CONDENSED MATTER PHYSICS

Lamplighter model of a random copolymer adsorption on a line

L.I. Nazarov[1], S.K. Nechaev[2,3,4], M.V. Tamm[1,4]

[1] Physics Department, M.V. Lomonosov Moscow State University, 119991 Moscow, Russia

[2] Université Paris-Sud/CNRS, LPTMS, UMR8626, Bât. 100, 91405 Orsay, France

[3] P.N. Lebedev Physical Institute, RAS, 119991 Moscow, Russia

[4] Department of Applied Mathematics, National Research University "Higher School of Economics", 101000 Moscow, Russia

We present a model of an AB-diblock random copolymer sequential self-packaging with local quenched interactions on a one-dimensional infinite sticky substrate. It is assumed that the A-A and B-B contacts are favorable, while A-B are not. The position of a newly added monomer is selected in view of the local contact energy minimization. The model demonstrates a self-organization behavior with the nontrivial dependence of the total energy, E (the number of unfavorable contacts), on the number of chain monomers, N: $E \sim N^{3/4}$ for quenched random equally probable distribution of A- and B-monomers along the chain. The model is treated by mapping it onto the "lamplighter" random walk and the diffusion-controlled chemical reaction of $X + X \to 0$ type with the subdiffusive motion of reagents.

Key words: *local optimization, heteropolymer folding, lamplighter random walk, subdiffusive chemical reaction*

PACS: *05.40.-a, 02.50.Ga, 87.15.Cc*

1. Introduction

In this letter we propose a simple one-dimensional model of stochastic dynamic system possessing a local optimization in quenched environment. The question of choosing an optimal strategy for a dynamic system if only *local* optimization is accessible, is a generic problem far beyond the scope of natural science. Let us imagine that someone knows nothing about the future and adapts his or her own behavior in each current time moment only on the basis of the knowledge about the best solution at a given narrow time slice. Such a behavior leads to an optimal local strategy when only a partial (current) knowledge is accessible, though it might be far from generic optimal one if the knowledge about the future is available.

As an example of physical problems sharing the properties of local and global optimization, we can mention the problems of protein secondary structure formation and DNA packaging in a viral capsid. If in the course of protein folding, the chain forms a "frozen" network of contacts, then the optimal conformation is determined only by a sequential step-by-step optimization of the added heteropolymer fragments. However, if the whole protein is capable of "adjusting" the structure of a contact network by breaking some bonds and creating other bonds, then the global optimization might essentially change the folding picture.

The system considered below should not be regarded as a model of any specific physical system (though it has some features of a protein folding), but it rather highlights the principles of local optimization of a specific one-dimensional stochastic system in quenched random environment, which could lead to a nontrivial self-organization.

Our model resembles the "lamplighter random walk" proposed in 1973 by A.M. Vershik in the afterword to the Russian edition [1] of the book "Invariant means on topological groups and their applications"

---

by F.P. Greenleaf as an example of nontrivial estimate for the growth of various numerical characteristics of groups. The object known as a "lamplighter group" (the name was also given by A.Vershik) became very popular among probabilists after the pioneering work by A. Vershik and V. Kaimanovich [2].

2. The model

The toy model presented below imitates the packaging of a random diblock copolymer with quenched sequence of monomers on a one-dimensional substrate. Schematically it can be viewed as a tissue folding getting out of the roll, as depicted in figure 1 (a), where the size of folds depends on local heterogenous interactions between different parts of the tissue. Such a schematic view might be compared with the specific secondary structure formation of a linear polypeptide chain in figure 1 (b), where two main types of secondary structure, α-helices and β-sheets, are shown and the secondary structure is stabilized by the hydrogen bonds.
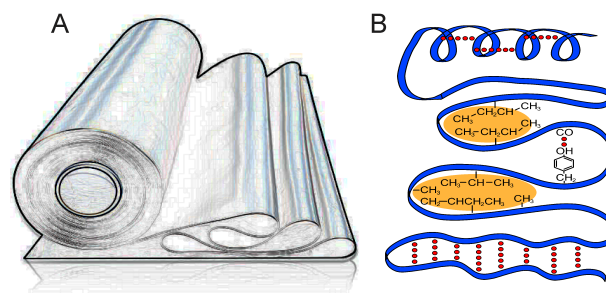

Figure 1. (Color online) (a) Folds of tissue getting out of the roll; (b) Secondary structure formation in a polypeptide heteropolymer chain.

To pass from these intuitive pictures to a more specific description, consider a N-monomer chain with quenched random primary sequence of two types of links, A and B, placed on an infinite line, which is sticky for links A and repulsive for links B. We suppose that the energy of the A-A or B-B contacts is 0, while the energy of the A-B contact is +1 (in dimensionless units), i.e., the monomers of the same types are indifferent to each other, while the monomers of different types repulse.

The heteropolymer folding (packaging) on an infinite line is an irreversible sequential process of adding monomers to the already existing frozen environment. Let us recursively describe the folding process. For definiteness, suppose that the line is initially uniformly covered with monomers of type A and let the first chain monomer be always of type A. Put the 1st monomer at a position $x_0 = 0$ on the line. The 2nd monomer could be placed either next to the 1st monomer on a line in the *trans* state, or it can be put on top of the 1st monomer in the *gauche* state, making a "hook" as it is shown in figure 2. The newly added monomer interacts with the one located below it in the projection to the line.
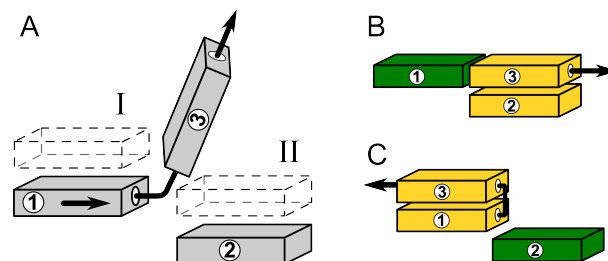
Figure 2. (Color online) Local rules of the monomers freezing to the structure. (a) The monomer 3 has two possible locations I and II. The selection of the position depends on the types of all three monomers. (b) Monomers 2 and 3 are identical, while the monomer 1 differs from 2 and 3. (c) The inverse situation with respect to (b).
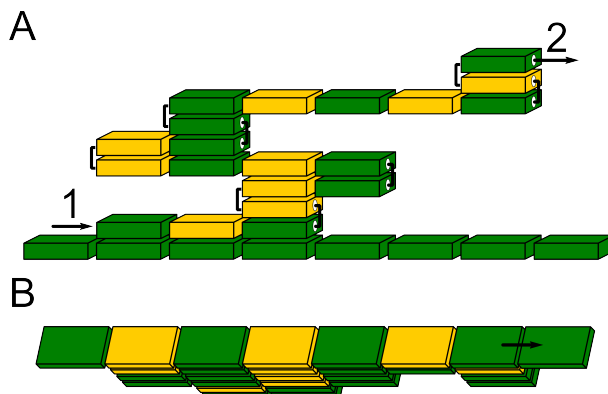
Figure 3. (Color online) The 20-step of block copolymer package on the plate with the uniform monomer type distribution [(a) The side view; (b) The top view].

The selection between *trans-* and *gauche-* states is made according to the following rules: i) if one of two possible new (*trans-* or *gauche-*) states is favorable and the other is not, then the favorable state is always selected; ii) if new *trans-* and *gauche-* states are both favorable or both unfavorable, then a new conformation (*trans-* or *gauche-*) is chosen randomly with the probability $1/2$. For any "defect" (i.e., unfavorable pairing) the energy penalty $+1$ is added. The sample structure of the first 20 steps is depicted in figure 3.

The position of a newly added monomer is selected in view of the irreversible local contact energy minimization, while for equal energies, the choice of *trans-* or *gauche-* states is random with equal probabilities. Let us emphasize that only the uppermost monomers or the "roof" of the structure [i.e., the ones visible from the top of the picture in figure 3 (b)] take part in the play: the monomers that are underneath do not participate in a further optimization process. Therefore, the folding is an effective Markovian process, whose states are specified by i) the set of monomers visible currently from the top of the structure (i.e., by the "roof"), and ii) the position and direction of the last step, which defines the positions where the next monomer can be added.

3. Results

In this section we consider the properties of the described locally optimal folding averaged over a set of completely random initial primary sequences of a polymer, i.e., we assume A and B monomers to appear in the chain with equal probabilities without any correlations between sequential letters. The most natural question to ask about this folding is how the average total energy penalty, $\langle E \rangle$, depends on the chain length, N, i.e., how many unfavorable A-B contacts one has on average in the N-monomer chain. Unexpectedly, it turns out that $\langle E \rangle$ grows sub-linearly with N, namely as $\langle E \rangle \sim N^{3/4}$ [see figure 5 (center)].

To understand this result, one should notice that in terms of the evolution of the roof of the structure defined above, the system has an adsorbing state. Indeed, if the roof is represented by an alternating sequence, i.e., $\ldots$-A-B-A-B-$\ldots$, every newly added monomer can be positioned on the top of this periodic sequence *without any energetic penalty*, and the resulting structure of the roof will stay unchanged. Moreover, if some part of the roof is alternating [as in figure 3 (b)], this part will remain alternating and any sequence can be adsorbed on it without penalty. Thus, the heteropolymer folding in our model can be considered as a dynamic process of approaching the adsorbing alternating state. This dynamics consists in sequential cancellation of "defects", represented by non-alternating pairs of A-A or B-B of consecutive monomers on the substrate as it is shown in figure 4 (a).

To formalize this idea, we introduce an Ising variable, $s(x)$, which labels the state of the roof in position x:

$$s(x) = \begin{cases} +1 & \text{for monomer A in position } x, \\ -1 & \text{for monomer B in position } x. \end{cases} \qquad (3.1)$$
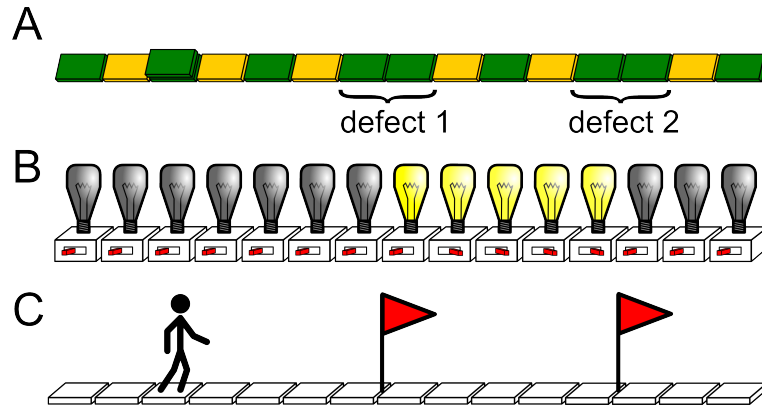
Figure 4. (Color online) The representation of two arbitrary packages in a context of 1D walks on an adjustable substrate. Periodic regions, shown in (a) are mapped (depending on the parity) to the domains of lamps in "off" or "on" positions, shown in (b). The white segments, shown in (c), correspond to the domains of lamps, while red flags mark the domain walls. The human figure designates the tip of the polymer chain, which performs a random walk and can push flags to the left or to the right when passing by them.

Now, we define another Ising variable, $\sigma(x)$:

$$\sigma(x) = (-1)^x s(x). \tag{3.2}$$

The change of variables form $s(x)$ to $\sigma(x)$ essentially simplifies the picture, because in terms of the variable, σ, the adsorbing state is homogenous. The state of the chain in terms of variables σ is depicted in figure 4 (b) by lamps switched on (if $\sigma = 1$) and off (if $\sigma = -1$). Each defect (i.e., the pair of segments A-A or B-B) is a "domain wall" separating two states: to the left of it, the lamps are off, while to the right of it, the lamps are on. A configuration with two defects is depicted in figure 4 (b). The defects (domain walls) are shown in figure 4 (c) by flags. In terms of variables σ, the uniform adsorbing state corresponds to the periodic adsorbing state of our initial Ising model, and the dynamics of folding in terms of σ means coarsening of the Ising domains.

We can easily formulate the dynamic rules describing the evolution of the system. The tip of the original polymer chain, shown as a human figure ("lamplighter") in figure 4 (c) performs a walk on a one-dimensional lattice. In the absence of any energy gain, the *trans*– and *gauche*–states are equiprobable. Since the concentrations of monomers of types A and B in the chain are the same, the lamplighter performs a symmetric random walk on a one-dimensional discrete line as it is shown in figure 4 (c).

The initial state for the spins σ is an infinite one-dimensional lattice with all domains of length 1. Using the notations of figure 4 (c), the flags are located at each boundary between adjacent monomers, while a single random walker (lamplighter) is initially positioned at $x = 0$. As the time (measured in the number of heteropolymer monomers, N) runs, the lamplighter performs a discrete symmetric random walk with the following properties: i) if the lamplighter stays within a domain of lamps which all are "on" or "off", he does nothing with the lamps, ii) if the lamplighter crosses the domain wall, marked by a flag, he may (with the probability 1/2) randomly shift the flag on one lattice site to the left or to the right. If two flags meet each other at one site, they annihilate. The energy of the folding configuration is equal to the number of flag shifts. As the flags annihilate and their concentration decreases, the energy of the folding grows slower.

To consider the dynamics more quantitatively, we have first checked our conjectures computing numerically the span, $\langle R^2(N) \rangle$, of the lamplighter's N-step walk. As shown in figure 5 (left), we have got

$$\langle R^2(N) \rangle \sim N \tag{3.3}$$

as expected for the standard Brownian random walk. In terms of the original problem, this means that the N-step random block-copolymer has a typical span of size of order $\sqrt{N}$ on the 1D substrate.
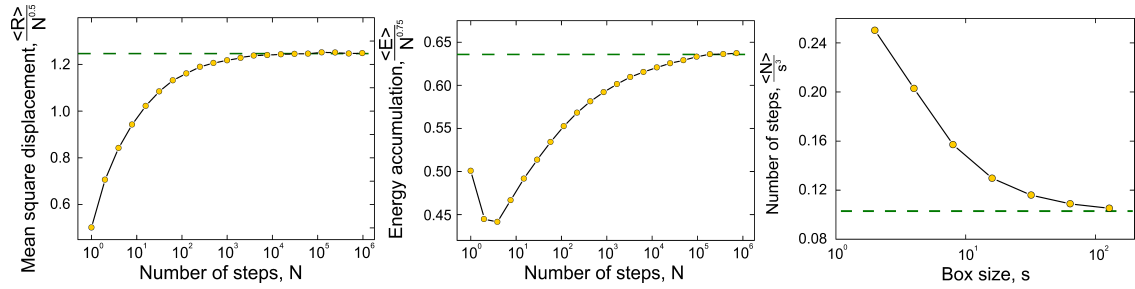
Figure 5. (Color online) (left): The average normalized span, $\langle R^2(N)\rangle$, of the folded trajectory on an infinite substrate, as a function of the number of steps, N; (center): The normalized cumulative folding energy, $E(N)/N^{3/4}$, as a function of the number of steps, N; (right): The typical normalized demixing time, τ/s^3, of the folded trajectory as a function of the bounding box size, s.

Now, let us pay attention to the dynamics of a single flag. We have already mentioned that the lamplighter passing across the flag, moves it on one lattice site to the left or to the right with equal probabilities. This means that the flag itself performs a random walk with the mean-square displacement $\langle \Delta x^2(s)\rangle \sim m$, where m is the number of jumps the flag has made (i.e., the number of times the lamplighter returned to the flag). Since the number of returns to a given point (the position of a flag), m, scales as $m \sim t^{1/2}$ for a one-dimensional random walk at large t, we immediately conclude that the mean-square displacement of the flag obeys the following scaling dependence

$$\langle \Delta x^2(t)\rangle \sim m \sim t^{1/2}. \tag{3.4}$$

The waiting times between sequential visits of a given flag by the lamplighter are distributed as the first-return times of a one-dimensional random walk, with a tail $t^{-3/2}$ at long times. This makes the movement of a single flag exactly the continuous-time random walk (CTRW) with the exponent 1/2 [3]. Note that since the movement of different flags is governed by a *single* lamplighter, they are correlated. However, in the first approximation one expects that it is possible to neglect these correlations and assume that the mean square distance, $d = \sqrt{\langle (x_{i+1}-x_i)^2\rangle}$ between two neighboring flags i and $i+1$ located at the points x_i and x_{i+1}, behaves as in (3.4), i.e. $d^2(t) \sim t^{1/2}$.

The process of annihilation of flags is described by the equation of a one-dimensional chemical kinetics $X + X \to 0$, where the reagent X designates a flag, and X experiences a sub-diffusive random walk at time t with the mean-square displacement $\langle \Delta x^2(t)\rangle \sim t^{1/2}$. Such a process has been widely studied in chemical kinetics — see, for example, [4].

Let $c(t)$ be the typical concentration of X at time t within a region visited by the lamplighter. We are interested in the survival probability of X in time, i.e., in the scaling behavior $c(t) \sim t^{-\alpha}$ at $t \to \infty$. Each time the two consecutive flags touch each other, they disappear. Since the flags cannot pass through each other, the typical distance between the reagents X (flags) at time t is of the order of the explored territory up to the time t. Therefore, $c(t) \sim d^{-1}(t)$ or

$$c(t) \sim t^{-1/4} \tag{3.5}$$

giving $\alpha = 1/4$.

The energy of the folded chain increases by one with the probability 1/2 every time the lamplighter visits a flagged site. There are two typical situations when it happens: a) when the lamplighter reaches the boundary of the domain visited (recall that outside the domain visited all sites are flagged), and b) when it encounters a flag in the bulk of the already visited region. It is easy to estimate the frequencies of both these events. Indeed, according to (3.3), the size of the domain visited grows as $t^{1/2}$, and the resulting energy increment scales as

$$\left.\frac{\mathrm{d}E(t)}{\mathrm{d}t}\right|_{\text{boundary}} \sim \frac{\mathrm{d}t^{1/2}}{\mathrm{d}t} \sim t^{-1/2}. \tag{3.6}$$

On the other hand, the probability for a random walk to meet a flag at time t in the domain's bulk, scales

as $c(t)$ and, therefore,

$$\left.\frac{dE(t)}{dt}\right|_{\text{bulk}} \sim d^{-1}(t) \sim t^{-1/4}. \tag{3.7}$$

We see that, at least on large timescales (i.e., for long chains), the second mechanism dominates, and the energy is mostly accumulated via interaction of the lamplighter with the flags in the bulk. Integrating (3.7) up to the total chain length, N, we end up with the scaling behavior of the average energy accumulation, $E(N)$ up to the length N ($N \gg 1$):

$$\langle E(N) \rangle \sim \int_0^N t^{-1/4} \, dt \sim N^{3/4}. \tag{3.8}$$

This scaling behavior is clearly seen in the numeric simulations shown in figure 5 (center), where we have plotted the normalized cumulative energy, $E(N)/N^{3/4}$, as a function of N. The results are obtained by averaging over 10^6 random block-copolymer packaging attempts on the infinite line.

Apart from the chain adsorption on an *infinite* one-dimensional substrate, it seems instructive to consider the case when the total size of a substrate is *finite*, e.g., the chain is trapped in a cage with periodic boundary conditions. In this case (for even total number of sites, L, in the cage), a long enough chain, once again, finally forms a state with completely regular . . . A-B-A-B . . . roof, which means, in terms of the Ising domain representation, that a single Ising domain is formed. The natural questions to ask are as follows: i) how long (on average) does it take for a perfect domain structure to be formed in the cage of L sites (i.e., which is the typical "demixing time", τ, as a function of L), and ii) what is the typical total energy, $\langle E(L) \rangle$ (the number of defects in the structure), as a function of the cage size, L. It turns out that the question ii) is easier to answer. Indeed, the energy E is the total number of the flag moves, regardless of the waiting times between them. Since the moves of the flags in both directions are equiprobable, the typical number of the moves needed for two adjacent flags to annihilate, scales as d^2, where d is the initial distance between them. Therefore, one expects the total accumulated energy to scale as

$$\langle E(L) \rangle \sim \sum_{n=0}^N d^2(n), \tag{3.9}$$

where n ($1 \leqslant n \leqslant L$) is the number of surviving flags at a given time (all numeric constants are omitted). Since the typical distance between the flags is $d(n) \sim L/n$, we can evaluate the sum in (3.9), getting finally

$$\langle E(L) \rangle \sim L^2 \tag{3.10}$$

for large L, which is perfectly confirmed by numeric simulations.

For the demixing time, we have not yet obtained the analytic scaling estimate, but according to the numeric simulations shown in figure 5 (right), the demixing time $\tau(L)$ scales as

$$\tau(L) \sim L^3. \tag{3.11}$$

Combining (3.10) and (3.11), we get a nontrivial prediction $\langle E(\tau) \rangle \sim \tau^{2/3}$, different from the result of the mean-field theory (3.8) in the infinite domain (τ plays the role of the chain length). Therefore, it turns out that in a bounded domain, the shifts of sequential flags are highly correlated since they are governed by a single lamplighter which cannot escape the cage. Indeed, when only several flags are left in a cage, the correlations between their shifts become very important. A more detailed consideration of this phenomenon will be provided elsewhere.

4. Discussion

We have already mentioned that our model is reminiscent to the so-called "lamplighter random walk", widely considered in the mathematical literature (see, for example, [1, 2, 5–11]). Here is the definition of the *lamplighter graph*, taken from [11]: ". . . Think of a (typically infinite) connected graph X, where in each vertex there is a lamp that may be switched off (state 0), or switched on with $q - 1$ different intensities (states $1, \ldots, q - 1$). Initially, all lamps are turned off, and a lamplighter starts at some vertex

of X and walks around. When he visits a vertex, he may switch the lamp sitting there into one of its q different states (including "off"). Our information consists of the position $x \in X$ of the lamplighter and of the finitely supported configuration $\eta : X \to \mathbb{Z}_q = \{0, \ldots, q-1\}$ of the lamps that are switched on, including their respective intensities. The set $\mathbb{Z}_q \wr X$ of all such pairs (η, x) can be in several ways equipped with a naturally connected graph structure, giving rise to a *lamplighter graph...*"

The key difference between the lamplighter model considered in the mathematical literature (math-LL) and the model analyzed in our work, consists in the following: in math-LL all lamps (governed by a single random walk) are immobile, are located at every lattice site, and do not interact with each other, while in our model, the flags (also governed by a single random walk) can perform random local shifts and interact (i.e., annihilate while in a contact). Despite the mentioned difference, we believe that some important questions, such as the computation of the "mixing time" [10] on lamplighter graphs, could be addressed in our model as well.

It seems interesting to compare the results of the local step-by-step optimization discussed at length in this work with the global optimization of the chain configuration. Namely, assume that after the described packaging, the frozen configuration of folds is annealed, i.e., the existing folds are allowed to retract and the new folds can be formed (for unchanged sequence of monomers). Clearly, the resulting ground state energy will be lower than the one obtained in our local optimization process. The partition function of such a system having an annealed secondary structure, apparently admits the recursive representation similar to the dynamic programming algorithm for RNA secondary structure [12]. In the forthcoming works we plan to provide a detailed analysis of the global optimization and its comparison with the results of local optimization for the lamplighter random walk.

Aknowledgements

References

1. Vershik A.M., In: Invariant Means on Topological Groups and Their Applications Greenleaf F.P., afterword to Russian edn., Mir, Moscow, 1973 (in Russian).
2. Kaimanovich V.A., Vershik A.M., Ann. Probab., 1983, **11**, 457–490; doi:10.1214/aop/1176993497.
3. Metzler R., Klafter J., Phys. Rep., 2000, **339**, 1–77; doi:10.1016/S0370-1573(00)00070-3.
4. Yuste S.B., Lindenberg K., Chem. Phys., 2002, **284**, 169–180; doi:10.1016/S0301-0104(02)00546-3.
5. Lyons R., Pemantle R., Peres Y., Ann. Probab., 1996, **24**, 1993–2006; doi:10.1214/aop/1041903214.
6. Grigorchuk R.I., Zuk A., Geometriae Dedicata, 2001, **87**, 209–244; doi:10.1023/A:1012061801279.
7. Kaimanovich V.A., Woess W., Ann. Probab., 2002, **30**, 323–363; doi:10.1214/aop/1020107770.
8. Revelle D., Ann. Probab., 2003, **31**, 1917–2300; doi:10.1214/aop/1068646371.
9. Revelle D., Electr. Commun. Probab., 2003, **8**, 142–154; doi:10.1214/ECP.v8-1092.
10. Peres Y., Revelle D., Electron. J. Probab., 2004, **9**, 825–845; doi:10.1214/EJP.v9-198.
11. Woess W., Comb. Probab. Comput., 2005, **14**, 415–433; doi:10.1017/S0963548304006443.
12. Nechaev S.K., Tamm M.V., Valba O.V., J. Phys. A: Math. Theor., 2011, **44**, 195001; doi:10.1088/1751-8113/44/19/195001.

Модель ліхтарника адсорбції випадкового кополімера на лінії

Л.І. Назаров[1], С.К. Нєчаєв[2,3,4], М.В. Тамм[1,4]

[1] Фізичний факультет, Московський державний університет ім. М.В. Ломоносова, Москва, РФ

[2] Університет Парі-Сюд/CNRS, Орсе, Франція

[3] Фізичний інститут ім. П.Н. Лєбєдєва, РАН, Москва, РФ

[4] Факультет прикладної математики, Національний дослідницький університет "Вища школа економіки", Москва, РФ

Ми представляємо модель послідовного самопакування AB-диблокового випадкового кополімера з локальними замороженими взаємодіями на одновимірній нескінченій липкій основі. Припускається, що контакти A-A і B-B є сприятливі, тоді як A-B є несприятливими. Положення нового мономера, що додається, вибирається з точки зору мінімізації енергії локального контакту. Модель демонструє самоорганізовану поведінку з нетривіальною залежністю загальної енергії, E (числа несприятливих контактів), від числа мономерів ланцюга, N: $E \sim N^{3/4}$ для замороженого хаотичного рівноймовірного розподілу A- і B-мономерів вздовж ланцюга. Модель розглядається шляхом зіставлення її з випадковим блуканням ліхтарника і дифузійно-контрольованою хімічною реакцією типу $X + X \to 0$ з субдифузійним рухом реагентів.

Ключові слова: *локальна оптимізіція, згортання гетерополімера, випадкове блукання ліхтарника, субдифузійна хімічна реакція*