

# Jump transition observed in translocation time for ideal poly- $X$ proteinogenic chains as a result of competing folding and anchoring contributions

José Antonio Vélez-Pérez\*

*Posgrado en Nanociencias y Nanotecnología, Centro de Investigación y de Estudios Avanzados del IPN, Ap. Postal 14-740, 07000 México, DF, México*

Luis Olivares-Quiroz†

*Colegio de Ciencia y Tecnología and Posgrado en Ciencias de la Complejidad, Universidad Autónoma de la Ciudad de México, CP 09760 México City, México*

(Received 18 October 2016; published 19 January 2017)

In this work we analyze the translocation of homopolymer chains poly- $X$ , where  $X$  represents any of the 20 naturally occurring amino acid residues, in terms of size  $N$  and single-helical propensity  $\omega$ . We provide an analytical framework to calculate both the free energy  $\mathcal{F}$  of translocation and the translocation time  $\tau$  as a function of chain size  $N$ , energies  $U$  and  $\epsilon$  of the unfolded and folded states, respectively. Our results show that free energy  $\mathcal{F}$  has a characteristic bell-shaped barrier as function of the percentage of monomers translocated. Inclusion of single-helical propensity  $\omega$  associated to monomer  $X$  and chain's native energy  $\epsilon$  in the translocation model increases the energy barrier  $\Delta\mathcal{F}$  up to one order of magnitude as compared with the well-known Gaussian chain model. Computation of the mean first-passage time as function of chain size  $N$  shows that the translocation time  $\tau$  exhibits a significant jump of several orders of magnitude at a critical chain size  $\mathcal{N}$ . This jump markedly slows down translocation of chains larger than  $\mathcal{N}$ . Existence of the transition jump of  $\tau$  has been observed experimentally at least in poly(ethylene oxide) chains [R. P. Choudhury, P. Galvosas, and M. Schönhoff, *J. Phys. Chem. B* **112**, 13245 (2008)]. Our results suggest the transition jump of  $\tau$  as a function of  $N$  may be a very well spread feature throughout translocation of poly- $X$  chains.

DOI: 10.1103/PhysRevE.95.012407

## I. INTRODUCTION

Translocation of biopolymers through membranes is a cornerstone in many cellular processes. Biopolymers involved in translocation usually include RNA, DNA, and proteins but translocation is not restricted to them. During translocation, biopolymers thread through a channel embedded in a membrane connecting two milieu called *cis* and *trans* regions. Protein translocation systems are diverse, such as the vesicle-embedded anthrax toxin channel [1], the virulent needle-like type three secretion system connecting the bacterial and host cell membranes [2], or the Sec complex channel embedded in the endoplasmic reticulum or the periplasmic membranes [3]. The general view suggests that in order to initiate translocation, a targeting signal has to be recognized and bound at the entrance of the channel. To accomplish translocation, unfolding, diffusion, and coupling of electrochemical forces are necessary steps [3–6]. Understanding translocation is thus a key step in the development of systems with potential applications involving biopolymer sequencing and delivery of proteins as therapeutic agents [7,8].

Most of the experimental and computational approaches characterizing translocation have focused on analyzing the free energy of translocation  $\mathcal{F}$  and the translocation time  $\tau$  in terms of the chain size  $N$ . Recent experiments of polynucleotide and polypeptide chains across  $\alpha$ -hemolysin channels and solid-state nanopores have provided both qualitative and quantitative

valuable information regarding functional forms for  $\Delta\mathcal{F}$  and  $\tau$  during translocation. For instance, umbrella sampling MD of Ubiquitin (PDB code 1UBI) have shown the existence of a free energy barrier  $\Delta\mathcal{F}$  with a characteristic umbrella-like profile [9] as a function of the percentage of monomers translocated. Results for confined, flexible charged polyelectrolytes crossing narrow nanopores also exhibit this type of umbrella-like profile [10]. Such a profile has been observed as well on MD studies on flexible hard-sphere chains through cylindrical holes [11] where some small-amplitude additional oscillations have also been observed. MD studies performed in other geometries have also shown these characteristic profiles [12]. In regard to the translocation time, adsorption-driven MD simulations have evaluated  $\tau$  as a function of size  $N$  and obtained that it follows a power-law  $\tau \simeq N^\xi$ , where  $\xi = 2.2$ , approximately [13]. Early translocation experiments done by Kasianowick *et al.* using ssDNA fragments through an  $\alpha$ -hemolysin pore, suggested initially that  $\tau$  scaled linearly with size  $N$  for driven translocation [14]. However, for non-driven translocation, Sung and Park [15] and Mutkhumar [16] have suggested that  $\tau \sim N^2$  based on Focker-Planck and Master Equation approaches.

In this work we shall introduce monomeric helical propensity  $\omega$ , energy penalty  $U$ , and the ground state energy  $\epsilon$  as relevant parameters to describe free energy barrier  $\Delta\mathcal{F}$  and translocation time  $\tau$ . These quantities have several advantages, since  $\omega$  for instance, takes into account the tendency of monomers to participate in helical structures, whereas  $U$  is a measure of the energy penalty per monomer in a non-native bond and  $\epsilon$  represents collective energy when all monomers have correct bonds. Following standard theoretical works on this subject [15,16], we use a coarse-grained description

\*Present address: Centro de Investigación y de Estudios Avanzados del IPN; josevelez@cinvestav.mx

†luis.olivares@uacm.edu.mx

focusing on poly- $X$  chain translocation. Our model is based on the calculation of the equilibrium partition function  $\mathcal{Z}$  for the  $N$ -monomer system. In contrast with other works, we shall consider two contributions for  $\mathcal{Z}$ . The first contribution accounts for the tendency of the chain to fold into a ground state conformation driven by energies  $U$ ,  $\epsilon$ , and helical propensity  $\omega$ , whereas the second contribution takes into account reduction of conformations by anchorage of the chain to the pore. We analyze in detail the phase-space of variables  $\Delta\mathcal{F}$  and  $\tau$  in terms of the set  $(N, \omega, \epsilon, U)$  focusing on poly- $X$  chains where  $X$  represents any of the 20 proteinogenic amino acid residues.

Our analytical results for the free energy of translocation show a characteristic bell-shape profile of  $\mathcal{F}$  as a function of the fraction of residues translocated. The free energy barrier  $\Delta\mathcal{F}$  is increased primarily by the value of the ground state energy  $\epsilon$  and the chain size  $N$ . The penalty energy  $U$  has also a drastic effect on the energy barrier, giving rise to two regimes defined here as the regimes of *flexible* and *rodlike* chains. Poly- $X$  chains in the flexible regime show the largest energy barriers in contrast to poly- $X$  chains in the rodlike regime. Results obtained in this work for non-driven translocation show that  $\tau$  exhibits a jump-like transition as a function of  $N$ . After such transition,  $\tau$  follows the exponential law  $\tau \simeq AN^\xi$ , where  $1.5 \leq \xi \leq 2.0$ . The specific value of  $\xi$  depends strongly on the energy penalty  $U$  and only slightly on the helical propensity  $\omega$ , as we shall discuss in detail. Constant  $A$  is in general, strongly dependent on parameters chosen here and it may, in fact, span several orders of magnitude for translocation time, from nanoseconds to seconds. Such a jump-like transition has been recently observed experimentally at least for poly(ethylene oxide) chains [17], therefore we expect that the type of transition predicted by our model might be very well spread in homopolymer chain translocation.

This paper is organized as follows. Section II provides details on the derivation of the model for poly- $X$  translocation from the equilibrium partition function. Section III describes results obtained for free energy of translocation  $\mathcal{F}$  and energy barrier  $\Delta\mathcal{F}$  whereas Sec. IV discusses the translocation time  $\tau$  profiles. Finally, Sec. V presents conclusions.

## II. MODEL FOR HOMOPOLYMER POLY- $X$ TRANSLOCATION

### A. Free energy of translocation

Let us consider poly- $X$  chains composed by  $N$  monomers of a single type  $X$  in which each monomer has a tendency to participate in a helical structure described by the helical propensity  $\omega$ . It will be assumed that translocation proceeds as the polymer threads a pore of small dimensions through a wall. During translocation, the polymer is considered as two chain segments joined at the pore with one end free and the other end anchored at the pore. The two chain segments, which we shall name *trans* and *cis*, are  $n$  and  $N - n$  monomers long, respectively, see Fig. 1. The free energy of the chain during translocation,  $\mathcal{F}(n)$ , is the sum of the free energies of the chain segment in the *trans* side,  $F_{trans}$ , and the chain segment in the *cis* side,  $F_{cis}$ . This is,

$$\mathcal{F}(n) = F_{cis}(N - n) + F_{trans}(n), \quad (1)$$

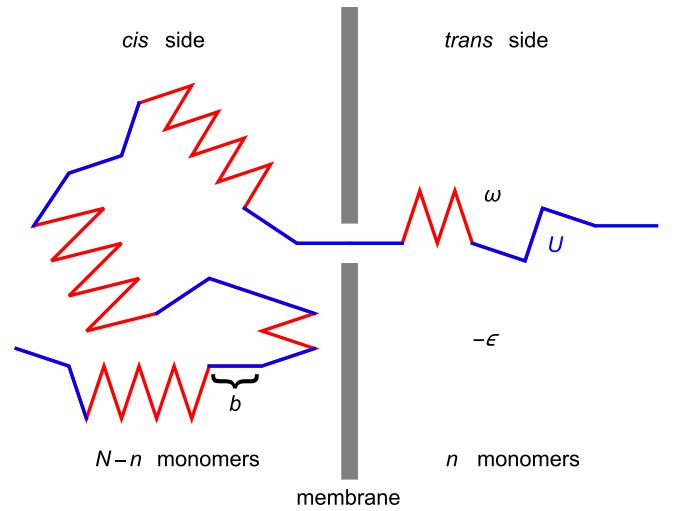


FIG. 1. Schematic representation of a poly- $X$  chain with helical propensity  $\omega$  in translocation. At time  $t$  there are  $n$  monomers in the *trans* side and  $N - n$  monomers in the *cis* side. Blue color is used to highlight monomers in non-native contacts with an energy penalty  $U$ . Red color indicates monomers in native contacts. Native energy when all monomers have native contacts is  $-\epsilon$ . Quantity  $b$  denotes bond length between two consecutive monomers.

where we have assumed that the same functional form for  $F$  applies at both sides of the wall.

The free energy  $F$  of each chain segment is obtained from the partition function  $\mathcal{Z}$ , which represents the effective number of available states at a temperature  $T$ . In free space, the number of states of one chain segment depends on the number of monomers in native and non-native contacts. In this work, we additionally consider the tendency of a single residue to participate in a helical structure. For such chains, the partition function in free space is  $Q_Z$ . Anchorage to the pore by one end of the chain reduces the available conformations as compared with the free-chain. Such reduction on the conformational space is accounted for by the steric factor  $Q_s$ . Both contributions give the total partition function  $\mathcal{Z} = Q_Z Q_s$ , therefore, the free energy of each chain segment is  $F = -k_B T \ln \mathcal{Z}$ , which can be written as  $F = F_Z + F_s$ , where  $F_Z = -k_B T \ln Q_Z$  and  $F_s = -k_B T \ln Q_s$ . Dimensions of the pore are considered negligible in such a sense that it only permits a single monomer to cross the boundary at a given time. In this way, the number  $n$  of monomers crossing the wall is the main dynamical variable describing the translocation, and computation of the free energy  $F$  relies on its functional form from  $F_Z$  and  $F_s$  via  $Q_Z$  and  $Q_s$ .

#### 1. Free energy of free-chain segments

To evaluate the free energy  $F_Z$  for a free-chain segment we shall use the idea proposed by Zwanzig [18] to compute the partition function  $Q_Z$ . In Zwanzig's model, a homopolymer chain poly- $X$  is composed of  $N$  identical monomers  $X$ . Each monomer has a number  $\nu + 1$  of potential conformations,  $\nu$  corresponds to non-native conformations, and there is a single native contact for each monomer. As a first approximation, we shall consider there is no correlation in conformations between close monomers, in this sense all monomer conformations are

independent. Let  $U > 0$  be the energy penalty per monomer in a non-native state and  $-\epsilon < 0$  be the overall energy of the chain when all monomers have native contacts,  $\epsilon$  is the ground state energy. For a given conformation of the chain with  $\alpha$  monomers having native contacts, it is proposed that the energy  $E_\alpha$  of the chain can be written as [18]

$$E_\alpha = (N - \alpha)U - \epsilon \delta_{N,\alpha} \quad \text{with} \quad 0 \leq \alpha \leq N. \quad (2)$$

It is worth realizing that in this description, the ground state of the polypeptide arises only as an *all-or-none* transition, i.e., only when all monomers have attained its native conformation. The degeneracy  $g_\alpha$  associated to energy level  $E_\alpha$  can easily be calculated to give [18]

$$g_\alpha = \nu^{N-\alpha} \binom{N}{\alpha}, \quad (3)$$

where  $\nu$  is the number of non-native potential conformations per monomer. From this, the equilibrium partition function  $\mathcal{Q}_Z$  can be obtained:

$$\mathcal{Q}_Z = e^{\beta\epsilon} + (1 + \nu e^{-\beta U})^N - 1, \quad (4)$$

from which the free energy is  $F_Z(n) = -k_B T \ln \mathcal{Q}_Z(n)$ , where  $n$  is the number of monomers translocated,  $k_B$  is the Boltzmann constant, and  $T$  is the temperature.

Recently, it has been shown that if we restrict to helix-coil transitions,  $\nu$  may be in fact related to the helical propensity  $\omega$  per monomer, which is associated with a conditional probability of a given monomer to participate in a helical state if its neighbor is already in a helical one [19]. The argument goes as follows. The free energy change per monomer  $\Delta f$  during a helix-coil transition can be written as

$$\Delta f = \Delta h - T \Delta s, \quad (5)$$

where  $\Delta h$  and  $\Delta s$  represent the enthalpy and entropy change per monomer between non-native and native conformations. During a helix-coil transition, the Lifson-Roig model states the helical propensity is  $\omega = -RT \ln(\Delta f)$  [20]. Assuming each monomer has  $\nu$  available non-native states, the change of entropy towards the native conformation is  $\Delta s = -RT \ln(\nu)$  [19]. With these two expressions and from Eq. (5), the relation between the helical propensity  $\omega$  and the number of non-native states  $\nu$  becomes [19]

$$\omega = \frac{1}{\nu} \exp(-\beta \Delta h), \quad (6)$$

where  $\beta = 1/RT$  is the Boltzmann factor. From the last expression,  $\omega$  is inversely proportional to the number of non-native conformations  $\nu$ . This implies that the number of non-native conformations  $\nu$  is reduced as the helical propensity  $\omega$  increases. This agrees with  $\omega$  viewed as a measure of the tendency to participate in a helical structure. A helical transition is favored by increasing the helical propensity or by decreasing the number of sampled conformations. The enthalpy change per monomer  $\Delta h$  also contributes positively or negatively to the helical propensity. For instance, if the transition is enthalpically favored,  $\Delta h < 0$ , the term  $e^{-\beta \Delta h}$  is an increasing function and therefore  $\omega$  is increased, an expected result if the helical transition is favored. Solving for  $\nu$  from Eq. (6) and by substitution in Eq. (4), gives

the full partition function  $\mathcal{Q}_Z$  as a function of the helical propensity [19]:

$$\mathcal{Q}_Z(n) = e^{\beta\epsilon} + \left(1 + \frac{1}{\omega} e^{-\beta(U+\Delta h)}\right)^n - 1, \quad (7)$$

from which the free energy contribution of a poly- $X$  chain of  $n$  monomers in thermal equilibrium with a given  $\omega$  is obtained as

$$F_Z(n) = -k_B T \ln \mathcal{Q}_Z(n). \quad (8)$$

## 2. Free energy of anchored-chain segments

For the second contribution  $F_s$ , we use the steric constraint factor  $\mathcal{Q}_s$ . Anchorage to the pore by one end of the chain segment blocks some of the potential conformations. In comparison with free space, there exists a reduction in the number of conformations. For the anchored chain, it can be shown that the corresponding partition function  $\mathcal{Q}_s$  is a power-law respect to the number of monomers which goes as  $\mathcal{Q}_s \sim n^{\gamma-1}$ , with  $\gamma > 0$  [15,16]. It has been shown that exponent  $\gamma$  can take values 0.5,  $\simeq 0.69$ , and 1, for a Gaussian, self-avoiding, and rodlike chain, respectively, [15,16]. Then, the contribution to the free energy of translocation due to the anchorage to the pore by one end of the chain segment can be written as [16]

$$F_s(n) = (1 - \gamma)k_B T \ln(n), \quad (9)$$

where  $n$  is the number of translocated monomers across the pore.

Summing up the last two free energy contributions, the free energy of one chain-segment can be written as  $F(n) = k_B T \ln \left(\frac{n^{1-\gamma}}{\mathcal{Q}_Z(n)}\right)$ . Considering both chain segments of the *cis* and *trans* sides the free energy of translocation of the poly- $X$  that possess helical propensity is

$$\mathcal{F}(n) = k_B T \ln \left\{ \frac{((N - n)n)^{1-\gamma}}{\mathcal{Q}_Z(N - n)\mathcal{Q}_Z(n)} \right\} \quad (10)$$

where  $n$  is the number of translocated monomers at a given time.

## B. Translocation time

Translocation time is computed following previous works where the dynamics of translocation is treated in terms of average first-passage times. Let  $P(n, t | n_0, 0)$  be the probability distribution function that  $n$  monomers have crossed the wall at time  $t$ , starting from  $n_0$  at  $t = 0$ . Such probability distribution  $P$  can be computed by solving a Fokker-Planck equation [21] in terms of the chain diffusivity during translocation  $D(n)$  and the free energy  $\mathcal{F}(n)$  [15]. In general,  $P$  satisfies a Fokker-Planck equation [15,22]:

$$\frac{\partial P}{\partial t} = \mathcal{L}_{\text{FP}}(n)P, \quad (11)$$

where  $\mathcal{L}_{\text{FP}}$  is the Fokker-Planck operator

$$\mathcal{L}_{\text{FP}}(n) = (1/b^2)(\partial/\partial n)D(n) \times \exp(-\beta\mathcal{F}(n))(\partial/\partial n)\exp(\beta\mathcal{F}(n)), \quad (12)$$

where  $b$  is the monomer length. From a series expansion, it can be shown that  $P$  also obeys the so-called backward

Kolmogorov equation in terms of the adjoint Fokker-Planck operator  $\mathcal{L}_{FP}^\dagger$  acting on variable  $n_0$  [21], i.e.,

$$\frac{\partial P}{\partial t} = -\mathcal{L}_{FP}^\dagger(n_0)P. \quad (13)$$

Knowledge of  $P$  permits to compute average quantities. However, its computation may be avoided if the translocation time is the goal of such calculation [21], as discussed below.

Translocation time  $\tau$  is defined as the average time to reach for the first time  $n = N - 1$ , starting at  $n = 1$ . In this formalism  $\tau$  corresponds to the mean first-passage time and can be calculated accordingly. Let  $\rho$  be its probability density, then the average time for translocation is the first moment of  $\rho$  given by

$$\tau(N, n_0) = \int_0^\infty t \rho(N, t | n_0, 0) dt. \quad (14)$$

In order to calculate  $\rho$ , we shall consider the probability  $q$  of not reaching  $n = N - 1$  at time  $t$ . This is,

$$q = \int_1^{N-1} P(n, t | n_0, 0) dn. \quad (15)$$

By its definition,  $q$  can also be computed by means of  $\rho$ . Integration of  $\rho$  on a time interval is the probability of reaching  $n$  at that interval, then, the complementary probability, the probability of not reaching  $n = N - 1$  at that interval, is  $q$ :

$$q = \int_t^\infty \rho(N, t | n_0, 0) dt. \quad (16)$$

The derivative of  $q$ , Eq. (16), over  $t$  gives  $\rho$  as a function of  $q$

$$\rho(N, t | n_0, 0) = -\frac{\partial q(n_0, t)}{\partial t} \quad (17)$$

and substitution of  $q$ , Eq. (15), in  $\rho$ , Eq. (17), gives  $\rho$  in terms of  $P$ :

$$\rho(N, t | n_0, 0) = -\int_1^{N-1} \frac{\partial P(n, t | n_0, 0)}{\partial t} dn. \quad (18)$$

The translocation time  $\tau$ , in terms of  $P$ , is obtained after substitution of  $\rho$  Eq. (18), back in Eq. (14), and after integration over  $t$ , to give

$$\tau(N, n_0) = \int_1^{N-1} \int_0^\infty P(n, t | n_0, 0) dt dn. \quad (19)$$

In order to obtain  $\tau$  in terms only of  $D(n)$  and  $\mathcal{F}(n)$ , the adjoint Fokker-Planck operator is applied to both sides of Eq. (19) giving

$$\mathcal{L}_{FP}^\dagger \tau(N, n_0) = \int_1^{N-1} \int_0^\infty \mathcal{L}_{FP}^\dagger P(n, t | n_0, 0) dt dn. \quad (20)$$

Substitution of the integrand, by the left-hand side of Eq. (13), and after integration, it results in a second-order differential equation for the translocation time  $\tau$  [21] given by

$$\mathcal{L}_{FP}^\dagger \tau(N, n_0) = -1. \quad (21)$$

Finally, integration of the last expression over  $n_0$ , Eq. (21), applying reflecting and absorbing boundary conditions at  $n_0 =$

1 and  $n_0 = N - 1$ , results in [15]

$$\tau = \int_1^{N-1} dn \frac{b^2}{D(n)} \exp\left(\frac{\mathcal{F}(n)}{k_B T}\right) \int_1^n dm \exp\left(-\frac{\mathcal{F}(m)}{k_B T}\right). \quad (22)$$

Here, the translocation time  $\tau$  depends on the following parameters:  $b$ ,  $N$ , and  $\omega$ ,  $U$ , and  $\epsilon$  by means of the free energy  $\mathcal{F}$ . Appropriate values for these parameters are discussed in the Appendix and throughout this work. Translocation time also depends on the chain diffusivity  $D(n)$  during translocation. In this work we shall consider  $D = 10 \text{ m}\mu^2/\text{s}$  which is a representative value within the order of magnitude of protein and protein-like molecules diffusing inside cells. For a non-constant  $D$  the interested reader is referred to [23].

### III. RESULTS

#### A. Free energy for homopolymer translocation: The energy barrier

We computed the free energy of translocation for poly- $X$  chains, where  $X$  represents any of the 20 proteinogenic amino acids. To compute the free energy from Eq. (10) first, we assigned values for the monomer length  $b$ , the helical propensity of amino acids  $\omega$ 's, the energy penalty  $U$ , and the native energy  $-\epsilon$ , as discussed in the Appendix. To show the general behavior of physical variables we choose a poly-T (poly-threonine) chain. Threonine helical propensity is  $\omega_{\text{THR}} = 0.14$ , a value in between the smallest and the largest helical propensities of amino acids proline  $w_{\text{PRO}} = 0.001$  and  $w_{\text{ALA}} = 1.64$ , according to the scale proposed in [20]. Results for poly-proline (poly-P) and poly-alanine (poly-A) are also analyzed as limiting cases. The free energy of translocation  $\mathcal{F}$  as a function of the percentage of translocated monomers  $n/N$  is shown in Fig. 2 for a poly-T chain of size  $N = 500$ . The  $\mathcal{F}$  profile obtained is in agreement with an umbrella-like profile reported in other works [10,11,15]. Associated with the free energy  $\mathcal{F}$  profile, an energy barrier  $\Delta\mathcal{F}$  can be defined

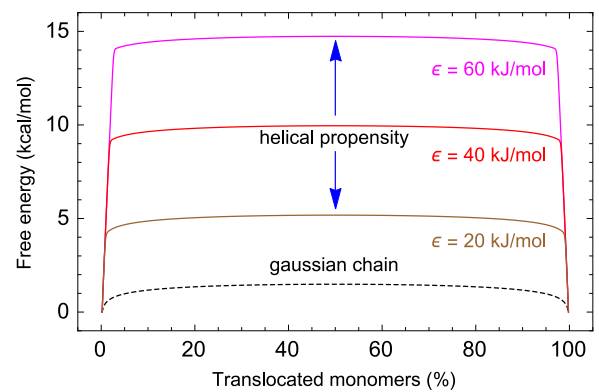


FIG. 2. Free energy  $\mathcal{F}$  of translocation as a function of the percentage of translocated monomers for a poly- $T$  chain of  $N = 500$  monomers.  $\mathcal{F}_{\min} = \mathcal{F}(1)$  was used as a reference point. We set the monomer energy penalty  $U = 1 \text{ kJ mol}^{-1}$  and the helical propensity  $\omega_{\text{THR}} = 0.14$ . Different values of the ground state energy  $\epsilon$  were considered. Dashed line shows the free energy  $\mathcal{F}$  for a Gaussian chain.

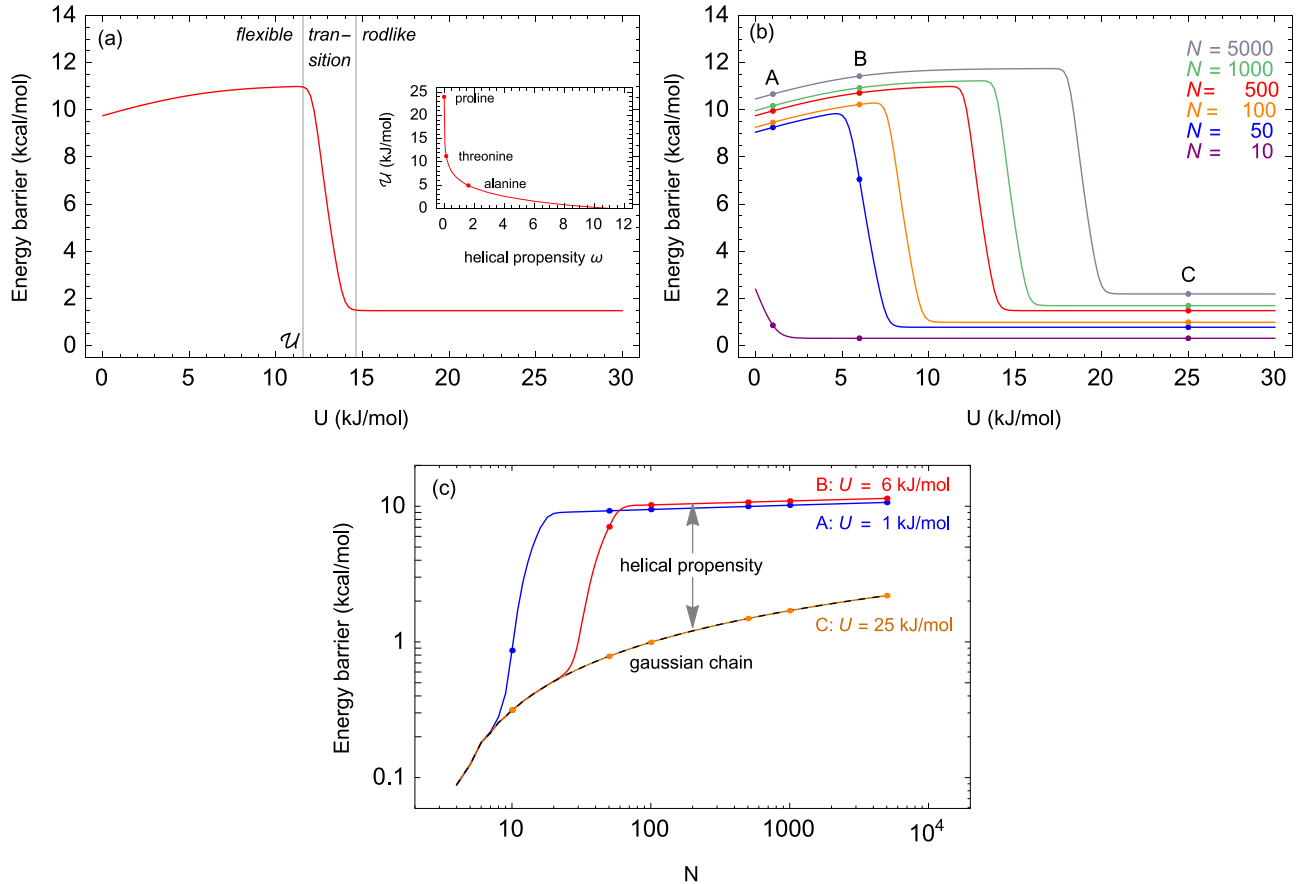


FIG. 3. (a) Free energy barrier  $\Delta\mathcal{F}$  as a function of energy penalty  $U$  for a poly-T chain of size  $N = 500$  monomers. Based on the profile of  $\Delta\mathcal{F}$ , we defined three regimes: *flexible*, *transition*, and *rodlike* (delimited by lines). A threshold energy penalty  $\mathcal{U}$  delimits the *flexible* regime where  $\Delta\mathcal{F}$  abruptly decreases. Inset shows  $\mathcal{U}$  as a function of the helical propensity  $\omega$ . (b) and (c) show the free energy barrier  $\Delta\mathcal{F}$  for poly-T chains for different chain size  $N$  and energy penalty  $U$ , respectively. In all plots we set the ground state energy  $\epsilon = 40 \text{ kJ mol}^{-1}$ . Dashed line shows the free energy barrier  $\Delta\mathcal{F}$  for a Gaussian chain.

as the difference between the maximum  $\mathcal{F}_{\max}$  and minimum  $\mathcal{F}_{\min}$  observed during translocation, where  $\mathcal{F}_{\min} = \mathcal{F}(1)$  is set as a reference and corresponds to free energy  $\mathcal{F}$  when only one residue has been translocated.

Results obtained here show that the energy barrier  $\Delta\mathcal{F}$  increases with ground state energy  $\epsilon$ . For the poly-T chain  $\Delta\mathcal{F}$  increases one order of magnitude, from 5.2 to 14.7 kcal/mol, as  $\epsilon$  increases from 20 to 60 kJ/mol. In comparison, the energy barrier observed for the Gaussian chain (GC) saturates approximately at 1.5 kcal/mol, which is almost one order of magnitude smaller than the ones predicted when the helical propensity is taken into account as shown in this work. It is worth to underline that the Gaussian chain is a single-parameter theory with the monomer length  $b$  as its only variable.

Results shown in Fig. 2 for a poly-T chain refer to the case when the energy penalty is  $U = 1 \text{ kJ/mol}$ . Changing the energy penalty  $U$  has a great impact on the energy barrier as shown in Fig. 3(a) for the ground state energy  $\epsilon = 40 \text{ kJ/mol}$ . Behavior observed for  $\Delta\mathcal{F}$  as a function of  $U$  suggests it is possible to identify three regimes depending on the value of variable  $U$ . For small  $U$ ,  $\Delta\mathcal{F}$  is an increasing function up till it reaches a maximum at a threshold  $U$ ,  $\mathcal{U}$ . From  $\mathcal{U}$ , in a range of  $\approx 3 \text{ kJ/mol}$ ,  $\Delta\mathcal{F}$  abruptly decreases up to a

constant value. For any  $U$  greater than  $\mathcal{U} + 3 \text{ kJ/mol}$ , the free energy barrier remains constant. Large values of  $U$  refer to stiff chains while small ones refer to flexible chains, therefore, according to the behavior of the free energy barrier  $\Delta\mathcal{F}$  we define these as the regimes of *flexible* (large  $\Delta\mathcal{F}$ ), *transition* (steep decreasing  $\Delta\mathcal{F}$ ), and *rodlike* (small  $\Delta\mathcal{F}$ ) poly-X chains. The main result here is that poly-X chains in the flexible regime would experience the highest energy barrier  $\Delta\mathcal{F}$ , in contrast to poly-X chains in the rodlike regime that would face the smallest energy barrier  $\Delta\mathcal{F}$ .

To know whether the presence of the three regimes in the free energy barrier is a feature of this model we analyzed if any helical propensity  $\omega$  gives rise to the  $\mathcal{U}$  that delimits the flexible regime. Figure 3(a) inset indeed shows there exists a  $\mathcal{U}$  in a wide range of helical propensities  $\omega$ 's, that includes proteinogenic chains and poly-X chains with larger helical propensities.  $\mathcal{U}$  is decreased by the helical propensity as  $\mathcal{U} \simeq \omega^{-1}$ , this means that for poly-X chains with larger helical propensities, like poly-alanine chains, the flexible regime is shortened to a smaller range of  $U$ 's. We also analyzed dependency on chain size. Figure 3(b) shows the free energy barrier for poly-T chains  $N = 10, 50, 100, 500, 1000, 5000$ . Chains of size  $N > 10$  possess the three regimes. The free energy barrier is strengthened by chain size, therefore the

flexible regime is lengthened. These results combined imply that poly- $X$  chains possess the three regimes of the free energy barrier and that they may be shortened or lengthened by specific combinations of  $(\omega, N)$ .

From the results of Fig. 3(b) we can also analyze the free energy barrier  $\Delta\mathcal{F}$  as a function of the chain size  $N$  at fixed  $U$ . Points indicated by A, B, and C in Fig. 3(b) refer to the free energy barrier for selected  $U$ 's, 1, 6 y 25 kJ/mol, respectively. As a function of  $N$ , points in A show a transition between the transition regime and the flexible regime for small  $N$ . The huge distance between points for small  $N$  indicates an abrupt increase in the free energy barrier. Points in B indicate that first there is a transition between the rodlike regime and the transition regime, followed by the transition to the flexible regime. The same change is expected as in points of case A. For points in C, there is only one regime, the rodlike regime, independent of the chain size. In this case, the distance between points indicates a smooth and small increase in the free energy barrier as a function of  $N$ . Computation of the free energy barrier  $\Delta\mathcal{F}$  as a function of  $N$  for cases A, B, and C, in Fig. 3(c), corroborates this behavior. Indeed, there is a huge increase for the energy barrier for cases A and B for chains

up to a few tens of monomers, and the increase of the energy barrier for case C is smooth and monotonic. Interestingly, it seems that the main contribution to the free energy for small chains comes from the anchoring of the chain to the pore, as the free energy barrier is similar to that of the GC model (dashed line). Similarly case C occurs for any chain size in the rodlike regime, in this case the main impediment for a rodlike chain to translocate is its anchoring to the pore. The choice of the penalty energy  $U$  has a deep effect on the free energy barrier  $\Delta\mathcal{F}$  as a function of the number of monomers  $N$ . The free energy barrier  $\Delta\mathcal{F}$  can display an increase up to two orders of magnitude for small poly- $X$  chains for  $N < 10$  residues in comparison to poly- $X$  chains with  $N \simeq 10^3$  residues. Such high increase is a central feature of our model, in comparison with the GC model which does not show such behavior. A more technical and detailed discussion on the role of penalty energy  $U$  and the values allowed for this parameter is added to the Appendix for further details.

A full perspective on the behavior of  $\Delta\mathcal{F}$  as a function of size  $N$  and helical propensity  $\omega$  is provided simultaneously by Fig. 4 in a three-dimensional plot. For a fixed  $N$ , Fig. 4(a) shows that  $\Delta\mathcal{F}$  increases slightly as the helical propensity

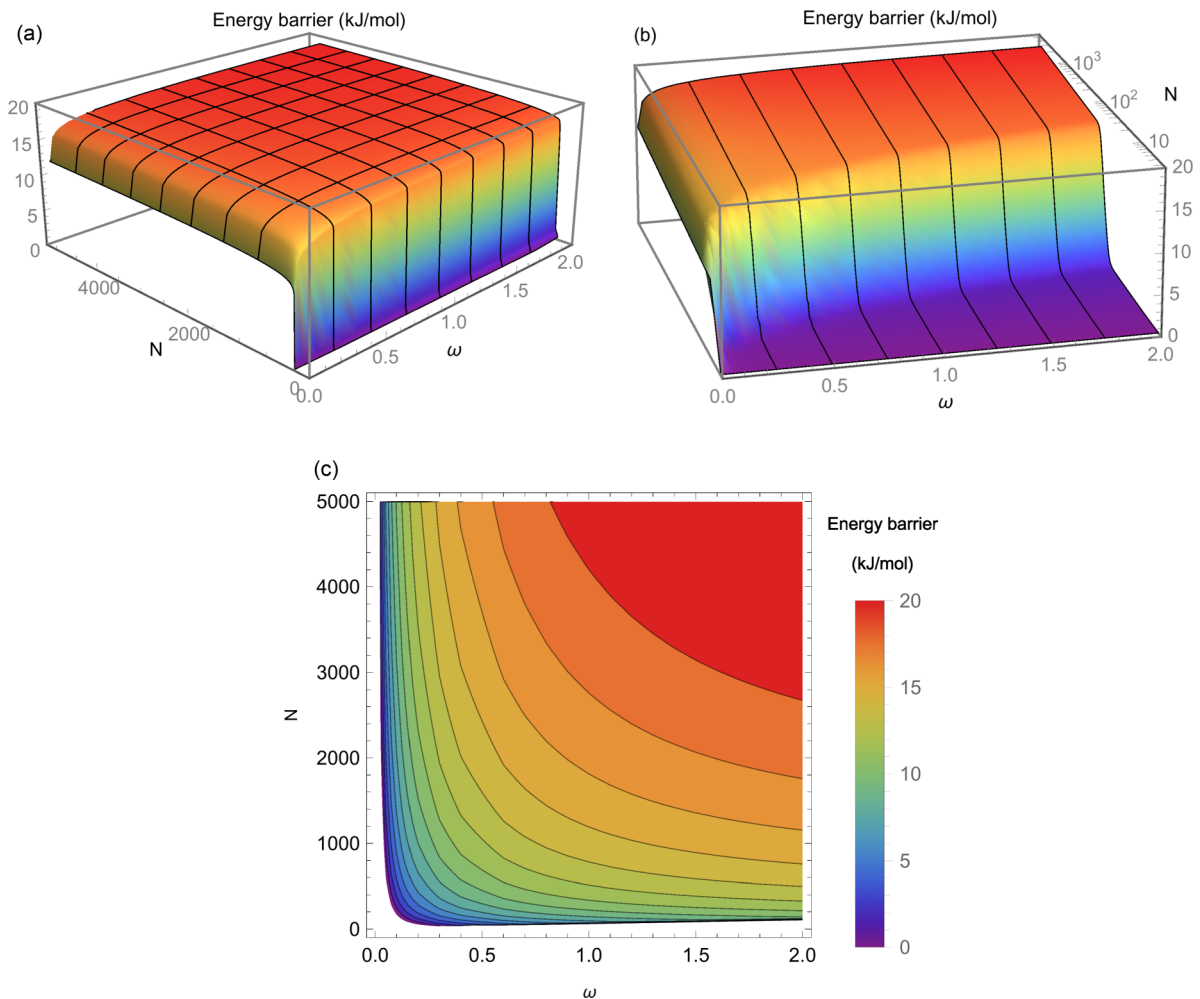


FIG. 4. (a) and (b) 3D plots of the free energy barrier  $\Delta\mathcal{F}$  as a function of helical propensity  $\omega$  and chain size  $N$  or  $\ln N$ , respectively. (c) Contour plot of the free energy barrier  $\Delta\mathcal{F}$  in terms of both relevant variables  $N$  and  $\omega$ . In all plots we set  $U = 1 \text{ kJ mol}^{-1}$  and  $\epsilon = 40 \text{ kJ mol}^{-1}$ .

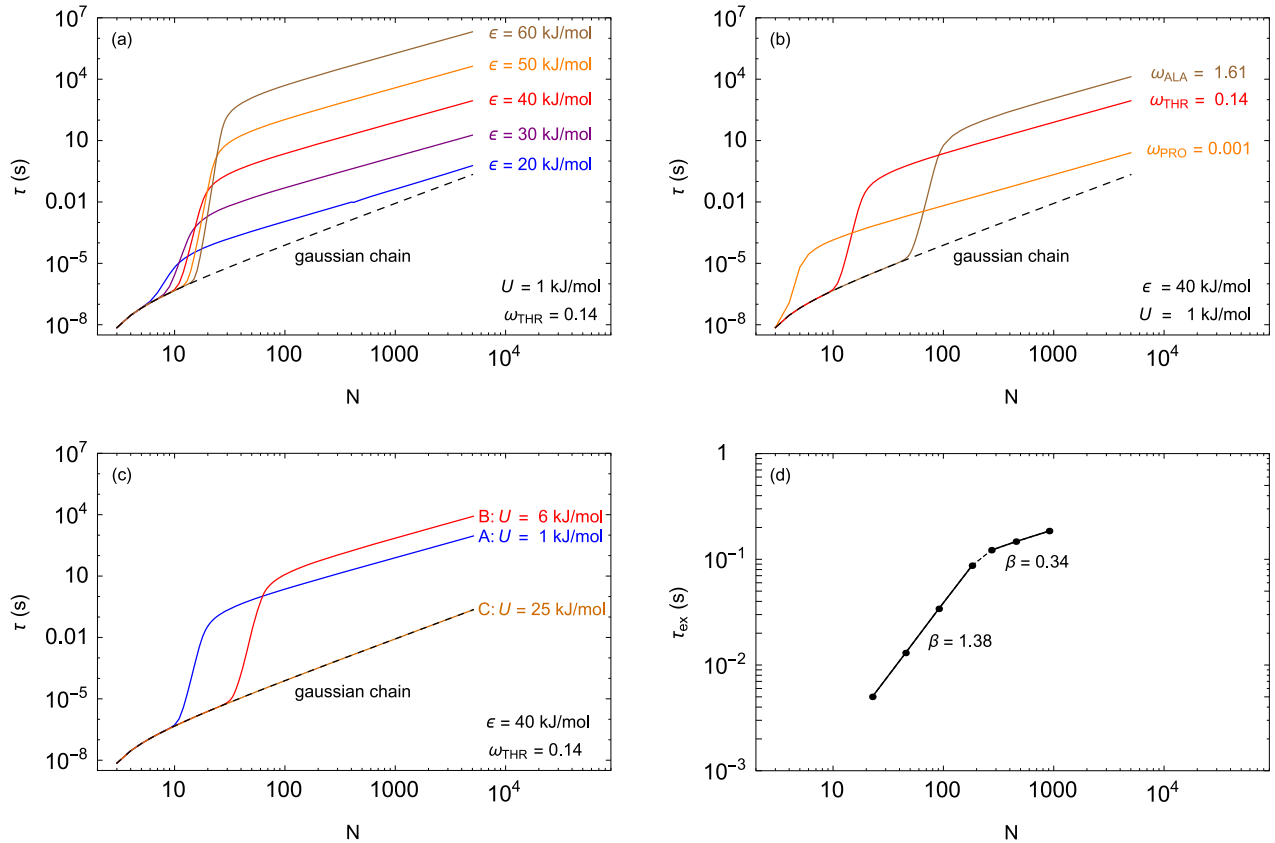


FIG. 5. Translocation time  $\tau$  as a function of poly- $X$  chain size  $N$  for different (a) ground state energy  $\epsilon$ , (b) helical propensity  $\omega$ , and (c) energy penalty  $U$ . (a) and (c) refer to poly-T chains at fixed  $U = 1 \text{ kJ mol}^{-1}$  and  $\epsilon = 40 \text{ kJ mol}^{-1}$ , respectively. (b) shows results for poly-ALA, poly-THR, and poly-PRO chains at fixed  $\epsilon = 40 \text{ kJ mol}^{-1}$  and  $U = 1 \text{ kJ mol}^{-1}$ . (d) experimental data of translocation times (dots) for poly(ethylene oxide) chains are from Ref. [17] along with the  $\beta$  exponents. Lines were drawn through points to highlight the transition jump predicted in our model.

also increases. The increase for an average size chain ranges from  $12 \text{ kJ mol}^{-1}$  up to  $18 \text{ kJ mol}^{-1}$  by changing the helical propensity from that of proline ( $\omega_{\text{PRO}} = 0.001$ ) to that of alanine ( $\omega_{\text{ALA}} = 1.64$ ). For small chains, the number of residues  $N$  also plays a significant role in  $\Delta F$  as Fig. 4(b) shows. The free energy barrier increases from a few  $\text{kJ mol}^{-1}$  for chains with  $N \simeq 10$  residues up to  $\simeq 20 \text{ kJ mol}^{-1}$  for chains with larger sizes  $N > 10^2 - 10^3$  residues. Interestingly, there is a plateau in the increase of the free energy barrier. Once an upper bound is reached in  $\Delta F$  for a particular poly- $X$  chain, a further increase in the number of residues  $N$  does not imply a significant further increase in  $\Delta F$ . Figure 4(b) also shows that threshold size  $N^*$  where this plateau is reached increases as the helical propensity  $w$  also increases. Figure 4(c) is a contour plot of the free energy barrier, it shows that a combination of  $(w, N)$  may modulate the free energy barrier. Larger values of  $\Delta F$  can arise as a combination of larger values of  $N$  and  $\omega$ . On the contrary, small poly- $X$  chains with less helical propensity have small free energy barriers and therefore would be expected to translocate faster.

**B. Translocation time  $\tau$  for homopolymers: The transition jump**

Translocation time  $\tau$  for ideal poly- $X$  chains is computed according to Eq. (22). Figure 5 shows the translocation time

$\tau$  as a function of chain size  $N$ . Panels (a), (b), and (c) show results for different values of one of the variables  $\epsilon$ ,  $\omega$ , or  $U$ , respectively, while the rest of the variables are kept fixed. The central result obtained in this work shows that  $\tau$  exhibits a transition, which is characterized by a sudden and steep increase of several orders of magnitude as  $N$  approaches a critical value  $\mathcal{N}$ . Location of the transition  $\mathcal{N}$  occurs at small  $N$ . Location  $\mathcal{N}$  for poly-T chains is  $\mathcal{N} \sim 10$ , and this value is barely increased by the ground state energy in the range  $20 < \epsilon < 60 \text{ kJ mol}^{-1}$ , see Fig. 5(a). Helical propensity  $\omega$  has a larger effect on the location of the transition. Location  $\mathcal{N}$  is increased from a few to a few tens of monomers whether the helical propensity is increased from the smallest,  $w_{\text{PRO}} = 0.001$ , to the highest,  $w_{\text{ALA}} = 1.64$ , Fig. 5(b). In comparing  $\epsilon$  and  $\omega$ , the helical propensity mainly determines the location of the transition.

The steep increase in the translocation time, i.e., the *jump* of the transition, may span several orders of magnitude. For poly-T chains at fixed  $\epsilon = 40 \text{ kJ mol}^{-1}$ , the jump from  $\tau \sim 10^{-6} \text{ s}$  to  $\tau \sim 10^0 \text{ s}$  spans six orders of magnitude, see Fig. 5(a). At fixed  $\epsilon = 60 \text{ kJ mol}^{-1}$ , the jump from  $\tau \sim 10^{-6} \text{ s}$  to  $\tau \sim 10^3 \text{ s}$  spans nine orders of magnitude. For poly- $X$  chains ranging from the smallest to the largest helical propensities, the jump spans 3 and 6 orders of magnitude, respectively, see Fig. 5(b). Both the ground state energy  $\epsilon$  and the helical propensity  $\omega$

TABLE I. Range of values for exponent  $\xi$  and constant  $A$ . Values of  $\xi$  and  $A$  were obtained by fitting the translocation time to the power law  $\tau = AN^\xi$  after the transition jump. Values of  $\xi$  and  $A$  were computed by changing one parameter while maintaining the rest fixed. For  $\omega$ , we fixed  $U = 1 \text{ kJ mol}^{-1}$  and  $\epsilon = 40 \text{ kJ mol}^{-1}$ . For  $U$  and  $\epsilon$ , we considered poly-T chains, at fixed  $\epsilon = 40 \text{ kJ mol}^{-1}$  or fixed  $U = 1 \text{ kJ mol}^{-1}$ , respectively. Arrows indicate variation of the parameter range or the variable range.

Parameter	Parameter range	Variable range	
		Exponent $\xi$	Constant $A$
$\omega$	0.001 $\rightarrow$ 1.61	1.50 – 1.53	65 $\mu\text{s}$ $\rightarrow$ 37 ms
$U$ (kJ/mol)	1 $\rightarrow$ 25	1.50 – 2.00	2.5 ms $\rightarrow$ 7.7 ns
$\epsilon$ (kJ/mol)	60 $\rightarrow$ 20	1.50 $\rightarrow$ 1.60	6 s $\rightarrow$ 0.6 $\mu\text{s}$

are related to the transition jump as well. Figure 5(c) shows the transition of  $\tau$  at a critical length  $\mathcal{N}$  may be overridden by large  $U$ . Poly-T chains at  $U = 25 \text{ kJ mol}^{-1}$  do not exhibit such transition. Indeed, poly-X chains in the rodlike regime lack such transition. For poly-T chains, the rodlike regime is defined by  $U \geq 20 \text{ kJ mol}^{-1}$ , as shown in Fig. 3(b), therefore  $U$ 's within this range will render translocation times without transition. The translocation time of the poly-T chain lacking the transition matches the translocation time of the GC model, dashed line in Fig. 5(c). This means that the main contribution to the translocation time of poly-T chains in the rodlike regime comes primarily from the anchoring to the wall. Similar results are expected for any poly-X chain in the rodlike regime.

Translocation time profiles show the transition separates two regions defined by  $N \leq \mathcal{N}$  and  $N \geq \mathcal{N}$ . Our results shown in Fig. 5 suggest that the translocation time  $\tau$  in each region follows an exponential law  $\tau \sim AN^\xi$ . Exponent  $\xi$  and constant  $A$  were computed by fitting the power law for each region. For  $N \leq \mathcal{N}$ , the translocation time of poly-X chains follows the power law of the Gaussian chain,  $\xi = 2$  [15]. According to the parameters used in this work, for the Gaussian chain  $A$  is 9 ps. For  $N \geq \mathcal{N}$ , the transition observed for the translocation time  $\tau$  determines the characteristic exponent  $\xi$ . Table I shows exponent  $\xi$  after the transition in varying  $\epsilon$ ,  $U$ , and  $w$ . Exponent  $\xi$  for poly-T chains ranges from 1.65 to 1.50 for  $20 \text{ kJ mol}^{-1} < \epsilon < 60 \text{ kJ mol}^{-1}$ , at  $U = 1 \text{ kJ mol}^{-1}$ . Changing  $U$  from small to large  $U$ , from the flexible to the rodlike regime, increases  $\xi$ . For poly-T chains at  $\epsilon = 40 \text{ kJ mol}^{-1}$ ,  $\xi$  ranges from 1.50 to 2.00. The helical propensity  $\omega$  barely changes the characteristic exponent, ranging from  $\xi = 1.50$ –1.53, in varying  $w$  from that of poly-P to that of Poly-A chains. For poly-T chains the exponent is  $\xi \sim 1.5$ , at fixed  $U = 1 \text{ kJ mol}^{-1}$  and  $\epsilon = 40 \text{ kJ mol}^{-1}$ . Exponent  $\xi$  depends primarily on ground state energy  $\epsilon$ , and penalty energy  $U$ , as shown in Table I, in comparing, the energy penalty  $U$  has the deepest effect on  $\xi$ .

Constant  $A$  also was analyzed as a function of  $\epsilon$ ,  $U$ , and  $w$ , as summarized in Table I.  $A$  spans several time scales according to a combination of the parameters used in this work. For the cases considered in Fig. 5(a), for poly-T chains of different  $\epsilon$ , from 20 to  $60 \text{ kJ mol}^{-1}$ , constant  $A$  is located between  $0.6 \mu\text{s} < A < 6 \text{ s}$ . This range covers the usual experimental polypeptide translocation times. Changing

$U$  shifts the time scale of  $A$  to range from nanoseconds to milliseconds, the smaller scale obtained for rodlike chains. Helical propensity has a lesser effect on  $A$ , changing the time scale from microseconds to milliseconds, with large  $\omega$ 's giving rise to slow translocation. The values predicted for  $A$  within our model span several orders of magnitude and are larger than the one predicted for the Gaussian chain model.

In our model,  $\tau$  spans several orders of magnitude due to a combination of helical propensity  $\omega$ , energy penalty  $U$ , and ground state energy  $\epsilon$ . Each variable has a substantial effect on the onset of the transition and the order of magnitude of the translocation time after the transition jump, a phenomenon that is absent in Gaussian chain models. In terms of variations in helical propensity  $\omega$ , constant  $A$  spans three orders of magnitude after the transition jump suggesting that poly-X chains of different residues  $X$  might exhibit observable variations for the translocation times. Additionally, poly-X chains of small  $\omega$  will be prone to translocate faster after the transition jump.

Results predicted by our model were compared with data from experimental work of diffusion-driven translocation of poly(ethylene oxide) (PEO) chains [17]. Data of permeation rate constants  $k_{\text{ex}}$  of PEO chains translocating through nanopores of hollow polymeric capsules showed two regimes as a function of size  $N$ . In each regime, the permeation constants seemed to follow a power-law dependence on  $N$ ,  $k_{\text{ex}} \sim N^\beta$ . We used the PEO permeation constants and plotted the translocation times as a function of  $N$  obtained from the exchange rates as  $\tau_{\text{ex}} = 1/k_{\text{ex}}(N)$ , see Fig. 5(d). Results exhibit the transition jump predicted by our model at small  $N$  followed by a monotonic increase according to the power law  $\tau \sim N^\beta$  as a function of  $N$ . Exponents  $\beta$  were obtained by fitting  $k_{\text{ex}}$  to the power law  $k_{\text{ex}} \sim N^\beta$  according to the data reported in [17]. The profile predicted for the translocation time as confirmed by PEO chains is a feature expected for poly-X chains in diffusion-driven translocation.

#### IV. CONCLUSIONS

In this work we proposed a model for poly-X translocation based on equilibrium statistical mechanics and the Focker-Planck equation. We computed the free energy  $\mathcal{F}$ , the free energy barrier  $\Delta\mathcal{F}$ , and the translocation time  $\tau$  of translocation for poly-X proteinogenic chains as a function of ground state energy  $\epsilon$ , energy penalty  $U$ , and helical propensity  $\omega$ . In this work,  $\epsilon$  refers to the energy of the chain when all monomers are in native contacts,  $U$  is the energy penalty for one monomer in a non-native contact, and  $\omega$  is a measure of the tendency of one monomer to participate in a helical structure. The model proposed includes the folding propensity of the chain and the reduction of configurational degrees of freedom due to anchorage of the chain to the pore through which it translocates. The inclusion of the helical propensity  $\omega$  and the energy scales  $U$  and  $\epsilon$  is a novel feature not present in other works on the subject.

The first result showed the free energy  $\mathcal{F}$  of translocation exhibits a characteristic umbrella-like profile as observed in other works. Such profile imposes an energy barrier for poly-X translocation. The free energy maximum and minimum were used to characterize the free energy barrier  $\Delta\mathcal{F}$  required for



translocation. To provide a detailed exploration of the phase space of this model, we analyzed how the energy barrier  $\Delta\mathcal{F}$  and the translocation time  $\tau$  depend on the ground state energy  $\epsilon$ , the energy penalty  $U$ , and the helical propensity  $\omega$ . As shown in Fig. 2, the free energy barrier  $\Delta\mathcal{F}$  is increased by the ground state energy  $\epsilon$ . This is an expected feature since in the partition function  $\mathcal{Q}$  for poly- $X$  proteinogenic chains the ground state energy  $\epsilon$  is a measure of its location within the folded energy landscape. The free energy barrier  $\Delta\mathcal{F}$  showed a particular sensitive dependence on  $U$ , which gives rise to three different regimes according to the profile of the free energy barrier. We have called these regimes *flexible*, *transition*, and *rodlike* regimes. The central feature observed is that poly- $X$  chains in the flexible regime show the largest energy barriers while those in the rodlike regime show the smallest. Occurrence of the flexible regime, limited to a range of small  $U$ 's, was modulated by  $\omega$  and  $N$ . The flexible regime was shortened by the helical propensity  $\omega$  and stretched by chain size  $N$ . The free energy barrier  $\Delta\mathcal{F}$  showed an increasing dependency on chain size  $N$ . For small  $N$ , there was a jump of almost two orders of magnitude as  $N$  approached a critical value  $\mathcal{N}$  reaching a plateau for  $N \gg \mathcal{N}$ . A direct implication of these results is that a combination of helical propensities  $\omega$ 's and chain size  $N$  could change drastically the free energy barrier  $\Delta\mathcal{F}$ . Since  $\omega$  is a measure of the tendency of a residue to be in a folded state, poly- $X$  chains with larger values of  $\omega$  would tend to be in the folded structure therefore requiring a higher energy to translocate.

Translocation time  $\tau$  was also analyzed in terms of the set  $(\epsilon, \omega, U)$  and  $N$ . Our results showed that  $\tau(N)$  exhibits a sudden transition of several orders of magnitude as  $N$  approaches a critical value  $\mathcal{N}$  similarly to the free energy barrier  $\Delta\mathcal{F}$ . As discussed along this work, the transition is enhanced when large values of  $\epsilon$  and  $\omega$  are considered, the critical length  $\mathcal{N}$  is increased and the transition jump is enlarged. This result implies that for two poly- $X$  chains of the same size  $N$ , poly- $X_1$  and poly- $X_2$ , with  $\omega_1 \neq \omega_2$ , the translocation times  $\tau_1$  and  $\tau_2$  will differ depending on whether the size of the chain is below or above the critical sizes  $\mathcal{N}_1$  and  $\mathcal{N}_2$ . If  $\omega_1 > \omega_2$  then  $\mathcal{N}_1 > \mathcal{N}_2$ , therefore, the translocation of poly- $X_1$  would be faster than that of poly- $X_2$  after the transition of poly- $X_2$ , only if the chain size is larger than  $\mathcal{N}_2$ . Energy penalty  $U$  had a drastic effect on the occurrence of the transition. Chains in the rodlike regime, large  $U$ , lack such transition, in this case the main contribution to the translocation time comes from the anchorage of the chain to the pore. As discussed in the Appendix,  $U$  must be a few  $k_B T$  to enable poly- $X$  chains to fold within physiological times, in order to avoid Levinthal's paradox [18].  $U$  cannot be zero nor have an exceedingly small value, which provides a criterion for the  $U$ 's used throughout this work.

The translocation time  $\tau(N)$  for large chains, after the transition jump, follows the exponential law  $\tau = A\tau^\xi$ , where exponent  $\xi$  and constant  $A$  depended on the transition. Main dependence of exponent  $\xi$  and constant  $A$  relied on the energy penalty  $U$  and the ground state energy  $\epsilon$ . The fitted exponent  $\xi$  ranges between 1.50–2.00 according to the parameters used. Constant  $A$  spanned several orders of magnitude according to a combination of the set  $(U, \epsilon, \omega)$ , from the nanoseconds to the seconds scale. The main feature of these results is

that the transition imposes the time scale of translocation for large chains indicated by constant  $A$ . These facts strongly suggest that the inclusion of helical propensity and adequate energy scales for the unfolded and folded states of the chain can provide a consistent formalism to explain the time scale observed in translocation experiments of polypeptide chains. By using experimental data from diffusion-driven translocation of poly(ethylene oxide) chains we confirmed the transition jump in the translocation time predicted by our model for small chains. We expect the transition in the translocation time predicted in our model is a feature spread through the translocation of any poly- $X$  chain.

## ACKNOWLEDGMENTS

L.O.-Q. thanks the support from Universidad Autonoma de la Ciudad de Mexico (UACM) and Secretaria de Ciencia e Inovacion Tecnologica (SECITE) through research Grant No. PI2011-14R (L.O.-Q.). J.A.V.P. thanks Consejo Nacional de Ciencia y Tecnología (CONACyT) for a postdoctoral fellowship (205293-290908) at CINVESTAV (México).

## APPENDIX: VALUES OF PARAMETERS

To analyze homopolymer translocation, we consider poly- $X$  chains, where  $X$  represents any of the 20 proteinogenic amino acids. In the case of Gaussian chains (GC), we used for the monomer length  $b$  the distance between consecutive monomers along the backbone chain. This is the distance between two consecutive  $C_\alpha$  atoms, set by the amide plane,  $b \approx 3.8 \text{ \AA}$ . The helical propensity was obtained according to the scale proposed in [20], which is an experimental scale measured taking as a reference alanine-based peptides. Helical propensities for the 20 proteinogenic amino acids are delimited by proline  $\omega_{\text{PRO}} = 0.001$  and alanine  $\omega_{\text{ALA}} = 1.61$ , which spans over three orders of magnitude. For most amino acids the helical propensity is a few times 0.1, therefore, as a reference over all orders of magnitude we use the helical propensity of threonine,  $\omega_{\text{THR}} = 0.14$ . In our calculations, we also include the upper and lower bounds  $\omega_{\text{PRO}}$  and  $\omega_{\text{ALA}}$ . Polymers of physiological interest that translocate are diverse in size. Proteins such as preprocecropin (63 amino acids) [24] and prepromielitin (70 amino acids) [25] are considered as small translocating proteins whereas proteins such as LifA (3223 amino acids) [26] and ApoB (4536 amino acids) [27] are among the largest. Average size of most proteins is a few hundred amino acids. We set  $N = 500$  as an average size for poly- $X$  chains.

Folded state energies for proteins usually range from  $-60$  to  $-20 \text{ kJ/mol}$ . As a reference, we use for the ground state energy  $\epsilon = 40 \text{ kJ/mol}$ . In analyzing the dependency on the native energy we use for the ground state energy  $\epsilon = 20, 30, 40, 50$ , and  $60 \text{ kJ/mol}$ . The energy scale for the penalty energy  $U$  is based on the Zwanzig *et al.* model [18]. Within this model, penalty energy  $U$  and the chain size  $N$  play a fundamental role in determining the mean first-passage time (MFPT) of folding  $\tau_f$ , i.e., the average time for a poly- $X$  chain to reach for the first time its native conformation [18]. Starting from a configuration with  $s$  incorrect bonds, the MFPT to reach the native state with all residues in their native location can be

written as [18]

$$\tau_f(s) = \frac{1}{k_0} (1 + K)^N K \int_0^1 dy \left( \frac{1 - y^s}{1 - y} \right) (1 + Ky)^{-N-1}, \quad (\text{A1})$$

where  $k_0$  represents the rate at which a native bond can be converted into a non-native bond and  $k_1$  is the rate at which non-native bonds convert into native bonds. The quantity  $K$  is the ratio  $K = k_0/k_1 = \nu \exp(-U/RT)$  and in general is different from unity. For a chain of size  $N = 100$  to have a  $\tau_f$  in the scale of seconds, the energy penalty must be  $U \approx 2 k_B T$  whether  $k_0 = 2 \times 10^9 \exp(-U/k_B T) \text{ s}^{-1}$  and the number of

incorrect native bonds is  $\nu = 2$  [18]. Larger chains require larger energy penalties or smaller rate constants for  $\tau_f$  to be in the same time scale. The role played by the energy penalty  $U$  becomes evident if there were no penalties for a non-native contact, i.e., if  $U = 0$ . In such a case, the MFPT of folding will scale as  $\tau_f \sim (1 + \nu)^N$  which is the mathematical formulation of the so-called Levinthal's paradox [18,28,29]. In this case  $\tau_f$  increases exponentially as the number  $N$  of residues. Thus,  $U$  plays a fundamental role in decreasing the time needed to reach a thermodynamically stable ground state. Following the results of the Zwanzig *et al.* model we set  $U$  to be a few  $k_B T$ .

- 
- [1] J. Jiang, B. L. Pentelute, R. J. Collier, and Z. H. Zhou, *Nature* **521**, 545 (2015).
- [2] A. G. Portalou, K. C. Tsois, M. S. Loos, V. Zorzini, and A. Economou, *Trends Biochem. Sci.* **41**, 175 (2015).
- [3] K. Denks, A. Vogt, I. Sachelaru, N.-A. Petriman, R. Kudva, and H.-G. Koch, *Mol. Membr. Biol.* **31**, 58 (2014).
- [4] Y. Akeda and J. E. Galán, *Nature* **437**, 911 (2005).
- [5] T. Minamino, Y. V. Morimoto, M. Kinoshita, P. D. Aldridge, and K. Namba, *Sci. Rep.* **4**, 7579 (2014).
- [6] S. L. Wynia-Smith, M. J. Brown, G. Chirichella, G. Kemalyan, and B. A. Krantz, *J. Biol. Chem.* **287**, 43753 (2012).
- [7] A. Balijepalli, J. Etedgui, A. T. Cornio, J. W. F. Robertson, K. P. Cheung, J. J. Kasianowicz, and C. Vaz, *ACS Nano* **8**, 1547 (2014).
- [8] H. A. Carleton, M. Lara-Terejo, X. Liu, and J. E. Galán, *Nat. Commun.* **4**, 1590 (2013).
- [9] F. Cecconi, U. M. B. Marconi, and A. Vulpiani, *Spectroscopy* **24**, 421 (2010).
- [10] R. Kumar and M. Muthukumar, *J. Chem. Phys.* **131**, 194903 (2009).
- [11] J. M. Polson, M. F. Hassanabad, and A. McCaffrey, *J. Chem. Phys.* **138**, 024906 (2013).
- [12] Y. Shen and L. Zhang, *Polymer* **48**, 3593 (2007).
- [13] A. Milchev and K. Binder, *Comput. Phys. Commun.* **169**, 107 (2005).
- [14] J. J. Kasianowicz, E. Brandin, D. Branton, and D. W. Reamer, *Proc. Natl. Acad. Sci. USA* **93**, 13770 (1996).
- [15] W. Sung and P. J. Park, *Phys. Rev. Lett.* **77**, 783 (1996).
- [16] M. Muthukumar, *J. Chem. Phys.* **111**, 10371 (1999).
- [17] R. P. Choudhury, P. Galvosas, and M. Schönhoff, *J. Phys. Chem. B* **112**, 13245 (2008).
- [18] R. Zwanzig, A. Szabo, and B. Bagchi, *Proc. Natl. Acad. Sci. USA* **89**, 20 (1992).
- [19] L. Olivares-Quiroz, *J. Phys.: Condens. Matter* **25**, 155103 (2013).
- [20] A. Chakrabarty, T. Kortemme, and R. L. Baldwin, *Protein Science* **3**, 843 (1994).
- [21] H. Risken, in *The Fokker-Planck Equation: Methods of Solution and Applications*, 2nd ed., edited by H. Haken, Springer Series in Synergetics Vol. 18 (Springer-Verlag, Berlin, 1996).
- [22] N. G. van Kampen, *Z. Phys. B: Condens. Matter* **68**, 135 (1987).
- [23] J. A. Vélez Pérez, O. Guzmán, and F. Navarro-García, *Phys. Rev. E* **88**, 012725 (2013).
- [24] A. K. K. Lakkaraju, R. Thankappan, C. Mary, J. L. Garrison, J. Taunton, and K. Strub, *Mol. Biol. Cell* **23**, 2712 (2012).
- [25] R. Zimmermann, M. Sagstetter, G. Schlenstedt, H. Wiech, B. Kaßbeckert, and G. Müller, *Membrane Biogenesis* (Springer-Verlag, Berlin, 1988), Chap. , pp. 337–350.
- [26] W. Deng, H. B. Yu, C. L. de Hoog, N. Stoynov, Y. Li, L. J. Foster, and B. B. Finlay, *Mol. Cell. Proteomics* **11**, 692 (2012).
- [27] Z. Yao, K. Tran, and R. S. McLeod, *J. Lipid Res.* **38**, 1937 (1997).
- [28] R. Zwanzig, *Proc. Natl. Acad. Sci. USA* **92**, 9801 (1995).
- [29] L. Olivares-Quiroz, *La Física Biología en México: Temas Selectos 2*, 1st ed. (El Colegio Nacional Mexico, Mexico, 2008), Chap. , pp. 141–171.