Tube Translocation of a Herniating Chain

S. Żurek, A. Drzewiński

Institute of Physics,
University of Zielona Góra, ul. Prof. Z. Szafrana 4a,
65-516 Zielona Góra, Poland
e-mail: S.Zurek@if.uz.zgora.pl; A.Drzewinski@proton.if.uz.zgora.pl

Abstract: Translocation is modeled with the Rubinstein-Duke rules for hopping reptons along a one-dimensional lattice (tube). Identification of some coupled states guarantees a semi-periodicity of the process. The chain is driven through the pore by a bias potential promoting the transition of stored length in one direction. Accounting exactly for all allowed states the translocation time of polymers up to the 10-links length is determined. The crossover from reptation to faster dynamics through gradually allowing hernia creation and annihilation is found. Some computational details are presented.
Key words: Rubinstein-Duke model, translocation, semi-periodic stochastic model

I. INTRODUCTION

Translocation of polymers across a narrow pore is ubiquitous in biology. Examples include the motion of DNA and RNA molecules across nuclear pores, the injection of viral DNA into a host cell, translocation of protein from cis to trans side of a membrane through channels [1]. Also in the industry, processes such as the separation and purification of synthetic and biological macromolecules, the recovery and separation of oil, the production of foods and drugs are related to the translocation of polymer chains.

Numerous experiments [2, 3] aimed at the translocation process examination were performed: the ionic current blocade duration in single-channel conductance measurements and the time distributions for a single-stranded DNA polymers during voltage-driven translocations were measured. In parallel to these experimental investigations, theoretical and computational studies (such as Monte Carlo, kinetic Monte Carlo, Langevin dynamics methods) aimed to model the translocation process and imitate the experimental results report only moderate success. For example, it has been shown recently that the forced translocation model using MC with basic Metropolis sampling gives the correct time dependence at moderate pore potentials but presents an artifact at large pore potentials [4].

This paper reports the results of the *exact* calculations of a polymer driven through the pore with the Rubinstein-Duke (RD) rules. The computational complexity limits the accessible chain length, but the details of the phenomena (repton-repton or repton-wall interactions, where a repton corresponds to one persistence length of the chain) investigated even for a short chains are interesting and give some additional insight into the translocation process.

II. MODEL

The 3D translocation phenomena is for simplicity considered here as an one-dimensional process. One of the justifications is that hernia creation and annihilation plays the same role in one-dimensional RD models as the sideways motion in the two- and three-dimensional versions [5].

Identical reptons building the chain are characterised by $N + 1$ position variables $x_0, ..., x_N$ (*self-avoidance* is neglected as it would make the one-dimensional version trivial). The two successive positions x_i and x_{i+1} are either the same or a nearest-neighbour distance apart. We call the difference $y_i = x_i - x_{i-1}$ a link. If the two successive reptons $i - 1$ and i are located in the same cell, the link is zero or *slack*. Otherwise, it is a nearest-neighbour distance and the link is *taut* one. The links determine the configuration of the chain. So, the configurations of the chain are given by vector $y = (y_1, ..., y_N)$. The number of configurations is 3^N, which is an enormous number for long polymers. In our problem the number of configurations is diminished by the presence of a wall preventing a chain from multipassing the

pore. A sketch of the chain in the lattice is given in the upper part of Fig. 1. The slack links represent elements of stored length, which are the basic ingredients for reptation. To accelerate the process, we allow the motion to be more general than reptation: when both links are slack, the middle repton can hop to any of its two neighbouring cells, thereby changing the slack links in a pair of opposite taut links, which is the creation of a hernia. The reverse process is a hernia annihilation. We give these hopping rates value h. When any repton moves through the pore, there is a bias which has the form of a Boltzmann factor $B = \exp(\epsilon/2)$ involving the energy difference between both sides, favouring jumps aligned with the direction of the external electric field E and discouraging jumps against the field direction[1].

Our model is, as all the hopping models, governed by the Master equation:

$$\frac{dP(y, t)}{dt} = \sum_{y'} \left[W\langle y | y' \rangle P(y', t) - W\langle y' | y \rangle P(y, t) \right] = \tag{1}$$
$$= \sum_{y'} H_{yy'} P(y', t),$$

where $P(y, t)$ indicates the probability of finding the polymer in a configuration y at time t and $W(y'|y)$ is a transition rate per unit of time from configuration y to configuration y'. Matrix H contains both the gain and loss terms and is stochastic in the sense that the sum over all columns vanishes, as required from the conservation of probability.
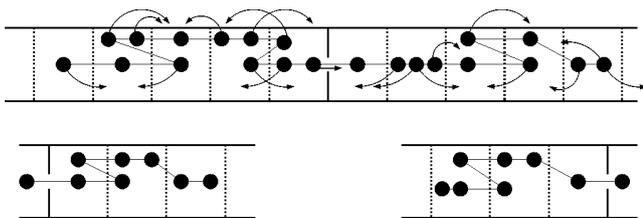

Fig. 1. Upper part: a picture of the polymer chain consisting of reptons and allowed repton motions. Bottom part: an example of coupled configurations ($\Delta X_i = 3$ here)

When a process is stationary, the time derivative at the left hand side of the Master Equation vanishes and a stationary state will be reached (corresponding to the zero eigenvalue of the Master Equation). Moreover, the elements of the stationary state are the probabilities $P(y_i)$. In

fact, the translocation is not a stationary process but can be modelled by a semiperiodic model by identification of some coupled states.

III. METHOD

Interchanges of slack and taut links, according to the Rubinstein rules, are the main moves in the model. Stored length moves through the pore (represented by a taut link) only when the slack link interchanges, as in a RD move, with the pore (the $W(y'|y)$ transition rates are then set to B if the move is aligned with the field direction and $1/B$ when the chain threads through the pore in a reverse direction). The pore is considered as fixed in space, but its position with respect to the numbering of the chain links changes. When the pore link is at the end of the chain, it may interchange with a slack link at the other end – this is the identification of (coupled) configurations (see Fig. 1) that keeps the process to run around forever. This interchange is done again with the bias B from left to right and with $1/B$ in the opposite direction. On the other hand, as long as the pore is not at an end of the chain, the end reptons move freely, thereby refreshing the tube configuration. By allowing hernia creation and annihilation, the motion is not reptative but crosses over to faster dynamics. The origin is that taut links are blocked by the pore. They have to be first converted into slack links by hernia annihilation and then pass through the pore as slack links and may be recreated at the other side by hernia creation. It is clear that the hernia creation and annihilation plays an important role in this translocation. So, it may be anticipated that the rate of hernia creation and annihilation functions as a cross-over parameter.

Drift velocity v (the velocity of the reptons constituting the chain) is the same for all reptons on the average (the formal expression for v is given by [7]). As we are interested in the translocation time, we need a quantity corresponding to the displacement along the x axis. The natural measure is a distance between the chain ends (ΔX) because when the pore link interchanges with a slack link at the other end, it is equivalent to the motion of the whole chain along the x axis on the same distance. Therefore, we can define

$$\langle \Delta X \rangle = \frac{\left(\sum_{\tilde{y}_i} w P(y_i) \Delta X_i \right)}{\sum_{\tilde{y}_i} w P(y_i)}, \tag{2}$$

[1] The reduced field $\epsilon = qEa/k_{\mathrm{B}}T$ is a dimensionless parameter, where the numerator represents the energy needed to move a repton a distance a against the electric force acting qE on it [6].

where $w = B$ or $1/B$ and a tilde means summing only over configurations with the pore link at an end on the condition that the opposite end link is taut. An example of such coupled configurations is presented in the bottom part of Fig. 1. Now we can define the translocation time as:

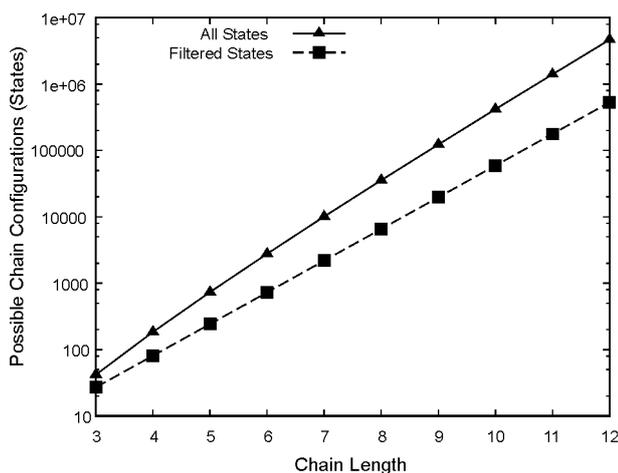$$t = \frac{\langle \Delta X \rangle}{\langle v \rangle} . \tag{3}$$

Fig. 2. The plot visualises the enormous gain of possible chain states along with the number of links that compose the chain (notice the semi-logarithmic scale) as well as the influence of the presence of the "wall" to the reduction of processed states

3.1. Computational details

The determination of translocation time t (Eq. 3) of a chain that is built up with N links is a nontrivial task. The enormous growth of the possible chain states (configurations) along with chain length (see Fig. 2) makes data processing a challenging problem.

In order to calculate the exact value of the translocation time of a chain that threads through the pore, one has to construct and evaluate the stochastic matrix H defined by elements that describe the probability of transition between the states of a chain (see Eq. 1). This process involves the following general steps:

A – constructing all possible states for a chain,
B – processing those states to exclude the forbidden ones (the wall condition),
C – processing the "proper" states to determine the transitions linking them,
D – constructing the sparse stochastic matrix,
E – calculating the matrix null-space and dependent quantities.

The steps listed above could be seen as *two abstract processes:* the system preparation (A, B, C) and the system evaluation/examination (D, E). While the former process goal is to create and store the general (symbolic) system information, the latter is intended to apply specific numerical values to the output of the former one and to perform algebraic calculations. The two processes meet at the data storage level which is driven by a SQL database engine.
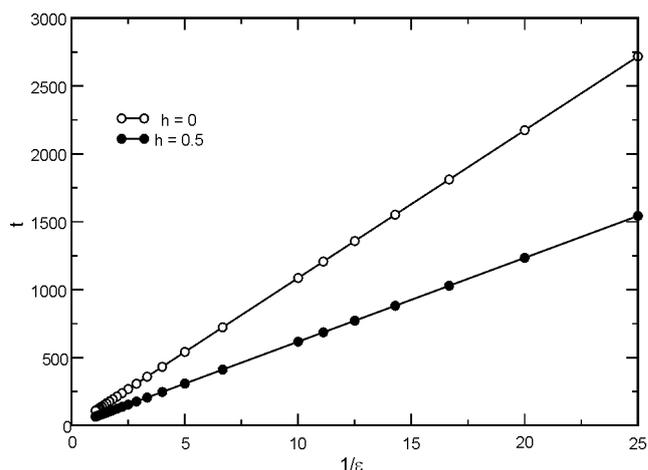

Fig. 3. The log-log dependence of the translocation time with respect to reduced field ϵ that drives the polimer motion through the pore. The calculations are done for the 8-link chain

The construction of possible chain states is performed with the tree approach, by a recursive links pairing (i.e. *taut* and *slack* states) that define the specific chain configuration. The outgoing configurations are next validated to check the "wall condition" – invalid states are excluded and the resulting states are saved in the SQL database to avoid further repetitive calculations. The validation procedure is implemented as a parallel process with OpenMPI [8] cluster interface: the chain configurations set is distributed to the computer cluster, each MPI process analyses the subset of the configurations and (with parallel access to the SQL server) stores the results. The results of such a filtering are presented in Fig. 2. To construct the stochastic matrix, the allowed states must be next related with each other by a *transition definition*. Each transition (i.e. hernia creation/annihilation, driven motion through a pore, thermal motion) can be assigned a specific value of probability rate that will be used to create the stochastic matrix cell value. The found transitions are next stored in the SQL database in a generic manner, i.e. the probability rate is encoded with a symbolic value to allow further flexible system evaluation and analysis. Each transition combines the two chain states.

When each transition has its symbolic probability rate substituted by a real number, the stochastic matrix is defined. This is a sparse matrix that we process using LU factorization (by means of the supernodal sparse LU factorization package SuperLU [9] embedded in the PySparse Python module) to calculate its *null-space*. The out-coming information describes the states probability, i.e. each eigenvector component value is directly related to the state probability. Eigenvector values combined with transitions information are next used to calculate translocation time according to the Eq. 3.
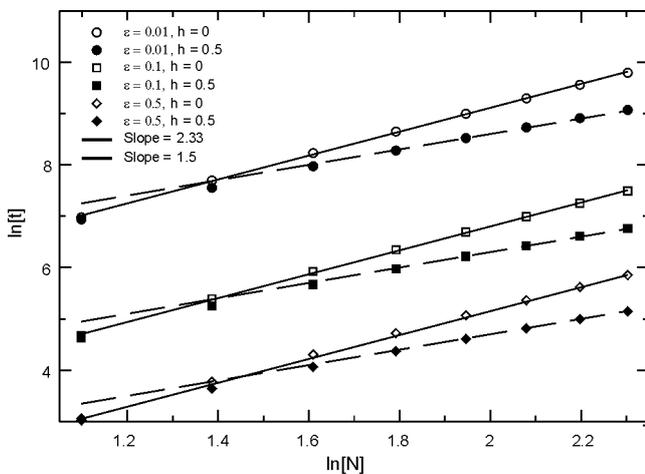

Fig. 4. The log-log dependence of the translocation time with respect to the number of links for different values of the reduced field ϵ. The $h = 0$ case corresponds to the pure reptation, whereas $h = 0.5$ is connected to the presence of hernia creations/annihilations

IV. RESULTS

First we determined the dependence of the translocation time on the field strength. Our results presented in Fig. 3 confirm the results of Kasianowicz et al. [2] that the channel blockade life-time is inversely proportional to the applied voltage.

In order to determine the N dependence of the translocation time, we performed a series of calculations on polymers with a length up to 10 at a various field strength. In order to check the influence of the hernia creation/annihilation, we assumed two values, zero and non-zero, for h.

Although our chains are too short to give concluding results, the plots presented in Fig. 4 indicate at substantial difference between the translocation time behaviour in the presence and in the absence of hernias.

V. CONCLUSIONS

The original translocation process has been formulated as a semi-periodic stochastic model with the Rubinstein-Duke rules for hopping reptons. The dynamics of short polymers were investigated using the *exact* calcucations, i.e. all allowed polymer states were used to build and evaluate the matrix H (see. Eq. 1). The crossover from reptation to faster dynamics has been found and the rate of hernia creation and annihilation is a cross-over parameter.

Acknowledgements

References

[1] B. Albert, D. Bray, J. Lewis, M. Raff, K. Roberts, J.D. Watson, *Molecular Biology of the Cell.* (Garland, New York, 1994).

[2] J.J. Kasianowicz, E. Brandin, D. Branton, D.W. Deamer, *Characterization of individual polynucleotide molecules using a membrane channel.* Proc. Natl. Acad. Sci. USA 93, 13770 (1996).

[3] A. Meller, L. Nivon, D. Branton, *Voltage-Driven DNA Translocations through a Nanopore.* Phys. Rev. Lett. 86, 3435 (2001).

[4] V.V. Lehtola, R.P. Linna, K. Kaski, *Critical evaluation of the computational methods used in the forced polymer translocation.* Phys. Rev. E78, 061803 (2008).

[5] A. Drzewiński, J.M.J. van Leeuwen, *Crossover from reptation to Rouse dynamics in a one-dimensional model*, Phys. Rev. E73, 061802 (2006).

[6] B. Widom, J.L. Viovy, A.D. Defontaines, *Repton model of gel electrophoresis and diffusion.* J. Phys. I France 1 (1991).

[7] J.M.J. van Leeuwen, A. Drzewiński, *Stochastic Lattice Models for the Dynamics of Linear Polymers.* Phys. Rep. 475, 53 (2009).

[8] E. Gabriel, G.E. Fagg, G. Bosilca, T. Angskun et al., *Open MPI: Goals, Concept, and Design of a Next Generation MPI Implementation.* Proceedings of the 11th European PVM/MPI Users' Group Meeting, 2004.

[9] X.S. Li, J.W. Demmel, *SuperLU DIST: A Scalable Distributed-Memory Sparse Direct Solver for Unsymmetric Linear Systems.* ACM Trans. Mathematical Software 29, 2 (2003).

MSC. SEBASTIAN ŻUREK works at the Institute of Physics of the University of Zielona Góra, Poland. His main areas of interest include the application of methods of computational physics and high-performance computing to the complex problems in physics and medicine, e.g. the heart rate variability analysis using nonlinear dynamics methods, a polymer translocation through pores, an analysis of ESR spectra with artificial intelligence methods, and molecular dynamics simulations.

DSC. ANDRZEJ DRZEWIŃSKI is Associate Professor at the Institute of Physics of the University of Zielona Góra, Poland. One of his current research is the modeling of the dynamics of polymers in the gel or liquid phase by means of stochastic processes. These studies were inspired by two physical phenomena often considered in the biological sciences: electrophoresis and translocation. Another subject of his research work are liquid-gas systems in confining geometries, closely related to microfluidic applications. The molecular magnetism is the next field, in particular some materials exhibiting extremely slow paramagnetic relaxation of their magnetization at low temperatures. The latter are potentially useful for magnetic media of very high density recording.