J Mol Model (2005) 11: 379–384
DOI 10.1007/s00894-005-0254-9

ORIGINAL PAPER

Andrzej Sikorski · Piotr Romiszowski

Computer simulation of polypeptide translocation through a nanopore

Received: 2 November 2004 / Accepted: 7 February 2005 / Published online: 2 April 2005

Abstract A simplified model of polypeptide chains was designed and studied by means of computer simulations. Chains were represented by a sequence of united atoms located at the positions of the α-carbons. A further assumption was the lattice approximation for the chains. We used a (310) lattice, which was found useful for studying properties of proteins. The force field used consisted of a long-range contact potential between amino-acid residues and a local preference for forming α-helical states. The chain consisted of two kinds of residues: hydrophilic (*P*) and hydrophobic (*H*) ones forming model helical septets *–HHPPHPP–* in a sequence. The chains were placed near an impenetrable surface with a square hole in it. The size of the hole was comparable or smaller than the size of a chain. The properties of these model chains were determined using the Monte-Carlo simulation method. During the simulations, translocation of the chain through the hole in the wall was observed. The influence of the chain length, the temperature differences on both sides of the wall and the force field on the chain properties were investigated. It was shown that the translocation time scales as $N^{2.2}$ and it was found that the presence of the local helical potential significantly slows down the process of translocation.

Introduction

The structure and the properties of polymer and polypeptide chains at interfaces are extremely complicated and thus interesting from the theoretical point of view. Some of them are the process of transport of chains (proteins, DNA, RNA and drug delivery) across a membrane, translocation of proteins through an endoplasmic reticulum or into mitochondria as well as the formation of signaling protein in a cellular membrane. These processes are very important because of their biological implications [1]. One has to remember that threading a chain through a pore is governed by the presence of an entropic barrier at least because of the limited number of chain conformations inside and near the pore, which slows the chain dynamics down. The influence of an electric field usually has also to be taken into consideration. The process of the translocation of polymer chains through a pore was recently the subject of numerous theoretical studies [2–8] and various computer simulations [9–11]. These studies were mainly focused on the mechanism of translocation and scaling properties of the translocation time. Most of the above models were based on one-dimensional diffusion [12]. It was pointed out that the translocation of a polymer chain can be coupled with many other phenomena like adsorption, polymerization, collapse, random coil-helix transition etc. [13]. Di Marzio and Kasianowicz [8] showed exact solutions for all possibilities of such couplings. Recently, Binder and co-workers presented theoretical considerations and computer simulation studies concerning the translocation induced by adsorption, which revealed the dynamics of this process [14]. They found that the translocation time's scaling exponent was 2.18, correcting previous theoretical [3] and simulation results [15].

It was recently shown that naïve and rather coarse models of a polypeptide were introduced some years ago could be a very useful tool for studying fundamental and general behavior of these chains [16]. The polypeptide chains were approximated by a sequence of united atoms and by a lattice representation. The Monte-Carlo simulations of these models showed the interplay between tertiary interactions and the formation of secondary structures. The dynamic behavior of chains built of different sequences and the thermodynamic description of the folding transition were also studied. These simple

A. Sikorski · P. Romiszowski (✉)
Department of Chemistry, University of Warsaw,
Pasteura 1, 02-093 Warszawa, Poland
E-mail: prom@chem.uw.edu.pl
Tel.: +48-22-8220211
Fax: +48-22-8225996

models were then extended by maintaining the proper level of the helical content and by improvement of some features of the folding transition [17].

In this work, we have developed a new model based on the (310) lattice representation of polypeptide chains. In this new model, polypeptide chains pass through a hole in an impenetrable surface. Beside this confinement, the chains had all the features of polypeptide chains from our previous models. Thus, the model system can be treated as a crude representation of a polypeptide translocating through a nanopore in a thin and rigid membrane. The properties of free polypeptide chains and polypeptides attached to a surface have already been determined within the frame of (310) lattice model [16–18]. Therefore, the results obtained for chains translocated through a pore could be compared with properties of 'free' polypeptides. The main goal of this work is to study the effect of various interactions on the properties of chains passing through a pore on a fundamental level. This paper was organized as follows. In the next section *The model and the simulation method* we outline the assumptions of the model and describe the simulation method used. In the section *Results and discussion* we present the results of simulations compared with results obtained for this model for a free chain as well as with other theoretical predictions and real experiments. In the last section *Conclusions* we present the most important and general conclusions as well as their meaning for studying protein-like systems.

The model and the simulation method

The model chain consisted of N residues. Because the times of translocation of single-stranded RNA and DNA through a membrane channel in real experiments are of the order of milliseconds it is impossible to simulate full atom-chain models on such time scale [19–21]. Therefore, we assumed that chains were built of structural elements that could be treated as an approximation of amino acid residues: each amino acid residue was represented by a single united atom located at the α-carbon position [16, 17]. An additional speed up of the calculations was gained by using the lattice approximation for model chains. We used a lattice which has been used frequently recently in simulations of proteins and polypeptides [16–18]. In this model united atoms of the chain are connected by vectors of the type [±3, ±1, ±1], [±3, ±1, 0], [±3, 0, 0], [±2, ±2, ±1] and [±2, ±2, 0]. This lattice is very flexible as its coordination number is 90 and the angles between the bonds reproduce conformations of real peptides and proteins with great accuracy. It was shown that the model chains were represented with the accuracy of 0.6–0.7 Å comparing with real polypeptides when the lattice unit was assumed to be 1.22 Å [16].

The properties of polypeptide chain are determined by their primary structure, i.e. the sequence of the amino acid residues in the chain. We limited our studies to simple heteropolymers and thus for the purpose of our investigations the model chains consisted of two kinds of residues only. We called them hydrophilic (or polar) residues, denoting them as P and hydrophobic (or non-polar), denoted as H. These two kinds of residues could be distinguished by an interaction potential. This potential was non-zero for distances between a pair of residues higher than 4.35 Å and lower than 7.40 Å ($3^{1/2}$ and 5 in lattice units, respectively). We chose the following set of interactions: HP interaction potential $\epsilon_{HP} = 0$, HH interaction potential $\epsilon_{HP} = -2$ kT, while for PP the potential was $\epsilon_{HP} = -1$ kT [16–18]. This particular choice was made in order to obtain a hydrophobic core for low-temperature collapsed chains. These values of potential were also chosen because they allow one to refer the results to a homopolymer with attractive potential −1 kT [16]. We would like to stress that these interactions should be precisely described as 'strongly hydrophobic' (H) and 'weakly hydrophobic' (P) [17]. It has been shown that this choice of the force field is more suitable for polypeptide chain models than the pure hydrophobic/hydrophilic potential [10]. The excluded-volume effect was also introduced into the model: a pair of non-bonded residues for distances shorter than 4.35 Å ($3^{1/2}$ in lattice units) interacted with the repulsive potential $\epsilon_{rep} = 5$ kT [16–18].

In the model we have also introduced the parameter ϵ_{loc}, which characterizes the local property of the chain to form a helical structure. It was previously shown that the existence of this helical potential led to the proper amount of helical structures for given amino acid sequences [16]. The helical state formed by three consecutive vectors can be identified by the specific vector expression as follows:

$$r^{*^2}_{i-1,i+2} = (v_{i-1} + v_i + v_{i-1})^2 \ \text{sign}((v_{i-1} \times v_i) \ v_{i+1}) \qquad (1)$$

where $v_{i\text{-}1}$, v_i, v_{i+1} stand for three consecutive vectors connecting residues from i-1-th to $i+2$-th. Since there are two possible helical structures (left-handed and right-handed helices), we decided to introduce the preference for the formation of right-handed helices only (as in real proteins). A right-handed α-helix conformation in our model chain corresponds to the values of $r^{*^2}_{i\text{-}1,i+2}$ located between 9 and 25 [16]. The appearance of a right-handed helical conformation in the chain during the simulation was associated with an energy loss equal to $\epsilon_{loc} < 0$.

The above-described model was studied by means of the Monte-Carlo simulations. The polypeptide chain was put into a Monte-Carlo box with the edges large enough not to use periodic boundary conditions ($L = 200$). One of the surfaces forming the box was impenetrable for the chain (for this purpose we chose the surface located at $x = 0$). A square hole with the edge d was made in the middle of this wall. The starting conformation of the chain was generated at random but one of chain ends was always located in a lattice vertex next to the hole in the impenetrable surface. Then, the confor-

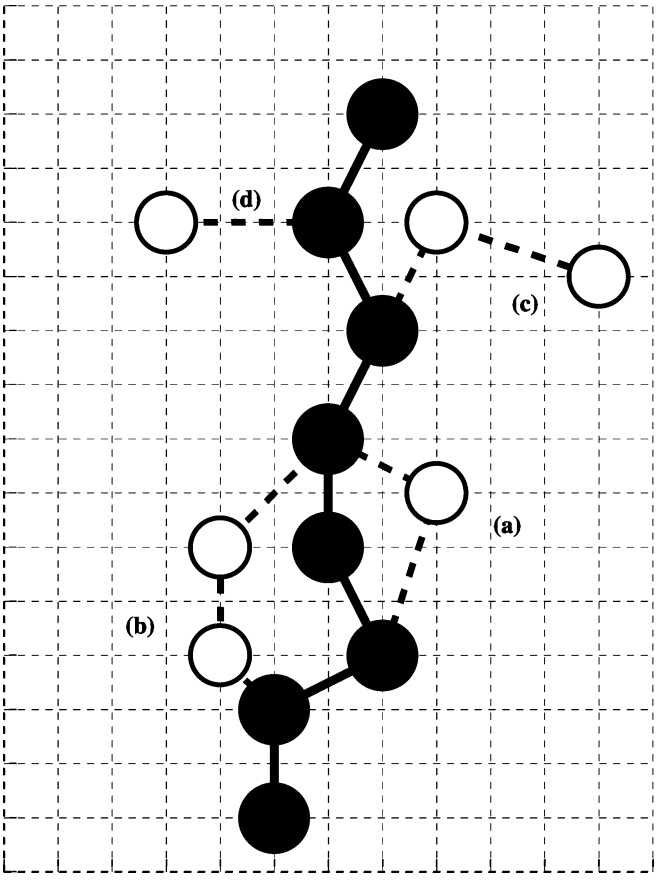

Fig. 1 Changes of chain's conformation used in the simulation algorithm: 1-residue flip **(a)**, 2-residues flip **(b)** and 2-residues end reorientations **(c)**

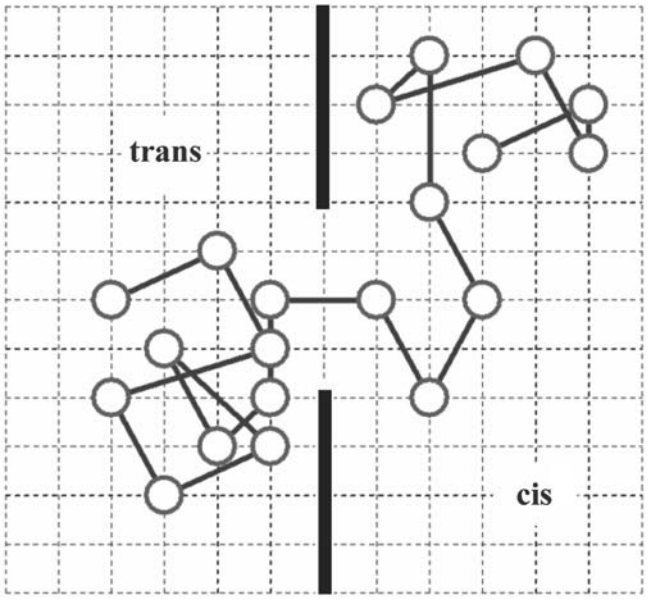

Fig. 2 The model chain passing through a hole in a rigid, thin and impenetrable *surface*

mation of the chain was randomly modified using the usual set of local motions: [16] 1-residue flip (a), 2-residues flip (b) and 2-residues end reorientations (c), which are shown schematically in Fig. 1. A new probe conformation was accepted with a probability proportional to $exp(-\Delta E/\mathrm{kT})$, where ΔE is an energy difference between a new and an old chain conformation. The driving force parallel to the x-axis and independent on the x coordinate was introduced into the model in order to emulate a static electric field. This potential was included into the Metropolis criterion only but was not counted in the total energy of the system. The value of this force was -0.2 kT what was negligible when compared with the total energy of chains (in order of 100 kT)

The initial conformation of the chain underwent a large number of micromodifications. The motion of the entire chain was monitored during the simulation. When it diffused away from the hole, the simulation was paused and the chain was placed again in the vicinity of the hole. When the entire chain passed through the hole, the process of simulation was repeated from the beginning starting from a different chain conformation. The simulations were continued until the given chain passed through the hole 100 times. It was previously shown that one attempt of each local motion per one amino acid residue defined a time unit. This choice has already been tested: It was shown that the chain's dynamics are Rouse-like, i.e. correct [14]. A model chain on the (310) lattice near the impenetrable wall passing a hole from the *cis* side (right) to *trans* side (left) is shown in Fig. 2.

Results and discussion

In this study, we present the simulation results for polypeptide chains built of $N = 10$–100 amino acid residues. All chains were built of typical helical septets -*HHPPHPP*-, which can be found in real helical proteins. The helical potential ϵ_{loc} was assumed to take values 0, -4 and -8. In our previous studies concerning free polypeptide chains, we showed that these values of the helical potential covered the range from fully flexible chains to very strongly helical ones, like in polyalanine [16]. It was also assumed that there is a difference in temperatures on the *cis* side (where chains are located initially) and on the *trans* side. On the *cis* side, the temperature was chosen to $T = 4$, corresponding to a random coil state of the chains, while it was low ($T = 2$) on the *trans* side. It was shown that at these temperatures chains were collapsed, forming a dense ordered globule [16–18]. The edge of the square hole was $d = 15$ lattice units corresponding to 18.3 Å. This choice means that the size of the chains was comparable or smaller than the hole, but much larger than the diameter of the α-helix.

In order to demonstrate the properties of the chain observed during the translocation, we present a series of flowcharts taken from the simulation trajectories. The flowcharts contain the data recorded at time intervals equal to two, so that for the trajectory of the length

40,000 time units we have 20,000 data. The trajectories shown in the flowcharts were taken for that fragment of the simulation for which the translocation took place. We decided to present the flowcharts for two values of the local potential $\epsilon_{loc} = 0$ and $\epsilon_{loc} = -8$, to demonstrate the dramatic changes in the properties of the system as this potential is introduced. In Fig. 3, we show the squared radius of gyration of the entire chain S^2 as a function of time for the total chain. This parameter describes the size of the polypeptide chain. One can observe that the radius of gyration value fluctuates. However these fluctuations are much larger for the case of the chain without the local potential. Its presence causes the formation of secondary structure, which is a stable fragment and therefore the dimensions of the chain remain also more stable. The introduction of the helical potential even stabilizes the size of chains on the *cis* side. It is seen especially when one compares the changes of the radius of gyration for the chain with $\epsilon_{loc} = -8$. At the moment when the translocation starts, the S^2 values begin to fluctuate. This is the case when one part of the chain remains in the *cis* position while the other is in *trans* state. The beginning of the translocation is denoted by an arrow (empty head). After the translocation is complete, the fluctuations of S^2 almost disappear for the case of the strong helical potential. This is the result of the presence of the helical secondary structure, which is a compact and stable. In contrast to the latter, the chain without the local potential almost does not change its S^2 qualitatively and its value fluctuates regardless of the translocation process.

The secondary structure of the chain plays an important role as the properties of the chains are discussed. Therefore, we analyzed of this parameter. A typical flowchart showing changes in helicity of the system during the passage through the nanopore can be seen in Fig. 4. The helicity, which we define as the fraction of all residues forming a helical structure, is

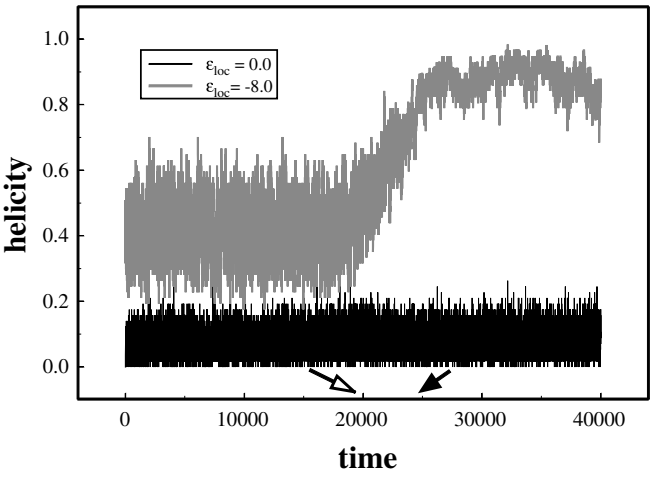

Fig. 4 The changes of the helicity of chains with time. The case of chain consisted of $N = 60$. The values of the local potential ϵ_{loc} are given in the *inset*

shown for the case of $\epsilon_{loc} = 0$ and $\epsilon_{loc} = -8$. One can observe that in the case of the presence of a strong helical potential in the *cis* state, the helicity fluctuates around the value 0.4, while after the translocation it increases up to 0.9. Also, the fluctuations in the *trans* state are much lower than in the *cis* one. The arrows denote the start and the completion of the translocation of the chain. The plot of the helicity for the case without a helical potential is also shown. The mean value is around 0.1 and does not change during the translocation process.

During the translocation through the nanopore the energy of the system changes along with the progress of the passage. Figure 5 presents the changes of the contact energy of the system for two cases: $\epsilon_{loc} = 0$ and $\epsilon_{loc} = -8$. One can see that for both cases the energy decreases with time. However, the changes in the presence of the

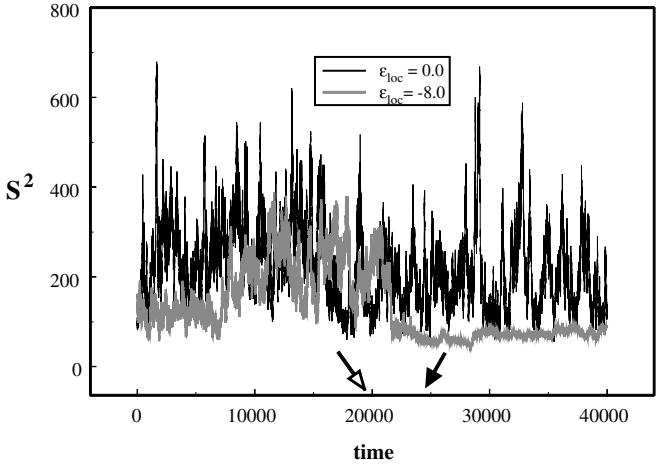

Fig. 3 The changes of the squared radius of gyration S^2 with time. The case of chain consisted of $N = 60$. The values of the local potential ϵ_{loc} are given in the *inset*

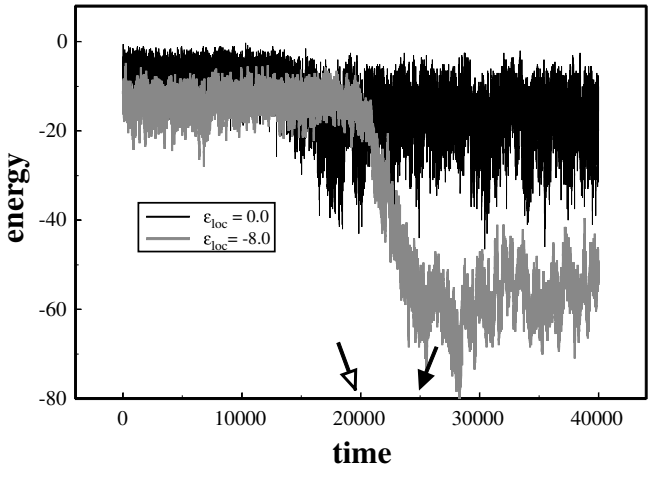

Fig. 5 The changes of the contact energy of chain with time. The case of chain consisted of $N = 60$. The values of the local potential ϵ_{loc} are given in the *inset*

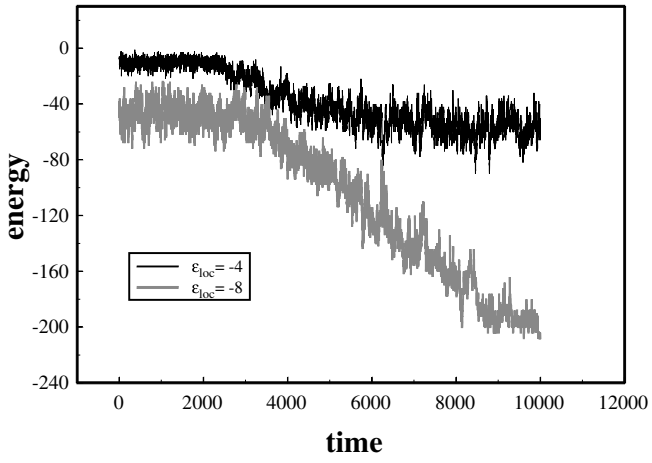

Fig. 6 The changes of the local energy of chain with time. The case of chain consisted of $N = 60$. The values of the local potential ϵ_{loc} are given in the *inset*

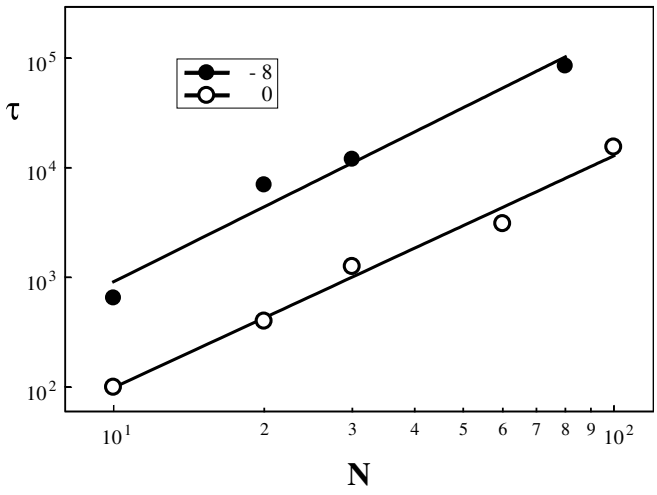

Fig. 8 The translocation times τ versus the chain length N. The values of the local potential ϵ_{loc} are given in the *inset*

local potential are much more pronounced than for the case $\epsilon_{loc} = 0$. One can also observe that the fluctuations of the energy are much larger for $\epsilon_{loc} = 0$ than for $\epsilon_{loc} = -8$. However, note that the substantial fluctuation of the total energy for the strong helical potential corresponds to very small fluctuations of chain's size (see Fig. 3). Figure 6 shows the changes of the local energy, i.e. the changes of the contribution to the total energy related to the helical potential. One can observe that there are no quantitative differences between strong and intermediate helical potential although the time of translocation are quite different. Comparing the value of the local energy with the contact energy (Fig. 5) for the strong helical potential ($\epsilon_{loc} = -8$) one can state that on the *cis* side, the local energy governs the behavior of the chain as it is three time lower than the contact one. This shows that the formation of the secondary helical structure could be a driving force of the translocation process [8].

In order to illustrate the chain during the translocation process, we show snapshots of the chain while it passes through the window. The two cases are presented in Fig. 7. The chains pass from the right side of the square window to the left. Without the local potential, the chain does not exhibit any helical structure on both sides of the window. The presence of the local potential changes the structure of the chain dramatically. The chain on *cis* position has no helical elements, while its part that passed the window is well organized and presents classical helices.

The translocation process is usually characterized by the mean translocation time. In Fig. 7 we show the dependence of the translocation times on the chain length for some helical potentials. One can see that this time scales with the chain length as N^{γ}. The scaling exponents for all cases under consideration are close one to each other: $\gamma = 2.11 \pm 0.15$ (for $\epsilon_{loc} = 0$) and $\gamma = 2.27 \pm 0.29$

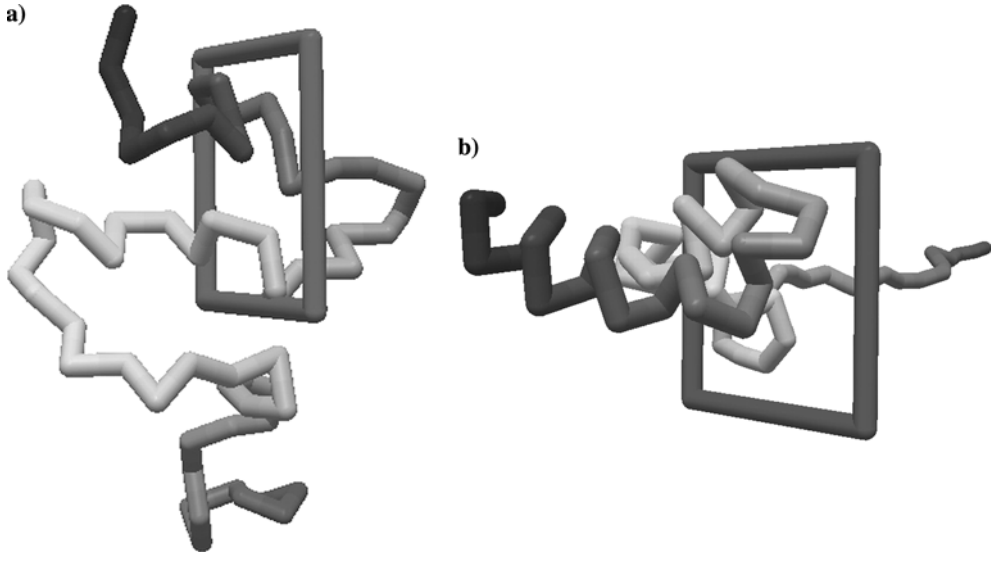

Fig. 7 The snapshots of typical chain's conformation obtained during the simulation for chain consisted of $N = 60$. The values of the local potential $\epsilon_{loc} = 0$ **(a)** and -8 **(b)**

(for $\epsilon_{loc} = -8$). These values are in a good agreement with theoretical findings that predict N^2 behavior [2, 14] (Fig. 8). The increase of strength of the helical potential led to considerably longer times of translocation through the nanopore. This can be explained by the fact that the secondary (helical) structures formed on the *trans* side of the box are stable and therefore the diffusion of the system is very slow in spite of the fact that the size of such chains is even smaller on the *cis* side.

Conclusions

Monte-Carlo simulations of lattice models of polypeptide chains were carried out in order to find the influence of the chain length and the force field on the passage of chain through a pore and its properties on both sides of the membrane. The reduced model of a polypeptide (united atoms in alpha carbon positions only) appeared to be sufficient for studying the properties of chains as a whole and therefore was employed for our calculations. The atoms representing residues interacted with a very simple binary square-well potential. Some local preferences in the chain's conformations were introduced in order to obtain the proper amount of helical conformations. The polypeptide consisted of two kinds of residues only: hydrophobic and hydrophilic. The chains were placed near an impenetrable wall with a hole in it. The intention of introducing this confinement was to mimic the translocation of polypeptide chains trough a nanopore in a thin and rigid membrane.

It was shown that the translocation of a polypeptide chains through a pore depends on the chain length and on the size of the pore. The translocation time scales with the chain length as N^{γ} with γ close to 2.2, in agreement with theoretical predictions [4, 5]. The strength of the helical potential did not change significantly the above scaling but affected the value of the translocation time. Translocation times for strong potential are an order of magnitude longer.

References

1. Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D, Darnell JE (2000) Molecular biology of the cell. WH Freeman, New York
2. Sung W, Park PJ (1996) Phys Rev Lett 77:783–786
3. Muthukumar M (1999) J Chem Phys 111:10371–10374
4. Lubensky DK, Nelson DR (1999) Biophys J 77:1824–1838
5. Elston TC (2000) Biophys J 79:2235–2251
6. Muthukumar M (2001) Phys Rev Lett 86:3188–3191
7. Berezhkovskii AM, Gopich IV (2003) Biophys J 84:787–793
8. DiManzio EA, Kasianowicz JJ (2003) J Chem Phys 119:6378–6387
9. Chem S-S, Cardenas AE, Coalson RD (2001) J Chem Phys 115:7772–7782
10. Zandi R, Reguera D, Rudnick J, Gelbart WM (2003) Proc Natl Acad Sci USA 100:8649–8653
11. Lansac Y, Maiti PK, Glaser MA (2004) Polymer 45:3099–3110
12. Meller A (2003) J Phys: Condens Matter 15:R581–R607
13. DiManzio EA (1999) Progr Polym Sci 24:329–377
14. Tian P, Smith GD (2003) J Chem Phys 119:11475–11483
15. Milchev A, Binder K, Bhattacharya (2004) J Chem Phys 121:6042–6051
16. Romiszowski P, Sikorski A (2000) Biopolymers 54:262–272
17. Sikorski A, Romiszowski P (2003) Biopolymers 69:391–398
18. Sikorski A, Romiszowski P (2004) J Chem Inf Comput Sci 44:387–392
19. Kasianowicz JJ, Brandin E, Branton D, Deamer DW (1996) Proc Natl Acad Sci USA 93:13770–13773
20. Szabo I, Bathori G, Tombola F, Brini M, Coppola A, Zoratti M (1997) J Biol Chem 272:25275–25282
21. Chen H, Hsu SL, Tirrell DA, Stidham HD (1997) Langmuir 13:4775–4778