scientific reports

Check for updates

OPEN

Discriminating protein tags on a dsDNA construct using a Dual Nanopore Device

Swarnadeep Seth[1], Arthur Rand[4], Walter Reisner[2], William B. Dunbar[4], Robert Sladek[3] & Aniket Bhattacharya[1]✉

We report Brownian dynamics simulation results with the specific goal to identify key parameters controlling the experimentally measurable characteristics of protein tags on a dsDNA construct translocating through a double nanopore setup. First, we validate the simulation scheme in silico by reproducing and explaining the physical origin of the asymmetric experimental dwell time distributions of the oligonucleotide flap markers on a 48 kbp long dsDNA at the left and the right pore. We study the effect of the electric field inside and beyond the pores, critical to discriminate the protein tags based on their effective charges and masses revealed through a generic power-law dependence of the average dwell time at each pore. The simulation protocols monitor piecewise dynamics at a sub-nanometer length scale and explain the disparate velocity using the concepts of nonequilibrium tension propagation theory. We further justify the model and the chosen simulation parameters by calculating the Péclet number which is in close agreement with the experiment. We demonstrate that our carefully chosen simulation strategies can serve as a powerful tool to discriminate different types of neutral and charged tags of different origins on a dsDNA construct in terms of their physical characteristics and can provide insights to increase both the efficiency and accuracy of an experimental dual-nanopore setup.

When a biopolymer is driven through a nanopore under an applied electric field, the molecule's passage creates a dynamic modulation of the trans-pore ionic current that can be used to deduce chemical, structural, and conformational properties of the translocating polymer[1–4]. Biopolymer transport through nanopores offers significant prospects for human health. Pores based on modified transmembrane proteins form the basis of a powerful label free DNA sequencing technology[3], and there is hope that solid-state pores (ss-pores) based on 2D nanomaterials[5] may eventually sequence with sufficiently high resolution to directly read current fluctuations from one nucleotide at a time passing through the pore.

In addition to their potential thinness, ss-pores are attractive because they can be fabricated with larger dimensions (>2.5 nm diameter) suitable for analyzing not just the translocation of pure ss- and ds-DNA, but DNA with bound molecular features of nanometric size that function as physical tags or have intrinsic biological functionality[6]. A wide-range of features on dsDNA have been detected with ss-pores, including proteins such as streptavidin labels[7,8], anti-DNA antibodies[9], DNA hairpins[10,11], protein nucleic acids[12] and aptamers[13,14]. Such bound molecular features can be readily detected as these features give rise to a secondary blockade riding on top of the underlying DNA blockade in the measured trans-pore ionic current. The duration (dwell-time) and amplitude of these feature associated blockades contain information concerning the physical characteristics of the tags. If the molecule translocation is linear (i.e., no folds are present), the feature blockades can also provide information regarding the feature's binding position with respect to the molecule's underlying sequence, or the relative distance of the given feature from other features[7,10]. For example, molecular features that bind specifically to repetitive sequence motifs (for example, the recognition sequence of nicking endonucleases), produce a barcode that can then be aligned genome scale[15]. In proposed DNA information storage applications, molecular features can be used to represent the position of '1' bits[11]. In functional genomics applications, there is a need to map a wide-range of overlapping transcriptional control mechanisms, for example arising from modified bases[16] and histone marks[17].

[1]Department of Physics, University of Central Florida, Orlando, FL 32816-2385, USA. [2]Department of Physics, McGill University, 3600 rue university, Montreal, Quebec H3A 2T8, Canada. [3]Departments of Medicine and Human Genetics, McGill University, Montreal H3A 0G1, Canada. [4]Nooma Bio, 250 Natural Bridge Dr, Santa Cruz, CA 95060, USA. ✉email: Aniket.Bhattacharya@ucf.edu

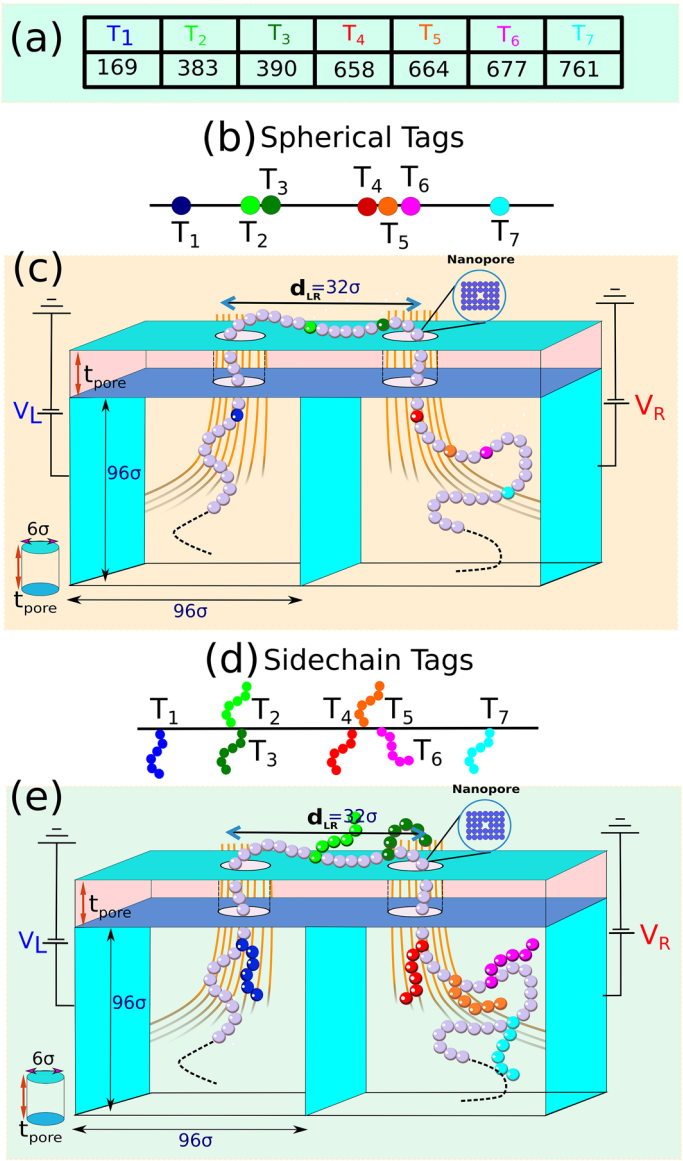

Figure 1. Schematics of the dual nanopore set up. (**a**) The simulated DNA construct with locations of the seven tags in a chain of length $L = 1024\sigma$, where σ is the diameter of each monomer of mass m. (**b**) The tags are modeled in the form of heavier beads of mass $6m$ (**c**)Simulated dual nanopore system: two nanopores of slightly different diameters 7σ (left pore) and 6σ (right pore) respectively separated by a distance $d_{LR} = 32\sigma$ are connected to two reservoirs of width 96σ. A spatially extended electric field is applied in the cavity-nanopore system. The voltage at the right pore V_R is kept at a constant value while the voltage at the left pore oscillates $V_L = V_R \pm \Delta V$ after every scan of the DNA which translocates from the left pore and vice-versa. The electric-field profile is calculated using the Finite-element method and the normalized color map shows the field strength (not in scale). (**d**) The sidechain motifs are modeled with with six monomers. (**e**) same as in (**c**) replacing beads by side-chains.

A core challenge in nanopore based feature sensing is developing techniques for accurate feature mapping and performing effective discrimination of different feature types. The double nanopore platform may have the potential to outperform devices based on single nanopores in this respect[18–23]. If a molecule is simultaneously captured at both pores in a dual pore device, applying opposing biasing to the pores will capture the molecule in a tug-of-war state where the molecule is extended between the pores[24] (Fig. 1). Such a state suppresses folding giving rise to predominantly linearized translocation traces[21]. In addition, if biasing can be independently adjusted at each pore, the two pore device can achieve controlled slow-down while maintaining high signal-to-noise current sensing[21]. This arises because the threading speed of the molecule between the pores is controlled by the difference in the potential biasing applied to each pore, while the signal is determined by the absolute bias level at each pore. Finally, when coupled to active logic that enables feedback between current measured at the pores and pore biasing, bipolar scanning can be achieved via flipping the differential bias[22]; this enables repetitive

scanning of a given genomic region (termed 'flossing')[18–22], enabling us to deduce the statistical distribution of feature characteristics and improve the measurement's statistical accuracy.

In this communication, we demonstrate that dual-pore translocation of tagged DNA in a tug-of-war regime has an intrinsic asymmetry: the dwell time of features passing through the entrance pore is on average higher than the dwell time at the exit pore. Using simulation performed using a coarse-grained (CG) model of the DNA-tag system used in dual-pore flossing experiments, we demonstrate that this aspect of dual-pore translocation is more than just a curiosity, but in fact, arises from physically distinct translocation physics specific to the dual-pore platform and may assist in discriminating between different feature classes. In a single pore device, the molecule always translocates in a direction aligned with the local electrophoretic force exerted at the pore. However, in a dual pore device with opposing biases applied to the pores, translocation can take place at the entrance pore in a direction opposite that of the local electrophoretic force exerted at this pore. In particular, the local electrophoretic force at the entrance pore acts to slow-down the passage of DNA through the entrance relative to the exit pore. When charged features are present on the DNA, the effect is to increase the passage time of the features through the entrance pore relative to the exit pore. The simulation results additionally reveal the subchain conformations and the dynamics in between the molecular features hard to access experimentally and bring out the subtleties of the interplay of the electric field inside and in the vicinity of the pore. This includes the evolution and unfolding of the non-equilibrium chain conformations due to the reversal of the electric field and the flossing direction; the inertial effect and the effect of the tension propagation (TP) along the chain backbone[25–29] and in particular how the propagation dynamics is affected by the presence of the tags that temporarily halt the tension propagation. The simulation results nicely capture the characteristics of different types of tags, neutral and charged, extended or localized, and display insights and physical understanding of the translocation process. Furthermore, a variation of the simulation parameters beyond those used in experiments enable us to understand the physical origin of the experimental uncertainties and more efficient design and analysis protocols of future multi-nanopore platforms.

Simulation and model details

Figure 1 shows the schematics of the system on which we carried out the Brownian dynamics (BD) simulation. A semiflexible chain of $N = 1024$ monomers (beads) of diameter σ is used to model a 48,500 base pair (bp) long λ-phage dsDNA construct of contour length $L = N\sigma \approx 16.5$ μm as used in the original experiment[22]. This translates to $\sigma \approx 47$ bp $\simeq 16$ nm for each bead (monomer). The details of the simulation methods are discussed in the Supplementary Materials I. In the actual experiment (please refer to the Supplementary Materials II) seven tags are placed along the chain whose relative positions are shown in Fig. 1a. The tags in the experiment consist of oligonucleotide flaps-ssDNA of 90 nt[22]. The charge of each bead of the dsDNA is chosen to be unity and the combined oligonucleotide flaps-ssDNA labels are partially charged. In order to understand the entropic and inertial effects, we have used two types of tags in the simulation. Figure 1b shows *spherical tags* in the form of beads of the same diameter σ but having a mass $m_{bead} = 6m$, where m is the mass of the individual chain monomers. Figure 1d shows *side-chain* consisting of 6 monomers, each of mass $m_{\text{sidechain}} = m$ and of the same diameter σ. We have also considered both *neutral* and *partially charged tags* in order to understand the effect of the electric field on the charged labels. The dsDNA is co-captured in (Fig. 1c,e) two closely spaced nanopores drilled on a common membrane and a tug-of-war situation is created by applying voltages V_L and V_R to the left (L) and right (R) reservoirs as shown. In compliance with the experiment we have chosen slightly different pore diameters of 7σ for the left pore and 6σ for the right pore respectively. The distance between the pores in the actual experiment is 550 nm which translates to $d_{\text{LR}} = 32\sigma$ in our simulation. We have also studied the cases for the symmetric L and R pores where $d_L^{pore} = d_R^{pore} = 6\sigma$ and observed that a 10–20% asymmetry in pore diameters does not make a large qualitative change (please refer to the Supplementary Materials III).

Flossing the captured dsDNA and the electric field in and around the pore

It is important to understand how the biases are applied in each pore in order to scan the dsDNA multiple times. In the experimental protocol[22] the voltage across the right pore $V_R = 300$ mV is kept constant while the voltage across the left pore V_L is switched from 150 to 650 mV for translocation to occur from the $L \rightarrow R$ and $R \rightarrow L$ respectively. However, small changes of ± 50 mV are sometimes made in V_L for $R \rightarrow L$ scans to maximize the data generation. We have used similar ratios for in the simulation and varied the bias at the left pore only as shown in Fig 2. We have translated 50 mV to one unit of applied bias so that the biases 150 mV, 300 mV, and 600 mV translates to 3, 6, and 12 simulation units (Fig. 2). This choice is later justified to reproduce similar Péclet number $\simeq 50$ for the simulation as well as the experiment. The co-captured dNA is scanned repeatedly by altering the voltage bias V_L at the left pore only while keeping the bias at the right pore V_R the same. Thus, $V_L = V_R + (\Delta V)_{L\rightarrow R} < V_R$ for the $L \rightarrow R$ translocation. Likewise, $V_L = V_R - (\Delta V)_{R\rightarrow L} > V_R$ for the $R \rightarrow L$ translocation. The electric field is calculated inside and in the vicinity of the nanopores exactly by solving the Maxwell's equation with the proper boundary condition (please refer to the Supplementary Materials IV for further details). The field is strongest inside the pore, extends but fades away quickly outside the pores. By varying the voltage at the left pore to drive the DNA back and forth makes the process inherently asymmetric as translocating beads face different energy barriers for the $L \rightarrow R$ and the $R \rightarrow L$ translocations. This is reflected in the dwell time distribution for the charged side-chains shown in Fig. 3. We have checked (please refer to the Supplementary Material V) that the charged spherical tags also exhibit asymmetric dwell time distributions. During flossing the chain conformations are compressed leading to a relatively faster translocation process compared to the relaxation process that takes place at a much longer time scale[30,31].

We first discuss the general characteristics of the $L \rightarrow R$ translocation. In this case a charged tag translocates through the left pore against the field but the field favors the translocation through the right pore (Fig. 2). Thus,

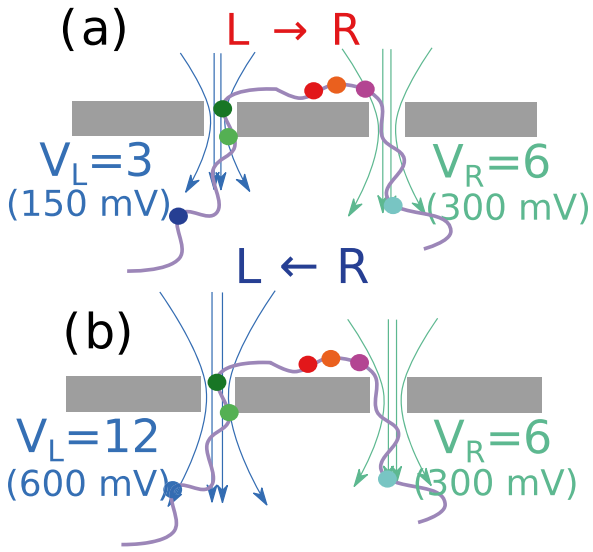

Figure 2. The schematic of the applied voltage to floss the co-captured dsDNA through the dual nanopore device. The voltage at the right pore V_R is kept constant while the voltage at the left pore $V_L = V_R + \Delta V_{L\to R}$ for $L \to R$ translocation and $V_L = V_R + \Delta V_{R\to L}$ for $R \to L$ translocation. The ratio of the voltages $V_L/V_R = 0.5$ and 2.0 for $L \to R$ and $R \to L$ translocations are the same as in the experiment[22] (Fig. 3a,b). $\Delta V_{L\to R} \neq \Delta V_{R\to L}$ in general. Other combinations are studied in Fig. 3c,d through Fig. 3k,l.

for the $L \to R$ translocation the dwell time should be broader at the left pore and sharper at the right pore (see the Supplementary Materials VI). The widths of the distributions are reversed for the $R \to L$ translocation. However, the strength of the opposing field at the R-pore is stronger for the $R \to L$ and the favorable field at the L-pore is weaker which makes the distributions for the $L \to R$ and $L \to R$ different as reflected in Fig. 3.

Each row in Fig. 3 shows the cumulative dwell time distributions for the $L \to R$ and $R \to L$ respectively. The first row shows the experimental dwell time distribution (Fig. 3a,b) where $V_L = 150$ mV and $V_R = 300$ mV for $L \to R$ translocation so that $V_R/V_L = 2.0$. Similarly, for the $R \to L$ translocation $V_L = 650$ and V_R is kept the same and in this case $V_R/V_L \approx 0.5$ (Please refer to the Supplementary Materials II). As expected, the dwell times are consistent with the above discussion so that $W_L^{L\to R}$ is broader compared to $W_R^{L\to R}$. Likewise, for the $R \to L$ translocation this order gets reversed. However, one notices that $W_R^{R\to L}$ is different than $W_L^{L\to R}$ as the biases are altered. During flossing the translocated chain at the left or right pore gets compressed to a different degree depending upon the strength of the downhill bias. Thus, when the voltage gets flipped, the degree of compression affects the unfolding and hence the speed of the translocation differently for $L \to R$ and $R \to L$ translocation. We have checked that a compressed configuration translocates faster than a fully equilibrated configuration[34]. In the experimental protocol, different combinations of voltages might be applicable in order to maximize the data generation and the scan times. Thus, we have used several different combinations of the biases in the simulation studies to check how the relative strength of the voltage at each pore affects the translocation process.

The next five rows of Fig. 3c,d,k,l are the dwell time distributions obtained from our simulation using charged side-chain tags. The second row (Fig. 3c,d) corresponds to the experimental parameters where we kept the ratio $V_R/V_L = 2$ and 0.5, the same for the $L \to R$ and $R \to L$ translocation as in the original experiment. Despite the simplicity of the model the simulation studies capture this asymmetry reasonably well. However, we observe that (1) for the $R \to L$ translocation the simulation resolves the asymmetry better than the experiment and (2) $W_L^{L\to R}$ obtained from the simulation is slightly broader than the experimental one. It is worth noting that there are some unknown factors those are not accounted for in the simulation, such as surface charges inside the pore, co-ion and counter ion movements, uncertainties in the applied voltage, and roughness of the pore. These factors could possibly reduce the experimental resolution. We have carried out another set of runs with $V_L = 13$ (650 mV) and $V_R = 6$ (300 mV) to mimic the experimental system as a possible replacement of Fig (d) and observe that this figure (shown in the Supplementary Materials VII) is very similar to Fig. 3d with the similar trend—that simulation resolves the dwell time distributions better. We have further explored other combinations also by first systematically reducing the bias at the left pore to $V_L^{L\to R} = 2$ (Fig. 3e,f), and $V_L^{L\to R} = 1$ We further observed that the asymmetry resembles closer to the experiment for the charged side chain tags than the charged spherical tags (Supplementary Materials V) (Fig. 3g,h). Reducing the bias at the left pore for the $L \to R$ translocation only changes the $W_L^{L\to R}$ without noticeably affecting the distributions at the right pore. While changing V_L for the $R \to L$ translocation causes the distributions to drift away from the experimental results. We have checked that the asymmetry observed in Fig. 3i,j almost disappears for the neutral tags as shown in Fig. 4 and for another combination of voltages which is shown in the Supplementary Materials VIII.

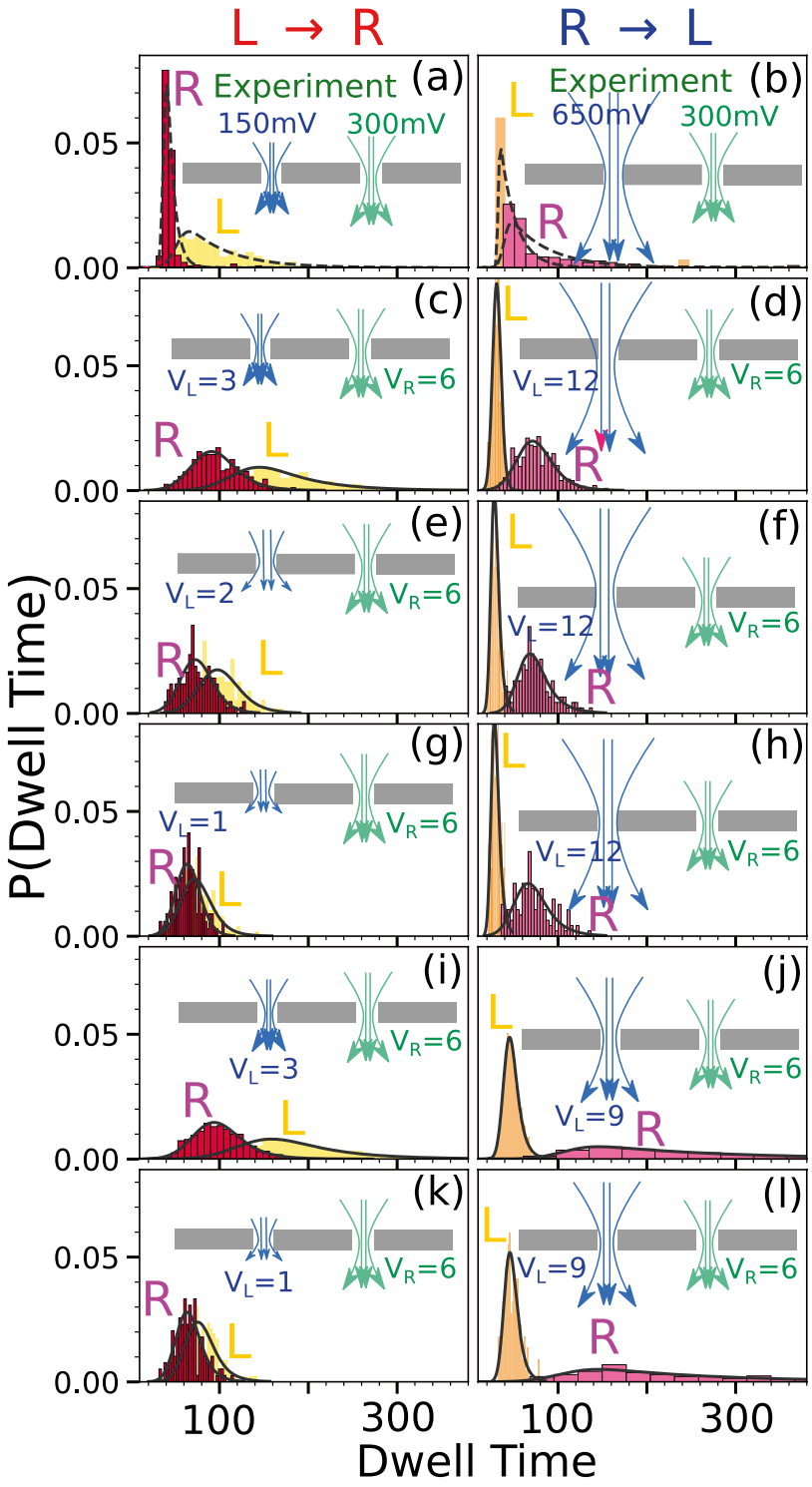

Figure 3. Cumulative dwell time distribution of the seven tags (sidechains) obtained from the experiment for (**a**) $L \rightarrow R$ and (**b**) $R \rightarrow L$ scans (1st row). The rest of the rows ((**c**)–(**d**)–(**k**)–(**l**)) (2nd–6th) are simulation dwell time data for the different combinations of voltages V_L and V_R applied across the left and the right pore for the seven sidechains placed exactly at the same locations as that of the experiment. Please note that each row consists of one flossing (two scans from $L \rightarrow R$ and $R \rightarrow L$ respectively with two distributions at left/right pores for each scan direction). In each row the yellow/red (left column) and the orange/magenta (right column) dwell time histograms are obtained from the left/right pore in $L \rightarrow R$ and $R \rightarrow L$ directions. Schematics of the electrostatic forces on the DNA in the left/right pore are shown by the blue/green arrows (not to scale). The black envelops represent the exponentially modified Gaussian distribution fit of the dwell time histograms.

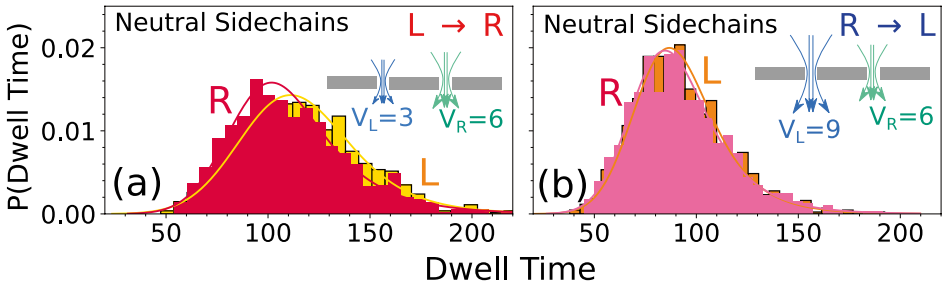

Figure 4. The cumulative dwell time distribution for (**a**) $L \rightarrow R$ and (**b**) $R \rightarrow L$ scans for the neutral sidechain tags. For the $L \rightarrow R$ scans left pore voltage is $V_L = 3$ and the right pore voltage is $V_R = 6$. For the $R \rightarrow L$ scans left pore voltage is changed to $V_L = 9$ while keeping the V_R fixed. Unlike the charged sidechain tags the asymmetry almost disappears.

How does the partial charge in side chains affect the dwell time?

From the discussions of Figs. 3 and 4 it is clear that the charge of the tags is one of the key factors that controls the asymmetry of the dwell time for the $L \rightarrow R$ and $R \rightarrow L$ translocation. In this section we show that it is the charge of the side-chain tags which is the dominating factor for the shape of the dwell time distribution. The charge lifts away the near degenerate $L \rightarrow R$ and $R \rightarrow L$ distributions for the neutral tags ($q_{tag} = 0$) (Fig. 4). Thus, we further explored the asymmetry of the dwell time distributions (Fig. 3) in terms of $\Delta W_{L\rightarrow R}(q_{tag})$ and $\Delta W_{R\rightarrow L}(q_{tag})$ (Eq. 1a) and its normalized counterpart (Eq. 1b) defined as

$$\Delta W_{L\rightarrow R}(q_{tag}) = \langle W_L^{L\rightarrow R}(q_{tag})\rangle - \langle W_R^{L\rightarrow R}(q_{tag})\rangle \tag{1a}$$

$$\Delta W^{scaled}_{L\rightarrow R}(q_{tag}) = \frac{\Delta W_{L\rightarrow R}(q_{tag})}{\Delta W_{L\rightarrow R}(q_{tag}=0)} - 1. \tag{1b}$$

for the charged sidechains shown in Fig. 5. Likewise, one can define $\Delta W_{R\rightarrow L}$ by interchanging L and R. The differential functions $\Delta W^{scaled}_{L\rightarrow R}(q_{tag})$ and $\Delta W^{scaled}_{R\rightarrow L}(q_{tag}) \rightarrow 0$ as the charge of the sidechain protein tag $q_{tag} \rightarrow 0$ and brings out more effectively the local effects of each pore on q_{tag}. The scattered plots of Figs. 5a–f clearly bring out how the isotropy ($q_{tag} = 0$) is broken and continuously evolves to acquire characteristics of the local effects from the left and the right pore as the charge of each sidechain is increased. This is an excellent demonstration of the importance of simulation studies to understand the corresponding experimental data shown in Fig. 5g. By comparing Fig. 5g with the set in Fig. 5a–f one can infer that not only the sidechain protein tags are charged, one can also estimate the partial charge content of the tags.

An important aspect of the experimental setup (Fig. 2) is that variation of the electric field occurs at the L-pore while the voltage at the R-pore is kept constant which results in the asymmetries of the dwell time. In Fig. 5j–m we further explore in detail the variations of the dwell time at the L/R pore as a function of the charge of the sidechain motifs and those can be easily understood by looking at the field directions (Fig 2) at the L/R for both $L/R \rightarrow R/L$ translocation. For example in Fig. 5j the E-field at the L/R pores are antiparallel/parallel for the $L \rightarrow R$ translocation that explains why $W^{scaled}_{L\rightarrow R}$ increases/decreases at the L/R-pore. Fig. 5k then immediately explains the monotonic increase of $\Delta W_{R\rightarrow L}(q_{tag})$. Likewise $W^{scaled}_{R\rightarrow L}$ increases/decreases at the R/L-pore (Fig. 5l) and explains the monotonic increase of the differential function $\Delta W_{R\rightarrow L}(q_{tag})$ (Fig. 5m). Furthermore, we find that the data in Fig. 5k,m can be fitted with a power law dependence $\Delta W_{L\rightarrow R}(q_{tag}) \sim A_{LR} q^{\alpha_{LR}}$ (and likewise, for $\Delta W_{R\rightarrow L}(q_{tag})$), where both the prefactors A_{LR} and A_{RL}, and the exponents α_{LR} and α_{RL} at the L/R pore are in general different and nonuniversal, and depends on the details of the parameters (please refer to the annotations in Fig. 5). We strongly believe this power law dependence can be potentially used to determine the effective charge of a motif in an experimental scan and can potentially discriminate tags by their charge contents.

How does the length of the side chains affect the dwell time?

In general, protein tags of different lengths can be present in a long DNA-strand. Then a dual nanopore device will be able to detect the presence of the different protein tags. In the previous section we studied the effect of the magnitude of the partial charge keeping the length of the sidechain protein tags the same. We have carried out a similar analysis replacing charge q_{tag} by the length of the tag l_{tag} in Eqs. (1a) and (1b) keeping the charge per bead the same $q_{tag} = 0.5q$. We observe the power law behaviors for both $W^{scaled}_{L\rightarrow R}(l_{tag})$ and $W^{scaled}_{R\rightarrow L}(l_{tag})$ (Fig. 6a,c) and monotonic increase for $\Delta W_{L\rightarrow R}(l_{tag})$ and $\Delta W_{R\rightarrow L}(l_{tag})$ (Fig. 6a,c).

It is worthwhile to note that significant deviations from the linearity as a function of q_{tag} and l_{tag} are observed. These are expected due to the combined effects of the tension propagation and varied degrees of inertial and extended electric field effects at the L/R pores for the $L \rightarrow R$ and $R \rightarrow L$ scans.

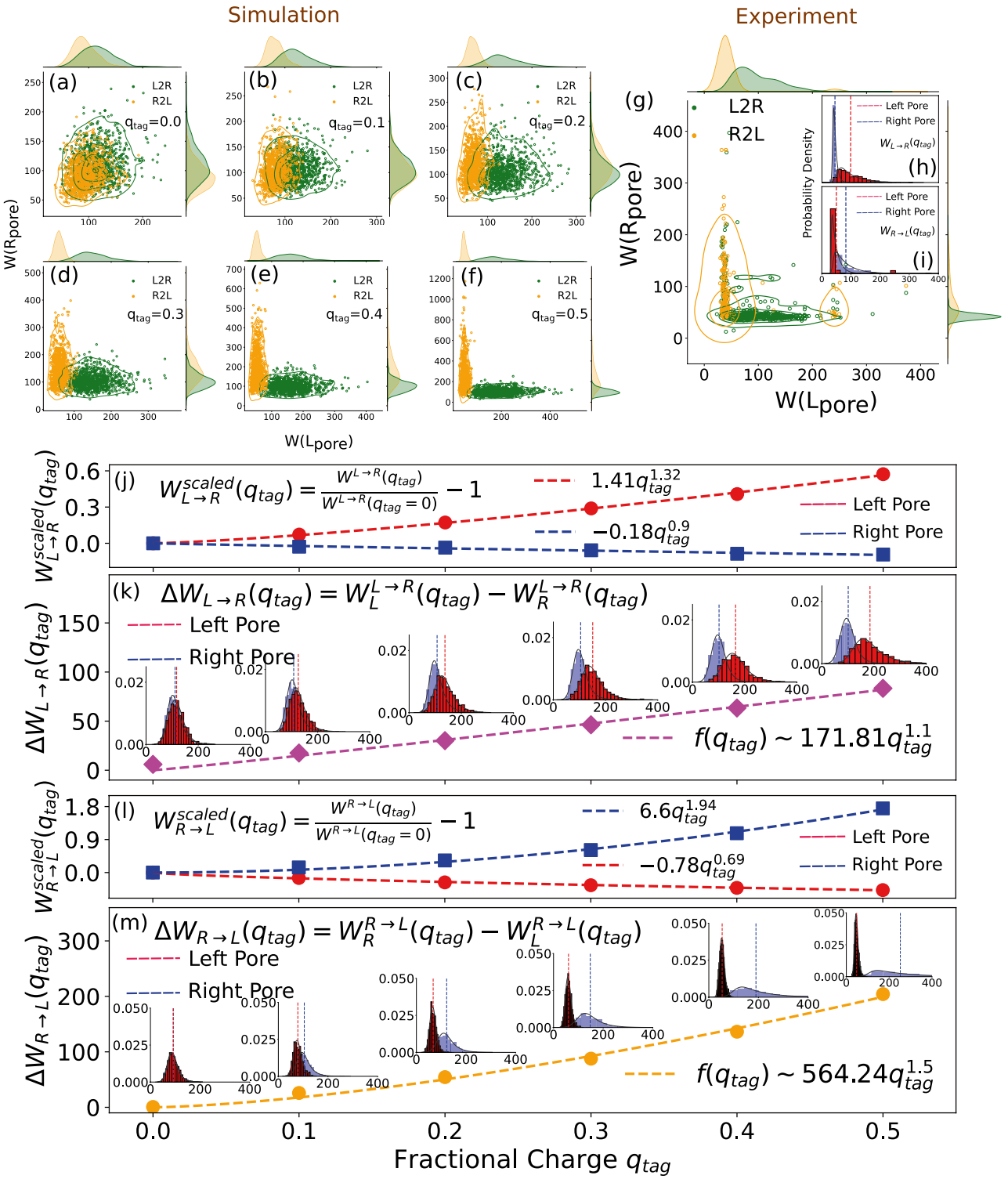

Figure 5. Scatter plots of the cumulative dwell time for the neutral (**a**) and the partially charged sidechain tags (**b**–**f**) in the L/R pores for $L \to R/R \to L$ scans (green/orange circles). Distributions of the corresponding dwell time are shown on the top and right hand axes. (**g**) the experimental scatter plot to be compared with (**a**–**f**). (**h**, **i**) are the corresponding experimental $\Delta W_{L\to R}$ and $\Delta W_{R\to L}$ (same as in Fig. 3a,b for comparison. (**j**) The scaled cumulative dwell time asymmetry $\Delta W_{L\to R}$ at the left (blue squares) and right (red circles) pore. The dotted lines (red and blue) are the corresponding nonlinear fits through the points which produce different exponents for the charge dependence. (**k**) The corresponding histograms of the points in (**j**) where the dotted vertical lines (red and blue) in each histogram represent the average dwell time. (**l**) and (**m**) are the same as in (**j**) and (**k**) except for the $R \to L$ translocation direction. The fitting exponents in (**l**) are also different than those in (**j**).

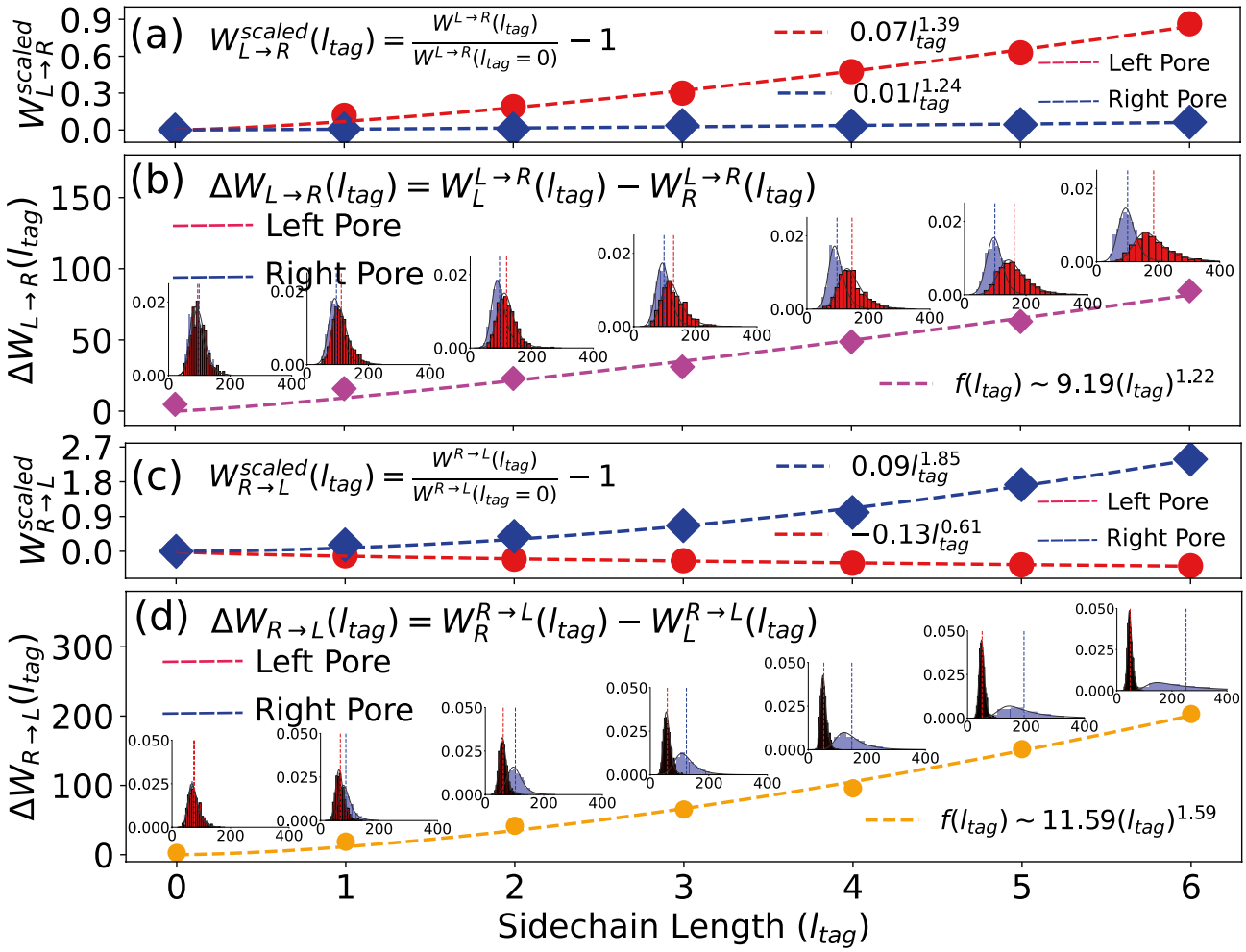

Figure 6. (**a**) The scaled cumulative dwell time asymmetry $\Delta W^{scaled}_{L\to R}$ at the left (red circles) and the right (blue squares) pore. The dotted lines are the nonlinear fits through the points which produce different exponents for the charge dependence. (**b**) The corresponding histograms of $\Delta W_{L\to R}(l_{tag})$ for different l_{tag} locations as in (**a**) where the vertical red and blue dotted lines in each histogram are the average values of the dwell time. (**c**) and (**d**) are the same as in (**a**) and (**b**) for $\Delta W^{scaled}_{R\to L}$.

Velocity of the tags from the time of flight (TOF) data

Compared to a single nanopore, in a dual nanopore setup the velocity of the tags are calculated more accurately from the TOF measurements[32] defined as the time taken by a monomer/tag of index m as it leaves one pore and reaches the other pore during its voyage across the pore separation d_{LR} defined as

$$\tau^{L\to R}(m) = t_R^{L\to R}(m) - t_L^{L\to R}(m), \tag{2a}$$

$$\tau^{R\to L}(m) = t_L^{R\to L}(m) - t_R^{R\to L}(m). \tag{2b}$$

here, $t_R^{L\to R}(m)$ and $t_L^{L\to R}(m)$ are the arrival and exit times at the right and left pore for the $L \to R$ translocation. By flipping L and R we get $\tau^{R\to L}(m)$. This is demonstrated in Fig. 7. The TOF flight velocity for a monomer/tag with index m then can be easily obtained from measurement of $\tau^{L\to R}$ and $\tau^{R\to L}$ using the known distance d_{LR}.

$$v^{tof}_{L\to R}(m) = d_{LR}/\tau^{L\to R}(m) \tag{3a}$$

$$v^{tof}_{R\to L}(m) = d_{RL}/\tau^{L\to R}(m) \tag{3b}$$

The TOF flight measurements can be experimentally obtained from the current blockade data. However, because the tags are in general of different mass, charge, and volume they introduce nonuniformity in the velocity along different portions of the chain which is difficult to access experimentally. The simulation studies show that

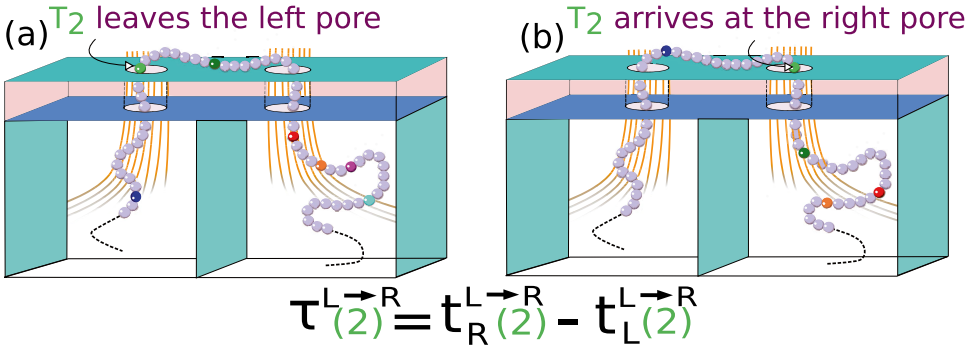

Figure 7. Illustration depicts the TOF of $\mathbf{T_2}$ is measured as the time taken to reach to left pore from right pore for $R \rightarrow L$ motion.

charged tags experience different forces compared to the dsDNA chain. This results in a nonuniform velocity profile along the chain which can be qualitatively understood using nonequilibrium tension propagation theory of Sakaue[25] and recently demonstrated in a single and dual nanopore setup[32,33]. In the following section we analyze how the mass and charge of the tags affect the velocity profile of the entire chain. One of the major goal of these simulation studies is to develop a fundamental understanding of the piecewise translocation process. Thus here in addition to mimicking the parameters used in the experiment, we study several other variations of the tag characteristics to decipher the effects of the mass and charge of the tags which create the nonuniform velocity profile along the chain shown in Fig. 8. We learn the following from a closer look at Fig. 8.

- As a reference first we show the results for the neutral and charged tags of the same mass as that of the dNA beads (Fig. 8a–f). Here, we observe that velocity increases/decreases as a function of the monomer index m for $L \rightarrow R$ and $R \rightarrow L$ translocation similar to a homopolymer[33]. Inclusion of the charge at protein tag locations does change the overall profile as expected but does not introduce significant nonuniformity in the velocity profile (Fig. 8d–f).
- It is only when tags are more massive compared to the monomer beads we observe a huge nonuniformity in the velocity profile of the entire chain with local minima roughly at the location of the indices of spherical tags (Fig. 8g–i,j–l). Thus we conclude that it is the inertia of the tags responsible for the local minima. A comparison of Fig. 8g–i,j–l shows that replacing neutral spherical tags by charged spherical tags does not alter the profile confirming the role of inertia for the case when tags are massive but have the same volume as that of the monomer beads.
- Replacing neutral spherical tags (Fig. 8g–i) by neutral side-chains of the same total mass reduce the nonuniformity of the velocity profile. This shows the interaction of the extended tags with the electric field beyond but in the vicinity of the pore and the entropy can make a significant difference.
- From an inspection of Fig. 8g–i,j–l we further observe that the TOF velocity has the capability of discriminating density distribution of the tags. There are two isolated tags at each end and two groups of tags—a group of two (T_2, T_3) and a group of three (T_4, T_5, T_6) in the system that we studied. For the *neutral* spherical tags for the $L \rightarrow R$ translocation the group of tags on which the tension front hits last lies in the minimum of the velocity profile (T_4 and T_2 in Fig 8g) while for the $R \rightarrow L$ translocation the location of the minima gets reversed (T_6 and T_3 in Fig 8h). This sequence of $v_{\mathrm{TOF}}(m) \sim m$ is reversed when *spherical neutral* tags become charged as we compare Figs. 8g,j.
- Finally, we compare the velocity profiles for charged spherical tags (Fig. 8j–l) and charged side-chain tags (Fig. 8p–r) which are relevant to understand the corresponding experimental data. Evidently, the velocity profile of the charged side-chains is resolved better compared to the spherical tags with detailed separation distances. This is due to the interaction of the side-chains which are more extended than the spherical tags hence interacting with a larger region containing the electric field. In this case not only the segment of the side-chain tag inside the pore sense the electric field inside the pore, but the segment lying outside in the immediate vicinity of the pore feels the electric field which extends from inside to the vicinity of the pore. This pre-sensing of the nanopore by the side-chains provides the velocity a more intricate structure as seen in Fig. 8p–r. It is also worth observing that the discrimination is more prominent for the $R \rightarrow L$ translocation which then can be used to our advantage to decipher the characteristics of the tags.

Sidechain length dependence of the velocity profile

We close the discussion of the velocity variations by showing its dependence on the length of the sidechains l_{tag}. Previously in Fig. 6 we showed the variation of the dwell time on l_{tag}. Since the velocity is determined experimentally from the TOF flight data, it is worthwhile to study how the sidechains alter the velocity profile. Fig. 9 shows systematic increase of the inertial effect as the tag length (and hence mass) is increased. To extract the variation we have subtracted out the velocity of the homopolymer chain without the protein sidechain tags.

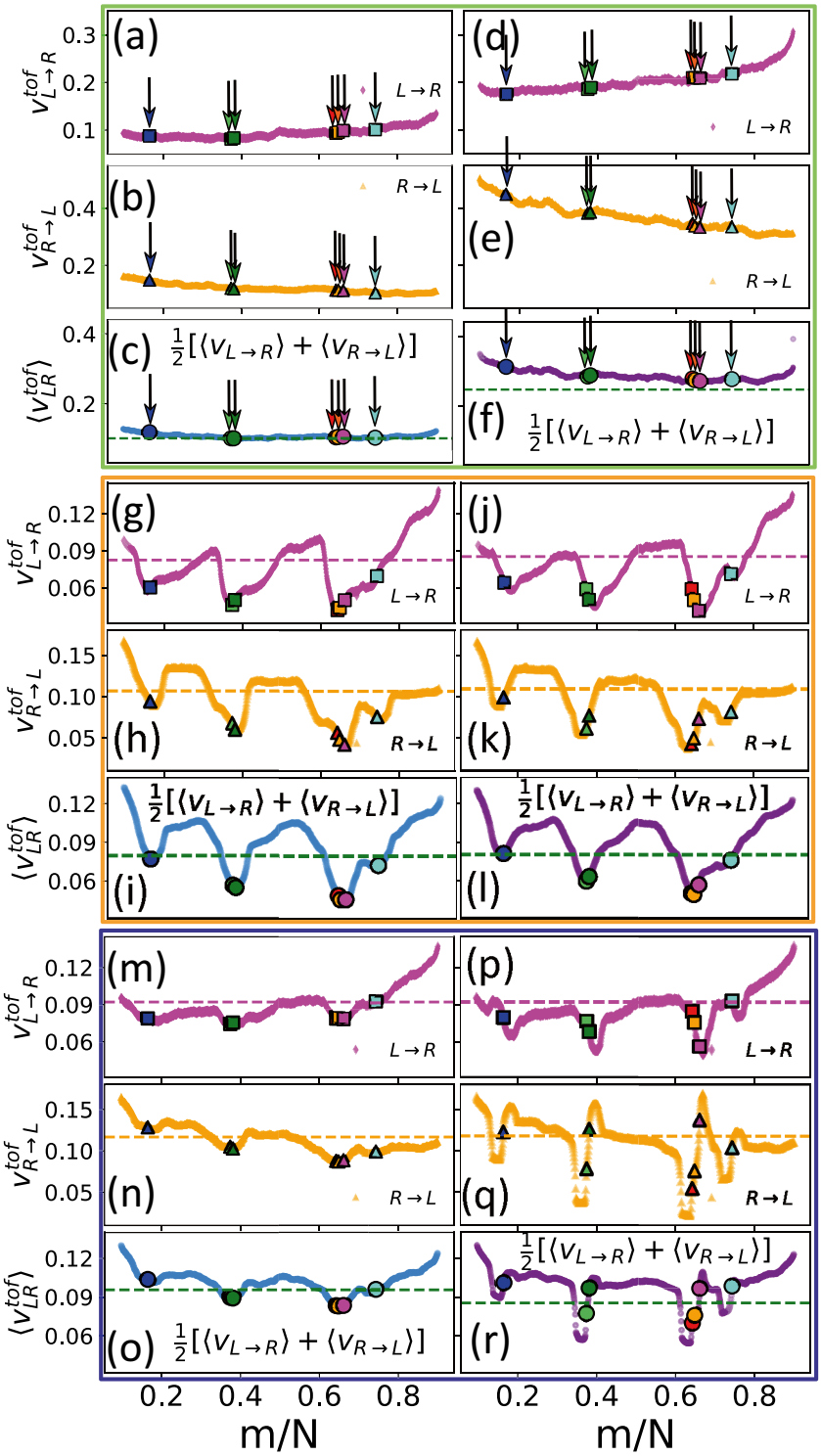

Figure 8. TOF velocity for neutral (**a**–**c**) and charged tags (**d**–**f**) with the same mass as that of the chain monomers. The locations of the tags are indicated by arrows. The positive slopes in $L \rightarrow R$ (**a**, **d**) and negative slope in $R \rightarrow L$ (**b**,**e**) denote the direction of the TP along the chain; (**g**–**a**) and (**j**)–(**l**) represent the same but for heavier (6m) neutral and charged *spherical* tags. The heavier tags introduces nonuniformity in the TOF velocity; (**m**)–(**o**) and (**p**)–(**r**) represent the same as that of (**g**)–(**i**) and (**j**)–(**l**) respectively but for the neutral and charged *side-chain* tags. The side-chains exhibit more discriminating features compared to the spherical tags of the same mass. The dashed lines in each case correspond to the average velocity of the chain.

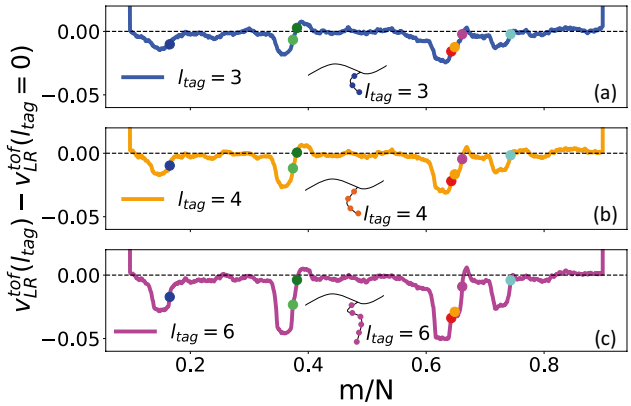

Figure 9. Variation of velocity profile of the entire chain as a function of sidechain tags of different length l_{tag}. To extract the dependence of the sidechain protein tags we have subtracted the velocity of the homopolymer ($l_{tag} = 0$). (**a**)–(**c**) correspond to $l_{tag} = 3, 4$ and 6 respectively. The increasing local variations around the position of each sidechain tag are evident. It is also worth observing that the last tag in all three cases has the same velocity of the entire chain (average velocity) and therefore, can be used as a reference velocity for the chain.

An important observation in Fig. 9a–c is that the very last tag has the average velocity of the entire chain. This information can be useful to extract the velocity data experimentally using the last tag as the reference. We also note that inertial effect become more prominent with increasing length of the sidechains l_{tag}. Thus we believe that the degree of velocity variation can be used to resolve protein tags of different mass and length (possibly of different origin) present along the dsDNA construct.

Simulation versus experimental time scale

In the Brownian Dynamics (BD) simulation we use a chain of $N = 1024$ beads ($L = 1024\sigma$) that corresponds to a 48,500 bp long dsDNA construct. This leads to the diameter of each bead $\sigma = 48,500/1024 \approx 47$ bp $\approx 47 * 0.34 = 16$ nm. We now calculate the unit of BD time scale to relate the velocity of the chain under bias to compare with those from experiments.

Mass of a base pair

The average mass of a DNA base-pair $m_{DNA} \simeq 650$ amu $= 1.67 \times 10^{-27}$kg $\times 650 \simeq 1.1 \times 10^{-24}$kg. This estimate can be obtained by looking at the chemical structures of the $A - T$ and $G - C$ bonds, and accounting for the molecular weights of deoxy-Ribose and the phosphate group both contributing to the dsDNA mass. This will allow us to get the time unit for the BD simulation from the following equation.

$$\tilde{t}_{BD} = \sqrt{\frac{m_{DNA}\sigma^2}{\epsilon}} \simeq 0.26 \text{ ns} \tag{4a}$$

$$\therefore \tilde{v}_{BD} = \frac{\sigma}{\tilde{t}_{BD}} = \sqrt{\frac{k_B T}{m_{DNA}}} \simeq 61.0 \text{ m/s}. \tag{4b}$$

here, we have used $\epsilon \simeq k_B T \simeq 4200 \times 10^{-24}$ J per monomer bead. Thus, $\tilde{v}_{BD}$ is independent of σ and only depends on the interaction strength and the mass of the beads. From simulation, typical dimensionless values of $\tilde{v}_{TOF} \approx 0.1$ (from Fig. 8) which translates to the actual velocity $= \tilde{v}_{TOF} \times 60$ m/s ≈ 6.0 m/s $= 600$ mm/s. It is well known that the BD simulation with implicit solvent makes the time scale faster depending upon the degree of coarse graining. Considering we have translated 48,500 bp to 1024 coarse-grained beads (a factor of ≈ 50), this simulation in actual solvent would translate to the velocity $\tilde{v}_{TOF}/50 \approx 10$ mm/s, which is the typical order of magnitude of velocity for the dual nanopore experiments[21,22]. This is the reason why the BD simulation is capable of reproducing the same experimental trend qualitatively. It is expected that if one would carry out a more expensive calculation with explicit solvents and with a longer chain the agreement will be similar. Roughly speaking the BD simulation captures the physical phenomena albeit at a faster time scale depending upon the degree of coarse-graining.

Translating simulation bias to the experimental bias & Péclet number

We have used the Péclet number[35–37]

$$P_e = \frac{\tau_{\text{relax}}}{\tau_{\text{trans}}} \tag{5}$$

to narrow down the applied bias used in the simulation. Here, $\tau_{\rm relax}$ and $\tau_{\rm trans}$ are the relaxation and translocation time for the translocating polymer, thus is a measure of diffusive versus the drift motion. It is expected then that if the Péclet number from the simulation $P_e^{sim} \approx P_e^{expt}$, where P_e^{expt} is the Péclet number obtained using the experiment data, the simulated system will acquire similar level of relative drift as observed in the experimental system. In the Supplementary Materials IX we provide the details of how to obtain P_e^{sim} and P_e^{expt} and demonstrate that $P_e^{sim} \simeq 63 \approx P_e^{expt} \approx 60$. Thus, this agreement of the Péclet numbers from the experiment and the simulation further justifies and closes the loop why the BD simulation studies capture the essential features of flossing in a dual nanopore device and give further confidence to use this model for analysis of a more complicated mixed system of tags *in silico*.

Concluding remarks

We have developed novel BD simulation strategies whose overarching goal is to extract the the underlying physics of the dual nanopore translocation at sub-nanometer length scales hard to obtain experimentally and hence improve the accuracy of the locations of protein tags on dsDNA constructs based on the details as revealed from the results obtained from the CG model. The simulation strategies are also capable of predicting possible variations of the device characteristics of the dual nanopore system to improve its accuracy. In the BD simulation we varied the magnitudes of local electric fields at each nanopore and demonstrated that both the average dwell time and degree of asymmetries due to opposing and favoring local fields follow power laws as a function of the charge as well as the length of the protein tags, albeit with different exponents and amplitudes. Establishment of such a result will be useful to analyze experimental data as one can study how the shapes of the dwell time distribution get altered under different electric fields as well as the characteristics of the protein tags. The time evolution of a flossed dsDNA subject to repeated scans needs to be understood in terms of nonequilibrium statistical mechanics. We have explained a variety of scenarios in terms of polymer physics concepts and nonequilibrium tension propagation theory, such as, how the fine structures of the velocity profile of the entire chain are altered due to the presence of the protein tags. Finally, a direct relevance of the model to experimental results is the observation that the coarse-graining length factor which is the ratio of the actual length to the simulation chain length provides a guide how to compare the experimental velocity to the simulation velocity. This argument is validated by comparing the Péclet numbers. Our studies demonstrate that protein tags of different biological origins can be discriminated in terms of their physical characteristics enabling the simulation protocols to have huge potential application in genomics.

Data availability

The datasets used and/or analyzed during the current study available from the corresponding author on reasonable request.

Received: 8 March 2022; Accepted: 9 June 2022
Published online: 04 July 2022

References

1. Deamer, D., Akeson, M. & Branton, D. Three decades of nanopore sequencing. *Nat. Biotechnol.* **34**, 518–524 (2016).
2. Muthukumar, M. *Polymer Translocation.* https://doi.org/10.1201/b10901 (2016).
3. Wanunu, M. Nanopores: A journey towards DNA sequencing. *Phys. Life Rev.* **9**, 125–158 (2012).
4. Palyulin, V. V., Ala-Nissila, T. & Metzler, R. Polymer translocation: The first two decades and the recent diversification. *Soft Matter* **10**, 9016–9037 (2014).
5. Farimani, A. B., Min, K. & Aluru, N. R. DNA base detection using a single-layer MoS2. *ACS Nano* **8**, 7914–7922 (2014).
6. Albrecht, T. Single-molecule analysis with solid-state nanopores. *Ann. Rev. Anal. Chem.* **12**, 371–387 (2019).
7. Chen, K. *et al.* Ionic current-based mapping of short sequence motifs in single DNA molecules using solid-state nanopores. *Nano Lett.* **17**, 5199–5205 (2017).
8. Kong, J., Bell, N. A. W. & Keyser, U. F. Quantifying nanomolar protein concentrations using designed DNA carriers and solid-state nanopores. *Nano Lett.* **16**, 3557–3562 (2016).
9. Plesa, C., Ruitenberg, J. W., Witteveen, M. J. & Dekker, C. Detection of individual proteins bound along DNA using solid-state nanopores. *Nano Lett.* **15**, 3153–3158 (2015).
10. Bell, N. A. W. & Keyser, U. F. Digitally encoded DNA nanostructures for multiplexed, single-molecule protein sensing with nanopores. *Nat. Nanotechnol.* **11**, 645–651 (2016).
11. Chen, K. *et al.* Digital data storage using DNA nanostructures and solid-state nanopores. *Nano Lett.* **19**, 1210–1215 (2018).
12. Singer, A., Rapireddy, S., Ly, D. H. & Meller, A. Electronic barcoding of a viral gene at the single-molecule level. *Nano Lett.* **12**, 1722–1728 (2012).
13. Sze, J. Y. Y., Ivanov, A. P., Cass, A. E. G., & Edel, J. B. Single molecule multiplexed nanopore protein screening in human serum using aptamer modified DNA carriers. *Nat. Commun.* **8** (2017).
14. Kong, J., Zhu, J., Chen, K. & Keyser, U. F. Specific biosensing using DNA aptamers and nanopores. *Adv. Functi. Mater.* **29**, 1807555 (2018).
15. Rand, A. *et al.* Electronic mapping of a bacterial genome with dual solid-state nanopores and active single-molecule control. *ACS Nano* **16**, 5258–5273 (2022).
16. Raiber, E.-A., Hardisty, R., van Delft, P., & Balasubramanian, S. Mapping and elucidating the function of modified bases in DNA. *Nat. Rev. Chem.* **1** (2017).
17. Carter, B. & Zhao, K. The epigenetic basis of cellular heterogeneity. *Nat. Rev. Genet.* **22**, 235–250 (2020).
18. Pud, S. *et al.* Mechanical trapping of DNA in a double-nanopore system. *Nano Lett.* **16**, 8021–8028 (2016).
19. Choudhary, A. *et al.* High-fidelity capture, threading, and infinite-depth sequencing of single dna molecules with a double-nanopore system. *ACS Nano* **14**, 15566–15576 (2020).
20. Zhang, Y. *et al.* Single molecule DNA resensing using a two-pore device. *Small* **14**, 1801890 (2018).
21. Liu, X., Zhang, Y., Nagel, R., Reisner, W. & Dunbar, W. B. Controlling DNA tug-of-war in a dual nanopore device. *Small* **15**, 1901704 (2019).
22. Liu, X. *et al.* Flossing DNA in a dual nanopore device. *Small* **16**, 1905379 (2020).

23. Choi, J. *et al.* Label-free identification of single mononucleotides by nanoscale electrophoresis. *Small* **17**, 2102567 (2021).
24. Yeh, J.-W., Taloni, A., Chen, Y.-L. & Chow, C.-F. Entropy driven single molecule tug-of-war at micro-nanofluidic intefaces. **12**, 1597–1602 (2012).
25. Sakaue, T. Nonequilibrium dynamics of polymer translocation and straightening. *Phys. Rev. E* **76**, (2007).
26. Ikonen, T., Bhattacharya, A., Ala-Nissila, T. & Sung, W. Influence of non-universal effects on dynamical scaling in driven polymer translocation. *J. Chem. Phys.* **137**, 085101 (2012).
27. Adhikari, R. & Bhattacharya, A. Driven translocation of a semi-flexible chain through a nanopore: A Brownian dynamics simulation study in two dimensions. *J. Chem. Phys.* **138**, 204909 (2013).
28. Bell, N. A. W., Chen, K., Ghosal, S., Ricci, M. & Keyser, U. F. Asymmetric dynamics of DNA entering and exiting a strongly confining nanopore. *Nat. Commun.* **8**, 380 (2017).
29. Chen, K. *et al.* Dynamics of driven polymer transport through a nanopore. *Nat. Phys.* **17**, 1043–1049 (2021).
30. Bhattacharya, A. & Seth, S. Tug of war in a double-nanopore system. *Phys. Rev. E* **101**, (2020).
31. Seth, S. & Bhattacharya, A. Polymer escape through a three dimensional double-nanopore system. *J. Chem. Phys.* **153**, 104901 (2020).
32. Seth, S. & Bhattacharya, A. DNA barcodes using a double nanopore system. *Sci. Rep.* **11**, 9799 (2021).
33. Seth, S. & Bhattacharya, A. DNA barcode by flossing through a cylindrical nanopore RSC. *Advances* **11**, 20781–20787 (2021).
34. Seth, S. & Bhattacharya, A. (unpublihsed).
35. Saito, T. & Sakaue, T. Process time distribution of driven polymer transport. *Phys. Rev. E* **85**, 061803 (2012).
36. de Haan, H. W., Sean, D. & Slater, G. W. Using a Peclet number for the translocation of DNA through a nanopore to tune coarse-grained simulations to experimental conditions. *Phys. Rev. E* **91**, 022601 (2015).
37. McMullen, A., de Haan, H. W., Tang, J. X. & Stein, D. Stiff filamentous virus translocations through solid-state nanopores. *Nat. Commun.* **5**, 061803 (2014).
38. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Molec. Graphics* **14**, 33–38 (1996).

Acknowledgements

The research has been supported by the Grant number 1R21HG011236-01 from the National Human Genome Research Institute at the National Institute of Health. All computations were carried out using STOKES High Performance Computing Cluster at UCF. The simulation movie is generated using the Visual Molecular Dynamics package[38].

Author contributions

A.B. directed the project. A.B. and S.S. planned the simulation studies and wrote the first draft. S.S. performed all the simulation studies and prepared all the graphs. A.R. and W.R. provided the experimental data. W.B.D., R.S., and W.R. contributed to the final draft of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-14609-9.

Correspondence and requests for materials should be addressed to A.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.