Nanotechnology

PAPER

Modeling single-molecule stochastic transport for DNA exo-sequencing in nanopore sensors

To cite this article: Benjamin Stadlbauer *et al* 2020 *Nanotechnology* **31** 075502

IOP Publishing Nanotechnology
Nanotechnology 31 (2020) 075502 (7pp) https://doi.org/10.1088/1361-6528/ab513e

Modeling single-molecule stochastic transport for DNA exo-sequencing in nanopore sensors

Benjamin Stadlbauer[1], Gregor Mitscha-Baude[1] and Clemens Heitzinger[1,2]

[1] Institute for Analysis and Scientific Computing, TU Vienna, A-1040 Vienna, Austria
[2] School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287, United States of America

E-mail: gregor.mitscha-baude@gmx.at

Received 24 June 2019, revised 27 August 2019
Accepted for publication 25 October 2019
Published 20 November 2019

Abstract

We present a simulation framework for computing the probability that a single molecule reaches the recognition element in a nanopore sensor. The model consists of the Langevin equation for the diffusive motion of small particles driven by external forces and the Poisson–Nernst–Planck–Stokes equations to compute these forces. The model is applied to examine DNA exo-sequencing in α-hemolysin, whose practicability depends on whether isolated DNA monomers reliably migrate into the channel in their correct order. We find that, at moderate voltage, migration fails in the majority of trials if the exonuclease which releases monomers is located farther than 1 nm above the pore entry. However, by tuning the pore to have a higher surface charge, applying a high voltage of 1 V and ensuring the exonuclease stays close to the channel, success rates of over 95% can be achieved.

Keywords: nanopores, DNA sequencing, narrow-escape problem, multiscale simulations

(Some figures may appear in colour only in the online journal)

1. Introduction

Nanopore sensors allow the label-free detection of proteins, single DNA nucleotides and other biomolecules for analysis and sequencing applications [1, 2]. Recent experimental progress [3–8] motivates the development of computer simulations of single-molecule transport to aid in the understanding of nanopore sensors and to enable their rational design.

The principle of a nanopore sensor is illustrated in figure 1(a). Target molecules entering the channel are detected by means of current disruption; variations in the height and length of the event signal allow for discrimination between species [2]. A specific binding site inside the nanopore may additionally enhance the current signal and help to distinguish target molecules [9].

Several attempts have been made to enable sequencing of biological macromolecules based on single-molecule sensing with nanopores [10, 11]. For example, in *exo-sequencing* [10] an exonuclease enzyme is used to cleave individual nucleotides from a DNA strand which are then free to translocate the pore; see figures 1(b), (c). It remains open whether this approach can actually be used for DNA sequencing [1]. Ideally, the nucleotides would have to reach the recognition site in their correct order. In particular, the chance that the nucleotide, once released by the exonuclease, escapes to the bulk and fails to enter the nanopore should be close to zero. Furthermore, after recognition, the nucleotide should reliably translocate to the *trans* side to prevent being detected twice.

In this work, we provide a probabilistic modeling framework based on physical simulations to shed light on the viability of exo-sequencing. We use a continuum description of water, ions and dielectric materials in the nanopore, coupled to a stochastic model for the motion of the nucleotide. A large number of simulated trajectories, where the nucleotide is released on the *cis* side of the nanopore, are analyzed to estimate the probability that it enters the

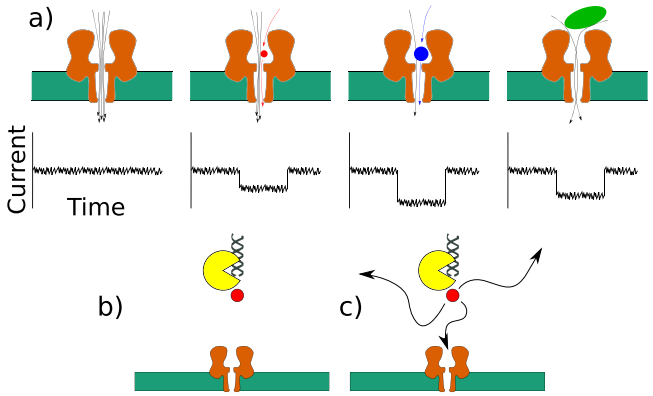

Figure 1. (a) Various molecules cause different reductions of the current. This is important for C, G, A, and T detection in exo-sequencing. (b) An exonuclease enzyme is located above the nanopore and cleaves off individual nucleobases. (c) The nucleobase wanders through the electrolyte and must enter the nanopore eventually in order to be detected.

nanopore and is detected. We also investigate whether the migration happens sufficiently fast, so that a nucleotide enters the pore before the next one is released and the order of the sequence is preserved. Our model not only accounts for electrostatic attraction or repulsion between target and sensor [12], but also for the effect of electro-osmotic flow near the nanopore [13, 14] and non-specific adsorption of nucleotides to the pore surface.

Continuum models based on the Poisson–Nernst–Planck or Poisson–Nernst–Planck–Stokes (PNPS) equations have been employed, for example, to determine the conductance and ion selectivity of nanochannels [15, 16] and to compute the electrophoretic force on a DNA strand [17, 18]. Recently, we combined PNPS theory with Brownian dynamics (BD) simulations to reproduce measurements of stochastic single-molecule transport in nanopore sensors [19]. Whereas the focus of that work was on the current trace and dwell time of proteins inside a nanopore, here we concentrate on the transport outside just before entering the channel. In both cases, the continuum description of ions and water enables to simulate the motion of analytes at macroscopic timescales and arbitrary distances to the nanopore, in contrast to low-level approaches such as all-atom molecular dynamics [20–22].

In our model setup (figures 1(b), (c)), an exonuclease enzyme is held above the channel protein α-haemolysin (αHL), which is a popular nanopore for sequencing applications [1, 11, 23, 24]. The DNA strand gets ratcheted to the enzyme and a nucleotide with one of the four bases (C, G, A, T) is cleaved from the DNA. The isolated monomer can now migrate to the nanopore, assisted by applied voltage and induced electro-osmotic flow, or be diverted from its path and disappear in bulk due to random diffusion.

The exonuclease releases nucleotides at a certain rate. We assume a rate of 275 nucleotides per second if we are operating with the *Escherichia coli* exonuclease I at 37 °C [25–27]. Since nucleotide cleavage is understood to be a first-order kinetic process [25, 26], the time until the next cleavage event is modeled as an exponentially distributed random variable with probability density $\lambda e^{-\lambda t}\,dt$ and rate parameter $\lambda = 275\ \mathrm{s}^{-1}$.

Our model is designed to predict the probability that the nucleotide successfully enters the nanopore within any given time. In mathematical terms, this is known as an *exit time problem* or *narrow-escape problem* [28–31].

As we will show, the exit probability depends on various parameters like cleaving position, applied voltage bias [32] and surface charge of the channel protein. Furthermore, adsorption of the nucleotide to the αHL surface may have an influence because it can prevent successful exit in time before the next nucleotide is released. In section 2 we describe our computational model for stochastic transport. Results and discussion are presented in sections 3 and 4, respectively.

2. Model

2.1. The Langevin equation

The membrane and pore protein are modeled as stiff, continuous solids whose positions are fixed. Motion of the nucleotide, which is modeled as a solid sphere, is governed by the Langevin equation [33]

$$\dot{x}(t) = \frac{D}{kT}F(x(t)) + \sqrt{2D}\,\zeta(t). \tag{1}$$

Here, $x(t)$ is the position of the nucleotide at time t, D is the diffusion coefficient, k the Boltzmann constant, and T the temperature. Motion is driven by the electrophoretic mean force $F(x)$ acting on the nucleotide, which depends on its position. The solvent (electrolyte) is modeled implicitly via the Gaussian noise term $\zeta(t)$—which represents stochastic collision forces—and the friction term $\dot{x}(t)$.

The Langevin equation (1) is discretized to produce the time-stepping scheme

$$x(t+\Delta t) = x(t) + \frac{D}{kT}F(x(t))\Delta t + \sqrt{2D\Delta t}\,\zeta, \tag{2}$$

where Δt is the size of the time step and ζ is a vector of three standard normally distributed random numbers. In order to simulate random trajectories of the nucleotide, the force $F(x)$ must be computed, at least for positions x in the vicinity of the nanopore. This is based on a continuum model, namely the PNPS equations.

2.2. The continuum model

The PNPS equations [34] are a set of partial differential equations (PDEs) that describe the interaction of mobile charge carriers (ions) with the electrostatic environment and induced electroosmotic flow of the background medium

(water). They can be derived from the Boltzmann equation [35–37]. We give a short overview of the computational model, as details can be found in our previous work [19, 38]. The geometry and parameters used for αHL are the same as in [19].

The electrostatic effects are described by the Poisson equation

$$-\nabla \cdot (\epsilon \nabla \phi) = q(c^+ - c^-) + \rho, \tag{3}$$

where ϕ is the electric potential, ϵ the material dependent permittivity, q the elementary charge and ρ the permanent charge density. We assume a 1:1 electrolyte with one positive and one negative ion species with concentrations c^+ and c^-, respectively. Ionic transport is captured by the Nernst–Planck equations

$$\nabla \cdot (-D^+ \nabla c^+ - \mu^+ c^+ \nabla \phi + c^+ u) = 0, \tag{4a}$$

$$\nabla \cdot (-D^- \nabla c^- + \mu^- c^- \nabla \phi + c^- u) = 0. \tag{4b}$$

Here D^+ and D^- are the diffusion coefficients of the cations and anions, μ^+ and μ^- are the corresponding mobilities subject to the Einstein relations $\mu^\pm = qD^\pm/kT$, and u is the fluid velocity. The convective term $c^\pm u$ provides coupling to the Stokes equations

$$-\eta \Delta u + \nabla p = -q(c^+ - c^-)\nabla \phi, \tag{5a}$$

$$\nabla \cdot u = 0, \tag{5b}$$

where η is the fluid viscosity and p is the pressure. Together, equations (3)–(5) form the PNPS equations, which have to be solved numerically to obtain the unknowns ϕ, $c^\pm$, u and p.

The potential ϕ and the fluid velocity u define the electrophoretic force $F(x)$ on the nucleotide in (2). The force

$$F(x) = F_{\text{el}}(x) + F_{\text{drag}}(x)$$

is the sum of two components. The first component, the electric force, stems from the action of the electric field on the charged nucleotide and is given by

$$F_{\text{el}}(x) = -Q\nabla \phi(x),$$

where $Q = -2q$ is the nucleotide charge. The second component is the drag force induced by the electro-osmotic flow of the background medium and is given by Stokes' law as

$$F_{\text{drag}}(x) = 6\pi \eta R u(x),$$

where R is the hydrodynamic radius of the nucleotide, which is estimated to be $R = 0.5$ nm.

The PNPS equations are solved on an axisymmetric computational domain which includes the pore protein, membrane and an electrolyte reservoir extending about 20 nm in each direction. The Poisson equation is solved on the entire domain, while the Nernst–Planck and Stokes equations are only solved on the fluid part of the domain.

The boundary conditions (BCs) for the unknown physical variables are specified as follows. The voltage bias is applied by fixing the potential at the top and at the bottom of the computational domain; charges on the pore walls are incorporated via interface conditions. The ion concentrations $c^\pm$ are fixed to the bulk concentration c_0 at the top and at the bottom of the reservoir and are subject to hard repulsion (i.e. homogeneous Neumann BCs) at the pore walls and the membrane. For the fluid velocity, the no-slip condition $u = 0$ is applied at fluid-solid interfaces and a stress-free BC is applied at the reservoir boundaries.

The nonlinear PNPS system is discretized on a finite-element mesh and solved iteratively with the method introduced in [38]. The solver, which was written by our group, is open source and available online[3]. To efficiently resolve regions of interest with high accuracy, we employ an adaptive mesh refinement algorithm tailored to the PDE system at hand.

3. Results

Figure 2 visualizes the two force-field components obtained from our continuum model. Based on this force field, we compute trajectories of the nucleotide using the discretized Langevin equation (2). The random walk starts at a certain height z_0 above the pore entry; the height is an important parameter that will be varied in our simulations below. The simulation ends in one of the following two cases: (1) the nucleotide arrives at a pre-defined *recognition site* inside the pore; (2) the nucleotide wanders too far in the wrong direction and leaves a sphere with a radius of 20 nm centered at the pore entry. In the first case, we count the trial as successful, while in the second case, it is considered failed. We are interested in the fraction of successful trials, i.e. the *exit probability*.

The role of the 20 nm cutoff is to limit the computational burden of calculating the force field around the nanopore. It is chosen large enough that only a tiny fraction of actually successful events are excluded, as our results for large starting distances will make evident.

The recognition site is where the nucleotide can be successfully detected based on the current disruption it causes. Inspecting the geometry of αHL, we define it as the cross-section at the first narrowing of the channel, about 1.3 nm below the upper pore entrance. To understand how this choice is motivated, see figure 3, where we show three simulated trajectories and the associated current traces. The simulation was continued here even after the nucleotide passed the recognition site; the current is an interpolation of multiple evaluations of our PNPS model [38]. The first example in the figure shows a failed attempt, where the nucleotide briefly entered the pore but changed direction and vanished into the bulk. Although it almost reached the recognition site (red line), the current disruption is markedly weaker than in the two other examples, where the nucleotide migrated through the entire pore. On the other hand, the two successful trials exhibit a very similar, down-down-up current pattern, so we expect that cases where the recognition site is reached can be detected reliably.

[3] https://github.com/mitschabaude/nanopores

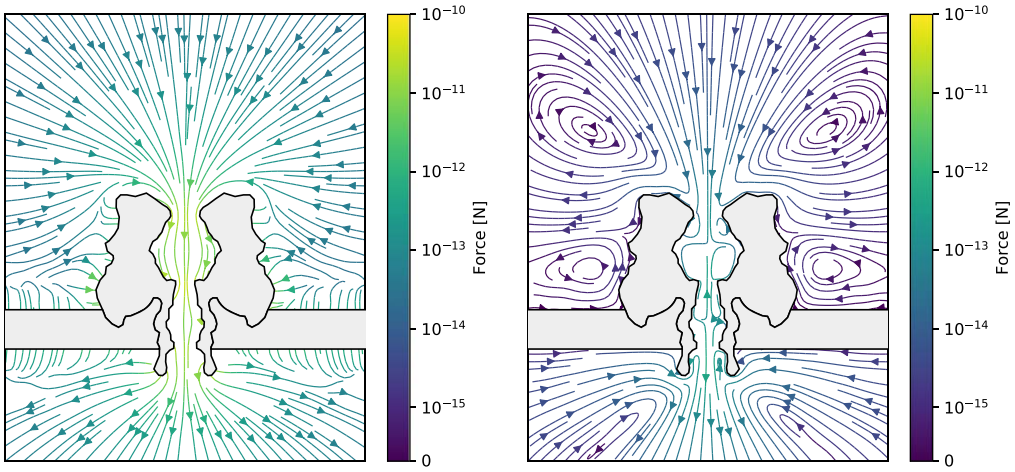

Figure 2. Stream-line plots of the electric force F_{el} (left) and the drag force F_{drag} (right) acting on the DNA nucleotide. The applied voltage is 0.5 V.

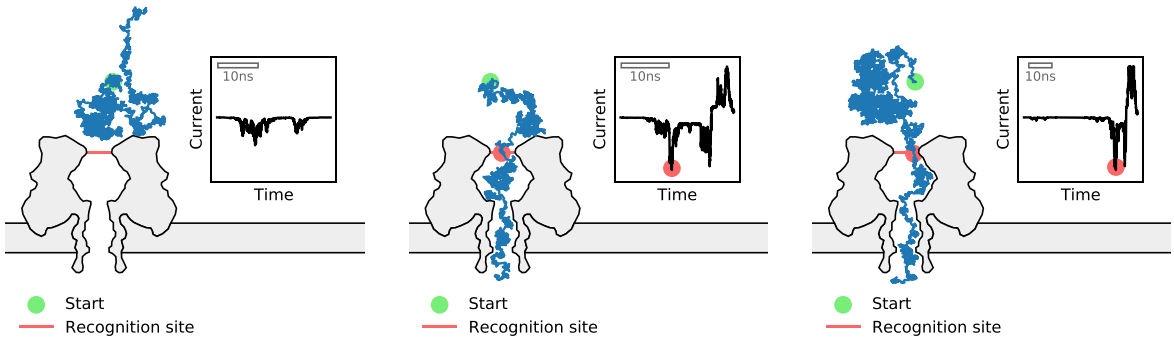

Figure 3. Brownian path of the nucleotide (blue lines) and the associated current trace (insets). We show one example of a failed migration attempt (left) and two examples of successful migration through the nanopore. A 2D projection of the path is shown, while the actual simulation is three-dimensional. The applied voltage is 0.5 V, and the height of the starting position is $z_0 = 4$ nm.

Algorithm 1. Random walk

Input: time step Δt, height of starting position z_0, mean binding duration τ_0.

Draw initial position $x = (x_1, x_2, x_3)$ at height $x_3 = z_0$ uniformly from flat disc with radius 1 nm in the (x_1, x_2) plane.
Load precomputed force field $x \mapsto F(x)$ from file.
Initialize time $t = 0$.
Initialize Boolean return variable, indicating successful exit: success := FALSE.
while TRUE **do**
  Draw standard normal vector ξ.
  Set $\Delta x := \frac{D}{kT} F(x)\Delta t + \sqrt{2D\Delta t}\,\xi$.
  if $x + \Delta x$ collides with nanopore wall **then**
  Find $\theta \in [0,1]$ such that $x + \theta\,\Delta x$ does not collide.
  Set $\Delta x := \theta \Delta x$.
  Draw binding duration $\tau \sim \mathrm{Exp}(\tau_0)$.
  Set $t := t + \tau$.
  Set $x := x + \Delta x$.
  Set $t := t + \Delta t$.
  if x has exited the domain **then**
  if exit is the recognition site **then**
  Set success := TRUE.
  break

Output: success, t.

3.1. Brownian dynamics algorithm with wall adsorption

The trajectories in figure 3 do not yet include adsorption of the nucleotide to the pore, and we can see that translocations occur on a timescale of several tens of nanoseconds. However, the current events observed in related experiments are much slower, with timescales ranging from microseconds to seconds for the translocation of DNA bases [39, 40] as well as proteins [9, 41]. Several researchers have hypothesized that non-specific adsorption of the molecule in question to the channel wall is responsible for such 'anomalous' dwell times [13, 40–42]. To enable our model to exhibit longer dwell times, we let the nucleotide adsorb spontaneously whenever it collides with the pore wall. The binding duration, called τ, for each of these binding events is exponentially distributed around a mean binding duration τ_0, i.e. its probability density is $\frac{1}{\tau_0} e^{-\tau/\tau_0}\, d\tau$. Per definition, a wall collision happens whenever the distance between nucleotide position and pore wall is smaller than the nucleotide radius. This is checked after every step of the random walk.

Clearly, the nucleotide cannot be allowed to penetrate the pore wall. Therefore, when a collision happens, we not only draw a random binding duration but also shorten the last step of the nucleotide so that it stops just *before* the pore wall. This

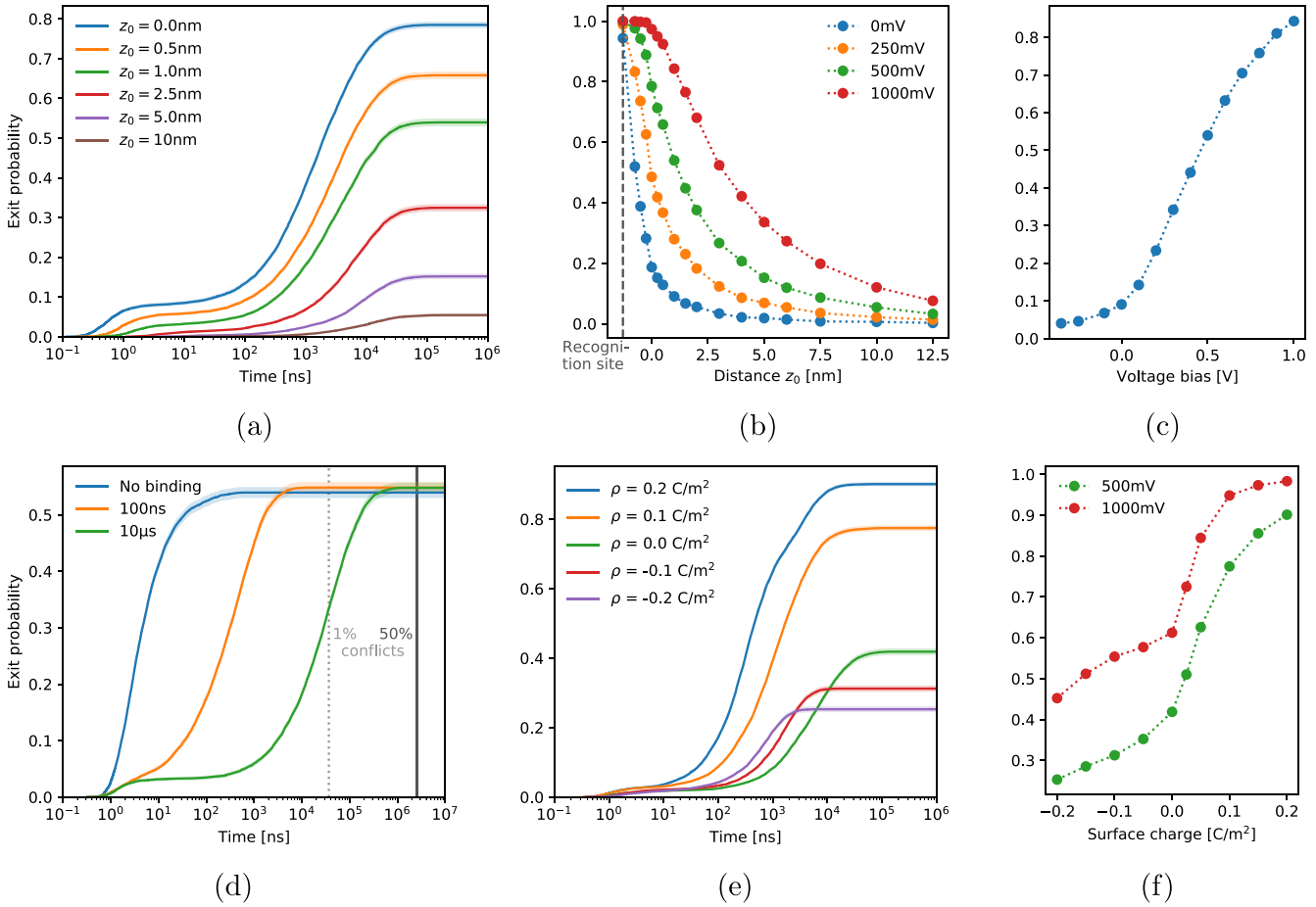

Figure 4. (a) Time-dependent exit probability for different starting positions. (b) Final exit probability depending on starting position, for different voltage biases. (c) Dependence on voltage bias. (d) Time-dependent exit probability for different binding durations. (e) Time-dependent exit probability depending on (homogeneous) surface charge. (f) Final exit probability for various surface charges at two different voltages.

is implemented efficiently via binary search. The entire BD algorithm is summarized as pseudo-code in algorithm 1.

As can be seen, the simulation depends on various input parameters. In addition, the force field depends on several physical variables such as the applied voltage. If not stated otherwise, we use the following default values:

- timestep $\Delta\, t$: 0.01 ns,
- height of starting position z_0: 1 nm,
- mean binding duration τ_0: 1 μs,
- applied voltage: 0.5 V,
- salt concentration: 1 M KCl.

The surface charges on αHL are, by default, designed to replicate the natural pattern of partial charges on the pore protein [19]. However, assuming the charges can be modified by chemical manipulation, as was done before with αHL [39, 43], we will also explore the influence of varying the surface charge. In that case, the chosen charge density ρ is distributed homogeneously over the pore surface.

3.2. Exit probability

Our results for the exit probability are collected in figure 4. For each displayed set of parameters, 10 000 random walks were performed. By plotting the cumulative exit probability of these 10 000 trials over the time needed to reach the exit, we obtain a smooth evolution profile that shows how the exit probability increases with time. Figure 4(a) shows six such evolution profiles for different starting heights z_0, modeling exonuclease enzymes anchored at different distances from the pore entry. The evolution profiles reach a steady state at about 100 μs, which means that after this time, all simulated nucleotides have either arrived at the recognition site or escaped far enough from the pore that the probability of returning is effectively zero. We see that for a starting distance of 1 nm, just about half of all nucleotides are detected. Even at 0 nm distance, which would require the DNA base cleaving to happen right at the pore entrance, over 20% of monomers fail to travel the additional 1.3 nm to the recognition site. For these results we assumed a moderately high voltage of 0.5 V.

The combined influence of distance and voltage can be seen in figure 4(b). Here, only the final exit probability is shown, i.e. the values (colored circles) correspond to steady states in the previous figure. Again, every data point represents the average exit probability of 10 000 simulation trials. For sufficiently large distances z_0, the exit probability falls off like z_0^{-1} in accordance with the mathematical theory of narrow-escape problems [28]. While for zero applied voltage, exit

probabilities outside the pore are very small, higher voltages increase the success rate considerably by forcing most nucleotides close to the pore to drift downwards. See also figure 4(c), where the voltage dependence for fixed $z_0 = 1$ nm is plotted.

Next, we assess the influence of mean binding duration on the time-dependent evolution of exit probability, see figure 4(d). The shaded areas beneath evolution profiles represent the 95% confidence interval. Although the three steady states for binding durations of 0, 100 ns and 10 μs appear to be slightly distinct, all three fall within the 95% interval of each other, which means the differences can be attributed to mere statistical error. Thus, as should be expected, the binding duration has no influence on the final, steady state exit probability. However, longer binding durations increase the time until steady state and can prevent the nucleotide to exit before the next one is released. To visualize the likelihood of such a conflict, we show the 1% and 50% quantiles of the time between cleavage events (figure 4(d), dotted and solid vertical lines).[4] This means that, for example, by the time a nucleotide hits the dotted line, there is a 1% chance that the next nucleotide has already been released. As can be seen, for binding durations of 0 and 100 ns, all nucleotides which exit successfully do so well before the 1% line, so the probability of out of order sequencing is very low. On the other hand, for a 10 μs binding duration, about a third of successful trials fall in the range between 1% and 50% conflicts. In this case, a certain amount of sequence shuffling would have to be expected.

Finally, in figures 4(e) and (f), we vary the surface charge of αHL. The resulting attraction or repulsion of the nucleotide —which, compared to its size, is highly charged—has a considerable influence on the exit probability. Just by increasing the surface charge from 0 to $+0.1$ C m^{-2}, the exit probability increases from 42% to 77% (at 0.5 V and 1 nm starting distance). By raising the voltage to 1 V, even higher exit probabilities become possible: 95% for $+0.1$ C m^{-2} and 98% for $+0.2$ C m^{-2} surface charge, respectively.

4. Discussion

Several insights can be taken away from our study for future attempts at workable exo-sequencing in nanopores. First, it is clear that reliable migration of the DNA nucleotide to the detection zone is not a given. In a sense, the intuition of particles getting drawn into the nanopore by the force field is misleading; particle trajectories are largely stochastic and dominated by the entropy barrier of entering a narrow opening from a large reservoir. The influence of the force field is very localized and does not extend more than one or two nanometers to the outside of the channel; only in this area, when the applied voltage is high, do the particle trajectories become largely deterministic. We predict that, for an exonuclease enzyme operating at larger distances such as 10 nm, exo-sequencing will not be possible.

For smaller starting distances, such as 1 nm, we have shown that voltage and surface charge can be tuned to introduce a large enough bias to downward migration that over 95% of released nucleotides will reach their target. Of course, if the exonuclease is very close to the pore mouth, volume exclusion effects are likely to play a role, which were not part of this simulation. Still, our findings suggest a possible line of attack in future exosequencing experiments. Bootstrapped by multiple read-outs of the same DNA, a 95% success rate could be sufficient for sequencing.

A factor that should not be neglected is non-specific adsorption of the nucleotide to the outer parts of the pore, where the current blockade is still too weak to detect them. While not influencing the final exit probability, the duration of random adsorption events is the main factor determining the overall migration speed, and can be responsible for changing the order in which DNA bases are detected. An interesting line of further research would be to precisely quantify the amount of base shuffling to be expected, and demonstrate the use of sequence alignment algorithms to recover the original DNA sequence from redundant readouts.

Apart from the parameters considered in this work, a major factor influencing the exit probability is the geometry of the exonuclease-pore complex. Further research could evaluate this influence and experiment with effective geometrical arrangements which limit the translational entropy of the nucleotide on its path to the detection zone.

The presented hybrid BD/continuum framework will be useful to inform further investigations into topics of stochastic transport where fine-grained simulation methods are too slow, especially in situations where thousands of trials are necessary to provide answers with sufficient statistical resolution.

[4] Since the times between cleavage events are exponentially distributed, the quantiles (in seconds) are calculated as $-\log(1-p)/\lambda$ for $p = 1\%, 50\%$, where $\lambda = 275$ s^{-1} is the cleavage rate.

Acknowledgments

The authors acknowledge financial support by the FWF (Austrian Science Fund) START Project No. Y660 *PDE Models for Nanotechnology*. The authors acknowledge discussions with Stefan Howorka (UCL).

Conflicts of interest

There are no conflicts to declare.

ORCID iDs

Gregor Mitscha-Baude https://orcid.org/0000-0003-2607-5695

References

[1] Bayley H 2015 Nanopore sequencing: from imagination to reality *Clin. Chem.* **61** 25–31
[2] Schneider G F and Dekker C 2012 DNA sequencing with nanopores *Nat. Biotechnol.* **30** 326–8
[3] Burns J R, Seifert A, Fertig N and Howorka S 2016 A biomimetic DNA-based channel for the ligand-controlled transport of charged molecular cargo across a biological membrane *Nat. Nanotechnol.* **11** 152–6

[4] Kukwikila M and Howorka S 2015 Nanopore-based electrical and label-free sensing of enzyme activity in blood serum *Anal. Chem.* **87** 9149–54

[5] Bell N A W and Keyser U F 2016 Digitally encoded dna nanostructures for multiplexed, single-molecule protein sensing with nanopores *Nat. Nanotechnol.* **11** 645

[6] Plett T, Shi W, Zeng Y, Mann W, Vlassiouk I, Baker L A and Siwy Z S 2015 Rectification of nanopores in aprotic solvents-transport properties of nanopores with surface dipoles *Nanoscale* **7** 19080–91

[7] Feng J, Liu K, Bulushev R D, Khlybov S, Dumcenco D, Kis A and Radenovic A 2015 Identification of single nucleotides in MoS 2 nanopores *Nat. Nanotechnol.* **10** 1070

[8] Larkin J, Henley R Y, Jadhav V, Korlach J and Wanunu M 2017 Length-independent DNA packing into nanopore zero-mode waveguides for low-input DNA sequencing *Nat. Nanotechnol.* **12** 1169

[9] Wei R, Gatterdam V, Wieneke R, Tampé R and Rant U 2012 Stochastic sensing of proteins with receptor-modified solid-state nanopores *Nat. Nanotechnol.* **7** 257–63

[10] Clarke J, Wu H-C, Jayasinghe L, Patel A, Reid S and Bayley H 2009 Continuous base identification for single-molecule nanopore dna sequencing *Nat. Nanotechnol.* **4** 265–70

[11] Astier Y, Braha O and Bayley H 2006 Toward single molecule dna sequencing: direct identification of ribonucleoside and deoxyribonucleoside 5-monophosphates by using an engineered protein nanopore equipped with a molecular adapter *J. Am. Chem. Soc.* **128** 1705–10

[12] Muthukumar M 2014 Communication: charge, diffusion, and mobility of proteins through nanopores *J. Chem. Phys.* **141** 16–21

[13] Firnkes M, Pedone D, Knezevic J, Döblinger M and Rant U 2010 Electrically facilitated translocations of proteins through silicon nitride nanopores: conjoint and competitive action of diffusion, electrophoresis and electroosmosis *Nano Lett.* **10** 2162–7

[14] Laohakunakorn N, Thacker V V, Muthukumar M and Keyser U F 2014 Electroosmotic flow reversal outside glass nanopores *Nano Lett.* **15** 695–702

[15] Noskov S Y, Im W and Roux B 2004 Ion permeation through the alpha-hemolysin channel: theoretical studies based on Brownian dynamics and Poisson–Nernst–Plank electrodiffusion theory *Biophys. J.* **87** 2299–309

[16] Vlassiouk I, Smirnov S and Siwy Z 2008 Ionic selectivity of single nanochannels *Nano Lett.* **8** 1978–85

[17] van Dorp S, Keyser U F, Dekker N H, Dekker C and Lemay S G 2009 Origin of the electrophoretic force on DNA in solid-state nanopores *Nat. Phys.* **5** 347–51

[18] Lu B, Hoogerheide D P, Zhao Q and Yu D 2012 Effective driving force applied on DNA inside a solid-state nanopore *Phys. Rev.* E **86** 011921

[19] Mitscha-Baude G, Stadlbauer B, Howorka S and Heitzinger C 2019 Protein transport through nanospace illuminated by high-throughput simulations unpublished

[20] Maffeo C, Bhattacharya S, Yoo J, Wells D and Aksimentiev A 2012 Modeling and simulation of ion channels *Chem. Rev.* **112** 6250–84

[21] Si W and Aksimentiev A 2017 Nanopore sensing of protein folding *ACS Nano* **11** 7091–100

[22] Chavent M, Duncan A L and Sansom M S P 2016 Molecular dynamics simulations of membrane proteins and their interactions: from nanoscale to mesoscale *Curr. Opin. Struct. Biol.* **40** 8–16

[23] Branton D *et al* 2008 The potential and challenges of nanopore sequencing *Nat. Biotechnol.* **26** 1146–53

[24] Hall A R, Scott A, Rotem D, Mehta K K, Bayley H and Dekker C 2010 Hybrid pore formation by directed insertion of [alpha]-haemolysin into solid-state nanopores *Nat. Nanotechnol.* **5** 874–7

[25] Werner J H, Cai H, Keller R A and Goodwin P M 2005 Exonuclease i hydrolyzes dna with a distribution of rates *Biophys. J.* **88** 1403–12

[26] Brody R S, Doherty K G and Zimmerman P D 1986 Processivity and kinetics of the reaction of exonuclease i from escherichia coli with polydeoxyribonucleotides *J. Biol. Chem.* **261** 7136–43

[27] Enderlein J 2007 Nucleotide specificity versus complex heterogeneity in exonuclease activity measurements *Biophys. J.* **92** 1556–8

[28] Bressloff P C and Newby J M 2013 Stochastic models of intracellular transport *Rev. Mod. Phys.* **85** 135

[29] Masoliver J and Porrà J M 1995 Exact solution to the mean exit time problem for free inertial processes driven by gaussian white noise *Phys. Rev. Lett.* **75** 189–92

[30] Bénichou O and Voituriez R 2008 Narrow-escape time problem: time needed for a particle to exit a confining domain through a small window *Phys. Rev. Lett.* **100** 168105

[31] Guérin T, Levernier N, Bénichou O and Voituriez R 2016 Mean first-passage times of non-Markovian random walkers in confinement *Nature* **534** 356

[32] Wilson J and Aksimentiev A 2018 Water-compression gating of nanopore transport *Phys. Rev. Lett.* **120** 268101

[33] Langevin P 1908 Sur la theorie du mouvement Brownien *C. R. Acad. Sci., Paris* **146** 530–3

[34] Rubinstein I 1990 *Electro-Diffusion of Ions* (Philadelphia, PA: SIAM)

[35] Heitzinger C and Ringhofer C 2011 A transport equation for confined structures derived from the Boltzmann equation *Commun. Math. Sci.* **9** 829–57

[36] Heitzinger C and Ringhofer C 2014 Hierarchies of transport equations for nanopores—equations derived from the Boltzmann equation and the modeling of confined structures *J. Comput. Electron.* **13** 801–17

[37] Khodadadian A and Heitzinger C 2015 A transport equation for confined structures applied to the OprP, Gramicidin A, and KcsA channels *J. Comput. Electron.* **14** 524–32

[38] Mitscha-Baude G, Buttinger-Kreuzhuber A, Tulzer G and Heitzinger C 2017 Adaptive and iterative methods for simulations of nanopores with the PNP-Stokes equations *J. Comput. Phys.* **338** 452–76

[39] Rincon-Restrepo M, Mikhailova E, Bayley H and Maglia G 2011 Controlled translocation of individual DNA molecules through protein nanopores with engineered molecular brakes *Nano Lett.* **11** 746–50

[40] Wiggin M, Tropini C, Tabard-Cossa V, Jetha N N and Marziali A 2008 Nonexponential kinetics of DNA escape from alpha-hemolysin nanopores *Biophys. J.* **95** 5317–23

[41] Oukhaled A, Cressiot B, Bacri L, Pastoriza-Gallego M, Betton J M, Bourhis E, Jede R, Gierak J, Auvray L and Pelta J 2011 Dynamics of completely unfolded and native proteins through solid-state nanopores as a function of electric driving force *ACS Nano* **5** 3628–38

[42] Sexton L T, Mukaibo H, Katira P, Hess H, Sherrill S A, Horne L P and Martin C R 2010 An adsorption-based model for pulse duration in resistive-pulse protein sensing *J. Am. Chem. Soc.* **132** 6755–63

[43] Maglia G, Restrepo M R, Mikhailova E and Bayley H 2008 Enhanced translocation of single DNA molecules through alpha-hemolysin nanopores by manipulation of internal charge *Proc. Natl Acad. Sci.* **105** 19720–5