A stochastic model for DNA translocation through an electropore

Miao Yu a, Wenchang Tan b, Hao Lin a,⁎

a Mechanical and Aerospace Engineering, Rutgers, The State University of New Jersey, 98 Brett Road, Piscataway, NJ 08854, USA
b State Key Laboratory for Turbulence and Complex Systems, Department of Mechanics and Engineering Science, College of Engineering, Peking University, Beijing 100871, China

1. Introduction

The phenomenon termed “electroporation” or “electropermeabilization” has wide applications in many fields [1–10]. The membrane becomes significantly permeabilized when the transmembrane potential (TMP) exceeds a critical threshold [11–18]. (A list of abbreviations is given in Table 1.) The generation of conducting electropores on the membrane [19] has been a widely-accepted theory to explain this phenomenon. The characteristics of these pores have been investigated by both theoretical models [20–26] and experimental measurements [27–30]. The delivery of small molecules is likely mediated by electrophoresis and diffusion through these transient openings. Based on this understanding, model studies [31–34] have been carried out to quantify the transfer of small ions into the cell, which generated predictions in good agreement with experimental data [29,35,36]. On the other hand, the transfer of large molecules such as DNA is presumably more complex. Although many studies have been carried out to optimize the transfection efficiency (TE) [37–45], the approach is primarily empirical and a comprehensive understanding has yet to be established [46].

One of the key differences between the delivery of DNA and small molecules is that the former stays much longer at the membrane, forming the so-called DNA–membrane complex [47–51]. The detailed structure of these complexes is still under debate. One theory hypothesizes that the internalization of DNA is facilitated by endocytosis, which we term the “endocytosis theory” [52–55]. In particular, a recent work by Wu et al. found that the TE is significantly decreased if certain endocytic mechanisms were inhibited, providing support for this hypothesis [53]. The results are, however, not yet conclusive. For example, a recent experiment by Pavlin et al. [56] indicates no significant correlation between the field strength and intracellular vesiculation, which is presumably required for endocytosis.

Meanwhile, other mechanisms for DNA electrotansfer have not been sufficiently explored. In this work, we focus on the so-called “translocation theory”. In this theory, macromolecules also go through the membrane like the small ones. However, the process is more complex due to their polymeric nature, and the translocation time can be long, which potentially explains the long coherence time of the DNA–membrane complex [47,50]. We are motivated by experimental observations from the literature. In Sukharev et al., the authors observed that the existence of DNA in the buffer solution greatly enhanced cell uptake for smaller pores, whereas the LV assists DNA translocation electrophoretically. The above arguments suggest that the translocation theory warrants further examination. In particular, the development of a predictive model is much needed, such as to generate quantitative data to directly correlate with experimental observations.

In what follows, we present a stochastic model to study DNA translocation across an electropore driven by finite-time electric pulses. The
model is based on previous developments [61–63] which were primarily used to study DNA sequencing with synthesized or protein nanopores. The model is capable of predicting the probability of successful translocation (PST) as a function of field strength, pulse length, and DNA size. Furthermore, many important insights are reviewed, including the effects of the electroforetic and the diffusive time scales, and a power-law correlation between the final probability of successful translocation (F-PST) and the governing parameters. The simulated results are compared with previous data to interpret the trends. In particular, the electric potentials of regions I and II are 
\[\Phi_1(t)\text{ and }\Phi_2(t)\], and the number of segments in region I is 
\[N_m\] segments pinned at the pore on one end. At each instant, the effective rate of change coefficient, and is given by the formula:
\[
k_0 = \frac{D_0}{L^2}.L = \frac{0.918\lambda_d}{N^0.4}.
\]
where \(D_0\) is the effective chain diffusivity during translocation, and \(L\) is a characteristic length derived from the radius of gyration, \(R_g\). 

**2. Model formulation**

A schematic describing the physical problem is shown in Fig. 1. The membrane is modeled as an infinitesimally thin plane separating two regions, namely, I and II. A pore is embedded on the membrane. The electric potentials of regions I and II are \(\Phi_1(t)\) and \(\Phi_2(t)\), respectively. Note that \(\Phi_1(t)\) and \(\Phi_2(t)\) are spatial constants, leading to a discontinuity known as the TMP, \(V_m = \Phi_1 - \Phi_2\). The DNA molecule is modeled as a charged polymer chain consisting of \(N\) segments, each with the Kuhn-length, \(l_k = 100\) nm for double-strand DNA, see Table 2 [64]. The polymer chain translocates through the pore as a single strand, i.e., we do not consider hairpin structures [65]. The number of segments in region II (the intracellular space) is denoted by \(m\). Correspondingly, the number of segments in region I is \(N - m\). In each region, the chain is assumed to be a random walk of \(m\) (or \(N - m\)) segments pinned at the pore on one end. At each instant, the chains are at thermodynamic equilibria.

Our model follows that of Sung et al. [61], which was initially used to simulate the translocation of a polymer through a membrane pore under a constant potential difference. This model has been subsequently developed by Muthukumar and other authors [62,63]. The above problem can be described by a one-dimensional Fokker–Planck equation governing the evolution of the probability density function (PDF), \(P\).

\[
\frac{\partial P(m,t)}{\partial t} = \frac{\partial}{\partial m} \left[ \frac{k_b T}{k_0} P(m,t) \frac{\partial f(m,t)}{\partial m} + k_0 \frac{\partial}{\partial m} P(m,t) \right]. \tag{1}
\]

Here \(P(m,t)\) is the probability density of having \(m\) segments in region II at time \(t\). The probability of having \(N\) segments in region II is the probability of successful translocation. \(k_b\) and \(T\) are the Boltzmann constant and temperature, respectively. \(k_0\) is the effective rate of change coefficient, and is given by the formula:

\[
\Delta u = \left\{ \begin{array}{ll}
\tilde{z} e V_m & \text{if } 0 < t < t_p, \\
0 & \text{if } t > t_p,
\end{array} \right.
\]

\(\tilde{z}\) is the effective charge number per segment, \(e\) is the electron charge, \(V_m\) is the TMP introduced above, and \(t_p\) is the pulse length. Note that our formulation deviates slightly from the previous work in two aspects. First, the effective rate of change coefficient, \(k_0\), is derived from the chain diffusivity \(D_0\) from a scaling analysis. Second, the effective charge number per segment, \(\tilde{z}\), is specified according to previous experimental measurements. The details of the derivations and arguments are found in Appendix A.

Eqs. (1)–(4) are solved numerically using a second-order finite-volume method. The convective term in Eq. (1) is discretized using an upwind scheme. A Crank–Nicolson algorithm is used to integrate the diffusive term in time. At the ends of the computational domain, \(m = 0\) and \(N\), we employ absorbing boundary conditions [61,67]. Namely,

\[
P(m = 0, N, t) = 0. \tag{5}
\]

In adopting Eq. (5), we assume that once the chain leaves the pore from either side, it will never return. This assumption is consistent with the fact that the entropic energy drops abruptly when the chain departs from the membrane. The flux at \(m = 0\) is collected, which we term the probability of unsuccessful translocation (PUT). The flux at \(m = N\) is collected and termed the probability of successful translocation (PST).

For initial condition, we assume a narrow-band Gaussian distribution approximating a delta function and satisfying the normalization condition \(\int_0^N P(m, t=0) dm = 1\). For most of the cases studied below, the center of the initial profile is located at \(m_0 = 0.2\). The standard deviation of the

---

**Table 1**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMP</td>
<td>Transmembrane potential</td>
</tr>
<tr>
<td>TE</td>
<td>Transfection efficiency</td>
</tr>
<tr>
<td>PST</td>
<td>Probability of successful translocation</td>
</tr>
<tr>
<td>F-PST</td>
<td>Final probability of successful translocation</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability density function</td>
</tr>
<tr>
<td>PUT</td>
<td>Probability of unsuccessful translocation</td>
</tr>
</tbody>
</table>

**Table 2**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
<th>Value/Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>(F)</td>
<td>Faraday constant</td>
<td>96,485 C/mol</td>
</tr>
<tr>
<td>(k_b)</td>
<td>Boltzmann constant</td>
<td>(1.38 \times 10^{-23}) J/K</td>
</tr>
<tr>
<td>(T)</td>
<td>Room temperature</td>
<td>298.15 K</td>
</tr>
<tr>
<td>(l_k)</td>
<td>The Kuhn-length of ds-DNA</td>
<td>100 nm [64]</td>
</tr>
<tr>
<td>(l_{ns})</td>
<td>DNA base-pair length</td>
<td>0.34 nm</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>Constant for a self-avoiding chain</td>
<td>0.69 [62]</td>
</tr>
</tbody>
</table>

**List of model parameters.**

- **Abbreviation**: PDF Probability density function
- **Definition**: \(\frac{\partial}{\partial t} P(m) = \frac{\partial}{\partial m} \left[ \frac{k_b T}{k_0} P(m) \frac{\partial f}{\partial m} + k_0 \frac{\partial}{\partial m} P(m) \right] \)

---

**Fig. 1.** A schematic of the problem. The membrane is an infinitesimally thin plain separating region I (the extracellular space) and region II (the intracellular space). The electric potential in each region is denoted by \(\Phi\).
Gaussian distribution is chosen to be 0.05. In choosing this initial condition, we assume that a small fractional segment of the polymer chain is already inserted into the pore. Other initial locations ($m_0 = 0.1,0.5, 1, \text{ and } 1.5$) are also studied to examine the effects of initial condition in Section 3. All model parameters are listed in Table 2.

3. Results

In what follows, we first present results using a DNA polymer chain of $N = 24$, which approximately corresponds to a linear length of 7.2 kbp. We assume the initial location is $m_0 = 0.2$. The effects of DNA length and initial location are studied later.

Fig. 2 demonstrates the typical evolution of the PDF, $P(m,t)$, subject to a TMP of $V_m = 0.2$ V and a pulse length of $t_p = 5$ ms. The pulsing parameters are chosen according to the experiments from Sukharev et al. [57] and a comparison with the data is presented later. During the pulse (Fig. 2a–c), the peak spreads by diffusion, and is pushed toward the right under the influence of the electrophoretic force. After the pulse ceases, the profile experiences diffusion only and spreads further toward the ends of the domain (Fig. 2d). At each instant, the PUT and the PST are collected and their values are presented at the artificial nodes of $m = -1$ and 25, respectively. However, they are only large enough to be visible in Fig. 2d. At a sufficiently long time, the PDF will become uniformly zero (not shown) due to the absorbing boundary conditions, and the sum of the PUT and the PST will reach one.

In Fig. 3, the results for a longer pulse of $t_p = 20$ ms are shown. All other parameters are identical to those used in Fig. 2. In contrast to Fig. 2, the electrophoretic force drives the PDF peak toward the right (Fig. 3a–c) until it reaches the end of the domain (Fig. 3d). At $t = 15$ ms, the PST is already reaching one. For this case, the longer pulse ensures the definitive success of translocation.

The evolution of the PST is more clearly shown in Fig. 4, where it is plotted as a function of time for the two cases studied above. For $t_p = 5$ ms, the PST is very small at the end of the pulse. Indeed, as indicated in Fig. 2c, at this time the PDF is centered at around $m = 11.75$. Most of the PST increment occurs by the slow diffusive process post-pulsation. This time scale is much longer ($s$) when compared with the pulse length. On the other hand, for $t_p = 20$ ms, we observe that the PST increases sharply to one when the pulse is still present. Fig. 4 demonstrates that two time scales may manifest during the translocation process. If the pulse is not able to completely translocate the chain during its presence, then translocation occurs on the slow diffusive time scale ($s$). On the contrary, if the pulse is sufficiently long, then translocation occurs on the much shorter electrophoretic time scale ($ms$).

For both of the cases in Fig. 4, the PST reaches a steady-state value given sufficient time. We term this value the “final probability of successful translocation” (F-PST), and examine its dependence on the pulse length in Fig. 5. In addition to $V_m = 0.2$ V, we also study two other cases, namely, $V_m = 0.4$ and $0.6$ V. For each value of the TMP, the F-PST increases until it saturates at a value very close to 1. Furthermore, the F-PST increases along with an increasing TMP. This trend is examined in detail in Fig. 6, where the contours of the F-PST are shown in the phase space of $V_m$ and $t_p$. The contour lines are linear and parallel, with slopes close to $-1$, suggesting that to reach the same value of the F-PST, $V_m$ and $t_p$ obey a reciprocal correlation. In other words, the product of $V_m$ and $t_p$ is constant along the contours. For the contour of F-PST = 0.99, $V_m \times t_p \approx 2.23$ V·ms. The latter value can be used to define a threshold of pulsing parameters for successful DNA delivery.
Because the drifting velocity has a linear dependence on the TMP proportional to the product of the electrophoretic velocity and time. At the end of the pulse, the distance that the peak travels is simply.

In fact, the F-PST depends strongly on the peak position of the PDF that electrophoresis is the main driving mechanism for translocation.

\[ \text{F-PST} = C \times \left( \frac{V_m t_p}{N^0} \right)^a \]  

where \( C = 45.0 \), \( a = 1.1 \), and \( b = 1.46 \). These constants are obtained by minimizing the fitting error. The coefficient of determination is \( R^2 \approx 0.999 \), indicating that the formula (Eq. (6)) accurately captures the data trend.

In generating the previous results, we have assumed \( m_0 = 0.2 \). This initial location corresponds to the electrophoretic insertion induced by a prior pulse of \( V_m = 1 \) V and \( t_p = 18.5 \) μs. This pulse is comparable to the first 10-μs strong pulse (HV) used in Sukharev et al. [57]. In Fig. 8b, the effect of a varying \( m_0 \) is studied. We run the simulations for \( m_0 = 0.1, 0.5, 1, \) and \( 1.5 \) for the same range of \( V_m, t_p, \) and \( N \) values considered in Fig. 8a. For each value of \( m_0 \), 106 cases are studied. We find that the change in \( m_0 \) does not cause appreciable deviation from the power-law behavior. The collection of data is best-fitted with the correlation, \( \text{F-PST} = 60.0 \times \left( V_m t_p \right)^{1.07} / N^{1.51} \). The powers \( a \) and \( b \) only slightly differ from those in Fig. 8a, and the coefficient of determination is \( R^2 = 0.946 \). We remark that as \( m_0 \) becomes large, eventual departure from Eq. (6) is expected (not shown). However, in the current work, we focus on studying DNA translocation with only a small segment initially inserted.

---

**Fig. 6.** Contour plot of the F-PST in the phase space of the TMP \( (V_m) \) and the pulse length \( (t_p) \). The slopes for the contour lines are approximately \(-1\), indicating that \( V_m \times t_p \approx \) Constant.

**Fig. 7a.** The F-PST as a function of the DNA segment number, \( N \), for \( t_p = 5,10, \) and 20 ms. The F-PST in general decreases as \( N \) increases and \( t_p \) decreases. The correlation between the F-PST and \( N \) is close to a power law, \( \text{F-PST} \sim N^{-1.5} \). In Fig. 7b, the F-PST is shown as contours in the phase space of \( N \) and \( t_p \). The contour space lines are again linear and parallel, with slopes close to 0.75, suggesting a power-law correlation between \( N \) and \( t_p \).

The results from Figs. 6 and 7 together suggest a similarity behavior of the F-PST as a function of \( V_m, t_p, \) and \( N \). We have run 568 simulations with \( V_m \) ranging from 0.1 to 1 V, \( t_p \) from 0.01 to 100 ms, and \( N \) from 17 to 134 (corresponding to DNA sizes from 5 kbp to 40 kbp). The data is shown in Fig. 8a and is well-fitted with the correlation, \( \text{F-PST} = 60.0 \times \left( V_m t_p \right)^{1.07} / N^{1.51} \) (solid). The coefficient of determination is \( R^2 = 0.946 \).

---

**Fig. 7.** The effect of DNA size on translocation probability. (a) The F-PST as a function of the DNA segment number, \( N \), for \( t_p = 5,10, \) and 20 ms. The descending part of the curves follows approximately the correlation, \( \text{F-PST} \sim N^{-1.5} \). (b) Contour plot of the F-PST in the phase space of \( N \) and \( t_p \). The contours are straight and parallel lines following the approximate correlation \( N \sim t_p^{-0.75} \).
4. Comparison with experiments

Direct, quantitative comparison with experimental data is difficult due to the lack of sufficient details in the previous measurements. In what follows, we compare and discuss the qualitative trends from our simulation and two of the most relevant experimental studies in the literature, namely, by Sukharev et al. [57] and Faurie et al. [50].

The study performed by Sukharev et al. used a two-pulse scheme to electroporate simian Cos-1 cells. The first pulse is 6 kV/cm in strength and 10 μs in duration. After a 100-μs delay, a second pulse of 0.2 kV/cm was applied with the duration varying between 0 and 10 ms. The TE was obtained as a function of the second pulse duration (see inset of Fig. 9). In general, a near linear dependence is observed. The simulated result using the current model is shown in Fig. 9. For this case we follow the setup in Section 3. The initial PDF is centered at \( m_0 = 0.2 \), and only the second pulse from the experiment is considered. By using this configuration, we assume that the first pulse is mainly responsible for pore creation, and the initial insertion of the DNA into the pore. The average DNA length is 7.2 kbp, or \( N = 24 \). The TMP is \( V_m = 0.2 \text{ V} \) which is obtained via the formula \( V_m = 1.25ER \), where \( E = 0.2 \text{ kV/cm} \) is the applied field strength of the second pulse, and \( R = 8 \mu \text{ m} \) is the cell radius. The simulated F-PST is plotted as a function of the pulse length, \( t_p \), in Fig. 9. Although the F-PST and the TE are not the same, they are both measures for the efficacy of DNA delivery, and we find comparable linear trends in the simulation and the experimental data in Fig. 9. Previous theory proposed that the increase in TE was induced by increased pore size and population, or increased permeabilized area due to prolonged field exposure [68]. Here we offer an alternative interpretation that longer pulses enhance the probability of translocation, which manifests itself as enhancement in TE.

Next, the experimental data by Faurie et al. [50] is examined. In these experiments, a train of six uni-directional (uni-polar) or alternating (bi-polar) pulses were applied to transfec CHO cells with plasmids of 4.7 kbp in average length. Each pulse was 1 ms in duration, with the delay of 0.013, 0.1, 1, and 10 s, corresponding to repetition frequencies of 77, 10, 1, and 0.1 Hz, respectively. The resulting fluorescence per cell is shown in Fig. 10a.

The simulated results with \( m_0 = 0.2 \), \( N = 16 \), and the same pulsing scheme as in Faurie et al. are presented in Fig. 10b. Although there are noticeable differences between the data and the simulation, the latter captures a few important features of the former. For the uni-polar pulse, the simulation predicts a curve with the similar shape to that in the experimental data. The F-PST begins to increase at approximately 1 Hz. For frequencies greater than 1 Hz, the F-PST reaches one asymptotically, suggesting a high probability of successful delivery. For the bi-polar pulse, the simulation captures the descending trend for frequencies higher than 1 Hz.

The main difference between the data and the simulation is at the lower frequencies. In the experiments, the bi-polar pulse results in a TE higher than the uni-polar one, whereas in the simulation, these two pulsing schemes lead to similar values of the F-PST. This effect is possibly attributed to the fact that DNA molecules can enter the cell from both sides under bi-polar pulses [39], enhancing the probability of delivery. This mechanism is not included in the current model, and is also likely responsible for the peak at 1 Hz shown in Fig. 10a.

The frequency dependence in the simulation is mediated by a diffusive time scale, which we argue is responsible for the trends observed in the experimental data. As we demonstrated in Section 3, if a single pulse is not able to complete the translocation by electrophoresis, then the translocation (either successful or unsuccessful) is governed by the diffusive drifting of the polymer chain post-pulsation. For the plasmids considered in Faurie et al., this time scale is \( R_g^2/D_0 = 5.5 \text{ s} \), using \( R_g = 250 \text{ nm} \), and \( D_0 = 1.13 \times 10^{-14} \text{ m}^2/\text{s} \) (see Appendix A). At low frequencies, the sufficient delay time between pulses ensures that the translocation is completed, such that there's no compounding effect between the individual pulses. For this case, the uni- and bi-polar pulses do not behave differently according to the model, although in reality the effect noted above may render the bi-polar pulse more advantageous. As the frequency increases and the delay time decreases below the threshold, additional uni-polar pulses help push-in the plasmids into the cell, whereas bi-polar pulses tend to reverse the translocation, hence causing the observed behavior.

However, we emphasize that in the experiments and between the pulses, diffusive drifting of the chain may not be the only mechanism at work. Endocytosis can be triggered at this stage, and the direct interaction of the DNA molecule and the membrane may also play a role. The complete intake of DNA may thus be a complex process involving all these aspects. On the other hand, regardless of the specific nature, the mechanism(s) needs to act on the characteristic time scale of a few seconds to be able to explain the frequency dependence observed.
Finally, we remark that although the above comparisons are only qualitative, our theory can be verified with well-defined quantitative experiments. For example, experiments similar to those by Sukharev et al. can be repeated, but with a wider and controlled range of $V_m \times t_p$ and $N$ values to validate the similarity behavior indicated by Eq. (6). In particular, pulses with the same product, $V_m \times t_p$, should result in similar values of TE. For uni-polar and bi-polar pulsing experiments, different sizes of plasmids can be used which leads to different diffusive time scales, such that they may be detected in the measurements. In addition, the length of each individual pulse can also be explored as a control parameter, such that the electrophoresis-dominant and diffusion-dominant regimes can be differentiated.

5. Conclusions

In this work, we presented a 1D Fokker–Planck simulation for the translocation of a DNA polymer through a membrane-bound nanopore, within the context of electroporation-mediated molecular delivery. The model provides a few important insights:

- The translocation may occur on two disparate time scales, namely, the electrophoretic time ($\sim$ms), and the diffusive time ($\sim$s). If the pulse is sufficiently long to complete the translocation via electrophoretic drifting, then the electrophoretic time scale is observed. Otherwise, translocation completes (either successfully or unsuccessfully) on a much longer diffusive time.
- The F-PST (the final probability of successful translocation) follows the correlation,

\[
\text{F-PST} \sim \left(\frac{V_m t_p}{N}\right)^a. 
\]

The values of $a$ and $b$ are close to 1 and 1.5, respectively, for small $m_0$-values, or for DNA chains with small initial insertion distances. The dependence of the F-PST on the product, $V_m \times t_p$, directly reflects that translocation is primarily driven by electrophoretic drifting.

The simulation results are compared with experimental data from previous studies. In particular, the diffusive time scale is proposed to explain the frequency dependency observed in electroporation experiments with uni- and bi-polar pulse trains. Another important contribution of the work is that the model predicts trends and correlations (such as Eq. (6)) that can be verified with well-defined experiments.

In this study, we begin with the initial condition that a small segment of the polymer chain is already inserted into the pore, and the insertion process itself is not considered. We remark that the latter is not a trivial process, and the polymer chain needs to overcome an entropic barrier to achieve this step. Muthukumar has used a continuum model to examine this process, and solved for the steady-state DNA capture rate in terms of the entropic barrier height and location [67]. A similar model can be developed, and be combined with the current one to improve the fidelity of prediction. For example, such model can be used to possibly interpret the formation of “DNA–membrane complexes” observed in experiments [47–51].

Finally, the complete uptake of DNA is likely a complex process involving many different mechanisms including electrophoresis, diffusion, and endocytosis. The authors of the present work believe that during the pulse, electrophoresis is the dominant mechanism in both aggregating the molecules near the membrane, and transporting them across. This modeling study thus may be harnessed to optimize DNA electrotransfer by best taking advantage of electrophoretic transport.

Acknowledgements

The authors acknowledge funding support from an NSF Award CBET-0747886 with Dr. William Schultz and Dr. Henning Winter as contract monitors.

Appendix A

The main difference between the current model and previous ones [61,62] is that we require a specific value for $k_0$, which we determine from $D_0$, the chain diffusivity. The latter can be obtained from experimental measurements. The relationship between $k_0$ and $D_0$ can be derived by considering the Fokker–Planck equation in the natural coordinate, namely, the center of mass of the entire chain, $x$:

\[
\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x} \left[ \frac{D_0}{R_g^2} \frac{\partial^2 P(x,t)}{\partial x^2} + D_0 \frac{\partial}{\partial x} P(x,t) \right]. \tag{7}
\]

Comparing Eq. (7) with Eq. (1), and considering $x = mL$, we obtain $k_0 = D_0/l^2$. Here, we take the characteristic length $l$ to be $2R_g/N$ (Fig. 11), where $R_g$ is the radius of gyration. In other words, the translocation of the complete $N$ segments is equivalent to a distance of $2R_g$ traveled by the center of mass. Further considering that $R_g = 0.459k_BT/\mu$, which is derived for a polymer chain with one end pinned on a rigid wall [66], we arrive at Eq. (2) in the proper text.

The bulk diffusivity of an $N$-segment DNA chain is found by the measurements of Dauty et al. [69].

\[
D_{\text{bulk}} = \frac{1.344 \times 10^{-11}}{N^0.6} \text{ m}^2/\text{s}. \tag{8}
\]

For the current model, we use a modified formula based on Eq. (8),

\[
D_0 = \alpha/b D_{\text{bulk}}. \tag{9}
\]

The factor $\alpha$ arises from a reduction due to the crowdedness of the cytoplasm. Extrapolating from the measurements by Dauty et al. (see Fig. 4B therein), we use $\alpha = 0.017$ for a chain length of 4.7 kbp, and $\alpha = 0.01$ for a chain length of 7.2 kbp.

The factor $\beta$ is the reduction effect due to the fact that the polymer chain is in the vicinity of a rigid wall. According to the measurements by Kihm et al. for solid nanoparticles [70], this factor typically ranges from 0 to 0.6 depending on the distance from the wall. Due to the lack of data for near-wall DNA particles, we simply use $\beta$ as a fitting parameter to generate the best comparison with data on DNA translation through synthesized nanopores [71]. From Fig. 12, we obtain $\beta = 0.322$ to achieve the best matching between our theoretical prediction and the measurements. Note to generate this comparison, we use $\alpha = 1$ because no cells are involved.

We remark that Eq. (9) gives a constant $D_0$ given a constant size of DNA. A more accurate model can be derived where $D_0$ is a function of the translocation coordinate, $m$ [61]. However, we have found no appreciable difference between the results following this approach and simply using Eq. (9) above. (The comparison is not shown here.

Figure 11. Upon the completion of translocation, the center of mass translates by $2R_g$, where $R_g$ is the radius of gyration.
for brevity.) The agreement results from the fact that the effective diffusivity of the DNA chain is dominantly controlled by its much reduced value in the cytoplasm. In this work, we therefore use Eq. (9) as an approximation to the more complete diffusivity model.

The effective charge per DNA segment, $z_s$, is given by the formula,

$$z_s = 0.5 \times \frac{L}{l_{bp}} = 135N^{-0.4}$$

The factor of 0.5 is obtained from the experiments by Keyser et al. [72], which suggest that the effective charge of a DNA base pair (2 electrons) is reduced by 75% within a pore. This reduction ratio corroborates with the theoretical prediction by Ghosal [73], which attributes the effects to viscous drags within the pore. Again, although these results are generated for solid-state nanopores, we employ them due to the lack of measurements for electropores.

References
