-\infty}^{\infty} V(i + \delta_i) + \frac{k}{2} \sum_{i=-\infty}^{\infty} (\delta_i - \delta_{i-1})^2, \quad (4)$$

where the second term describes the stiffness of the string which we characterize by elasticity module k . As in the previous case, E is a periodic function of $\vec{\delta}$ with unit period. We compute ΔE_{TR} using the following procedure. We set constant the position of a charge M which is far enough from the potential well V so that the perturbation of relative positions of charges is negligible, i.e. $\delta_M - \delta_{M-1} \approx 0$. The positions of all other charges $\delta_{i \neq M}$ are computed by numerical minimization of Equation (4). ΔE_{TR} , computed as an energy barrier for translocation the string from state $\delta_M = 0$ to $\delta_M = 1$, is shown on the Figure 6 as a function of string elasticity. We use maximum relative displacement of neighboring charges $\max_i |\delta_i - \delta_{i-1}|$ to characterize elasticity instead of k since the use of this variable provides better insight. Most importantly, *the elasticity of the string leads to ΔE_{TR} decrease*. This decrease is less pronounced for shorter W , and it also depends on the polarity of V_0 .

III. Nanopore-based trap

In what follows we assume that the proposed device operates on single stranded (ss) DNA molecules. This selection is due to the practical importance of ssDNA in biochemistry and molecular biology, e.g. in DNA sequencing (see Section V). We assume the helix radius is $r_h = 0.65 \text{ nm}$ while all other parameters are similar to that of double stranded DNA. We also ignore the flexibility of the molecule and consider it as a solid helix.

Figure 7 depicts the diagram of DNA transistor. A nanopore of radius r_0 slightly exceeding r_h penetrates through a membrane that separates two reservoirs. The membrane is a ‘‘sandwich’’ of three metal electrodes separated by insulators. The thickness of the central electrode is W , and the distance between electrodes is S . The voltage difference between central and side electrode is kept at V_0 . An independent voltage difference applied between two reservoirs drives the ssDNA inside the nanopore. The phosphate charges are in close proximity to the walls of the nanopore. Setting aside for a while the perturbations introduced by the nanopore, the potential on the wall of the nanopore is the same as in our model problem. To estimate the trapping energy, we assume that the metal-dielectric sandwich is fabricated using methods typical in microelectronics, and the magnitude of electric field between the electrodes could be as high as in modern Metal-Oxide-Semiconductor Field Effect Transistors (MOSFET). For modern MOSFETs the thickness of gate oxide approaches 1 nm and the gate voltage is 1 V . Substituting $\mathcal{E} \approx 1 \text{ V} / \text{nm}$ into (2) we have the estimate of trapping energy $\Delta E_{TR} = 170 \text{ meV} \approx 13 E_T$.

We have to revise our estimate to account for field distortion introduced by the nanopore. We assume the nanopore has cylindrical geometry and compute the field distribution inside the nanopore using a 2D finite element solver. We also make the following assumptions. Silicon dioxide with dielectric constant $\epsilon_d \approx 4$ is used for electrode insulator. The macroscopic dielectric constant of water $\epsilon_w \approx 80$ can still be used inside the nanopore.

Our simulations for varying dielectric constants do not show any important dependence on their exact values. Finally, we assume that the separation between phosphate charges and the electrodes is much smaller than Debye length: $r_0 - r_h \ll l_D$ so that the screening by counterions can be neglected. From the estimate $l_D \approx 0.3 \text{ nm} / \sqrt{M}$, where M is molarity of the buffer, we see that this condition can be satisfied by decreasing concentration of the buffer.

Coaxial nanopore and polymer

We start our analysis using a simplifying assumption that the axes of the nanopore and the polymer coincide. In practice this symmetry can be “enforced” by keeping side electrodes negative, so that any drift of the negatively charged polymer toward the walls of the nanopore leads to the energy increase. The problem is axially symmetric, and the energy depends only on the displacement in z-direction. The potential acting on charges inside the nanopore depends on how far these charges are from the walls of the nanopore (see Figure 8a). For a “tight” nanopore ($r_0 / r_h \approx 1$), $V(z)$ is almost trapezoidal and the sidewalls are well defined. For “loose” (larger r_0 / r_h) nanopore, the “corners” are smooth while the depth of the trap decreases. The potential energy of the molecule (Equation (1)) is very sensitive to the shape of sidewalls (see Figure 8b). For a loose nanopore the energy practically does not change with the molecule’s displacement.

We note that ΔE_{TR} is somewhat sensitive to the exact shape of the trapping potential. From Figure 8b we see that $\Delta E_{TR} \approx 26E_T$ at $r_0 / r_h = 1$ while for the same parameters “trapezoidal” estimate (2) yields $\Delta E_{TR} \approx 13E_T$.

Polymer near the wall of nanopore

The coaxial approximation is simple but not necessarily stable. We now consider a more realistic case when the axes of polymer and nanopore are parallel and separated by distance $\Delta r \leq r_0 - r_h$. In this approximation, the polymer is allowed to drift toward the wall of the nanopore. We refer to this as “off-axial” approximation. We model ssDNA as a set of charges positioned at locations

$$z_i = \delta + d \cdot i, \quad x_i = \Delta r + r_h \cos\left(\frac{2\pi}{10}i + \varphi\right), \quad y_i = r_h \sin\left(\frac{2\pi}{10}i + \varphi\right), \quad (5)$$

where angle φ describes rotation of the molecule about its axis. For simplicity we ignore the flexibility of the molecule. Now the energy is a function of three variables - $E(\delta, \Delta r, \varphi)$. As an example, we consider geometry $S = 3.5$, $W = 4$, which is optimal for the one-dimensional model trap. Figure 9 shows the results of the numerical computation of $E(\delta) = \min_{\Delta r, \varphi} E(\delta, \Delta r, \varphi)$ for nanopores of increasing radius r_0 . Similar to the model trap, the sign of V_0 determines positions of minima – integer for minus, half-integer for plus. In contrast, the trapping energy now depends on the sign of V_0 . For $V_0 < 0$, ΔE_{TR}^- drops much slower with the increase of r_0 / r_h , as compared with ΔE_{TR}^+ , corresponding to $V_0 > 0$. The splitting of trapping energy into ΔE_{TR}^+ , ΔE_{TR}^- is a result of helicoidal charge distribu-

tion along the polymer. When $V_0 > 0$ the “crest” is attracted to the central electrode, when $V_0 < 0$ the “groove” covers this electrode to maximize the distance to nearby charges.

Figure 10 compares trapping energies of coaxial and off-axial solutions. It shows that coaxial solution puts rather serious constraints on the diameter of nanopore - ΔE_{TR} drops below E_T at $r_0/r_h \approx 1.15$ while $\Delta E_{TR}^-/E_T \approx 7$ when $r_0/r_h = 1.8$.

The above example shows that even a “loose” nanopore can serve as efficient polymer trap. Complete analysis of off-axial or more realistic approximations as well as the influence of molecular flexibility is beyond the scope of this work.

IV. Device performance

Digital electrophoresis in the presence of Brownian motion

In Section II we discussed an implementation of digital electrophoresis in the context of a deterministic model (see Figure 5). In this section we extend the concept of digital electrophoresis to a Brownian environment.

In a Brownian context, the equation governing the distributions of positions of one of the reference bases in the DNA (taken to be the one closest to $x=0$) is the Smoluchowsky equation [2]

$$\frac{\partial P(x,t)}{\partial t} = L_S P(x,t), \quad (6)$$

where L_S is the Smoluchowsky operator. When applied to the position distribution $P(x,t)$, the operator L_S is defined as [3]:

$$L_S P(x,t) \equiv D \frac{\partial^2 P(x,t)}{\partial x^2} - \frac{\partial}{\partial x} [\mu F(x,t) P(x,t)]. \quad (7)$$

In equation 7, $F(x,t)$ is the force acting on the DNA molecule at time t and position x , μ and D are the mobility and the diffusivity of the DNA molecule in the nanopore. The force $F(x,t)$ applied to the DNA, changes depending on the phase of the digital electrophoresis protocol. The implementation of digital electrophoresis in a Brownian environment is schematized in Figure 11, in which we show three phases: the equilibrium phase (Figure 11a), the drag and diffusion phase (Figure 11b) and the trapping phase (Figure 11c), which we discuss next.

Equilibrium phase. Assume that the DNA polymer is in thermal equilibrium in the trapping potential as shown in Figure 11a. For the reasons discussed in the previous sections, we choose a trapping potential with $V_0 > 0$, for which even in the “loose” nanopore configuration, we have a trapping energy barrier $E_{TR} \sim 10 k_B T$ (see Figure 10). The functional form of the trapping potential for $-d/2 < x < d/2$ is $U_{TR}(x) = \frac{2E_{TR}}{d} |x|$. Even though the po-

tential is periodic, with periodicity equal to the distance d between bases, the energy barrier is sufficiently high that we can assume a V-shaped potential well. Under this assump-

tion the force $F(x,t)$ in equation (7) is equal to $-\frac{\partial U_{TR}(x)}{\partial x} = -\text{sgn}(x) \frac{2E_{TR}}{d}$. In equilibrium the time derivative in equation (7) vanishes, and we have $L_S P_{eq}(x,t) = 0$, or explicitly:

$$D \frac{\partial P_{eq}(x,t)}{\partial x} + \mu \text{sgn}(x) \frac{2E_{TR}}{k_B T} P_{eq}(x,t) = \text{const} \quad (8)$$

with solution (after determining the constant by normalization of the distribution)

$$P_{eq}(x) \approx \frac{1}{2\xi} \exp\left(-\frac{|x|}{\xi}\right) \quad (9)$$

where the length scale $\xi = \frac{d}{2} \frac{kT}{E_{TR}}$ is the spread due to thermal motion of our reference

DNA nucleotide. In equation (9) (which is the equilibrium Boltzmann distribution), the approximate sign indicates that we have approximated the periodic potential by a triangular potential well. Observe, however, that in the off-axial loose pore regime, $\xi \approx d/20$, a tight length scale which is consistent with our approximation of the trapping periodic potential as a V-shaped potential well.

Drag and diffusion phase. At $t = 0$ the trapping potential is deactivated (i.e., the V_0 of Figure 2 is set to 0), and the DNA begins its random walk in the direction imposed by the external electric field. We assume the nanopore to be a few nm long. When a voltage of a few Volts is imposed between the reservoirs at each side of the nanopore, an electric field $\mathcal{E}_{dr} \approx 1V/nm$ is established which will drive the drag of the DNA molecule. The force is exerted on the DNA therefore $F_{dr} = qN_{eff} \mathcal{E}_{dr}$, where qN_{eff} is an effective number of un-screened charges within the nanopore multiplied by the elementary charge unit q . The drift velocity is computed according to $v_{dr} = \mu F_{dr}$, where the mobility μ can be computed using the fluctuation-dissipation theorem as $\mu = D/k_B T$, with D being the diffusivity of the DNA along the axis of the nanopore. We define the drift time, τ_{dr} , as the time it takes the DNA to drift a length d (the V-shaped potential period) by the imposed drag force. That is $\tau_{dr} = d/v_{dr}$. Putting together the previous estimates, we obtain

$$\tau_{dr} = \frac{d}{v_{dr}} = \frac{d}{D} \frac{k_B T}{N_{eff} q \mathcal{E}_{dr}} = \frac{d^2}{D} \frac{1}{N_{eff}} \frac{3k_B T}{1eV} = \frac{d^2}{D} \frac{1}{r N_{eff}} \quad (10)$$

where we have used that $1nm \sim 3d$ and defined $r = \frac{1eV}{3k_B T}$. For $T = 300K$, $r \approx 10$. The

parameter r is directly proportional to the potential difference across the nanopore. If the potential difference between the two reservoirs is reduced by 2, then $r \approx 5$. In these estimates we have ignored entropic forces of the order of $k_B T/l_{pers}$ $k_B T/l_{persistence}$ —where l_{pers} is the persistence length of the ssDNA— arising due to the lowering of the entropy needed to disentangle part of the random coil when entering the nanopore.)

As soon as the trapping potential is deactivated, the equilibrium distribution of Equation (6) starts to get distorted by the dragging forces (due to the external electric field \mathcal{E}_{dr}), and diffusion (due to Brownian motion), as shown in Figure 11b. Under these conditions, the time dependent Smoluchowsky equation ruling the dynamics of the DNA molecule position distribution is given by equation (6) with the drag velocity v_{dr} replacing the generic drift $\mu F(x)$ in the Smoluchowsky operator:

$$\frac{\partial P(x,t)}{\partial t} = D \frac{\partial^2 P(x,t)}{\partial x^2} - \frac{\partial}{\partial x} [v_{dr} P(x,t)]. \quad (11)$$

The solution of (11) with initial condition given by the equilibrium distribution $P(x, t=0) = P_{eq}(x)$ is given by

$$P(x,t) = \int_{-\infty}^{\infty} dx' \frac{1}{\sqrt{4\pi Dt}} \exp\left[-\frac{(x-x'-v_{dr}t)^2}{4Dt}\right] P_{eq}(x'). \quad (12)$$

At $t = \tau_{dr}$, the peak of the distribution is at $x=d$ (Figure 11b).

Trapping phase. At $t = \tau_{dr}$, we re-establish the trapping potential (Figure 11c). The distributions of positions will now evolve from the initial condition $P(x, \tau_{dr})$, equation (12), to the stationary distribution in the trapping potential $P_{eq}(x)$, equation (9), according to the time dependent Smoluchowski equation (6) in the trapping potential of force

$U_{TR}(x) = \frac{2E_{TR}}{d}|x|$. The important fact about the solution to equation (6) with this potential

is that the characteristic time scale τ_{TR} to re-obtain the trapping equilibrium distribution is given by the smallest (in absolute value), nonzero eigenvalue of the equation $L_{FP}P = \lambda P$. It can be shown [3] that the smallest nonzero λ is D/ξ^2 . We will assume that trapping equilibrium obtains after, say, 5 characteristic times and take the trapping phase time to be $\tau_{TR} \approx 5\xi^2/D$. At this time we have re-trapped the DNA to its equilibrium state. Notice that the reference nucleotide, which was close to $x=0$ at $t=0$ has now translocated to $x=d$. After a time $\tau_{dr} + \tau_{TR}$ has elapsed, we are back in the situation depicted in Figure 11a, and we have completed a cycle of digital electrophoresis. At this point another cycle of oscillations start, that will translocate the DNA one more period to the right. This process of directed motion by turning on and off a trapping potential well has been previously described in [4], in the absence of an external drag force, but with an asymmetric potential well.

Frequency of trap oscillation

A cycle of digital electrophoresis consists of a lapse τ_{dr} during which the DNA drifts and diffuses and a lapse τ_{TR} where it gets re-trapped. Therefore the frequency of oscillation of

the potential well is $f = 1/(\tau_{dr} + \tau_{TR})$. Recalling our estimates of $\tau_{dr} = \frac{d^2}{D} \frac{1}{rN_{eff}}$ and

$\tau_{TR} = \frac{d^2}{D} \frac{1}{80}$, using the value $D = 10^{-12} \text{ m}^2/\text{s}$ [5] and $d=0.34 \text{ nm}$, and estimating $r=10$

and $N_{eff} = 15$ (N_{eff} is the number of DNA charges inside the nanopore, whose length is $\sim 3\text{nm}$), we obtain $\tau_{dr} \approx 0.7\text{ns}$ and $\tau_{TR} \approx 1.2\text{ns}$, from where $f \approx 500\text{MHz}$. Such frequencies are well with the capabilities of modern electronics. In practice f can be decreased by using larger τ_{TR} . We will see in the next subsection that τ_{dr} cannot be much larger than this estimate without degrading the accuracy of 1 translocation per cycle. On the other hand, the trapping time cannot be smaller than τ_{TR} if we want the translocations to be well defined.

Accuracy of operation for one translocation step

There are a few possible scenarios in which the accuracy of the operation of the digital electrophoresis imagined in Figure 11 can be less than perfect. First, if the polymer escapes the trap (especially if it is near to the top of the trap in the drift and diffusion step at the time of re-establishing the trap), in which case the equilibrium will be established in another minimum at the right or the left of the intended one. Second, the translocation of a polymer by one monomer for each oscillation of the trapping potential is not guaranteed: e.g., the diffusion wave shown in Figure 11b can be spread enough that at the time of re-establishing of the trapping potential the DNA molecule is not located in the intended wall. The two effects are linked: if there is little leakage due to the second reason, then there will be little chance that the DNA molecule will jump a barrier, simply because the probability of being close to the top of the barrier is small. Therefore we compute the accuracy assuming that if at time t the DNA molecule is in the basin of attraction of a given minimum, then it will reach equilibrium in that same potential well. With these considerations in mind, we define the accuracy of the DNA transistor as the probability of correct translocation by one monomer after a drift and diffusion time t :

$$A(t; N_{eff}, r) = \int_{d/2}^{3d/2} dx P(x, t) \quad (13)$$

The dependence of A on t for different values of N_{eff} and r is shown in Figure 12. For $r = 10$ (and both values of $N_{eff} = 15$ and 20, black lines in Figure 12) the accuracy is very close to 1 for values of t in the range $0.75\tau_{TR} \leq t \leq 1.25\tau_{TR}$, reaching the maximum at $t = \tau_{TR}$. For smaller values of r (i.e. lower drift fields), the accuracy degrades, and a calibration of the drift force becomes necessary to stay at the value $t = \tau_{TR}$ that maximizes the accuracy.

V. Discussion

Applications of the DNA transistor to DNA sequencing

The technologies that make reading DNA fast, cheap and widely available have the potential to revolutionize bio-medical research and herald the era of personalized medicine. The sequencing of a first single human genome by the international team of scientists from the Human Genome Project took 15 years and \$3 billion to complete, culminating in 2003 with the announcement of the publicly available rough draft by former US President Bill Clinton and British Prime Minister Tony Blair. The continuing refinement of

sequencing techniques brought the price of human genome sequence down to about \$20 million, making it affordable for large research organizations. A human genome sequencing capability affordable for individuals is the ultimate goal of DNA sequencing industry, and is commonly referred to as “\$1000 genome”. \$1000 genome will enable comparative studies of variations between individuals in both sickness and health. Ultimately it can improve the quality of medical care by identifying patients who will gain the greatest benefit from a particular medicine, and those who are most at risk of adverse reactions. The research on \$1000 genome is encouraged by the National Institutes of Health funding programs for “Revolutionary Genome Sequencing Techniques” challenging scientists to achieve \$100,000 human genome by 2009 and \$1000 genome by 2014. A highly visible private X-prize [6] has been established to reward the first group to achieve the \$1000 genome milestone.

Sanger sequencing

We see an immediate application of our device in *reducing the cost of conventional DNA sequencing*. The workhorse of sequencing technologies, sequencing by separation, was developed in 1970s by Frederic Sanger. Nowadays this technique (1) breaks DNA into fragments of different lengths, (2) adds fluorescent tags to fragment’s terminals, (3) separates different fragments by length using capillary electrophoresis, (4) identifies fragments by fluorescent tags, and finally (5) assembles complete genome from identified fragments using computers. Capillary electrophoresis is an important contributor to the cost of sequencing. It limits the decrease of sequencing cost in two ways. First, the present day separation is still a slow process (tens of base pairs per second), resulting in significant operation costs. This issue is being addressed by capillary arrays, which significantly increase the throughput of the technique. Second, separation still requires significant amount of DNA material (femtomole), hence reagent cost. This fact is considered the main obstacle in reaching the \$1,000 target.

Our device seems to be capable of determining the length of a single DNA molecule by counting the number of trapping pulses required to translocate the molecule through the nanopore. The presence of the molecule inside the nanopore can be detected, for example, by monitoring the ion current through the nanopore. Taken alone, the single molecule length separation can bring down the amount of reagents by eight orders of magnitude. A somewhat similar approach to measuring the length of single DNA molecule based on translocation time through a nanopore is suggested in [7].

For single molecule length separation the original Sanger approach (four parallel extension reactions for different nucleotide types) is more suitable since the monitoring ion current does not seem capable of resolving the type of terminating nucleotide. We emphasize the fact that in this case the need for optical detection is eliminated by using electronic detection.

Additionally, we envision that the device can be fabricated using mainstream microelectronics methods which allow a high level of device integration, further decreasing sequencing costs.

Finally, unlike conventional electrophoresis, digital electrophoresis seems to be applicable to DNA fragments of arbitrary large lengths, provided that the accuracy of translocation is sufficiently close to 1.

Nanopore sequencing

The interest in using nanopores was sparked by the experimental detection of single RNA and DNA molecules translocation through a biological nanopore formed by a membrane-bound protein [8]. The detection of single molecules is based on applying DC voltage (fraction of Volts) between two reservoirs and monitoring the resulting ion current (fraction of nA) through the nanopore. The translocation of a few hundred nucleotides long polymer molecule through the nanopore causes almost complete blockade of the current for a fraction of ms which can be easily detected. The original motivation of this and subsequent experiments [9, 10] was ultra-fast single DNA molecule sequencing by discriminating nucleotide types based in their ion current signature. After more than 10 years of research this goal is still out of reach.

Soon DNA translocation experiments were repeated using synthetic nanopores formed in materials such as Si_3N_4 [11] or SiO_2 [12]. The transition to synthetic nanopores was determined in large part by their reliability and the desire to use well-developed and understood microelectronics fabrication methods instead of protein engineering, which, by standards of microelectronics, is still in its infancy. Additional motivation for synthetic nanopores is a possibility to place metal electrodes in the vicinity of translocating polymer molecule, a feature that is completely out of reach for known biological nanopores. According to computer simulations metal electrodes may facilitate nucleotide type discrimination by measuring transverse electron transport [13] or capacitance variations [14]. Similar to ion current based efforts, these methods have not yet produced a single base resolution.

The failure to discriminate individual nucleotides is attributed to excessively fast DNA translocation speed, in the range of tens nucleotides per microsecond. Attempts to slow down this speed by controlling electrolyte temperature, buffer concentration and viscosity resulted in the order of magnitude slower speeds [15].

A less immediate but nonetheless promising application for our device is nanopore sequencing. We believe that our device could help these methods achieve their goal by *localizing* the DNA inside the nanopore with single nucleotide accuracy, thus giving enough time to perform required measurements.

Other applications

Other applications of this technology include the possibility of quantification of the concentration of specific segments of DNA or RNA by counting the number of DNA fragments with a given sequence. If the segment is long enough (say longer than 15 to 20 nucleotides) then the segment will give a high specificity for detection of organism with known genomes. This would make it possible, e.g., to detect pathogens in biological samples. Furthermore, if the technology is fast enough, measuring for a long time would allow for a sufficient sampling of the genomic material present in a sample to detect with good sensitivity even small traces of pathogens.

Implementation issues

Modern thin film deposition technologies, such as Atomic Layer Deposition (ALD), seem to be well suited for metal-dielectric sandwich fabrication. For example, ALD is capable of depositing metal or dielectric layers with single atom precision. The most challenging part now seems to be defining a nanopore through the sandwich. The neces-

sity to place metal electrodes in close proximity with DNA renders unpractical a number of fabrication methods currently used in nanopore sequencing research such as ion-beam sculpturing of Si_3N_4 nanopores [11] or electron beam shrinking of SiO_2 nanopores [12]. These methods start with a large (tens of nm) pore and narrow it by using some type of material reflow or deposition. The later process will not preserve the required sandwich-like structure of the membrane. The level of resolution needed for our device seems to be within the reach of such techniques as electron beam decomposition and sputtering [16], focused ion beam or emerging technologies like Scanning Probe Lithography.

Possible problems

We see the flexibility of ssDNA molecule as a possible problem. Our simple model of ssDNA assumes that the distance between charges on DNA backbone is constant, while in real life the persistence length of ssDNA is just few nm. The confinement inside the nanopore is likely to improve the “stiffness” of the molecule. Also, using very thin central electrode (few atomic distances) could be of help.

Another possible problem could come from the predicted increase of counterion condensation on DNA molecule in nanopore [17]. This effect may also decrease the depth of electrostatic trap.

Finally, the bead and spring type models used in this work are obviously over-simplified and can not address in detail the problems raised above. The use of more realistic models, such as an elastic rod model (e.g. see [18]) or even the full scale molecular dynamics simulations of DNA in nanopore such as [19, 20], are the next step in the analysis of the device.

VI. Conclusion

The possibility of controlling the position of a polynucleotide molecule with a single nucleotide resolution is interesting both from fundamental and practical points of view. The discreteness of charge distribution on phosphate groups of polynucleotide backbone suggests the use of sharp-shaped electrostatic potential wells to accomplish this goal. Our DNA transistor is a nanopore defined in an atomically precise sandwich of metal electrodes and insulators. Application of external voltages to the electrodes creates a potential well inside DNA transistor that is capable of trapping the polymer molecule with a trapping energy in the range of a few $k_B T$ range. Variation of applied voltages leads to translocation of the DNA molecule by one nucleotide with high probability. We called this effect digital electrophoresis and estimate that it can run at ~ 500 MHz range frequencies. Digital electrophoresis can find immediate application in bringing down the cost of conventional DNA sequencing down by minimizing the amount of required chemicals and speeding up DNA fragment separation.

We outline a number of detrimental effects such as translocation accuracy, flexibility of nucleotide molecules, and increased counterion condensation. It is difficult to estimate the seriousness of these problems without referring to more advanced models such as elastic rod model of DNA or full-scale molecular dynamics simulations, which we are addressing in ongoing studies.

VII. Acknowledgements

We thank useful conversations with Joe Jasinski (who suggested the name DNA transistor), Ajay Royyuru, Kumar Wickramasinghe, Yuhai Tu, Alan Grossfield , and Tom Theis.

References:

1. Saenger, W., *Principles of Nucleic Acid Structure* 1988: Springer.
2. Lubensky, D.K. and D.R. Nelson, *Driven polymer translocation through a narrow pore*. Biophys J, 1999. **77**(4): p. 1824-38.
3. Risken, H., *The Fokker-Planck equation: Methods of solution and applications*. 2 ed. 1989, Berlin: Springer-Verlag.
4. Faucheux, L.P., et al., *Optical Thermal Ratchet*. Phys. Rev. Lett, 1995. **74**: p. 1504 - 1507.
5. Dill, K. and S. Bromberg, *Molecular driving forces*. 2002.
6. *X-Prize Foundation Web-Site*. Available from: <http://www.xprize.org/>.
7. Heng, J.B., et al., *Sizing DNA using a nanometer-diameter pore*. Biophys J, 2004. **87**(4): p. 2905-11.
8. Kasianowicz, J.J., et al., *Characterization of individual polynucleotide molecules using a membrane channel*. Proc. Natl. Acad. Sci. USA, 1996. **93**: p. 13770-13773.
9. Akeson, M., et al., *Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules*. Biophys J, 1999. **77**(6): p. 3227-33.
10. Meller, A., L. Nivon, and D. Branton, *Voltage-Driven DNA Translocations through a Nanopore*. Physical Review Letters, 2001. **86**(15): p. 3435.
11. Li, J., et al., *Ion-beam sculpting at nanometre length scales*. Nature, 2001. **412**: p. 166-169.
12. Storm, A.J., et al., *Translocation of double-strand DNA through a silicon oxide nanopore*. Physical Review E, 2005. **71**(5).
13. Lagerqvist, J., M. Zwolak, and M. Di Ventra, *Fast DNA sequencing via transverse electronic transport*. Nano Letters, 2006. **6**(4): p. 779-782.
14. Gracheva, M.E., A. Aksimentiev, and J.P. Leburton, *Electrical signatures of single-stranded DNA with single base mutations in a nanopore capacitor*. Nanotechnology, 2006. **17**(13): p. 3160-3165.
15. Fologea, D., et al., *Slowing DNA Translocation in a Solid-State Nanopore*. Nano Lett., 2005. **5**(9): p. 1734 – 1737.
16. Ho, C., et al., *Electrolytic transport through a synthetic nanometer-diameter pore*. Proc Natl Acad Sci U S A, 2005. **102**(30): p. 10445-50.
17. Rabin, Y. and M. Tanaka, *DNA in nanopores: Counterion condensation and coion depletion*. Physical Review Letters, 2005. **94**(14).
18. Olson, W.K. and V.B. Zhurkin, *Modeling DNA deformations*. Curr Opin Struct Biol, 2000. **10**(3): p. 286-97.
19. Heng, J.B., et al., *The electromechanics of DNA in a synthetic nanopore*. Biophys J, 2006. **90**(3): p. 1098-106.
20. Aksimentiev, A., et al., *Microscopic Kinetics of DNA Translocation through synthetic nanopores*. Biophys J, 2004. **87**(3): p. 2086-97.

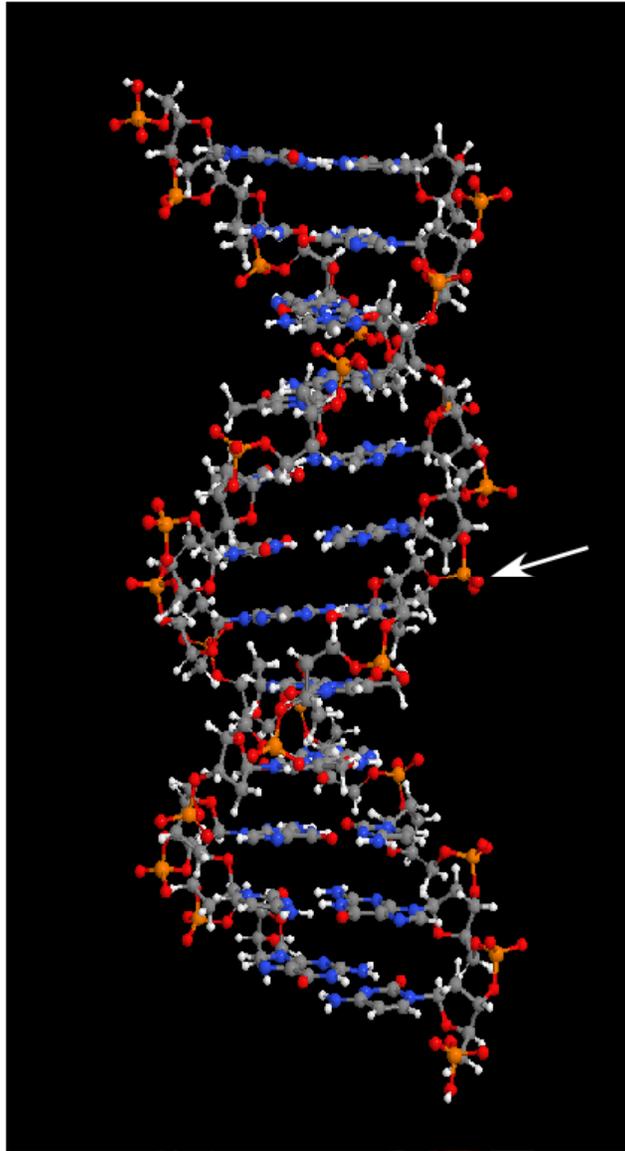


Figure 1: A 3D visualization of double stranded DNA molecule. Location of phosphate group (phosphorus – orange; oxygen - red) is shown by arrow.

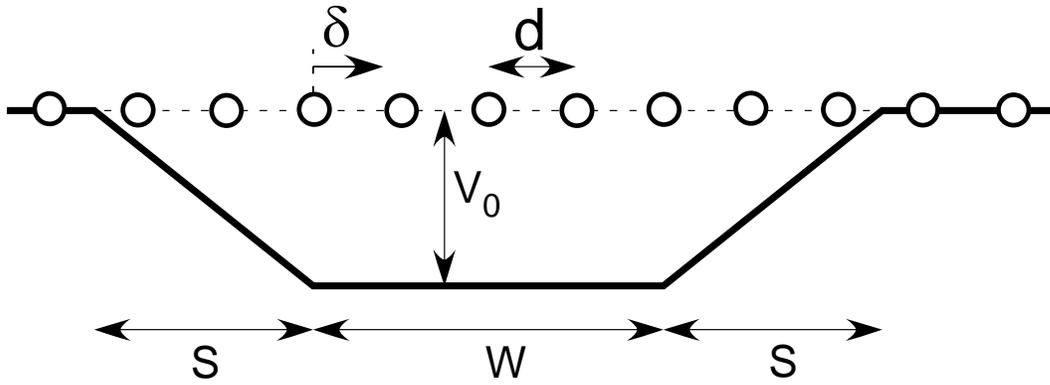


Figure 2: Polymer string inside trap potential. S is the size of the sidewalls. W is the length of the potential well.

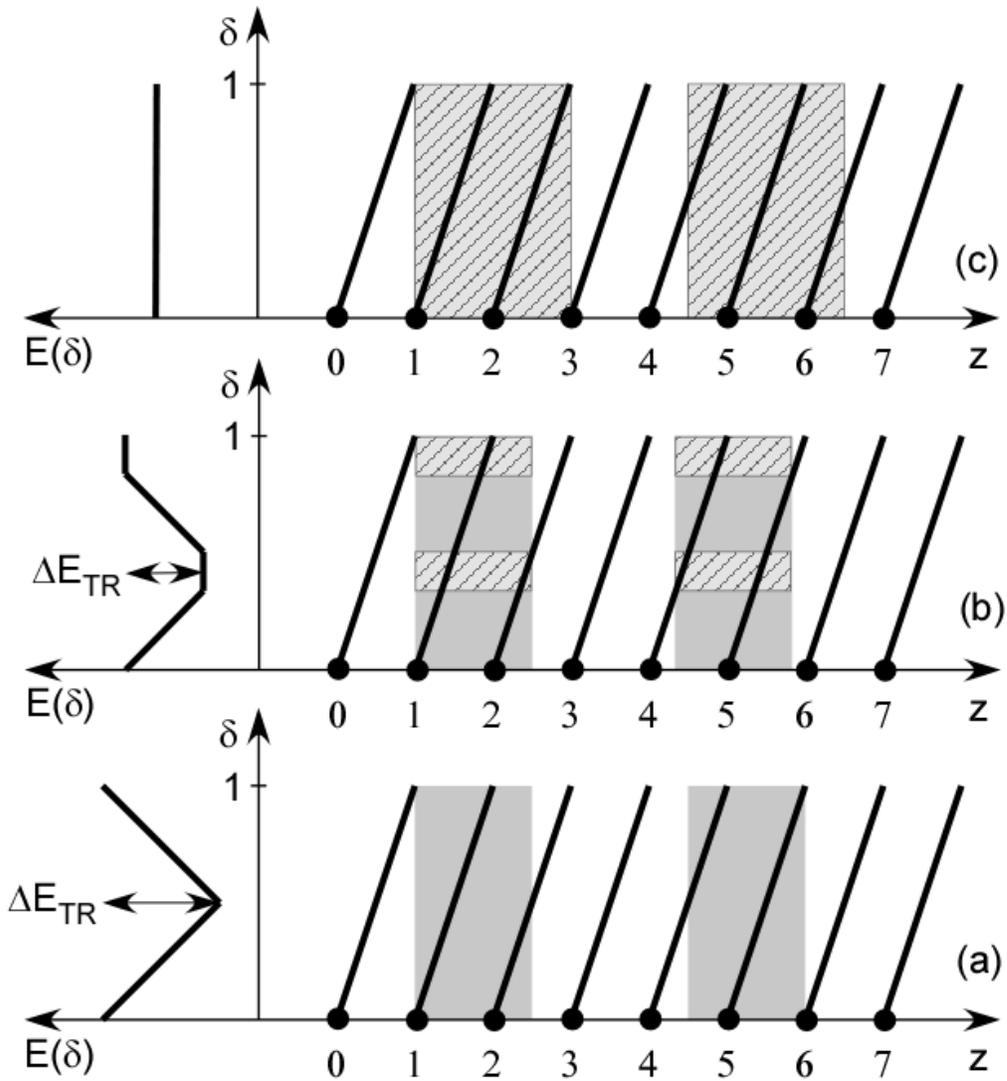


Figure 3: Energy of the string: (a) optimum case - half-integer sidewall, integer well; (b) suboptimal case - half-integer sidewall, non-integer well; (c) worst case - integer sidewall.

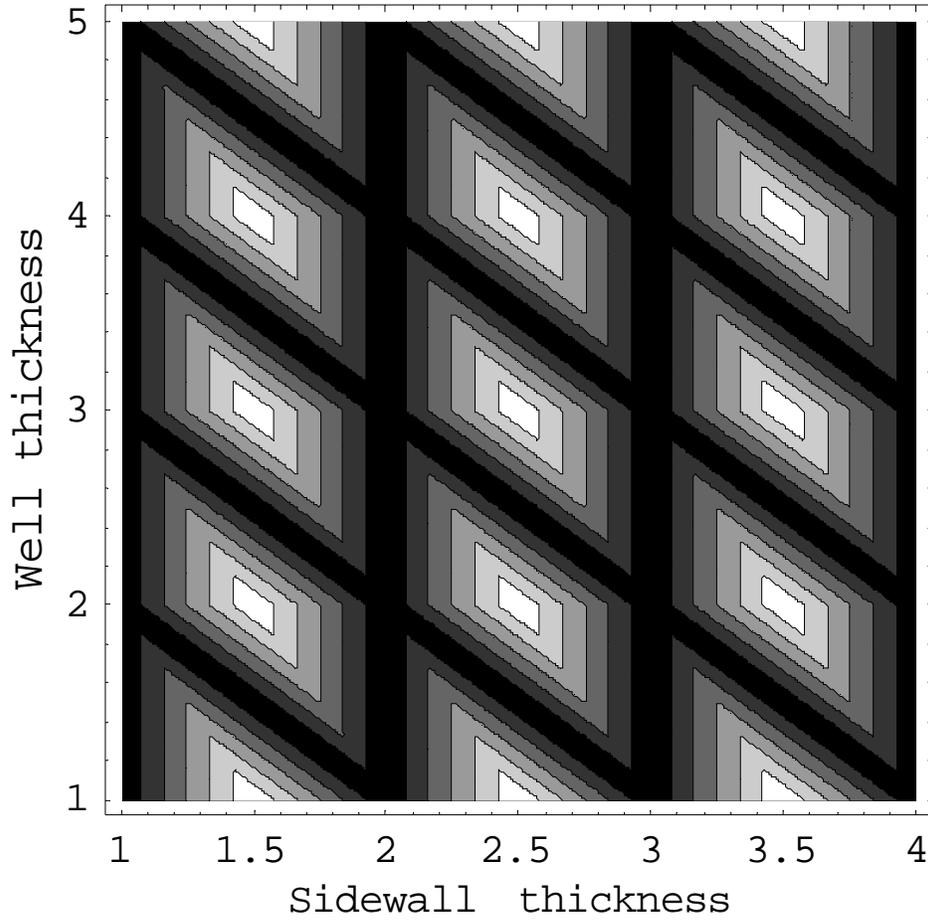


Figure 4: Dependence of ΔE_{TR} geometry of trapping potential. The field inside sidewalls (V_0 divided by sidewall thickness) is kept constant. Lighter regions correspond to higher values of ΔE_{TR} .

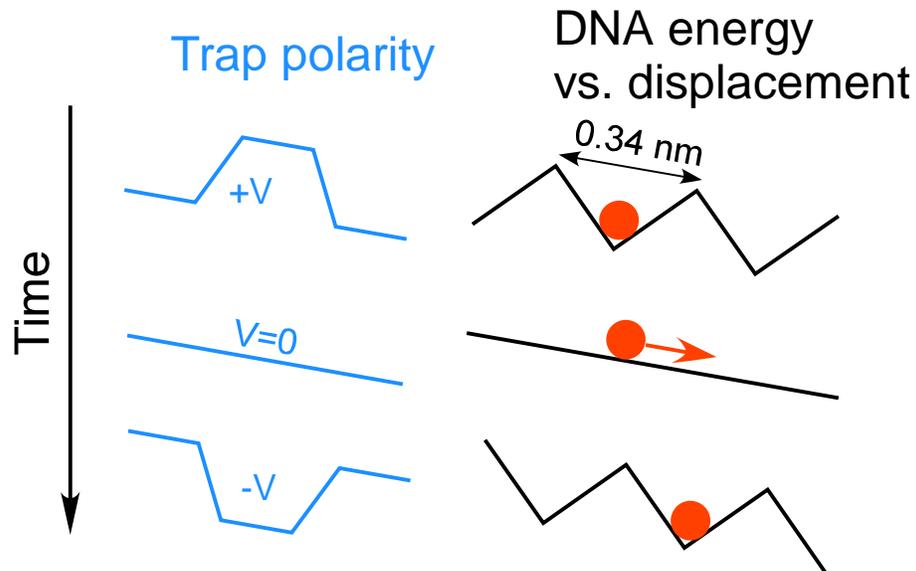


Figure 5: The concept of digital electrophoresis.

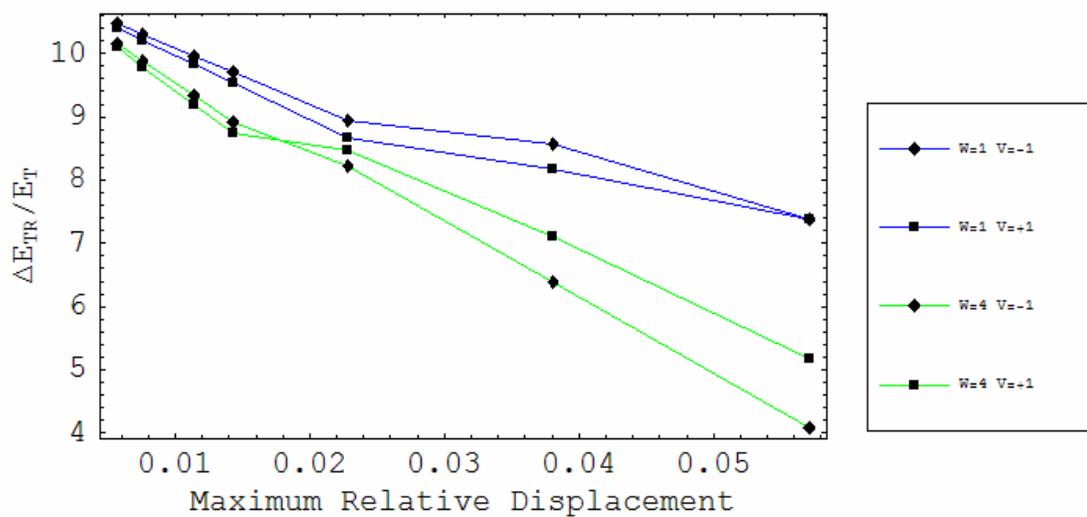


Figure 6: Trapping energy of an elastic string. Different colors correspond to different width of the potential well $W = 1, 4$. Different symbol shapes correspond to different polarities of the trapping-potential $V_0 = +1V, -1V$. Sidewall thickness is $S = 3.5$.

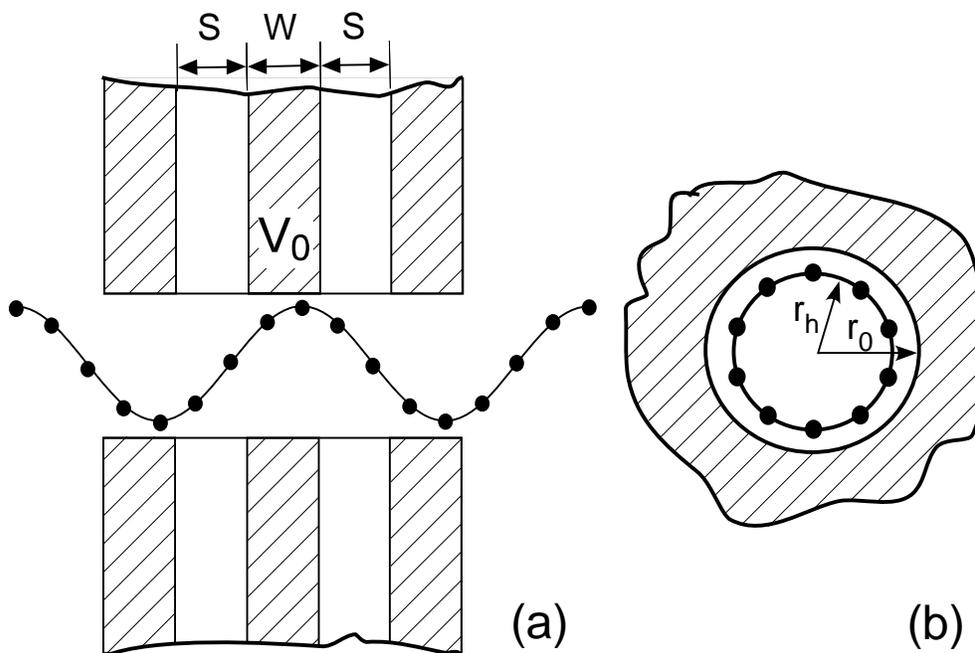
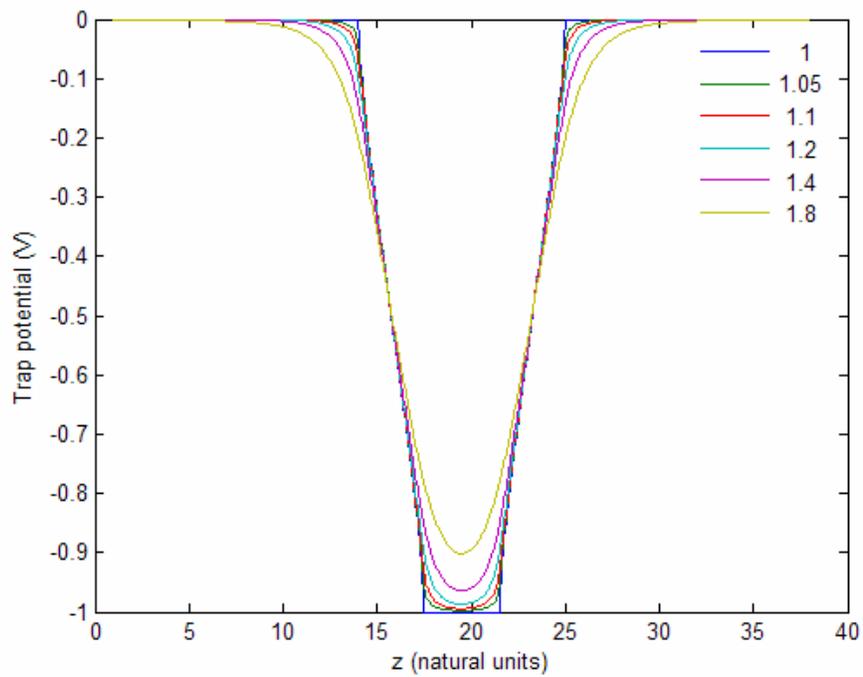
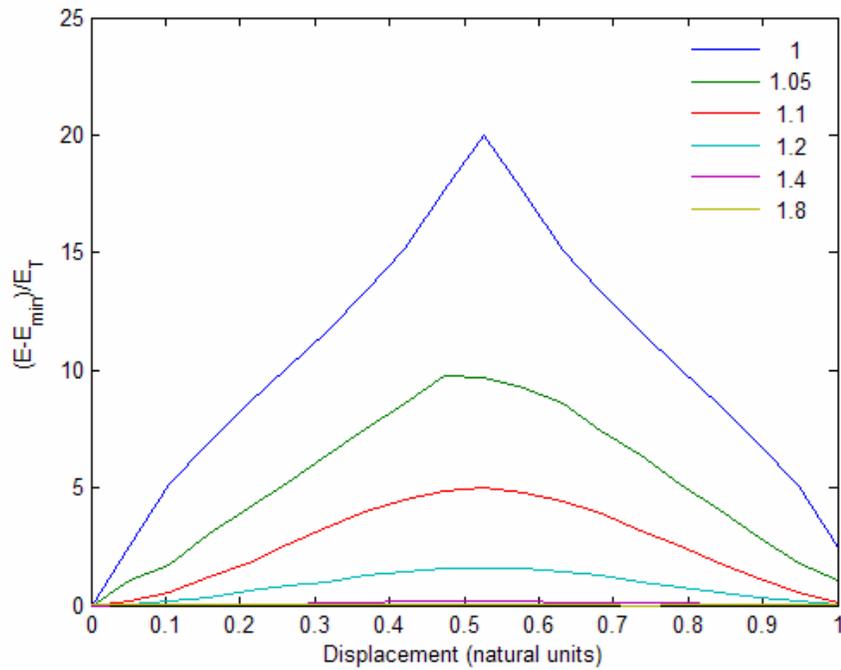


Figure 7: Nanopore-based trap device. (a) Cross-Section. (b) Top view.



(a)



(b)

Figure 8: Influence of nanopore radius on trap parameters: (a) Potential acting on DNA charges; (b) Dependence of DNA energy on displacement. Different curves correspond to different r_0/r_h . All curves correspond to optimal trap parameters $S = 3.5$, $W = 4$. Trap depth is kept at $V_0 = 1V$.

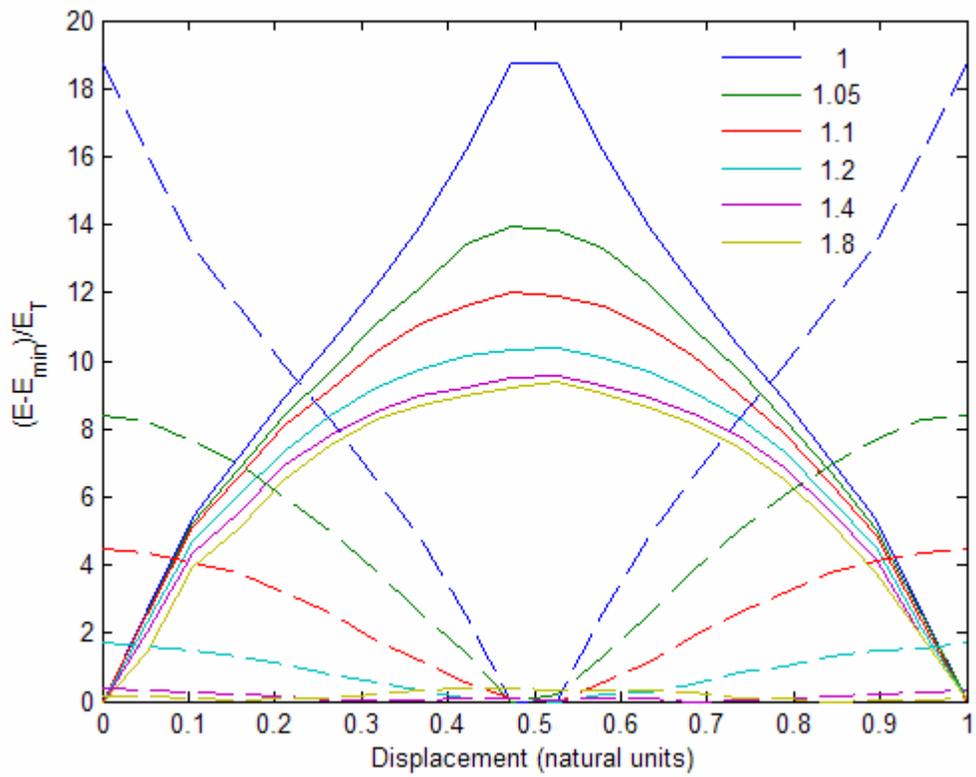


Figure 9: Dependence of energy on displacement using off-axial approximation ($S = 3.5$, $W = 4$). Different curves correspond to varying parameter r_0 / r_h . Solid line - $V_0 = -1V$, dashed line - $V_0 = +1V$.

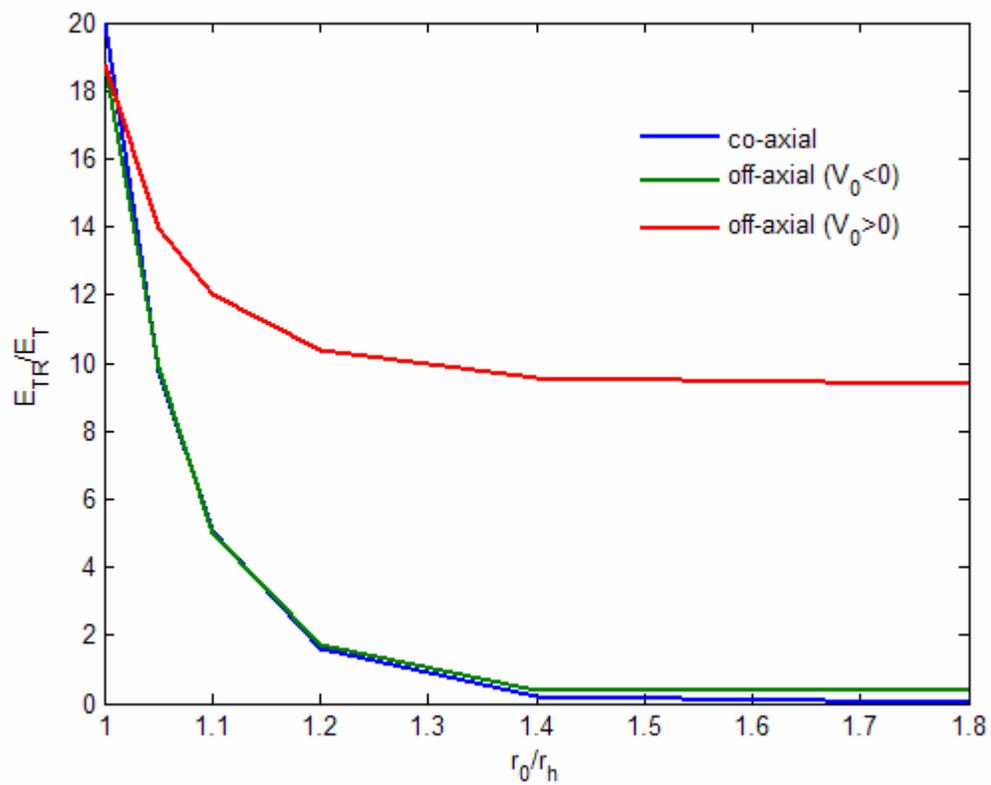


Figure 10: Trapping energy dependence on nanopore radius. Trap parameters are the same as on Figure 8

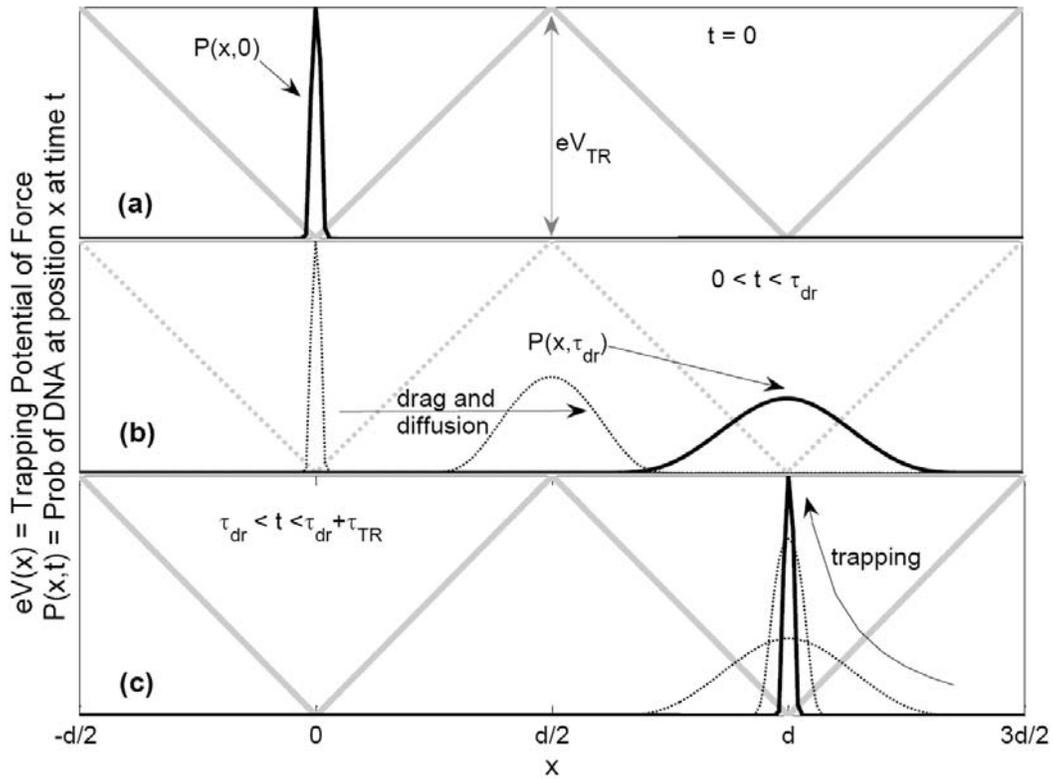


Figure 11: Time evolution of DNA position distribution during one cycle of digital electrophoresis.

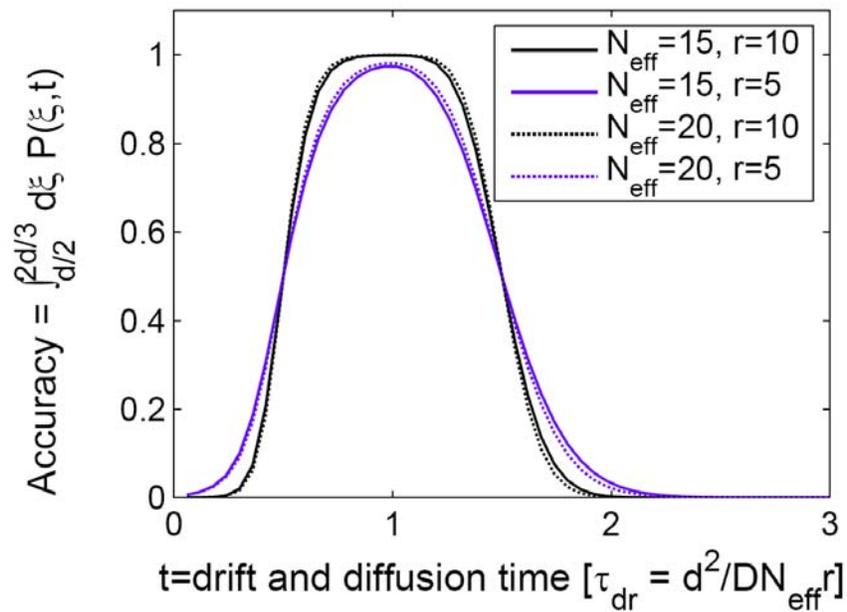


Figure 12: Accuracy of digital electrophoresis as a function of drift and diffusion time for the values of N_{eff} and r indicate in the legend.