Anomalous scaling in nanopore translocation of structured heteropolymers

Malcolm McCauley[1], Robert Forties[1], Ulrich Gerland[2] and Ralf Bundschuh[3]

[1]Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, Ohio 43210-1117, USA

[2]Arnold-Sommerfeld Center for Theoretical Physics and Center for Nanoscience (CeNS), LMU München, Theresienstrasse 37, 80333 München, Germany

E-mail: gerland@lmu.de

[3]Departments of Physics and Biochemistry and Center for RNA Biology, The Ohio State University, 191 West Woodruff Avenue, Columbus, Ohio 43210-1117, USA

E-mail: bundschuh@mps.ohio-state.edu

Abstract. Translocation through nanopores has emerged as a new experimental technique to probe physical properties of biomolecules. The question of how the typical translocation time for a single unstructured polymer depends on its length has already triggered many theoretical and computational studies. Here, we address the same question, but for structured RNA molecules where the breaking of base-pairing patterns is the main barrier for translocation. Within a simple model, we calculate the typical time for single-stranded RNAs with random sequences to translocate through an idealized nanopore. It is believed that with respect to secondary structure formation, such a random RNA is typically either frozen into a single dominant structure (glassy phase) or many different structures with similar total energies coexist (molten phase), depending on the temperature and the base-pairing energetics. We find that these two phases can be clearly distinguished by their translocation behavior in the absence of an external voltage bias. In both cases, the typical translocation time depends on the sequence length as a power law with an exponent that exceeds the naively expected exponent of two for purely diffusive translocation. However, whereas in the molten phase the exponent is constant, at a value of 5/2, the exponent increases rapidly with decreasing temperature in the glassy phase. We explain the behavior in the molten phase and the qualitative trend in the glassy phase theoretically.

PACS numbers: 87.15.ak, 87.15.La, 87.14.G-, 87.15.bd

Submitted to: *Phys. Biol.*

Keywords: nanopores, RNA, translocation

1. Introduction

Nanopore technology has opened a completely new window for probing the properties of polymers in general and biopolymers in particular [1, 2, 3]. In a nanopore setup two macroscopic chambers filled with a buffer solution are separated from each other by a wall. Embedded into this wall is a single nanopore, i.e., a hole with a diameter in the few nanometer range, connecting the two chambers. When charged polymers are added into one chamber, an electric field applied across the nanopore can drive these polymers through the pore one by one. Drops in the induced counter ion current due to the occlusion of the pore by the translocating polymer allow the translocation dynamics of individual polymers to be observed. In recent years, this technique has been applied extensively to study DNA [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11] and RNA [12] molecules as well as proteins [13, 14].

The emergence of this new experimental technique has also spurred a lot of activity on the theoretical side. There has been particular interest in understanding the nonequilibrium statistical mechanics associated with the translocation of unstructured, linear polymers, e.g., single-stranded DNA in which all nucleotides are the same [15, 16, 17, 18, 19, 20, 21, 22, 23]. The quantities of interest are the (experimentally measurable) distribution of translocation times, and the asymptotic behavior of the typical translocation time as the polymers become very long.

On the simplest level of description, the translocation of a linear polymer is hindered by an entropic barrier. An entropic barrier emerges since the wall separating the two chambers effectively separates the polymer into two sections: the *trans* section which has already translocated and the *cis* section which yet has to translocate. Each of these sections is constrained in its motion by the wall, and the constraint is most severe when the polymer has translocated half way through the pore. More quantitatively, if a polymer with sequence length N is divided into sections of length m and $N-m$, respectively, the total number of configurations available to this polymer is reduced (compared to a free polymer) by the power law factors $m^{-\gamma_u}$ and $(N-m)^{-\gamma_u}$ [24]. Here, the exponent γ_u depends on the asymptotic statistical properties of the polymer that are affected only by the spatial dimensionality and a possible self-avoidance interaction ($\gamma_u = 1/2$ for an ideal, noninteracting chain). As a consequence, the entropic barrier experienced by the translocating polymer (i.e., the difference in free energy between a polymer that just entered the pore ($m = 0$) and a polymer with m bases on the *trans* side) has the shape

$$F(m) = \gamma\, k_B T\, \ln[(N-m)m/N] \tag{1}$$

with $\gamma = \gamma_u \lessapprox 1/2$ (up to some irrelevant finite size regularization at $m \approx 0$ and $m \approx N$). The maximum of this barrier at $m = N/2$ depends logarithmically on N, with γ_u as a prefactor. Modeling the translocation process as a one-dimensional diffusion across this entropic barrier is an appropriate description, if the translocation process is adiabatically slow, e.g. due to friction at the pore, such that the polymer ends on each side sample many different configurations during the time required to translocate

a macroscopic portion of the polymer. It has been established that if entropy reduction is the only barrier, translocation is purely diffusive (i.e., the translocation times scale as N^2) in the limit of zero voltage and ballistic (i.e., the translocation times scale as N for long polymers) at finite voltages independent of the characteristics of the polymer model (i.e., independent of the precise value of the exponent γ_u) [23]. However, there is still an ongoing debate what effect the actual polymer dynamics may have on the translocation time distributions under conditions where the adiabatic approximation breaks down such that the polymer dynamics is directly coupled to the translocation dynamics [18, 21, 25, 26, 27, 28, 29].

Here, we focus on a different, but similarly challenging theoretical question, which arises when the translocating polymers are structured heteropolymers. This issue has obtained some experimental [5, 6, 7] and theoretical [7, 30, 31] attention but far less than the case of unstructured molecules. In particular, we consider a polynucleotide, an RNA or a single-stranded DNA, consisting of a specific sequence of individual nucleotides, i.e., A, C, G, and U for the case of RNA. For simplicity, we will loosely use 'RNA' to refer to both RNA and single-stranded DNA in this article, as the biochemical difference between these polynucleotides is insignificant for the questions we address. RNA molecules have a strong propensity to form intramolecular Watson-Crick, i.e., G–C and A–U, base pairs. The formation of such base pairs forces the molecules to fold into sequence-dependent structures, which are characterized by their base-pairing pattern. The naturally evolved sequences of structural RNA's, e.g. ribosomal RNA, are biased to stably fold into particular, functional structures, whereas the sequences of many other RNA's, e.g. most messenger RNA's, primarily encode information, not structure. The structural features of this latter class can be modelled via the ensemble of random RNA sequences [32, 33, 34]. Here, we characterize the translocation dynamics of this class, focusing on the slow translocation limit. We identify nontrivial translocation behavior, and study the physical origin of this behavior.

Even with a random sequence, a single RNA molecule may spend most of the time in a dominant base-pairing pattern ('glassy behavior' [32]). Or else it may sample a promiscuous array of alternative structures with different shapes [35]. The transition between these two types of behavior occurs as a function of temperature, with low temperatures favoring glassy behavior [34, 36, 37]. It is interesting to ask whether this transition is reflected also in the translocation behavior, and if so, how?

Generally, if a folded molecule is to translocate through a very narrow pore that allows only single strands to pass, it has to break its base pairs in the process. This yields a coupling between the observed translocation dynamics and the base-pairing properties of the molecule [31]. In this system, the separation of the polymer into a *cis* and a *trans* section has an additional effect, namely that bases on each side of the pore can only pair with bases on the same side of the pore thus limiting the possible pairing partners. On average, this restriction in the base-pairing pattern again is believed to lead to a free energy barrier that is logarithmic in the length of the polymer (see below and [34]). Thus, at least superficially, the problem of a structured RNA molecule translocating through a

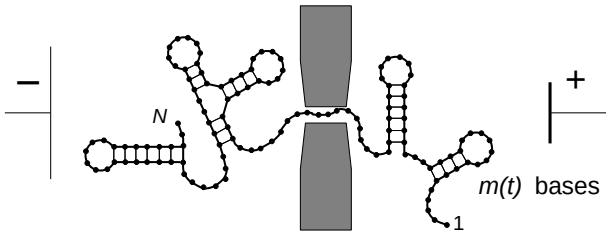

Figure 1. Sketch of a structured RNA molecule translocating through a narrow pore, which allows single but not double strands to pass. Translocation can be driven by an applied voltage acting on the negative charges of the RNA backbone. An appropriate reaction coordinate for the translocation process is the number of bases m that have reached the *trans* side. If the translocation is sufficiently slow, for instance due to molecular friction at the pore or energetic barriers caused by base-pairing, m becomes the only relevant degree of freedom. In this slow translocation limit, there is sufficient time for the base-pairing patterns on the *cis* and *trans* sides to reoptimize whenever m changes.

nanopore is mathematically similar to the problem of homopolymer translocation, even though the physical origin of the logarithmic barrier is completely different in nature. However, the problem is deeper than this analogy suggests: while the logarithmic barrier is insignificant for the translocation of homopolymers (see above), we will see below that for structured heteropolymers the translocation dynamics is drastically affected. This is a consequence of the fact that in the structured case, the prefactor γ of the logarithmic barrier is both bigger in magnitude (such that it exceeds a critical threshold) and dependent on temperature.

The rest of this manuscript is organized as follows. In Sec. 2, we lay out our model assumptions and the general theoretical framework used here to describe the translocation dynamics, review the relevant aspects of the statistical physics of RNA folding, and then link the folding and translocation characteristics of random RNA. In Sec. 3 we first explore the translocation dynamics of random RNA sequences numerically, and identify an anomalous scaling of the typical translocation time with the length of the RNA. Then, we provide some theoretical insight into the origin of this anomalous scaling in the discussion. Sec. 4 summarizes our results and provides an outlook to future work.

2. Materials and methods

2.1. Translocation dynamics: general framework

As illustrated in figure 1, we consider a polynucleotide translocating from the *cis* to the *trans* side of a pore in a membrane. The pore is so narrow that only a single strand of the polynucleotide can pass through, and hence only unpaired bases can enter the pore. If an external electric voltage V is applied across the pore, translocation is biased towards the positive terminal, since RNA has a negatively charged backbone. The translocation process has a natural "reaction coordinate": the number of bases

$m(t)$ that have reached the *trans* side at time t. For simplicity we will consider an ideal pore with a negligible depth, i.e., we assume that the remaining $N - m$ bases are all exposed on the *cis* side and none reside within the pore. In general, the translocation process cannot be described solely by a dynamic equation for the coordinate m, since the spatial and base-pairing degrees of freedom of the polymer are coupled to $m(t)$, see [31]. However, under conditions where the translocation process is sufficiently slow, the translocation dynamics becomes effectively one-dimensional, as $m(t)$ reduces to a stochastic hopping process in an appropriate one-dimensional free energy landscape $F(m)$. Such a description is appropriate if the base-pairing patterns on the *cis* and *trans* sides have sufficient time to reoptimize whenever m changes. Slow translocation arises when the molecular friction at the pore is large, the voltage bias V is small, and the energetic barriers due to base-pairing are significant. Throughout the present paper, we focus entirely on this slow translocation limit.

With the above assumptions, the stochastic translocation process is described by a master equation for $P(m,t)$, the probability to find an RNA molecule with a given sequence in a state with m bases on the *trans* side at time t. This master equation takes the general form

$$\begin{aligned}\partial_t P(m,t) = {} & k_+(m-1)\,P(m-1,t) + \\ & + k_-(m+1)\,P(m+1,t) + \\ & - [k_+(m) + k_-(m)]\,P(m,t)\end{aligned} \tag{2}$$

with a set of "hopping" rates $k_+(m)$ and $k_-(m)$ that depend explicitly on the translocation coordinate m. Here, $k_+(m)$ is the rate to translocate the base with index $m+1$ from the *cis* to the *trans* side, whereas $k_-(m)$ is the rate at which base m translocates back from the *trans* to the *cis* side. The hopping rates also depend on the voltage bias V, the temperature T, and the nucleotide sequence of the RNA. In other words, at a given voltage bias and temperature, we need to obtain a set of $2N$ hopping rates for each RNA sequence such that (2) describes the translocation dynamics. We then want to characterize the translocation behavior for the ensemble of random sequences.

If the m-dependence of the hopping rates is dropped, $k_+(m) \equiv k_+$ and $k_-(m) \equiv k_-$, (2) describes a homogeneous drift-diffusion process and becomes equivalent to the Fokker-Planck equation

$$\partial_t P(x,t) = D\,\partial_x^2 P(x,t) - v\,\partial_x P(x,t) \tag{3}$$

in the continuum limit, where m is replaced by a continuous reaction coordinate $0 < x < N$. Here, D and v are the effective diffusion constant and drift velocity, respectively. As was shown by Lubensky and Nelson [38], past translocation experiments with unstructured single-stranded polynucleotides are quantitatively consistent with (3): The experimental distribution of translocation times $p(\tau)$ is well described by the corresponding distribution from (3), which is determined by the probability current into the absorbing boundary at $x = N$.

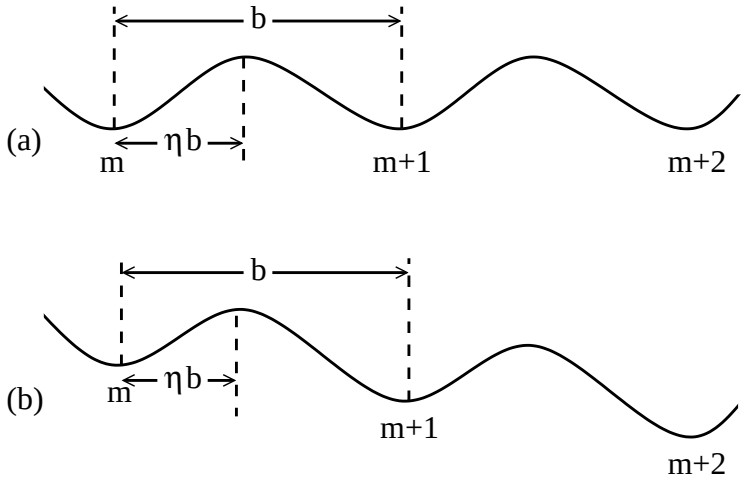

Figure 2. Illustration of the voltage-dependence of the translocation rates in (4) of the main text. Even in the absence of secondary structure, the translocation of a single base is envisaged as a barrier crossing process. The coarse-grained, discrete reaction coordinate m (the number of translocated bases) then corresponds to the minima of a continuous microscopic free energy landscape. The distance of the minima reflects the base-to-base distance b of the RNA. The position of the transition state, at a fractional distance η from the minimum to the left, is an unknown microscopic parameter which determines how the biasing effect of the applied voltage is split between the forward and reverse translocation rates: When the unbiased landscape of (a) is tilted by the applied voltage as shown in (b), the reduction in the free energy barrier for forward translocation is proportional to η, while the increase in the barrier for reverse translocation is proportional to $(1-\eta)$.

For structured RNA's, we express the hopping rates of (2) more explicitly in the form

$$k_+(m) = k_0 \cdot w_{cis}(m) \cdot \exp\left(\eta \, \frac{q_{\text{eff}} V}{k_B T}\right) \tag{4}$$
$$k_-(m) = k_0 \cdot w_{trans}(m) \cdot \exp\left((\eta - 1) \, \frac{q_{\text{eff}} V}{k_B T}\right) \ .$$

Here k_0 denotes the basic "attempt" rate for the translocation of a single unpaired base, while $w_{cis}(m)$ and $w_{trans}(m)$ denote the probability that the base attempting to translocate is indeed not paired. The exponential (Arrhenius) factors account for the voltage bias V across the pore, which acts on the effective charge q_{eff} of a nucleotide [9, 39] (note that the applied voltage drops primarily directly across the pore, while the nucleotides do not experience a significant electrostatic force on either side). The dimensionless factor η is a measure for the position of the microscopic transition state that limits the rate for the crossing of a single nucleotide. More precisely, η is the relative distance of this transition state from the entrance of the pore; for a symmetric pore, $\eta = 1/2$ (see figure 2).

For a fixed but arbitrary set of hopping rates $k_+(m)$, $k_-(m)$, the (thermal) average of the translocation time can be calculated analytically using the mean first passage

time formalism [40]. One obtains

$$\langle \tau \rangle = \sum_{m=0}^{N-1} \sum_{l=0}^{m} \frac{1}{k_+(l)} \prod_{j=l+1}^{m} \frac{k_-(j)}{k_+(j)} \,. \tag{5}$$

This equation assumes that at time $t = 0$, the translocation process starts at translocation coordinate $m(0) = 0$, i.e., it does not consider the entrance stage (with possible failed attempts) of the translocation process. Our focus is only on the detailed passage dynamics of the successful translocation events. Equation (5) assumes reflecting boundary conditions at $m = 0$, i.e., the molecule is only allowed to exit the pore on the *trans* side, as in previous theoretical studies [18, 23, 25, 26]. Experimentally, this corresponds to a situation where, e.g., a protein or a small bead is attached to the *trans* end of the molecule, preventing exit to the *cis* side. In particular at low driving voltages, such a "road block" will be experimentally required, since otherwise it would not be possible to separate failed translocation attempts from full translocation events to the *trans* side. At larger driving voltages, the boundary condition at $m = 0$ is expected to be less relevant, since molecules are then unlikely to exit the pore on the *cis* side once they are inserted into the pore. At the other end, $m = N$, (5) assumes an absorbing boundary, i.e., the translocation time τ is defined as the time when the state $m = N$ is first reached.

To determine the hopping rates (4) for an RNA molecule with a given sequence, we first need to calculate the probabilities $w_{cis}(m)$ and $w_{trans}(m)$. To this end, we review in the following section the physics of RNA folding and the characteristics of random RNA sequences, before we return to link these characteristics to the translocation dynamics in Sec. 2.3.

2.2. Folding of random RNA sequences

In this section, we will review the aspects of the statistical physics of structures of random RNA molecules that are relevant for our study. We will follow the bulk of the previous literature and exclusively focus on RNA secondary structures [32, 33, 36, 37, 41, 42, 43]. An RNA secondary structure is the collection of all base pairs formed by a molecule. Formally, it can be described as a set $\mathcal{S} = \{(i_1, j_1), \ldots, (i_n, j_n)\}$ of all pairs of indices (i_k, j_k) (with $i_k < j_k$) of bases that are paired. A pairing configuration is only considered to be a valid secondary structure if it fulfils two conditions: (i) Each base is paired with at most one other base. (ii) If (i, j) is a base pair and (k, l) is a base pair with $i < k$, they have to be either nested, i.e., fulfil $i < k < l < j$, or independent, i.e., fulfil $i < j < k < l$. Forbidden base-pairing configurations with $i < k < j < l$ are called pseudo-knots. Restricting the allowable secondary structures to only those that contain neither base triplets nor pseudo-knots is an approximation since both structural elements do occur in actual structures. However, the approximation is reasonable since base triplets and pseudo-knots are believed to be rare in natural structures and can be effectively suppressed by performing experiments in the absence of multi-valent ions, such as Mg^{2+} [44].

The energy $E[\mathcal{S}]$ of a given structure $\mathcal{S}$ depends on the sequence $b_1 \dots b_N$ of the RNA molecule. For quantitative analyzes such as the prediction of the actual secondary structure of an RNA molecule [45, 46], very detailed energy models with hundreds of parameters have been developed [47]. Since we are interested in more generic questions such as the scaling behavior of translocation times, we will use a strongly simplified energy model that focuses on the base-pairing alone. More specifically, we will assign an energy solely derived from the base pairs formed in the structure, i.e.,

$$E[\mathcal{S}] = \sum_{(i,j)\in\mathcal{S}} \varepsilon_{i,j} \tag{6}$$

where $\varepsilon_{i,j}$ is the energy for the formation of a base pair between base b_i and b_j. Following [34, 37] we will even ignore the differences between the stability of different Watson-Crick base pairs and use the simplest possible model

$$\varepsilon_{i,j} = \begin{cases} -\varepsilon_m & b_i \text{ and } b_j \text{ are a Watson-Crick pair} \\ \varepsilon_{mm} & \text{otherwise} \end{cases} \tag{7}$$

where the match and mismatch energies ε_m and ε_{mm} are positive constants. Such a simplified energy model clearly is not suitable for the quantitative prediction of the behavior of an individual RNA molecule. However, the universal properties of the RNA folding problem, such as the thermodynamic phases, the topology of the phase diagram, and the critical exponents characterizing these phases in the thermodynamic limit are expected to be correctly captured.

For this as well as other more complicated energy models, the partition function of an RNA molecule of a given sequence $b_1 \dots b_N$ can be calculated exactly in polynomial time [48]. This can be done by introducing as an auxiliary quantity the partition function $Z_{i,j}$ for the substrand $b_i \dots b_j$ of the original molecule. The jth base can either be unpaired or paired with the kth base, where k takes all of the possibilities from i to $j-1$. If the jth base is unpaired, the allowable structures are exactly the allowable structures for the substrand $b_i \dots b_{j-1}$. If the jth base is paired with the kth base, the exclusion of pseudo-knots implies that in the presence of the (k,j) base pair, any structure is possible on the substrand $b_i \dots b_{k-1}$ and on the substrand $b_{k+1} \dots b_{j-1}$ but base pairs between these two substrands are forbidden. That yields the recursion equation

$$Z_{i,j} = Z_{i,j-1} + \sum_{k=i}^{j-1} Z_{i,k-1} e^{-\beta\varepsilon_{k,j}} Z_{k+1,j-1} \tag{8}$$

where $\beta = (k_B T)^{-1}$. Since the substrands referred to on the right hand side of this equation are shorter than the substrands referred to on the left hand side, this recursion equation can be used to start from the trivial single and two base substrands and calculate the partition functions for the increasingly larger substrands. The partition function $Z_{1,N}$ is then the partition function of the whole molecule. Since in this process $O(N^2)$ of the $Z_{i,j}$ have to be calculated with each calculation requiring one summation over the index k, the total computational complexity for this calculation is $O(N^3)$.

Through various numerical and analytical arguments it has been established that RNA secondary structures undergo a glass transition between a high temperature molten and a low temperature glassy phase [32, 34, 36, 37, 42, 49, 50]. In the (high temperature) molten phase the energetic differences between different structures become irrelevant and configurational entropy is the main contributor to the free energy of the structural ensemble [35] (it is to be noted that our simplified model of RNA secondary structures does not show a denaturation transition where base-pairing itself becomes unfavorable and the molecule becomes completely unstructured. Thus, "high temperature" in terms of real RNA molecules refers to temperatures still below the denaturation temperature, but close enough so that the energetic differences between different base pairs are smeared out). In the glassy phase, one or a few structures (determined by the specific sequence of the molecule) become dominant in the thermal ensemble — the molecule "freezes" into those structures.

The molten (high temperature) phase of RNA secondary structures is completely understood analytically [35]. Since in the molten phase by definition the base-pairing energetics do not play a role any more, the behavior of the molten phase can be determined by setting all base-pairing energies equal, i.e., by choosing $\varepsilon_{i,j} = -\varepsilon_0$ with some positive ε_0. Under this choice the partition functions $Z_{i,j}$ no longer depend on the nucleotide sequence and thus become translationally invariant, i.e., $Z_{i,j} \equiv Z(j-i+1)$. The recursion equation (8) then simplifies to

$$Z(N+1) = Z(N) + q\sum_{k=1}^{N} Z(k-1)Z(N-k) \tag{9}$$

where $q \equiv \exp(\beta\varepsilon_0)$ is the Boltzmann factor associated with a base pair. This recursion equation can be solved in the limit of large N and yields [35, 51]

$$Z(N) \approx AN^{-\gamma_m} z_0^N \tag{10}$$

where A and z_0 depend on the Boltzmann factor q. The exponent $\gamma_m = 3/2$, however, is universal and is characteristic of the molten phase.

2.3. Translocation of random RNA sequences

In the context of polymer translocation, it is necessary to determine what effect the pore has on the possible secondary structures of the molecule. If direct interactions with the pore are ignored, the only effect of the pore is that it divides the molecule into two segments, namely the *trans* part with m bases and the *cis* part with $N-m$ bases. Each part of the molecule can still form RNA secondary structures, but base pairs between a base on the *trans* side and a base on the *cis* side become impossible. This constraint results in a free energy cost. In the entropically dominated molten phase a reduction in the number of possibilities for base-pairing will decrease the entropy; in the energetically dominated glassy phase, a reduction in the number of possibilities to find well matching substrands will increase the energy. In both cases, the free energy cost provides a barrier to the translocation process, and we refer to the cost as the pinch

free energy $F(m)$. The pinch free energy depends explicitly on our reaction coordinate m and hence constitutes a free energy landscape for the translocation process (To keep the notation concise, we suppress the dependence of the pinch free energy $F(m)$ on the total sequence length N.)

With the help of the partition function $Z_{i,j}$ introduced in the previous section, the pinch free energy can be easily calculated: The partition function for the RNA molecule at position m in the pore has the product form $Z_{1,m}Z_{m+1,N}$ (the structures on the *cis* and *trans* sides are uncorrelated), whereas the partition function of the unconstrained RNA in solution is $Z_{1,N}$. The free energy difference between these states is the pinch free energy,

$$F(m) = -k_B T \left[\ln\left(Z_{1,m}Z_{m+1,N}\right) - \ln Z_{1,N}\right] \ . \tag{11}$$

Using the definition of the partition function, we can also establish the explicit link of the pinch free energy landscape to the translocation dynamics model of Sec. 2.1. To this end, we need to determine the probabilities $w_{cis}(m)$ and $w_{trans}(m)$ in (4). Since $Z_{i,j}$ represents the total statistical weight of all permitted base-pairing patterns for the RNA substrand from base i to base j, the probability $w_{cis}(m)$ for the base immediately in front of the pore on the *cis* side to be unpaired is given by

$$w_{cis}(m) = \frac{Z_{m+2,N}}{Z_{m+1,N}} \ . \tag{12}$$

Similarly, the probability for the base immediately in front of the pore on the *trans* side to be unpaired is given by

$$w_{trans}(m) = \frac{Z_{1,m-1}}{Z_{1,m}} \ . \tag{13}$$

Together, (2), (4), (8), (12), and (13) fully specify the translocation dynamics of structured RNA molecules within our model. The general form (5) for the average translocation time then simplifies to [31]

$$\begin{aligned} k_0\langle\tau\rangle &= e^{-\eta\frac{q_{\rm eff}V}{k_BT}} \sum_{m=0}^{N-1}\sum_{\ell=0}^{m} e^{-(m-\ell)\frac{q_{\rm eff}V}{k_BT}} \frac{Z_{1,\ell}Z_{\ell+1,N}}{Z_{1,m}Z_{m+1,N}} \\ &= e^{-\eta\frac{q_{\rm eff}V}{k_BT}} \sum_{m=0}^{N-1}\sum_{\ell=0}^{m} e^{\frac{F(m)-F(\ell)-(m-\ell)q_{\rm eff}V}{k_BT}} \end{aligned} \tag{14}$$

using the free energy $F(m)$ as defined in (11). It is now evident that the translocation dynamics of Sec. 2.1 corresponds to a random walk in the pinch free energy landscape which is tilted by the applied voltage.

Equation (11) can be used to compute the free energy landscape for a specific RNA sequence. To characterize the typical translocation behavior of structured RNA molecules, we need to generate such landscapes for a large sample from the ensemble of random sequences. We will take this numerical approach in section 3.1. However, using (10), we can analytically determine the typical form $F_{\rm molten}(m)$ of the landscape in the

molten phase,

$$\begin{aligned} F_{\text{molten}}(m) = & -k_B T \ln \frac{Z(m)Z(N-m)}{Z(N)} \\ \approx & -k_B T \ln \frac{m^{-\gamma_m}(N-m)^{-\gamma_m}}{N^{-\gamma_m}} \\ = & \gamma_m k_B T \ln[m(N-m)/N] \; . \end{aligned} \tag{15}$$

This is formally the same logarithmic free energy landscape as for the translocation of unstructured polymers, (1). However, its physical origin is completely different (namely, the structural entropy of base-pairing configurations rather than the positional entropy of the backbone), and its prefactor $\gamma_m = 3/2$ is larger, which will be important below.

It is interesting to note that the logarithmic behavior of (15) and the value of the prefactor can be physically understood by realizing that the ensemble of secondary structures in the molten phase corresponds to the ensemble of (rooted) branched polymers: The number of possible configurations of a rooted branched polymer of molecular weight m is known to scale like $m^{-3/2}$ [52] (in addition to the non universal extensive factor) and thus the pinch free energy landscape of a translocating RNA molecule in the molten phase is the same as the landscape generated by cutting a branched polymer of molecular weight N into two rooted branched polymers of molecular weights m and $N-m$, respectively [34].

In the glassy phase the situation is much less clear, since there are no analytical calculations of the typical pinch free energy for the ensemble of random RNA sequences. Furthermore, different numerical studies [34, 42, 43], which examined the maximal pinch (at $m = N/2$), disagree whether $F(N/2)$ scales logarithmically or as a small power with the sequence length N. One numerical argument in favor of a logarithmic dependence is that different choices of the sequence disorder yield different prefactors of the logarithm or different exponents in the power law. While different prefactors of the logarithm are not problematic, exponents that depend on the choice of the disorder contradict the notion that exponents should be universal. In [34] the dependence of the maximal pinch free energy on sequence length and temperature was studied in detail and it was found that the dependence of the maximal pinch free energy on sequence length can be described rather well by a logarithmic law for all temperatures. At high temperatures, the prefactor $a(T)$ of the logarithmic dependence is $\frac{3}{2}k_B T$, as expected. However, at low temperatures this prefactor ceases to be proportional to temperature and converges toward a finite value at zero temperature. Thus, if we assume that the entire averaged pinch free energy landscape still has the logarithmic form of (1) in the glassy phase, the logarithmic dependence of its maximum on sequence length implies that the effective prefactor

$$\gamma(T) = \frac{a(T)}{k_B T} \tag{16}$$

equals 3/2 at high temperatures and diverges as the temperature is lowered below the glass transition temperature. For random sequences with equal probability for all four bases and energies $\varepsilon_m = \varepsilon_{mm}$ this behavior is numerically illustrated in figure 3.

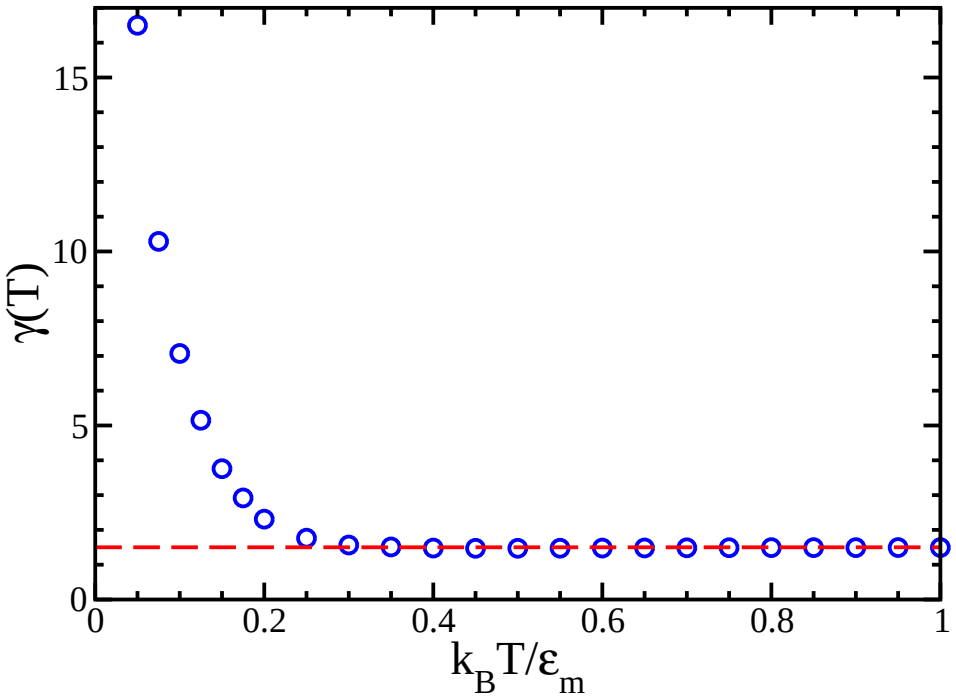

Figure 3. Numerically determined prefactor $\gamma(T)$ of the logarithmic free energy landscape (1) as a function of temperature (most of the data from [34]). The statistical error of the data is on the order of the symbol size. It can be seen that the prefactor is constant $\frac{3}{2}$ in the high temperature (molten) phase. In the low temperature (glassy) phase the prefactor becomes temperature dependent and diverges. The prefactors were determined by generating many random RNA sequences with equal probability of the four bases of lengths $N = 160$, 320, 640, and 1280, calculating the restricted partition functions for the energy model (7) with $\varepsilon_m = \varepsilon_{mm}$ via (8) and extracting the pinch free energies at $m = N/2$ via (11). The prefactor $a(T)$ of the logarithmic law for the pinch free energy was determined by fitting such a logarithmic law to the numerical data and the corresponding prefactor of the logarithmic free energy landscape $\gamma(T)$ was calculated from (16).

3. Results and discussion

3.1. Numerical analysis

From the arguments in section 2, one may expect that the translocation of structured RNA molecules can be described as a one-dimensional diffusion process in the logarithmic energy landscape (1) with a temperature dependent, potentially large prefactor γ. Of course, the scaling of the translocation time with sequence length in such a landscape can be derived analytically as a function of the prefactor γ. However, there are several uncertainties in this description. First, different numerical studies of the pinch free energy of random RNA molecules do not even agree on the question if the maximum of the landscape at $m = N/2$ scales logarithmically or with a small power of the sequence length in the glass phase [34, 42, 43]. Second, even if the maximum of the landscape scales logarithmically, it has not been established that the whole average

landscape follows the simple shape (1). Third, even if the *average* landscape has the suggested shape, the landscape of any given RNA molecule can differ significantly from the average landscape. Thus, it is not obvious how the ensemble of translocation times of actual landscapes is related to the translocation time over the average landscape.

In light of all the uncertain approximations going into the energy landscape arguments outlined above, we perform a detailed numerical study of the translocation dynamics of random RNA molecules on the basis of the model defined in sections 2.1 and 2.3. We want to point out that the only approximation this numerical study is based on is that the dynamics of an RNA molecule can truly be described as motion in a one-dimensional energy landscape, i.e., that the base-pairing degrees of freedom have time to equilibrate for every position m of the whole molecule relative to the pore.

In order to study the exact translocation times of random RNA molecules, we generate explicit free energy landscapes for 2500 different RNA molecules of length $N = 1600$ using the partition function recursion (8), and the definition (11) of the pinch free energy (During the calculation, the auxiliary partition functions $Z_{i,j}$ are rescaled to avoid numerical overflows due to the large exponential factors at low temperatures.) The RNA sequences are random and uncorrelated, with equal probabilities of 1/4 for each of the four bases. We use the energy model (7) with $\varepsilon_m = \varepsilon_{mm}$ and quote all energies in units of ε_m.

First of all, our ensemble of free energy landscapes allows us to directly inspect the shape of the average landscape. Figure 4 shows the pinch free energy landscape $F(m)$ averaged over the 2500 realizations of the random sequences (symbols). Superimposed as lines are logarithmic energy landscapes as given by (1) with prefactors γ that are directly obtained by multiplying the values shown in figure 3 by k_BT. These energy landscapes are shifted by fitted offsets which reflect the behavior of the pinching free energy at very small m and which are irrelevant for the scaling behavior of the translocation dynamics. The comparison indicates that the overall shape of the average free energy landscapes is indeed the logarithmic one, even in the glassy temperature regime of figure 3 where $\gamma(T)$ is significantly larger than $\frac{3}{2}$.

Next, we examine the translocation times. For each sequence, we calculate the thermal average of the translocation time $\langle\tau\rangle$ using the exact expression (14). The most straightforward quantity to extract from the 2500 translocation times thus obtained would be the ensemble averaged translocation time $\overline{\langle\tau\rangle}$, where the bar denotes averaging over the sequence ensemble. However, as frequently observed in disordered systems, the distribution of characteristic times develops a fat tail at low temperatures, which renders the ensemble averaged translocation time ill defined (see figure 5). Instead, we must use a definition for the typical value that does not rely on the existence of the mean value. For instance, the median or the average of the logarithm both provide a well-defined typical value even for fat tailed distributions. Here, we use the average of the logarithm of the translocation time, which can be interpreted as a typical effective energy barrier.

Figure 6 shows the resulting ensemble averages of the logarithms of the translocation times for different sequence lengths of $N = 50$, 100, 200, 400, 800, and 1600 and for

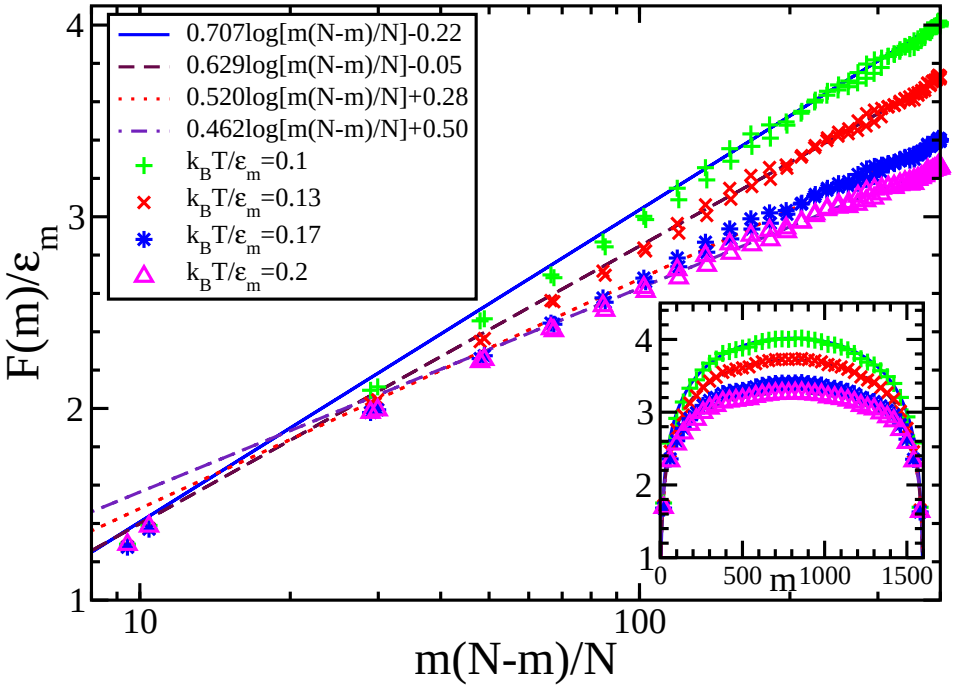

Figure 4. Numerically determined average free energy landscapes for translocation of RNA molecules through a nanopore. The main graph shows the average free energies as a function of $m(N-m)/N$ on a logarithmic scale to allow a direct comparison with the logarithmic shape (1). The inset shows the same data as a function of the linear position m of the molecule in the pore. For clarity, the numerically determined free energies are averaged over ranges of the reaction coordinate m of size 20. The statistical errors on the data are smaller than the size of the symbols. From the straight lines in the main graph it can be seen that the average free energy landscape follows up to irrelevant additive constants the logarithmic shape (1) with prefactors $\gamma k_B T$ where γ is taken from figure 3.

temperatures of $k_BT/\varepsilon_m = 0.1$, 0.13, 0.17, 0.2, 0.3, 0.6, 0.8, and 1.0. Two features of figure 6 are immediately obvious. First, at all temperatures, the double logarithmic plot of translocation times as a function of sequence length is perfectly linear over the whole range of sequence lengths studied. Second, the slope of these lines is independent of temperature for large temperatures and increases sharply as the temperature is lowered.

To quantify the power laws and their temperature dependence, we apply linear regression to our logarithmic data. The resulting slopes (i.e., exponents of the power law) are shown in table 1. It can be seen that *all* exponents are larger than two, i.e., the trivial diffusive scaling $\tau \sim N^2$ does not describe the translocation dynamics. In the high temperature regime, the exponent is clearly independent of temperature, while it becomes large and very sensitive to temperature in the low temperature regime. This salient feature in the translocation behavior of structured RNA molecules implies highly anomalous sub-diffusive dynamics for the translocation process. In the next section, we will discuss this behavior from a theoretical perspective.

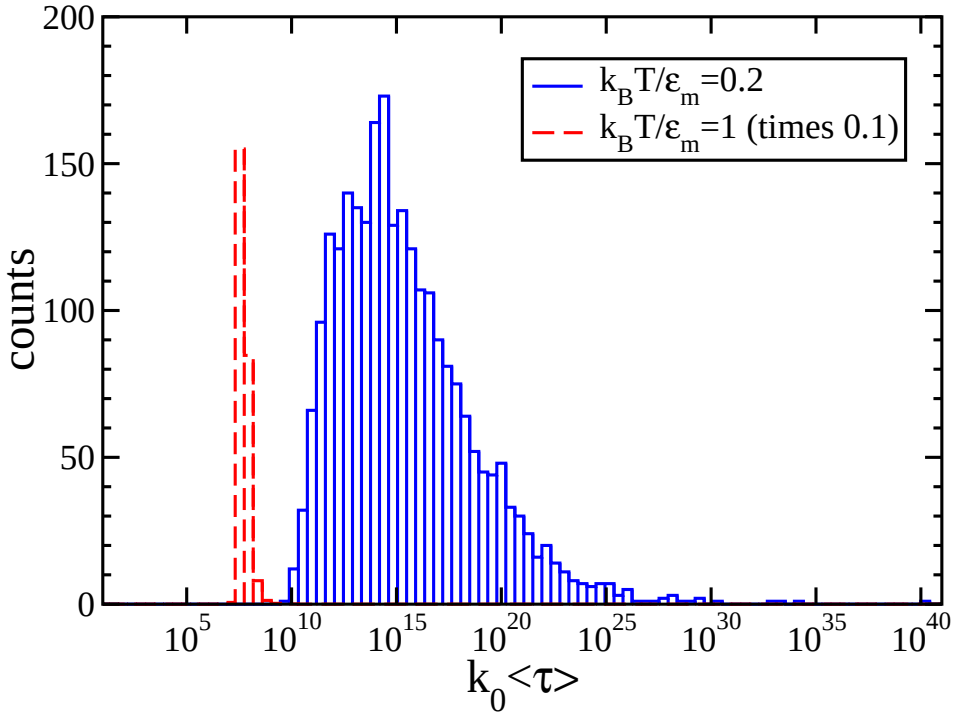

Figure 5. Histogram of thermally averaged translocation times for 2500 random sequences of length $N = 1600$ at temperatures $k_BT/\varepsilon_m = 0.2$ and $k_BT/\varepsilon_m = 1$. Note, that counts in the histogram for $k_BT/\varepsilon_m = 1$ are rescaled by a factor of 10 to fit into the same plot as the histogram for $k_BT/\varepsilon_m = 0.2$. It can be seen that the distribution already at $k_BT/\varepsilon_m = 1$ spans several decades. At the low temperature $k_BT/\varepsilon_m = 0.2$ the distribution develops a very fat tail consisting of few sequences with very long translocation times.

3.2. Anomalous scaling of the translocation times

Given that the translocation dynamics on a free energy landscape of the logarithmic form (1) was previously studied, and its scaling behavior was found to be normally diffusive [23], our finding of anomalous scaling in the present case is surprising. Our numerical computation of the *average* free energy landscape shown in figure 4 indeed confirmed that the typical landscape for the translocation of structured RNA molecules has the logarithmic shape, as the theoretical arguments of Sec. 2.3 had suggested. To resolve the apparent contradiction, we now revisit the arguments of reference [23] that led to the diffusive scaling.

Chuang, Kantor, and Kardar considered a continuum description of the translocation process, based on a Fokker-Planck equation similar to (3), but with the drift velocity v replaced by the local gradient of the free energy landscape,

$$\frac{\partial}{\partial t}P = D\,\frac{\partial^2}{\partial x^2}P + \frac{D}{k_BT}\frac{\partial}{\partial x}\left(P\,\frac{\partial}{\partial x}F(x)\right)\ ,$$

with $F(x) = \gamma\, k_BT \ln[(N-x)x/N]$. They note that the polymer length N and the diffusion constant can be eliminated from this equation by introducing a rescaled time

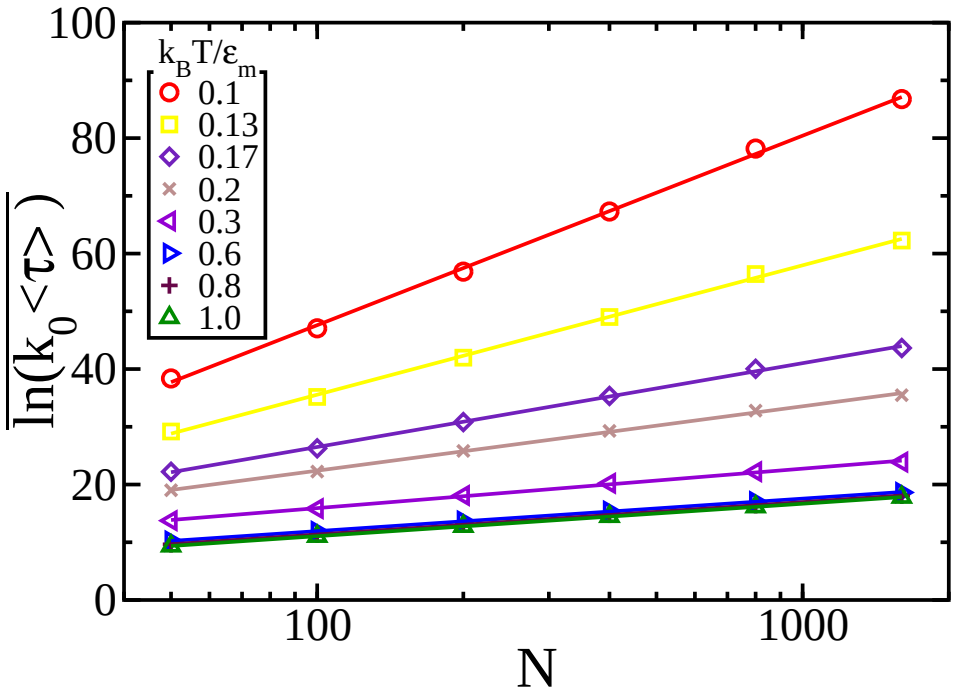

Figure 6. Dependence of the typical translocation time on sequence length for several different temperatures. To avoid problems with fat tails of the distribution of the translocation times at low temperatures, the ensemble average of the logarithms of the translocation times is taken to determine the typical translocation time. The statistical errors on the translocation times are on the order of the size of the symbols. It can be seen that the typical translocation time has a very clean power law dependence over the whole range of sequence lengths and for all temperatures. The translocation time is independent of temperature for large temperatures and becomes very sensitive to temperature at low temperatures.

$\tau = tD/N^2$ and translocation coordinate $s = x/N$,

$$\frac{\partial p}{\partial \tau} = \frac{\partial^2 p}{\partial s^2} + \gamma \frac{\partial}{\partial s}\left(\frac{1-2s}{(1-s)s}\, p\right) , \tag{17}$$

where $p = p(s,\tau)$ now is the probability distribution in the rescaled variables. Consequently, the authors then argue that the solution of this dimensionless equation may be converted back to real time by multiplying the time axis by N^2/D, resulting in a diffusive scaling of the translocation time. Indeed, as the authors point out, the argument is independent of the value of γ. Application of the argument to the present case, with $\gamma \geq 3/2$, would suggest that the secondary structure of the RNA is irrelevant in the slow translocation limit considered here.

However, we will now see that this conclusion cannot be drawn. The argument rests on the tacit assumption that the probability distribution $p = p(s,\tau)$ develops no structure on a microscopic scale. For instance, the continuum Fokker-Planck description breaks down if most of the probability is localized on one or a few points along the translocation coordinate m. Indeed, such a localization transition occurs, if γ exceeds a threshold value of one: Assuming a quasi-stationary solution to (17) which is localized

Table 1. Exponents of the power law dependence of the typical translocation time on the length of the molecule. These exponents are determined by linear regression of the data in figure 6.

k_BT/ε_m	exponent
0.1	14.05 ± 0.08
0.13	9.62 ± 0.03
0.17	6.23 ± 0.04
0.2	4.79 ± 0.06
0.3	2.94 ± 0.06
0.6	2.44 ± 0.02
0.8	2.43 ± 0.01
1.0	2.44 ± 0.01

at the $s = 0$ border is a self-consistent ansatz, if p behaves as $p \sim s^{-\gamma}$ for small s. For $\gamma > 1$ the integral of this distribution diverges at the boundary, i.e., the free energy barrier to translocation becomes strong enough for the quasi-stationary distribution to localize at the boundary.

In the regime $\gamma > 1$ where the argument for the independence of the translocation time on γ breaks down, there are two ways to obtain the correct scaling behavior of the translocation time. The first and more intuitive possibility is to use the standard Kramers rate theory for thermally-induced barrier crossing [40, 53]. In the present case, this approach yields

$$\tau \sim \frac{e^{F(N/2)/k_BT}}{\sqrt{F''(N/2)}} \sim N^{\gamma+1} \,, \tag{18}$$

where it is important to note that the barrier height itself according to (1) only yields a power law of N^γ and that the additional power of N results from the denominator which is often ignored in applications of Kramers rate theory.

The second, exact but less intuitive possibility to obtain the scaling behavior of the translocation time in the regime $\gamma > 1$ is to directly insert the landscape (1) into the expression (14) for the mean translocation time. In the absence of a driving voltage this yields

$$k_0\langle\tau\rangle = \sum_{m=1}^{N-1}(N-m)^\gamma m^\gamma \sum_{\ell=1}^{m}(N-\ell)^{-\gamma}\ell^{-\gamma} \tag{19}$$

where we changed the lower limits on the summations from zero to one in order to avoid the artificial divergence caused by the absence of a regularization of the landscape expression (1) for small m. The symmetry of the landscape with respect to $m = N/2$ allows us to rewrite this expression as

$$k_0\langle\tau\rangle = \frac{N-1}{2} + \frac{1}{2}\left(\sum_{m=1}^{N-1}(N-m)^\gamma m^\gamma\right)\left(\sum_{\ell=1}^{N-1}(N-\ell)^{-\gamma}\ell^{-\gamma}\right). \tag{20}$$

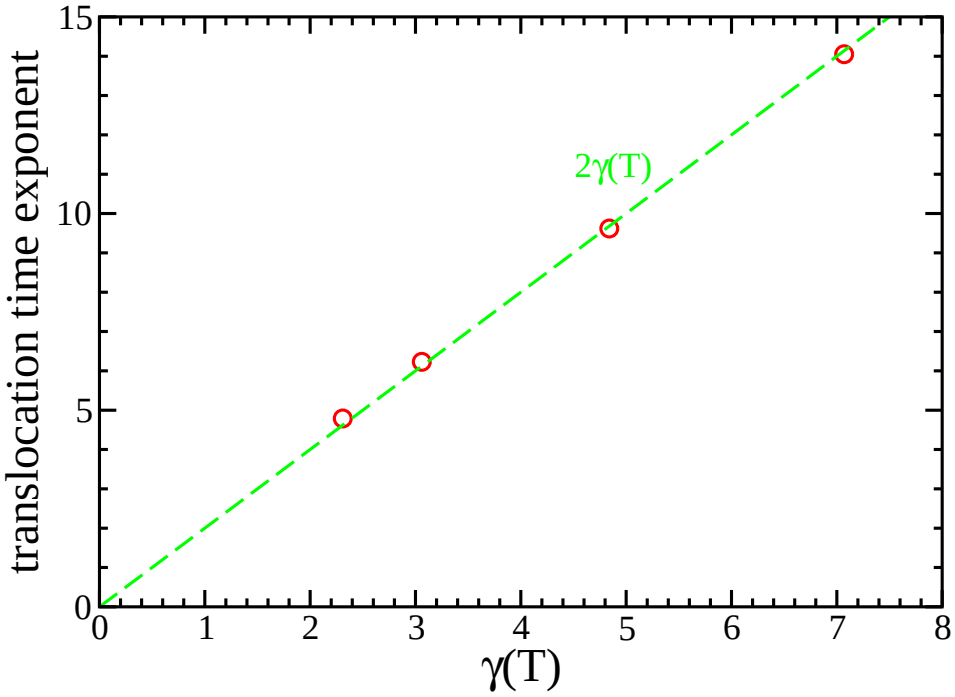

Figure 7. Comparison of the landscape prefactors $\gamma(T)$ from figure 3 and the numerically determined translocation time exponents from table 1 for the temperatures k_BT/ε_m where RNA is expected to be in the glassy phase. It can be seen that the observed translocation time exponents empirically behave like $2\gamma(T)$.

In the limit of large N (and for $\gamma > 1$) the second sum is dominated by the contributions from the boundaries and can be approximated as

$$\left(\sum_{\ell=1}^{N-1}(N-\ell)^{-\gamma}\ell^{-\gamma}\right) \approx 2N^{-\gamma}\sum_{\ell=0}^{\infty}\ell^{-\gamma} = 2\zeta(\gamma)N^{-\gamma} \tag{21}$$

where it is important to notice that although the main contributions to this sum come from the boundaries the scaling with N comes from the part of the expression that is insensitive to the specific small m regularization of the landscape. The first sum in (20) is bracketed by $2(N/2)^{\gamma}\sum_{m=1}^{N/2} m^{\gamma}$ and $2N^{\gamma}\sum_{m=1}^{N/2} m^{\gamma}$ and thus scales like $N^{2\gamma+1}$. Taking everything together confirms the $N^{\gamma+1}$ scaling of the translocation time obtained from Kramers rate theory in (18).

If we apply (18) to the molten phase where $\gamma = \frac{3}{2}$, this yields $\tau \sim N^{2.5}$ in good agreement with our numerical estimates, see table 1. Furthermore, (18) rationalizes the sharp increase of the scaling exponent as the temperature is lowered into the glass phase, i.e., for $k_BT/\varepsilon_m \leq 0.2$. However, in that regime the quantitative comparison of the translocation time exponents in table 1 and the barrier heights in figure 3 shown in figure 7 reveals that the translocation time exponents increase even more dramatically than the increase of the landscape prefactors suggests, namely approximately like $2\gamma(T)$. This indicates that the typical translocation of individual molecules in the glass phase of RNA is *not* well approximated by translocation in the average landscape but rather

must be dominated by the fluctuations of the free energy landscape around the average.

4. Conclusion and outlook

In conclusion, we see that our numerical observation that the scaling of the typical translocation time is drastically affected by the secondary structure is qualitatively well in accord with theoretical expectation but quantitatively even exceeds the magnitude of the effect expected from the theory. For translocation in a logarithmic landscape it is clear from (18) that $\gamma = 1$ constitutes a threshold for a change in the translocation behavior: The regime $\gamma < 1$ is marked by an insignificant barrier, diffusive translocation, and failure of the Kramers approximation, which assumes a "reaction-limited" process and ignores the time required to diffuse from the starting to the end point. In contrast, for $\gamma > 1$ the barrier dominates the translocation dynamics and leads to the sub-diffusive scaling (18). Importantly, for unstructured polymers where the logarithmic free energy landscape is only due to the configurational entropy of the polymer, we have $\gamma < 1$ even if self-avoidance is included. Here, we found that the case of structured RNA molecules is always in the opposite regime of $\gamma > 1$. Thus, despite the similarity in the form of the free energy landscape, (1), the translocation behavior of unstructured and structured polynucleotides is quite different.

The anomalous scaling of translocation times found in our study is only observable in the absence of an external voltage. In the presence of an external voltage the gain in electrostatic energy due to moving $N/2$ bases into the pore is linear in N and thus for large N always overcomes the logarithmic barrier (1) leading to a linear dependence of the translocation time on sequence length. However, for finite but small voltages the anomalous scaling could still be observed in an intermediate regime of sequence lengths where $Nq_{\rm eff}V/2 < \gamma(T)$.

Our empirical finding of even stronger anomalous scaling in the glassy phase than expected from the average free energy landscape indicates that translocation in the glassy phase is strongly affected by the fluctuations and the free energy landscapes of the individual RNA molecules. Understanding these fluctuations and the origin of the intriguing empirically found $2\gamma(T)$ law for the translocation time exponent will be subject of future research.

Acknowledgments

We gratefully acknowledge stimulating discussions with Yariv Kafri and Julius Lucks. This work was supported in part by the Petroleum Research Fund of the American Chemical Society through Grant No. 42555-G9 to RB, by the National Science Foundation through grant No. DMR-0706002 to RB, by the Deutsche Forschungsgemeinschaft via SFB 486 to UG, and by the German Excellence Initiative via the program "Nanosystems Initiative Munich (NIM)" to UG.

Glossary

Nanopore. An opening connecting two macroscopic chambers that has a diameter in the nanometer range. If the diameter is only about one nanometer, double stranded RNA and DNA molecules have to separate into single-stranded configurations in order to pass through such a pore.

Translocation time. The time from the arrival of a molecule on one side of a nanopore until the departure of the molecule from the other side of the nanopore.

Scaling. In general the behavior of one physical quantity as a function of another physical quantity as a power law. Here, specifically the behavior of the translocation time of a molecule as a function of the length of the molecule. In such a power law the exponent itself is typically insensitive to microscopic details of the model while the prefactor depends on such microscopic details. Thus, the exponent is the main object of interest.

Pinch free energy. Difference in free energy between an RNA molecule that is allowed to explore all possible base-pairing configurations and an RNA molecule that is separated (pinched) into two sections each of which only allows base-pairings within the section. The free energy difference can have entropic origins from a reduction in the number of base-pairing configurations or energetic origins because particularly energetically favorable base-pairing configurations may be disallowed after pinching.

Kramers rate theory. A theory describing the motion of a diffusing particle over a single energy barrier. The theory states that the time for a particle to leave a potential well on one side of the barrier and to reach the other side of the barrier depends only on the height of the barrier and on the shape of the potential at the starting potential well and at the top of the barrier while other details of the potential are irrelvant.

References

[1] J J Kasianowicz, E Brandin, D Branton, and D W Deamer. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. USA*, 93(24):13770–13773, Nov 1996.

[2] A Meller, L Nivon, and D Branton. Voltage-driven DNA translocations through a nanopore. *Phys. Rev. Lett.*, 86(15):3435–3438, Apr 2001.

[3] A J Storm, J H Chen, H W Zandbergen, and C Dekker. Translocation of double-strand DNA through a silicon oxide nanopore. *Phys. Rev. E*, 71:051903, May 2005.

[4] A Meller, L Nivon, E Brandin, J Golovchenko, and D Branton. Rapid nanopore discrimination between single polynucleotide molecules. *Proc. Natl. Acad. Sci. USA*, 97:1079–1084, Feb 2000.

[5] W Vercoutere, S Winters-Hilt, H Olsen, D Deamer, D Haussler, and M Akeson. Rapid discrimination among individual DNA hairpin molecules at single-nucleotide resolution using an ion channel. *Nat. Biotechnol.*, 19:248–252, Mar 2001.

[6] Jerome Mathé, Hasina Visram, Virgile Viasnoff, Yitzhak Rabin, and Amit Meller. Nanopore unzipping of individual DNA hairpin molecules. *Biophys. J.*, 87(5):3205–3212, Nov 2004.

[7] A F Sauer-Budge, J A Nyamwanda, D K Lubensky, and D Branton. Unzipping kinetics of double-stranded DNA in a nanopore. *Phys. Rev. Lett.*, 90:238101, Jun 2003.

[8] J Mathé, A Arinstein, Y Rabin, and A Meller. Equilibrium and irreversible unzipping of DNA in a nanopore. *Europhys. Lett.*, 73(1):128–134, JAN 2006.

[9] Ulrich F Keyser, Bernard N Koeleman, Stijn van Dorp, Diego Krapf, Ralph M M Smeets, Serge G Lemay, Nynke H Dekker, and Cees Dekker. Direct force measurements on DNA in a solid-state nanopore. *Nature Physics*, 2:473–477, Jul 2006.

[10] M Wanunu, J Sutin, B McNally, A Chow, and A Meller. DNA translocation governed by interactions with solid-state nanopores. *Biophys. J.*, 95:4716–4725, Nov 2008.

[11] M Wanunu, B Chakrabarti, J Math, D R Nelson, and A Meller. Orientation-dependent interactions of DNA with an alpha-hemolysin channel. *Phys. Rev. E*, 77:031904, Mar 2008.

[12] M Akeson, D Branton, J J Kasianowicz, E Brandin, and D W Deamer. Microsecond time-scale discrimination among polycytidylic acid, polyadenylic acid, and polyuridylic acid as homopolymers or as segments within single RNA molecules. *Biophys. J.*, 77:3227–3233, Dec 1999.

[13] R Stefureac, Y T Long, H B Kraatz, P Howard, and J S Lee. Transport of alpha-helical peptides through alpha-hemolysin and aerolysin pores. *Biochemistry*, 45:9172–9179, Aug 2006.

[14] C P Goodrich, S Kirmizialtin, B M Huyghues-Despointes, A Zhu, J M Scholtz, D E Makarov, and L Movileanu. Single-molecule electrophoresis of beta-hairpin peptides by electrical recordings and Langevin dynamics simulations. *J. Phys. Chem. B*, 111:3332–3335, Apr 2007.

[15] W Sung and P J Park. Polymer Translocation through a Pore in a Membrane. *Phys. Rev. Lett.*, 77:783–786, Jul 1996.

[16] E A DiMarzio and A J Mandell. Phase transition behavior of a linear macromolecule threading a membrane. *J. Chem. Phys.*, 107(14):5510–5514, OCT 8 1997.

[17] M Muthukumar. Polymer translocation through a hole. *J. Phys. Chem.*, 111:10371–10374, 1999.

[18] M Muthukumar. Translocation of a confined polymer through a hole. *Phys. Rev. Lett.*, 86:3188–3191, Apr 2001.

[19] Y C Chen, C Wang, and M B Luo. Simulation study on the translocation of polymer chains through nanopores. *J. Chem. Phys.*, 127:044904, Jul 2007.

[20] Elena Slonkina and Anatoly B Kolomeisky. Polymer translocation through a long nanopore. *J. Chem. Phys.*, 118(15):7112–7118, 2003.

[21] A Milchev, K Binder, and A Bhattacharya. Polymer translocation through a nanopore induced by adsorption: Monte Carlo simulation of a coarse-grained model. *J. Chem. Phys.*, 121:6042–6051, Sep 2004.

[22] A Meller. Dynamics of polynucleotide transport through nanometre-scale pores. *J. Phys.: Condens. Matter*, 15:R581–R607, 2003.

[23] Jeffrey Chuang, Yacov Kantor, and Mehran Kardar. Anomalous dynamics of translocation. *Phys. Rev. E*, 65:011802, 2002.

[24] Erich Eisenriegler. *Polymers near surfaces.* World Scientific, 1993.

[25] R Metzler and J Klafter. When translocation dynamics becomes anomalous. *Biophys. J.*, 85:2776–2779, Oct 2003.

[26] Y Kantor and M Kardar. Anomalous dynamics of forced translocation. *Phys. Rev. E*, 69:021806, Feb 2004.

[27] A Matsuyama. Kinetics of polymer translocation through a pore. *J. Chem. Phys.*, 121(16):8098–8103, OCT 22 2004.

[28] Debabrata Panja, Gerard T Barkema, and Robin C Ball. Anomalous dynamics of unbiased polymer translocation through a narrow pore. *J. Phys.: Condens. Matter*, 19(43), OCT 31 2007.

[29] C Chatelain, Y Kantor, and M Kardar. Probability distributions for polymer translocation. *Phys. Rev. E*, 78:021129, Aug 2008.

[30] Ulrich Gerland, Ralf Bundschuh, and Terence Hwa. Translocation of structured polynucleotides

through nanopores. *Phys. Biol.*, 1(1-2):19–26, Jun 2004.

[31] Ralf Bundschuh and Ulrich Gerland. Coupled dynamics of RNA folding and nanopore translocation. *Phys. Rev. Lett.*, 95(20):208104, Nov 2005.

[32] P G Higgs. Overlaps between RNA secondary structures. *Phys. Rev. Lett.*, 76:704–707, Jan 1996.

[33] P G Higgs. RNA secondary structure: physical and computational aspects. *Q. Rev. Biophys.*, 33:199–253, Aug 2000.

[34] R Bundschuh and T Hwa. Statistical mechanics of secondary structures formed by random RNA sequences. *Phys. Rev. E*, 65:031903, Mar 2002.

[35] P G de Gennes. Statistics of branching and hairpin helices for the dAT copolymer. *Biopolymers*, 6(5):715–29, 1968.

[36] A Pagnani, G Parisi, and F Ricci-Tersenghi. Glassy transition in a disordered model for the RNA secondary structure. *Phys. Rev. Lett.*, 84:2026–2029, Feb 2000.

[37] R Bundschuh and T Hwa. Phases of the secondary structures of RNA sequences. *Europhys. Lett.*, 59:903–909, 2002.

[38] D K Lubensky and D R Nelson. Driven polymer translocation through a narrow pore. *Biophys. J.*, 77(4):1824–38, 1999.

[39] J Zhang, A Kamenev, and B I Shklovskii. Conductance of ion channels and nanopores with charged walls: a toy model. *Phys. Rev. Lett.*, 95:148101, Sep 2005.

[40] Crispin W Gardiner. *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences.* Springer, 2004.

[41] E Marinari, A Pagnani, and F Ricci-Tersenghi. Zero-temperature properties of RNA secondary structures. *Phys. Rev. E*, 65:041919, Apr 2002.

[42] F Krzakala, M Mézard, and M Müller. Nature of the glassy phase of RNA secondary structure. *Europhys. Lett.*, 57(5):752–758, MAR 2002.

[43] S Hui and L-H Tang. Ground state and glass transition of the RNA secondary structure. *Europ. Phys. J. B*, 53(1):77–84, Sep 2006.

[44] I Tinoco and C Bustamante. How RNA folds. *J. Mol. Biol.*, 293:271–281, Oct 1999.

[45] I L Hofacker, W Fontana, P F Stadler, S Bonhoeffer, M Tacker, and P Schuster. Fast folding and comparison of RNA secondary structures. *Monatshefte f. Chemie*, 125:167–188, 1994.

[46] Michael Zuker. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.*, 31(13):3406–3415, Jul 2003.

[47] D H Mathews, J Sabina, M Zuker, and D H Turner. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.*, 288(5):911–940, May 1999.

[48] J S McCaskill. The equilibrium partition function and base pair binding probabilities for RNA secondary structure. *Biopolymers*, 29(6-7):1105–19, 1990.

[49] A K Hartmann. Comment on ıglassy transition in a disordered model for the rna secondary structureȷ. *Phys. Rev. Lett.*, 86(7):1382, Feb 2001.

[50] A Pagnani, G Parisi, and F Ricci-Tersenghi. Pagnani, Parisi, and Ricci-Tersenghi reply:. *Phys. Rev. Lett.*, 86(7):1383, Feb 2001.

[51] R Bundschuh and T Hwa. RNA secondary structure formation: a solvable model of heteropolymer folding. *Phys. Rev. Lett.*, 83:1479–1483, 1999.

[52] T C Lubensky, J Isaacson, and S P Obukhov. Field-theory for ARB2 branched polymers. *J. Physique*, 42(12):1591–1601, 1981.

[53] H A Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7:284–304, 1940.