Disentangling Steric and Electrostatic Factors in Nanoscale Transport Through Confined Space

Steven Buchsbaum,[1,⊥] Nick Mitchell,[2,⊥] Hugh Martin,[3] Matt Wiggin,[4] Andre Marziali,[4] Peter V. Coveney,[3] Zuzanna Siwy,[1,] Stefan Howorka[2,*]*

[1] School of Physical Sciences, University of California, Irvine, CA 92697, [2] Department of Chemistry, Institute of Structural Molecular Biology, University College London, London WC1H 0AJ, U.K., [3] Department of Chemistry, Centre for Computational Science, University College London, London WC1H 0AJ, U.K., [4] Department of Physics & Astronomy, University of British Columbia, Vancouver BC V6T 1Z1, Canada

[⊥]These authors contributed equally to the work.

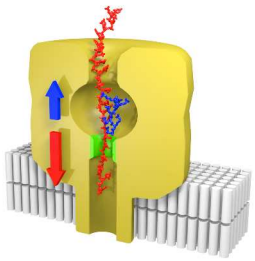
* Corresponding authors: s.howorka@ucl.ac.uk, zsiwy@uci.edu

Phone: 0044 20 7679 4702, Fax: 0044 20 7679 7463

TITLE RUNNING HEAD: Dis-entangling Nanoscale Transport

ABSTRACT

The voltage-driven passage of biological polymers through nanoscale pores is an analytically, technologically, and biologically relevant process. Despite various studies on homopolymer translocation there are still several open questions on the fundamental aspects of the pore transport. One of the most important unresolved issues revolves around the passage of biopolymers which vary in charge and volume along their sequence. Here we exploit an experimentally tunable system to disentangle and quantify electrostatic and steric factors. This new, fundamental framework facilitates the understanding of how complex biopolymers are transported through confined space and indicates how biopolymer translocation can be slowed down to enable future sensing methods.



SYNOPSIS: The voltage-driven translocation of complex biopolymers through a protein nanopore is biophysically modeled to disentangle and quantify electrostatic and steric factors which can oppose electrophoretic movement.

KEYWORDS: Single-molecule, nanopore, DNA, peptide, kinetics, biophysical model

INTRODUCTION

The transport of biopolymers through nanoscale channels is of relevance in nanopore analytics, nanofiltration,[1] and biology. In nanopore analytics, single molecules are detected via the passage-induced changes in ionic current.[2-7] The approach can be implemented with protein[4, 8-10] and solid-state nanopores[10-16] and has gained popularity[17] due to its simplicity, the wide range of accessible analytes,[2, 13, 18-20] and the prospect of DNA sequencing[21-23] based on nucleic acid sensing.[24-27] The translocation of protein and DNA strands through protein channels is also a biologically relevant process in pathogenic bacteria and human cells.

Reflecting the technological applications, a fundamental understanding of nanoscale transport including a biophysical model is key. Detailed insight has been gained from studies on pores of known molecular structure and the electrophoretically-driven movement of individual DNA and RNA strands.[3, 11, 21, 23, 28-30] Several biophysical parameters have been determined such as threading frequency,[31-34] orientation of strands,[35-39] velocity of DNA transport,[40] influence of pore geometry,[41] and interactions with the pore wall.[39] In addition to nucleic acids, translocation has also been investigated for peptides,[42-45] proteins,[46-48] and peptide-oligonucleotide conjugates,[49] which, unlike nucleic acid strands, frequently feature inhomogeneous charge distributions and vary considerably in diameter along the polymer sequence.

Here, we ask the fundamental question of how charge and steric factors of biopolymers influence their transport through nanopores. The relevance of this topic is highlighted by the increasing role of polypeptides[50] and DNA strands with point mutations[24, 26, 27] in electrical

sensing and, in more general terms, the need to develop a generic biophysical model applicable to many different molecules and pores. Current models for homogeneously charged DNA strands do not account for biopolymers where charge and size vary along the sequence. In this study we address this shortcoming by using an interdisciplinary approach comprising chemically defined biopolymers, nanopore recordings, biophysical modeling, and computational molecular simulations. By examining the translocation of biopolymers as a function of tunable diameter, length and charge, as well as applied potential, we generate a biophysical model able to disentangle and quantify steric and electrostatic factors in pore translocation, something hitherto not achieved.

RESULTS AND DISCUSSION

Defining the components of the experimental system. We determined the influence of charge and size in the voltage-driven biopolymer translocation using the α-hemolysin (αHL) pore. This natural membrane protein is of known structure (Figure 1A)[51] and a widely accepted reference standard in nanopore studies. With regard to biopolymer translocation, the structurally most important part of αHL channel (Figure 1A, yellow) is its 1.3 nm-wide inner constriction (Figure 1A, green). The narrow bottleneck limits the passage to single stranded DNA (ssDNA) molecules and is the site where the majority of an applied electric field drops off.[37, 52, 53] The electric field is the electrophoretic driving force which moves a negatively charged ssDNA molecule (Figure 1A, red) from e.g. the *cis* to the positive pole on the *trans* side of the pore. Given the key role of the inner constriction, we surmised that a non-homogeneous ssDNA molecule with a wide central segment of tunable charge and length (Figure 1A, blue) would help elucidate the influence of sterics and electrostatics in pore translocation. In particular, one could foresee that a strand with a central wide segment of positive charge would be slowed down in the inner constriction due to the steric bulk and the lack of electrophoretic force (Figure 1A, arrows).

In order to disentangle these two factors, we synthesized DNA molecules with a wide segment, tunable length and charge as shown schematically in Figure 1B. A common DNA oligonucleotide of 27 nt carries a positively charged oligoarginine tag of three, five or seven residues covalently attached at an internal base position (R_3-DNA, R_5-DNA, R_7-DNA; Figure 1B, blue) or a negatively charged hexa-aspartate tag (D_6-DNA; Figure 1B, magenta). The attachment is mediated by a linker (Figure 1B, black); the detailed chemical structure of DNA-

peptide conjugates is shown in Figure 1C. The strategy for the chemical synthesis of the conjugates[49] and evidence for their successful chemical preparation is provided in Supporting Information, Figures S-1, S-2, S-3 and Table S-1.

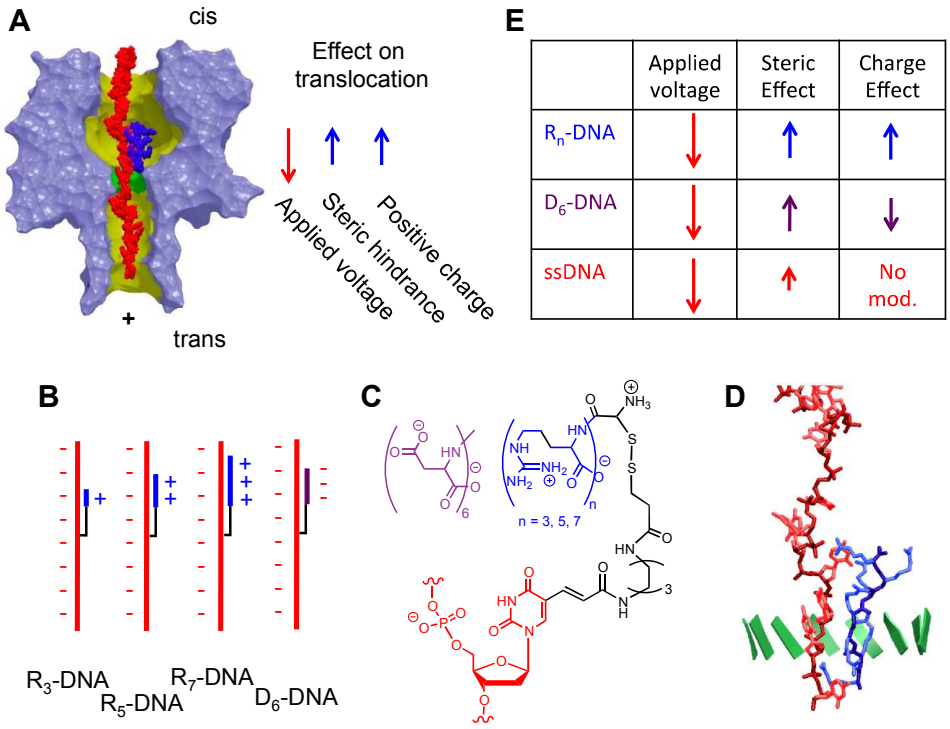


Figure 1. Bio-chemical components for examining pore translocation. (A) The αHL protein pore features a channel (yellow) with the inner constriction (green).[51] A DNA strand (red) with an attached peptide (blue) encounters steric hindrance upon translocation through the constriction. (B) Four DNA oligonucleotides R_3-DNA, R_5-DNA, R_7-DNA, and D_6-DNA carry an oligoarginine tag (blue) of varying length or a hexa-aspartate tag (purple) attached to an internal base position. (C) Chemical structures of the oligoarginine and hexa-aspartate tags, linker, and DNA oligonucleotide. (D) Snap-shots of energetically equilibrated computer simulation of the R_7-DNA conjugate threaded inside the narrow inner constriction represented by the green-colored stripes (residues 112, 113, 146 and 147). (E) Summary table showing how steric and

electrostatic factors support or oppose the translocation of peptide-tagged DNA strands through αHL.

We determined the width, length and net charge of the central segment of the conjugate threaded inside αHL using molecular models and simulations[54, 55] as detailed experimental data are not available. The molecular model of the αHL pore was based on its X-ray coordinates whereas the DNA oligonucleotide and the R_7-peptide were assembled *in silicio* based on the known monomer dimensions and molecular modeling software. In the computational simulations, the R_7-DNA conjugate was threaded into the narrow inner constriction of the pore via constant velocity-steered molecular dynamics simulations followed by equilibration for 4 ns to obtain an energetically representative structure (Supporting Information, Figure S-4). As illustrated by the close-up view of R_7-DNA (Figure 1D), the parallel aligned and slightly coiled peptide strand is intertwined with DNA. Together, the biopolymers have an average diameter of 1.63 nm – in agreement with experimental studies[56, 57] – but feature a smaller value of 1.39 nm at the inner constriction (measured from the hard-sphere surfaces) (Supporting Information, Figure S-5). This thickness is very close to the maximum width of 1.35 nm of the inner constriction which has a corrugated surface at the measured residue position 113. Furthermore, the R_7-peptide within the constriction has a pitch of 0.29 nm (C_α-C_α) which is lower than the value of 0.52 nm for the elongated DNA (C3′- C3′). The mismatch implies that the R_n-DNA segment has a positive charge of 0.8 per DNA residue.

Based on its width and positive charge, the central segment of the R_n-conjugates was expected to slow down electrophoresis-driven translocation (Figure 1E). Conjugate D_6-DNA served as a control due to its negative net charge yet similar diameter (Figure 1E).

The translocation of R_3-DNA, R_5-DNA, R_7-DNA, and D_6-DNA through αHL. The pore-passage of peptide-modified DNA oligonucleotides was characterized using single-channel current recordings. Under standard electrolyte conditions[3, 35, 49] at 100 mV, the blank αHL pore exhibited a conductance of 1930 ± 190 pS (n = 4, number of independent recordings), in line with literature.[49, 58, 59] Addition of the peptide-tagged strands R_3-DNA, R_5-DNA, and R_7-DNA to the *cis* side gave rise to high-amplitude current blockades (Figure 2A-C, respectively). The blockades most likely arise when individual net negatively charged DNA molecules are electrophoretically driven through the inner constriction towards the positively polarized *trans* side. For quantitative analysis, we only considered type I events (Figure 2A) which start with the high-blockage state and likely represent a conjugate where DNA and peptide are aligned in parallel (Figure 2A) with DNA threaded first into the pore. The other type II events had a preceding mid-blockade level (Figure 2A) which probably stems from misfolded strands that reside in the internal cavity[35] but eventually thread into the inner constriction to generate an almost complete blockade.[34, 49] Similarly, D_6-DNA blockades (Figure 2D) also showed two types of events, and the noisier mid-blockade of type II events can be explained in a related fashion as for the arginine-bioconjugates.[60]

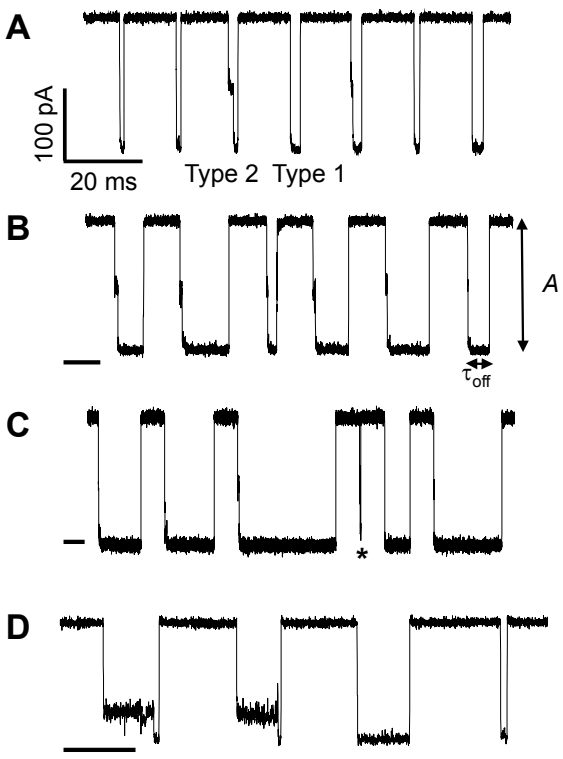


Figure 2. Representative single-channel current traces of α-HL with strands (A) R_3-DNA, (B) R_5-DNA, (C) R_7-DNA, and (D) D_6-DNA recorded at a transmembrane potential of +100 mV and in 2 M KCl, 20 mM Tris, pH 8.0. The current traces were acquired in the variable-length mode of Clampex where intermittent stretches of the open-channel current are omitted. A small and short current blockades in panel C (indicated by an asterisk) likely stem from a non-modified DNA strand. Impurities of DNA oligonucleotides without a peptide tag were very rare in the R_7-DNA traces as 98% of all events were of long duration and high-amplitude.

The amplitude of the blockades, *A*, (Figure 2B), obtained by all-point histogram analysis, were: R_3-DNA: 98.2 ± 0.5 %; R_5-DNA, 98.7 ± 0.5 %; R_7-DNA, 99.3 ± 0.4 %; and D_6-DNA, 97.4 ± 1.43 % of the open channel current (Figure 2). The amplitudes are high compared to non-

modified DNA oligonucleotides (91.7%),[49] suggesting that the peptide tag almost completely blocks the inner pore by steric and/or electrostatic factors. The peptide tags also affected the translocation duration, τ_{off}, (Figure 2B). Its characteristic average, τ^*, was obtained as described in the Methods section (Supporting Information) using on average 2500 events per condition (Supporting Information, Table S-2) and fitting to the Becquerel decay function[39, 61] (Supporting Information; Figure S-6). The characteristic duration for R_3-DNA, R_5-DNA and R_7-DNA events was 2.4 ± 0.3 ms, 10.9 ± 1.2 ms and 19.1 ± 7.8 ms, respectively (Figure 2A-2C). The dependence of translocation time on Arg-tag length is also visualized in Figure 3A at another voltage. By comparison, the translocation time for D_6-DNA was 2.0 ± 0.2 ms (Figure 2D). The durations for all of the noted peptide-modified DNA strands are one to two orders of magnitude larger than the value of 0.12 ms for non-modified DNA of similar length.[49]

Several qualitative interpretations can be drawn from the observations. Firstly, the correlation between R_n-DNA event duration and peptide length strongly supports the notion that the tags fully translocate through the pore. Secondly, the dramatic slowing down of the R_n-DNA compared to non-modified DNA implies the influence of electrostatic factors. In a molecular representation, a net positively charged R_n-DNA segment moving into the constriction where the electric field drops off will experience a weaker or even reversed electrostatic force. The presence of electrostatic factors is also supported when the R_5-DNA duration is compared to D_6-DNA data. The translocation caused by the positively charged R_5 tag is an order of magnitude slower than for the negatively charged D_6 tag, even though both peptide tags have approximately the same size and volume.

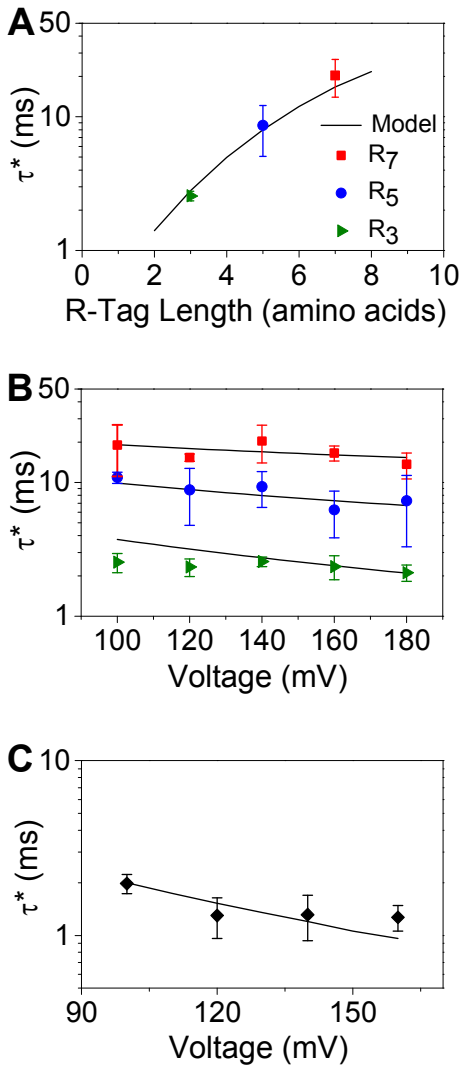


Figure 3. Quantitative analysis of pore translocation for DNA-peptide conjugates. Plots show experimental data points and fits for τ^*, the dominant characteristic time scales of the Becquerel decay function. Characteristic time scales for translocation of (A) R_3-DNA (green triangles), R_5-DNA (blue circles), and R_7-DNA (red squares), as a function of applied potential, (B) for the R_n-tag as a function of arginine tag length at 140 mV, and (C) for D_6-DNA as a function of applied voltage. Error bars are calculated as the average standard deviation in τ^* obtained from three or more independent experiments at each voltage with the exception of data points for R_5-DNA at 120 and 180 mV with two recordings but with an above-average total number of 2598 and 3145

events, respectively. The fits are derived from the biophysical model and provide the parameters of $q_1 = 0.31e$, $q_2 = -0.013e$, $k_1 = 1387\ s^{-1}$, $k_0 = 10^5\ s^{-1}$ for data in panels A and B, and $q_2 = 0.48$ for data in panel C.

However, thirdly, the slowing down of conjugates is also partly caused by steric hindrance when the wide DNA-peptide segment passes through the narrow inner constriction. The role of this steric factor is evident as the negatively charged and hence electrophoretically attracted D_6-DNA with a wider central segment is one order of magnitude slower than the non-modified DNA strand.

In order to better discriminate electrostatic and steric factors, we increased the data set by measuring the translocation of all four peptide-DNA conjugates as a function of voltage as illustrated for the R_n-DNA strands (Figures 3B) and D_6-DNA (Figure 3C). Almost all data points were obtained by averaging a total of 2500 events from three independent recordings (Supporting Information, Table S-2; legend to Figure 3). The quality of data for R_3-DNA and R_7-DNA was high as indicated by the small error bars (Figure 3B). For some R_5-DNA data points (Figure 3B, 120 to 180 mV) the errors are lager even though the number of events was very high and ranged between 2500 and 3100 events, implying that obtaining event durations can be associated with a larger data spread, as observed by others.[62] Irrespectively, two major trends are apparent. Firstly, longer arginine tag lengths lead to longer translocation times at all voltages (Figure 3B), thereby confirming and extending the previously mentioned observations at an isolated voltage (Figure 3A). Furthermore, the event durations decrease with increasing magnitude of the potential. This is the case for R_5-DNA and R_7-DNA (Figure 3B) as well for D_6-DNA (Figure 3C). While the first trend on tag length suggests the influence of sterics and potentially electrostatic forces, the second trend is most probably due to faster electrophoresis at

higher voltages. To better discriminate electrostatic and steric factors we developed a biophysical model whose good fit to the data serves to validate it.

Quantitative analysis using a biophysical model helps define the role of sterics and electrostatics. A biophysical model was developed to describe τ^* as a function of V and arginine/aspartate tag lengths. The model is based on a previously developed approach[63] and now considers the electrostatic and steric effects of modified ssDNA translocating the internal constriction of the αHL pore (Figure 4). In the model, the DNA-peptide conjugate is divided into a central, chemically modified and two non-modified flanking regions which differ with regard to effective charge and steric-dependent speed of translocation (Figure 4A). The force for translocation is the applied potential which drops off at the narrow inner constriction (Figure 4B). In order to obtain expected translocation times, our model calculates the free energy of translocation and the associated hopping rates for a DNA strand which is considered to slide forward and backward for each nucleotide, depending on steric and electrostatic factors (Figure 4C).

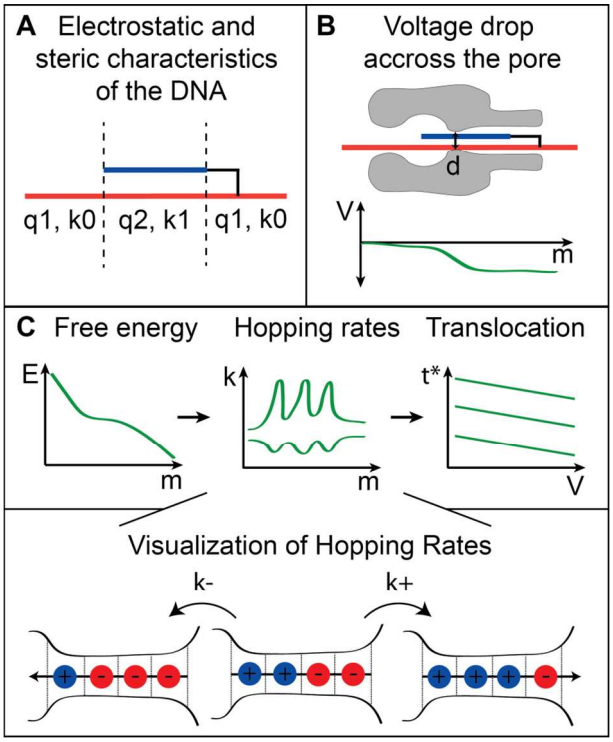


Figure 4. Scheme illustrating our biophysical model. (A) The charge sequence and steric features of the peptide-modified DNA strand are defined. (B) The strand translocation is analyzed for the αHL pore which has an inner constriction that is 4 nucleotides in length and 1.3 nm in diameter. The voltage drop is given as a function of DNA position, m, which measures how many nucleotides have entered the pore. (C) After deriving the free energy change for the translocation as a function of m, the hopping rates are calculated for the peptide-modified region (k_1) and the unmodified region (k_0) of the DNA to account for the steric factors caused by an increase in diameter. By combining both voltage and diffusion based translocation, the predicted event duration for a specific DNA strand is calculated.

Each nucleotide segment of the DNA oligomer is characterized by a net charge, Q, and a basal hopping rate, k (Figure 4A). In the unmodified region, the charge Q1 is written as the product of the electron charge (e) and the partial charge constant q_1. The unmodified basal hopping rate, k_0,

is known from previous experiments, and describes the passage of ssDNA through the αHL pore. The charge in the chemically modified region of the nucleotide, Q2, is the product of e and the constant q_2 which accounts for the combined contributions from the ssDNA and modification. Due to the effective enlargement of the molecule diameter, the modified basal hopping rate, k_1, will be smaller than k_0.

As charge q and rate k depend on the nucleotide position of the DNA-conjugate, functions q(m) and k(m) are defined in Eq. 3 and 4, where parameter m defines how many nucleotides have entered the pore.

$$q(m)=\begin{cases} q1 * e & m < 18 \\ q2 * e & 18 \leq m \leq (17+R) \\ q1 * e & m > (17+R) \end{cases} \quad (3)$$

$$k(m)=\begin{cases} k0 & m < 18 \\ k1 & 18 \leq m \leq (17+R) \\ k0 & m > (17+R) \end{cases} \quad (4)$$

A value of m = 1 means that one nucleotide is located in the pore while the rest of the ssDNA has yet to enter. m takes values from 1 to 27, and the chemical peptide modification is set to begin at base 18 of the DNA molecule (Figure 4B). This is different to the actual covalent attachment point at nucleotide 14 but accounts for the charge-neutral and thin aliphatic linker chain. Moving the ssDNA within the αHL pore is associated with a free energy of translocation, F(m) (Figure 4C). The free energy is assumed to contain only one term[63] caused by the interaction with the applied potential. As a consequence of the small diameter of αHL's inner

constriction, the potential is assumed to drop linearly and completely over this region (Figure 4B). F(m) is calculated by summing up the energy contribution of each nucleotide segment for each position, m, as the ssDNA translocates the pore (Figure 4C). The result is given by

$$\mathrm{F(m)}=\begin{cases} \sum_{l=0}^{m-1} V(l) * q(m-l) & m \le L \\ V * \sum_{l=1}^{m-L} q(l) + \sum_{l=0}^{L-1} V(l) * q(m-l) & L < m < 27 \\ V * \sum_{l=1}^{m-L} q(l) + \sum_{l=m-27}^{L-1} V(l) * q(m-l) & m \ge 27 \end{cases} \quad (5)$$

where L is the maximum number of nucleotides that can be contained within the inner constriction; a value of L = 4 is used. The position-dependent hopping rates, k(m), represent the likelihood that the ssDNA will shift in the forward (+) or backward (-) direction (Figure 4C). An expression for the hopping rates is given by

$$\begin{aligned} k_+ &= k(m) * e^{-\alpha * \Delta(m+1)} \\ k_- &= k(m) * e^{(1-\alpha) * \Delta(m+1)} \end{aligned} \quad (6)$$

and depends on Δ(m) that gives the difference in energy between two consecutive states and is defined by

$$\Delta(\mathrm{m}) = \frac{\mathrm{F(m)-F(m-1)}}{k_b * t} \quad (7)$$

The variable α represents any possible asymmetry present in the physical transition point between two successive values of m. Previous experiments have shown that α ~ .6 for ssDNA

translocating αHL in the 5′ to 3′ direction.[64] Using the hopping rates, the total theoretical translocation times <t> can be calculated using

$$< t > (m) = \sum_{l=1}^{N-1} \sum_{i=1}^{l} \frac{w_{trans}(i)}{k_+(i)} \prod_{j=i+1}^{l} \frac{k_-(j)}{k_+(j)} - \frac{1}{w_{trans}(m)} \sum_{l=1}^{m-1} \sum_{i=1}^{l} \frac{w_{trans}(i)}{k_+(i)} \prod_{j=i+1}^{l} \frac{k_-(j)}{k_+(j)} \quad (8)$$

where the function $w_{trans}(m)$ which represents the probability of exiting the trans side is given by

$$w_{trans}(m) = \frac{1+\sum_{l=1}^{m-1} \prod_{j=1}^{l} \frac{k_-(j)}{k_+(j)}}{1+\sum_{l=1}^{N-1} \prod_{j=1}^{l} \frac{k_-(j)}{k_+(j)}} \quad (9)$$

Combining Eq. 3 – 7 and the known value $k_0 \sim 10^5\ s^{-1}$,[40] Eq. 8 can be expressed as a function solely of q_1, q_2, k_1, and the voltage. This function can now be fitted to the entire data set covering the voltage range from 100 mV to 180 mV and the arginine tag length of 3, 5 and 7. As illustrated in Figure 3B and an overall correlation coefficient, R^2, of 0.98, the fit achieves an excellent match to data of the R_7-DNA and R_5-DNA set, and a good match for R_3-DNA. Possible reasons for lower fit quality at low voltages for R_3 are discussed in the Conclusions. All fits were obtained using the following parameter values: $q_1 = 0.31$, $q_2 = -0.013$, and $k_1 = 1387\ s^{-1}$ (Table 1). The value for q_1 is in very good agreement with previously published values which range from 0.1 to 0.4.[40] The variable q_2 is of opposite charge, in line with a net positively charged DNA-arginine segment predicted from the computational simulations. Finally, k_1 is several orders of magnitude lower than k_0 ($10^5\ s^{-1}$), realistically implying that the modified region faces significantly larger steric hindrance when compared to the unmodified parts of the DNA strand.

The model was tested for the translocation of D_6-DNA. This peptide tag is similar in size to arginine but of opposite charge. It was therefore expected that the fit to the data generates a q_2 larger than q_1 because the former includes the negative charges of the D_6 peptide and nucleic acid while the latter is only for the DNA strand. Indeed, the best fit to the data (Figure 3C) resulted in a value of 0.48 for q_2 (Table 1), matching our expectation.

Further analysis shows that the biophysical model strongly supports the dual role of steric and charge in the determining the translocation time. Good fits of the model to the data are obtained only when both factors are considered; ignoring either one results in inadequate fits. For example, DNA-Arg_5 at 160 mV has an experimental translocation time of ~ 8 ms which is matched by a fit when both steric and charge are considered. If the steric effects of the modification are ignored (i.e. k_1 is the same as k_0), the best fit results in an inadequate translocation time of 0.2 ms. Similarly, accounting for the increased diameter caused by the peptide, but setting a homogeneous charge over the entire molecule to q_1, the predicted value is too low with a value of 1.5 ms. When both effects are entirely eliminated, a value of 0.1 ms is obtained which satisfactorily matches the translocation speeds of unmodified ssDNA through αHL.[49]

CONCLUSIONS

In this manuscript, we have examined how a non-homogenous charge distribution influences the electrophoretic transport of linear biopolymers through a nanopore. Using chemically tunable DNA and peptide strands, computational simulations, and single-molecule analysis of the

reference pore α-hemolysin, we have developed a powerful biophysical model which is able to disentangle electrostatic and steric factors. The model provides excellent to very good fits even though several simplifying assumptions were made. (i) The voltage was assumed to drop off linearly over the inner constriction of the αHL pore even though in reality the potential profile is likely to be more complicated.[41, 52] (ii) Only the steric interactions between the pore and the translocating molecule were taken into account, while other factors such as electrostatic influence of local charges at the pore walls were neglected. This simplified approach may be valid for large voltages but as the potential decreases, the electrostatic pore-DNA interactions could become more prominent. Indeed, the lower fit quality for the R_3-DNA at lower voltages could be accounted for by such factors. The mismatch between fit and data occurs for the R_3 but not the longer peptides likely because pure electrophoresis and sterics become less important for short tags thereby increasing the relative influence of pore interactions. (iii) Finally, our model does not explicitly consider the effects of electro-osmotic flow (EOF) which has previously been shown to be important in determining DNA translocation kinetics.[50, 65, 66] However, it is unlikely that EOF plays a major role in our experiments because the average flux of the ions inside the pore blocked by the peptide-DNA conjugate is very low.

Within these boundaries, our model has advanced the field of nanopore research and established that it is possible to separate and quantify how inhomogeneous charge and sterics along the primary sequence affect biopolymer translocation. We are confident that a refined model can be expanded and fine-tuned to biophysically analyze and aid the sensing of DNA strands with single point mutations and also of polypeptide strands which differ in size and charge at every residue position. This is an important task as the examination of proteins and

possible future sequencing of polypeptides has received a major boost.[67, 68] In order to adapt the biophysical model to polypeptides, the basal hopping rate would have to be varied along the sequence (currently it is a constant), and the free energy would have to include more terms such as the pore-biopolymer interactions. The varied hopping rate could be obtained following previous studies by extrapolating the translocation speed to zero voltage.[40] If done for homo-polymers, the corresponding diffusion coefficient could then be assembled to model to hetero-polymers under voltage following Boltzmann statistics. Similarly, the inclusion of pore-biopolymer interactions into an expanded model can benefit from previous work in this area.[64, 69] We note that our results on the charge-induced slowing down of DNA strands may also be applicable to several nanopore-based DNA strand sequencing techniques in which it is often difficult to achieve high signal-to-noise ratios for speedily translocating DNA strands. Our concept of partial charge neutralization for delayed DNA translocation could furthermore be extended to nucleic acid derivatives that carry a different partially neutral DNA backbone as well as the detection of DNA mutations.[24, 27] Finally, our results apply to the use of nanopores as model systems to study biopolymer transport through biological pores. Though *in vivo* transport is in most cases not driven by electric fields, detection in these model systems is most easily accomplished by applying an electrostatic potential, and observing ionic current blockades during translocation. Understanding the effects of electrostatic forces in these systems is therefore essential to interpretation of translocation results. In conclusion, our model and data are of interest to nanopore analytics and for the basic understanding of driven transport through confined space.

ASSOCIATED CONTENT

Supporting Information. Methodological details about the chemical synthesis of the DNA-peptide conjugates, the computational modeling of the conjugates inside the nanopore, the acquisition of nanopore data and their analysis, and the building of a biophysical model, as well as data on the chromatographic and mass spectrometric analysis of the conjugates, results on the analysis and fitting of the nanopore data, and snapshots of the computational simulations on the pore translocation of DNA-peptide bio-conjugates. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

* Corresponding authors: s.howorka@ucl.ac.uk, zsiwy@uci.edu

Phone: 0044 20 7679 4702, Fax: 0044 20 7679 7463

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. SB analyzed the nanopore data and carried out the biophysical modeling in cooperation with ZS, NM synthesized and chemically characterized the DNA-bioconjugates and analyzed nanopore recordings, HM conducted the computational simulations under the supervision of PC, MW analyzed the nanopore data to obtain characteristic time scales under the supervision of AM, and SH acquired nanopore data, supervised the project, and, together with SB, wrote the manuscript.

Funding Sources

This work has been supported by the Engineering and Physical Science Research Council, The Royal Society of Chemistry (Journals Grant), The Leverhulme Trust (RPG-170), and the Department of Chemistry at University College London. NM held a Scholarship from the UCL Graduate School. PVC and HM acknowledge funding by the 2020 Science: EPSRC programme EP/I017909/1EPSRC, FP7 funds INBIOMEDvision with grant agreement no 270107 and VPH-NoE with agreement no 223920, and XSEDE/Teragrid allocation TG-MCB090174. ZS and SB recognize the support of the National Science Foundation CHE 1306058 and UC National Lab Fee Program 12-LF- 236772

There are no competing financial interests.

ACKNOWLEDGMENT

We thank A. B. Tabor for use of a peptide synthesizer.

REFERENCES

1. Basore, J. R.; Lavrik, N. V.; Baker, L. A. Adv. Mater. 2010, 22, (25), 2759-2763.

2. Howorka, S.; Siwy, Z. Chem. Soc. Rev. 2009, 38, (8), 2360-2384.

3. Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W. Proc. Natl. Acad. Sci. U S A 1996, 93, (24), 13770-13773.

4. Bayley, H.; Cremer, P. S. Nature 2001, 413, (6852), 226-230.

5. Martin, C. R.; Siwy, Z. S. Science 2007, 317, (5836), 331-332.

6. Movileanu, L. Trends Biotechnol. 2009, 27, (6), 333-341.

7. Bezrukov, S. M.; Vodyanoy, I.; Parsegian, V. A. Nature 1994, 370, (6487), 279-281.

8. Bayley, H.; Braha, O.; Cheley, S.; Gu, L. Q., Engineered nanopores. In Nanobiotechnology: Concepts, applications and perspectives, Niemeyer, C. M.; Mirkin, C. A., Eds. Wiley-VCH: Weinheim, 2004.

9. Bayley, H.; Jayasinghe, L. Mol. Membr. Biol. 2004, 21, (4), 209-220.

10. Rhee, M.; Burns, M. A. Trends Biotechnol. 2007, 25, (4), 174-181.

11. Healy, K. Nanomed. 2007, 2, (4), 459-481.

12. Dekker, C. Nat. Nanotechnol. 2007, 2, (4), 209-215.

13. Howorka, S.; Siwy, Z., Nanopores: generation, engineering and single-molecule applications. In Handbook of single-molecule biophysics, Hinterdorfer, P.; Van Oijen, A. M., Eds. Springer: New York, 2009.

14. Healy, K.; Schiedt, B.; Morrison, A. P. Nanomed. 2007, 2, (6), 875-897.

15. Gracheva, M. E.; Melnikov, D. V.; Leburton, J. P. ACS Nano 2008, 2, (11), 2349-2355.

16. Gracheva, M. E.; Vidal, J.; Leburton, J. P. Nano Lett. 2007, 7, (6), 1717-1722.

17. Griffiths, J. Anal. Chem. 2008, 80, (1), 23-27.

18. Nakane, J.; Wiggin, M.; Marziali, A. Biophys. J. 2004, 87, (1), 615-621.

19. Tropini, C.; Marziali, A. Biophys. J. 2007, 92, (5), 1632-1637.

20. Sexton, L. T.; Horne, L. P.; Sherrill, S. A.; Bishop, G. W.; Baker, L. A.; Martin, C. R. J. Am. Chem. Soc. 2007, 129, (43), 13144-13153.

21. Branton, D.; Deamer, D. W.; Marziali, A.; Bayley, H.; Benner, S. A.; Butler, T.; Di Ventra, M.; Garaj, S.; Hibbs, A.; Huang, X. H.; Jovanovich, S. B.; Krstic, P. S.; Lindsay, S.; Ling, X. S. S.; Mastrangelo, C. H.; Meller, A.; Oliver, J. S.; Pershin, Y. V.; Ramsey, J. M.; Riehn, R.; Soni, G. V.; Tabard-Cossa, V.; Wanunu, M.; Wiggin, M.; Schloss, J. A. Nat. Biotechnol. 2008, 26, (10), 1146-1153.

22. Clarke, J.; Wu, H. C.; Jayasinghe, L.; Patel, A.; Reid, S.; Bayley, H. Nat. Nanotechnol. 2009, 4, (4), 265-270.

23. Marziali, A.; Akeson, M. Annu. Rev. Biomed. Eng. 2001, 3, 195-223.

24. An, N.; Fleming, A. M.; White, H. S.; Burrows, C. J. Proc. Natl. Acad. Sci. USA 2012, 109, (29), 11504-11509.

25. Wanunu, M.; Dadosh, T.; Ray, V.; Jin, J.; McReynolds, L.; Drndic, M. Nat. Nanotechnol. 2010, 5, 807-814.

26. Wang, Y.; Zheng, D.; Tan, Q.; Wang, M. X.; Gu, L. Q. Nat. Nanotechnol. 2011, 6, (10), 668-674.

27. Wanunu, M.; Cohen-Karni, D.; Johnson, R. R.; Fields, L.; Benner, J.; Peterman, N.; Zheng, Y.; Klein, M. L.; Drndic, M. J. Am. Chem. Soc. 2011, 133, (3), 486-492.

28. Wanunu, M.; Soni, G. V.; Meller, A., Single-molecule studies on nucleic acid interactions using nanopores. In Handbook of Single-Molecule Biophysics, Hinterdorfer, P.; Van Oijen, A. M., Eds. Springer: New York, 2010; pp 265-291.

29. Vercoutere, W.; Akeson, M. Curr. Opin. Chem. Biol. 2002, 6, (6), 816-822.

30. Deamer, D. W.; Branton, D. Acc. Chem. Res. 2002, 35, (10), 817-825.

31. Henrickson, S. E.; Misakian, M.; Robertson, B.; Kasianowicz, J. J. Phys. Rev. Lett. 2000, 85, (14), 3057-3060.

32. Meller, A.; Branton, D. Electrophoresis 2002, 23, (16), 2583-2591.

33. Nakane, J.; Akeson, M.; Marziali, A. Electrophoresis 2002, 23, (16), 2592-2601.

34. Maglia, G.; Restrepo, M. R.; Mikhailova, E.; Bayley, H. Proc. Natl. Acad. Sci. U S A 2008, 105, (50), 19720-19725.

35. Butler, T. Z.; Gundlach, J. H.; Troll, M. A. Biophys. J. 2006, 90, (1), 190-199.

36. Wang, H.; Dunning, J. E.; Huang, A. P.; Nyamwanda, J. A.; Branton, D. Proc. Natl. Acad. Sci. U S A 2004, 101, (37), 13472-13477.

37. Mathe, J.; Aksimentiev, A.; Nelson, D. R.; Schulten, K.; Meller, A. Proc. Natl. Acad. Sci. U S A 2005, 102, (35), 12377-12382.

38. Wanunu, M.; Chakrabarti, B.; Mathe, J.; Nelson, D. R.; Meller, A. Phys. Rev. E 2008, 77, (3), 031904 (031901-031905).

39. Wiggin, M.; Tropini, C.; Tabard-Cossa, V.; Jetha, N. N.; Marziali, A. Biophys. J. 2008, 95, (11), 5317-5323.

40. Meller, A.; Nivon, L.; Branton, D. Phys. Rev. Lett. 2001, 86, (15), 3435-3438.

41. Howorka, S.; Bayley, H. Biophys. J. 2002, 83, (6), 3202-3210.

42. Sutherland, T. C.; Long, Y. T.; Stefureac, R.; Bediako-Amoa, I.; Kraatz, H. B.; Lee, J. S. Nano Lett. 2004, 4, (7), 1273–1277.

43. Movileanu, L.; Schmittschmitt, J. P.; Scholtz, J. M.; Bayley, H. Biophys. J. 2005, 89, (2), 1030-1045.

44. Stefureac, R.; Long, Y. T.; Kraatz, H. B.; Howard, P.; Lee, J. S. Biochemistry 2006, 45, (30), 9172-9179.

45. Goodrich, C. P.; Kirmizialtin, S.; Huyghues-Despointes, B. M.; Zhu, A.; Scholtz, J. M.; Makarov, D. E.; Movileanu, L. J. Phys. Chem. B 2007, 111, (13), 3332-3335.

46. Huang, L.; Kirmizialtin, S.; Makarov, D. E. J. Phys. Chem. B 2005, 123, (12), 124903.

47. Oukhaled, G.; Mathé, J.; Biance, A. L.; Bacri, L.; Betton, J. M.; Lairez, D.; Pelta, J.; Auvray, L. Phys. Rev. Lett. 2007, 98, (15), 158101.

48. Talaga, D. S.; Li, J. J. Am. Chem. Soc. 2009, 131, (26), 9287–9297.

49. Mitchell, N.; Howorka, S. Angew. Chem. Int. Ed. 2008, 47, (30), 5476-5479.

50. Firnkes, M.; Pedone, D.; Knezevic, J.; Döblinger, M.; Rant, U. Nano Lett. 2010, 10, (6), 2162-2167.

51. Song, L.; Hobaugh, M. R.; Shustak, C.; Cheley, S.; Bayley, H.; Gouaux, J. E. Science 1996, 274, (5294), 1859-1866.

52. Aksimentiev, A.; Schulten, K. Biophys. J. 2005, 88, (6), 3745-3761.

53. Kathawalla, I. A.; Anderson, J. L.; Lindsey, J. S. Macromolecules 1989, 22, (3), 1215-1219.

54. Martin, H.; Jha, S.; Howorka, S.; Coveney, P. J. Chem. Theory Comput. 2009, 5, (8), 1955-2192.

55. Maffeo, C.; Bhattacharya, S.; Yoo, J.; Wells, D.; Aksimentiev, A. Chem. Rev. 2012, 112, (12), 6250-6284.

56. Cantor, C. R.; Schimmel, P. R., Biophysical chemistry. Part III. The behavior of biological macromolecules. W. H. Freeman and Co.: New York, 1980; p 1283.

57. Puglisi, J. D.; Chen, L.; Blanchard, S.; Frankel, A. D. Science 1995, 270, (5239), 1200-1203.

58. Howorka, S.; Movileanu, L.; Braha, O.; Bayley, H. Proc. Natl. Acad. Sci. U S A 2001, 98, (23), 12996-13001.

59. Howorka, S.; Cheley, S.; Bayley, H. Nat. Biotechnol. 2001, 19, (7), 636-639.

60. The mid-blockade level of type II events (58% of all) was noisier and longer (around 10-20 ms) than Rn. For these events, the negatively charged D6 peptide is most probably threaded into the inner constriction of αHL; the resulting DNA hairpin-like loop cannot pass the narrow inner constriction until it retracts so that one DNA terminus threads first.

61. Berberan-Santos, M. N.; Bodunov, E. N.; Valeur, B. Chem. Phys. 2005, 316, 171-182.

62. Ho, C.; Qiao, R.; Heng, J. B.; Chatterjee, A.; Timp, R. J. Proc. Natl. Acad. Sci. U S A 2005, 102, (30), 10455-10450.

63. Schink, S.; Renner, S.; Alim, K.; Arnaut, V.; Simmel, F. C.; Gerland, U. Biophys. J. 2012, 102, (1), 85-95.

64. Lathrop, D. K.; Ervin, E. N.; Barrall, G. A.; Keehan, M. G.; Kawano, R.; Krupka, M. A.; White, H. S.; Hibbs, A. H. J. Am. Chem. Soc. 2010, 132, (6), 1878-1885.

65. Luan, B. Q.; Aksimentiev, A. Phys. Rev. E 2008, 78, (2), 021912.

66. van Dorp, W. F.; Someren, B.; Hagen, C. W.; Kruit, P.; Crozier, P. A. Nano Lett. 2005, 5, 1303-1307.

67. Nivala, J.; Marks, D. B.; Akeson, M. Nat. Biotechnol. 2013, 31, (3), 247-250.

68. Rodriguez-Larrea, D.; Bayley, H. Nat. Nanotechnol. 2013, 8, (4), 288-295.

69. Lubensky, D. K.; Nelson, D. R. Biophys. J. 1999, 77, (4), 1824-1838.