

УДК: 577.323

О компьютерных экспериментах Касмана

А. А. Гриневич^{1,2,a}, Л. В. Якушевич¹

¹ Институт биофизики клетки Российской академии наук — обособленное подразделение
Федерального государственного бюджетного учреждения науки «Федеральный исследовательский центр
«Пушкинский научный центр биологических исследований Российской академии наук»,
Россия, 142290, Московская обл., г. Пушкино, ул. Институтская, д. 3

² Федеральное государственное бюджетное учреждение науки
«Институт теоретической и экспериментальной биофизики Российской академии наук»,
Россия, 142290, Московская обл., г. Пушкино, ул. Институтская, д. 3

E-mail: ^a grin_aa@mail.ru

Получено 25.02.2019, после доработки — 04.04.2019.

Принято к публикации 08.04.2019.

В 2007 году Касман провел серию оригинальных компьютерных экспериментов с кинками уравнения синус-Гордона, движущимися вдоль искусственных последовательностей ДНК. Были рассмотрены две последовательности. Каждая состояла из двух частей, разделенных границей. Левая часть первой из последовательностей содержала повторяющиеся триплеты ТТА, кодирующие лейцины, а правая часть содержала повторяющиеся триплеты СGC, кодирующие аргинины. Во второй последовательности левая часть содержала повторяющиеся триплеты СТG, кодирующие лейцины, а правая часть содержала повторяющиеся триплеты АGА, кодирующие аргинины. При моделировании движения кинка в этих последовательностях был обнаружен интересный эффект. Оказалось, что кинк, движущийся в одной из последовательностей, останавливался, не достигнув конца, а затем «отскакивал», как будто ударялся об стенку. В то же время в другой последовательности движение кинка не прекращалось в течение всего времени проведения эксперимента. В этих компьютерных экспериментах, однако, использовалась простая модель ДНК, предложенная Салерно. Она учитывает различия во взаимодействиях комплементарных оснований внутри пар, но пренебрегает различием в моментах инерции азотистых оснований и расстояниях между центрами масс оснований и сахарно-фосфатной цепочкой. Вопрос о том, сохранится ли эффект Касмана при использовании более точных моделей ДНК, до сих пор остается открытым. В настоящей работе мы исследуем эффект Касмана на основе более точной модели ДНК, которая учитывает оба эти различия. Мы получили энергетические профили последовательностей Касмана и построили траектории движения кинков, запущенных в этих последовательностях при разных начальных значениях энергии. Результаты наших исследований подтвердили существование эффекта Касмана, но только в ограниченном интервале начальных значений энергии кинков и при определенном направлении движения кинков. В других случаях этот эффект не наблюдался. Мы обсудили, какие из исследованных последовательностей энергетически были более предпочтительны для возбуждения и распространения кинков.

Ключевые слова: компьютерное моделирование, динамика ДНК, последовательности из кодонов ДНК, энергетический профиль, траектории кинков

UDC: 577.323

On the computer experiments of Kasman

A. A. Grinevich^{1,2,a}, L. V. Yakushevich¹

¹ Institute of Cell Biophysics, Russian Academy of Sciences,
3 Institutskaya st., Pushchino, Moscow oblast, 142290, Russia

² Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences,
3 Institutskaya st., Pushchino, Moscow oblast, 142290 Russia

E-mail: ^a grin_aa@mail.ru

Received 25.02.2019, after completion — 04.04.2019.

Accepted for publication 08.04.2019.

In 2007 Kasman conducted a series of original computer experiments with sine-Gordon kinks moving along artificial DNA sequences. Two sequences were considered. Each consisted of two parts separated by a boundary. The left part of the first sequence contained repeating TTA triplets that encode leucines, and the right part contained repeating CGC triplets that encode arginines. In the second sequence, the left part contained repeating CTG triplets encoding leucines, and the right part contained repeating AGA triplets encoding arginines. When modeling the kink movement, an interesting effect was discovered. It turned out that the kink, moving in one of the sequences, stopped without reaching the end of the sequence, and then “bounced off” as if he had hit a wall. At the same time, the kink movement in the other sequence did not stop during the entire time of the experiment. In these computer experiments, however, a simple DNA model proposed by Salerno was used. It takes into account differences in the interactions of complementary bases within pairs, but does not take into account differences in the moments of inertia of nitrogenous bases and in the distances between the centers of mass of the bases and the sugar-phosphate chain. The question of whether the Kasman effect will continue with the use of more accurate DNA models is still open. In this paper, we investigate the Kasman effect on the basis of a more accurate DNA model that takes both of these differences into account. We obtained the energy profiles of Kasman's sequences and constructed the trajectories of the motion of kinks launched in these sequences with different initial values of the energy. The results of our investigations confirmed the existence of the Kasman effect, but only in a limited interval of initial values of the kink energy and with a certain direction of the kinks movement. In other cases, this effect did not observe. We discussed which of the studied sequences were energetically preferable for the excitation and propagation of kinks.

Keywords: computer simulations, DNA dynamics, sequences consisting of DNA codons, energy profile, kink trajectories

Citation: *Computer Research and Modeling*, 2019, vol. 11, no. 3, pp. 503–513 (Russian).

1. Introduction

The question of whether there is a relationship between the functional and dynamic properties of DNA has occupied the minds of researchers for a long time. Despite this, there is still no complete answer to it, as well as there are not enough simple and convincing examples confirming or disproving this relationship.

In 2007 at the IMA Workshop “Mathematics of DNA Structure, Function, and Interactions” [IMA Workshop..., 2007] Alex Kasman presented one of the examples. To study the relationship and might be to explain the existence of codon bias [Kasman, LeMesurier, 2007], he proposed to model numerically the movement of DNA kinks in the two artificial sequences shown in Fig. 1.

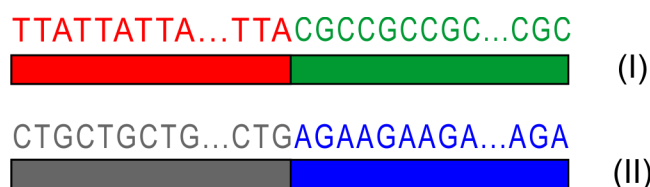


Fig. 1. Two artificial sequences used in the Kasman's computer experiments

Each of the sequences shown in Fig. 1 consists of two parts separated by a boundary. In the sequence (I), the left part contains repeating triplets TTA coding leucines, and the right part contains repeating triplets CGC coding arginines. In the sequence (II), the left part contains repeating triplets CTG coding leucines, and the right part contains repeating triplets AGA coding arginines. Thus, the sequences (I) and (II) have different primary structures but the same functional properties. It would be interesting to elucidate if the sequences have identical dynamic properties. If not, it would be interesting to answer the question: can the difference in the dynamic properties of the kinks in the sequences (I) and (II) explain why some of the codons have privileges in comparison with the other?

To answer these questions, Kasman conducted a series of original computer experiments with the sine-Gordon kinks moving along the sequences (I) and (II) from the right parts consisting of triplets coding arginines to the left parts consisting of triplets coding leucines. As a result, he found that the kink moving in one of the sequences stopped before reaching the end. Then the kink “bounced” and reversed direction as if it hit a wall. At the same time the kink moving in the other sequence did not stop. It travelled all the way through. The conclusion was: the dynamic properties of the sequences (I) and (II) were different.

To imitate numerically the kinks movement, Kasman used the model of Salerno [Salerno, 1991] which described rotation oscillations of nitrous bases in the inhomogeneous DNA double chain. By that time there have existed already several nonlinear models of DNA including the models of Englander et al. [Englander et al., 1980], Yomosa [Yomosa, 1983], Takeno and Homma [Takeno, Homma, 1983], Yakushevich [Yakushevich, 1989], but only the model proposed by Salerno took into account the differences in the interactions between nitrous bases inside the base pairs (two bonds in the AT pair and three bonds in the GC pairs) that permitted to apply it to imitate kinks dynamics in the inhomogeneous DNA double chain.

From that time when the model of Salerno has been published, many other more realistic mathematical models imitating kinks dynamics in the DNA sequences have been proposed [Gaeta et al., 1994; Yakushevich, 2004; Peyrard, 2004; Peyrard, Dauxois 2014]. In our recent works [Grinevich, Yakushevich, 2015; Grinevich et al., 2015] we also have proposed the model where not only the differences in the interactions between complementary bases inside the pairs, but also the differences in the moments of inertia of nitrous bases and in the distances between the centers of mass of bases and the sugar-phosphate chain were taken into account. This model permitted to imitate kinks dynamics in any inhomogeneous sequences, to calculate the energy profiles of the sequences, to launch kinks at different energy/velocity values and to obtain their trajectories. The goal of this work is to apply just this model to study kink dynamics in the Kasman's sequences (I) and (II).

In the next section, we describe both models: the Salerno–Kasman model and our (GRY) model. Then, we describe the main approximations and artificial sequences used in the computer experiments. After that, we present the results: the energy profiles of the sequences and the trajectories of the kinks launched at different initial energies in these sequences. In the final section, we compare the results obtained in the frameworks of the two models, discuss advantages and disadvantages of our approach as well as conclude which of the studied sequences are more preferable energetically for the excitation and propagation of kinks.

2. Two models of the DNA kink dynamics

2a. Salerno–Kasman model

Model equations applied by Kasman to describe DNA kink dynamics, have the form [Kasman, LeMesurier, 2007]:

$$\frac{d^2\phi_n}{d\tau^2} = (\phi_{n+1} - 2\phi_n + \phi_{n-1}) - \frac{2e_n}{5}\sin(\phi_n), \quad (1)$$

where $\phi_n(t) = \theta_n(t) - \psi_n(t)$. $\theta_n(t)$ and $\psi_n(t)$ are the angular deviations of the n -th nitrous bases in the first and second polynucleotide chains, respectively. $e_n \in \{2, 3\}$. Thus $e_n = 2$ if adenine (A) or thymine (T) is placed at the n -th point, and $e_n = 3$, if guanine (G) or cytosine (C) is placed at the n -th point. $n = 1, 2, \dots, N$.

These equations are similar to those proposed by Salerno [Salerno, 1991]:

$$I \ddot{\phi}_n = K(\phi_{n+1} - 2\phi_n + \phi_{n-1}) - V_n \sin(\phi_n), \quad (2)$$

where K denotes the rigidity of the sugar-phosphate chains, I is the moment of inertia of the individual bases, V_n characterizes the strength of hydrogen bonds between complementary bases in the n -th pair, which is determined as

$$V_n = \lambda_n \beta,$$

where $\lambda_n = 2$ if it refers to A-T or T-A base pairs, $\lambda_n = 3$ if it refers to G-C or C-G base pairs, with β is a free parameter to be determined later.

Indeed, let us introduce a new variable:

$$\tau = \alpha t,$$

where $\alpha = \sqrt{K/I}$. Then the model equations (2) transform to:

$$\frac{d^2\phi_n}{d\tau^2} = (\phi_{n+1} - 2\phi_n + \phi_{n-1}) - \frac{\beta}{K}\lambda_n \sin(\phi_n). \quad (3)$$

Eqs. (3) coincide with Kasman's equations (1) in the case $\beta/K = 2/5$. This is why we use the name Salerno–Kasman model for both equations: Eqs. (1) and Eqs. (3).

2b. GRY model

The GRY model proposed by Grinevich, Ryasik and Yakushevich, takes into account the differences (1) in the interactions of complementary nitrous bases in pairs, (2) in the moments of inertia of the bases and (3) in the distances between the bases and the sugar phosphate chain [Grinevich, Yakushevich, 2015; Grinevich et al., 2015]. In this model, angular oscillations of nitrous bases are described by the following discrete equations:

$$I_n \frac{d^2\theta_n}{dt^2} + V_n \sin(\theta_n) - KR_n(R_{n+1}\theta_{n+1} - 2R_n\theta_n + R_{n-1}\theta_{n-1}) = 0. \quad (4)$$

Here $\theta_n(t)$ is the angular deviation of the n -th base from its equilibrium position. I_n is the moment of inertia of the n -th base, R_n is the distance from the center of mass of the n -th base to the sugar-phosphate chain, K is the rigidity of the sugar-phosphate chain, V_n is the coefficient characterizing interactions (hydrogen bonds) between the bases inside the n -th base pair. The values of these parameters are presented in Table 1.

Table 1. Coefficients of the model Eqs. (4)

n -th base	$I_n \times 10^{-44}$ (kg·m ²)	$V_n \times 10^{-20}$ (J)	$K'_n = KR_n^2 \times 10^{-18}$ (J)	$R_n \times 10^{-10}$ (m)
A	7.61	2.09	2.27	5.8
T	4.86	1.43	1.56	4.8
G	8.22	3.12	2.20	5.7
C	4.11	2.12	1.50	4.7

Quasi-homogeneous approximation. In the quasi-homogeneous approximation, proposed in [Yakushevich, Krasnobaeva, 2008], the model Eqs. (4) take the form:

$$\bar{I} \frac{d^2\theta_n}{dt^2} + \bar{V} \sin(\theta_n) - K\bar{R}^2(\theta_{n+1} - 2\theta_n + \theta_{n-1}) = 0, \tag{5}$$

with the averaged coefficients are equal to:

$$\begin{aligned} \bar{I} &= I_A \frac{N_A}{N} + I_T \frac{N_T}{N} + I_G \frac{N_G}{N} + I_C \frac{N_C}{N}, \\ \bar{V} &= V_A \frac{N_A}{N} + V_T \frac{N_T}{N} + V_G \frac{N_G}{N} + V_C \frac{N_C}{N}, \\ \bar{R} &= R_A \frac{N_A}{N} + R_T \frac{N_T}{N} + R_G \frac{N_G}{N} + R_C \frac{N_C}{N}. \end{aligned}$$

This approximation allows taking into account the dependence of the coefficients on the composition of the inhomogeneous sequence (but not on the location of nucleotides along the polynucleotide chain).

Continuum approximation. If we suggest that the solutions of Eqs. (5) are rather smooth, the continuum approximation can be applied:

$$\begin{aligned} a &\rightarrow 0, \\ z_n &\rightarrow z, \\ \theta_n(t) &\equiv \theta(z_n, t) \rightarrow \theta(z, t), \end{aligned}$$

where a is the distance between the nearing base pairs and z_n is the coordinate of the n -th base. Then Eqs. (5) take the form:

$$\bar{I} \frac{d^2\theta}{dt^2} + \bar{V} \sin(\theta) - \bar{K}' a^2 \frac{d^2\theta}{dz^2} = 0, \tag{6}$$

where $\bar{K}' = K\bar{R}^2$.

The kink-like solution of the Eq. (6) is:

$$\theta(z, t) = 4 \arctan \left\{ \exp \left[\frac{\eta}{d} (z - vt - z_0) \right] \right\},$$

where v is the kink velocity, $d = a(\bar{K}'/\bar{V})^{1/2}$ is the kink size, $\eta = (1 - (v^2/C_0^2))^{-1/2}$, $C_0 = a(\bar{K}'/\bar{I})^{1/2}$ is the sound velocity in DNA, z_0 is an arbitrary constant.

The total kink energy is defined by formula

$$\bar{E} = \frac{\bar{E}_0}{\sqrt{1 - \frac{v^2}{\bar{C}_0^2}}}, \quad (7)$$

where $\bar{E}_0 = 8\sqrt{\bar{K}'\bar{V}}$ is the rest energy of the kink.

3. The sequences under consideration

In addition to the Kasman's sequences (I) and (II) let us consider four artificial DNA sequences ((III)–(VI)) consisting of periodically repeating triplets of nucleotides (codons). They are shown in Fig. 2. The two of them ((III) and (IV)) consist of triplets coding arginine, the other two ((V) and (VI)) consist of triplets coding leucine.

The averaged coefficients \bar{I} , \bar{K} , and \bar{V} for the sequences (III)–(VI), as well as the parameter a , used further for numerical calculations, are given in Table 2.

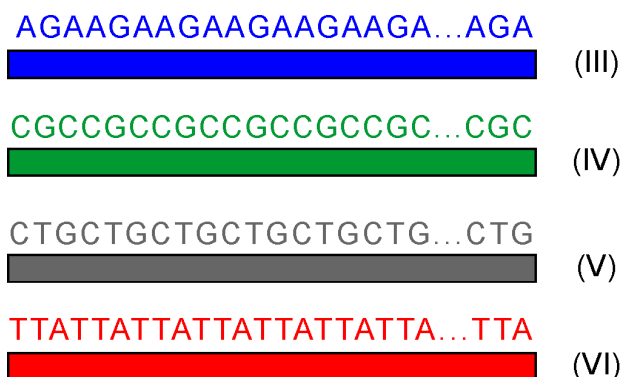


Fig. 2. Four artificial sequences consisting of periodically repeating triplets of nucleotides

Table 2. Averaged coefficients for the sequences (III)–(VI) (Fig. 2)

Sequences	$\bar{I} \cdot 10^{-43}$ (kg·m ²)	$\bar{K}' \cdot 10^{-17}$ (J)	$\bar{V} \cdot 10^{-19}$ (J)	$a \cdot 10^{-10}$ (m)
(III)	0.781	0.232	0.243	3.4
(IV)	0.548	0.179	0.245	3.4
(V)	0.573	0.181	0.222	3.4
(VI)	0.578	0.186	0.165	3.4

4. Results

4a. Rest energy of the kinks in the sequences (III)–(VI)

Suppose that at the initial moment the kink velocity is zero. If the external fields are absent, the kink velocity remains zero, and the kink energy is equal to the rest energy that is determined by the following formula:

$$\bar{E}_0 = 8\sqrt{\bar{K}'\bar{V}}.$$

In the quasi-homogeneous approximation, this value of the rest energy will not change if we shift the kink to any other point of the chain. Therefore, the graph of the kink energy is a straight horizontal line. In Fig. 3 we present the four straight lines corresponding to the rest energies of the kinks in the sequences (III)–(VI).

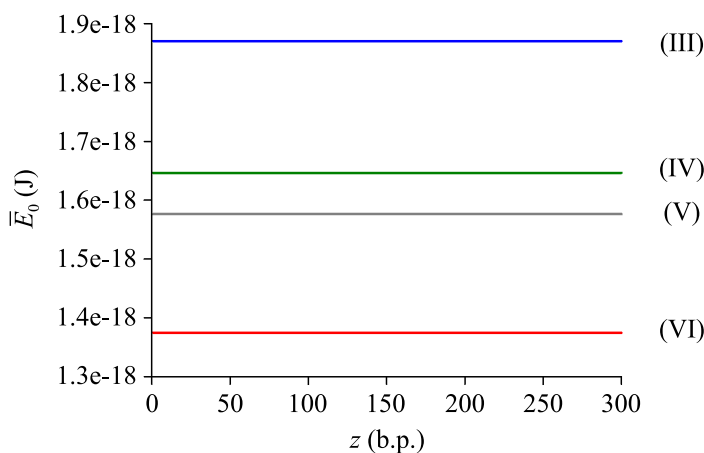


Fig. 3. Rest energy levels

From Fig. 3 it can be seen that to synthesize a chain of leucines, it is more advantageous (energetically) to use the sequence (VI), consisting of TTA triplets, than the sequence (V), consisting of CTG triplets. Similarly, for the synthesis of a chain of arginines, it is preferable to use the sequence (IV) consisting of CGC triplets, than the sequence (III) consisting of AGA triplets.

4b. Energy profile of the sequences (I)–(II)

To describe the change of the coefficients (\bar{I} , \bar{K}' , \bar{V}) at the crossing of the boundaries in the Kasman’s sequences (I)–(II), it is convenient to use sigma functions:

$$\begin{aligned}
 I(z) &= \bar{I}_{left} + \frac{\bar{I}_{right} - \bar{I}_{left}}{1 + \exp((z_b - z)/\sigma)}, \\
 K(z) &= \bar{K}'_{left} + \frac{\bar{K}'_{right} - \bar{K}'_{left}}{1 + \exp((z_b - z)/\sigma)}, \\
 V(z) &= \bar{V}_{left} + \frac{\bar{V}_{right} - \bar{V}_{left}}{1 + \exp((z_b - z)/\sigma)},
 \end{aligned}
 \tag{8}$$

where the index “left” means that this value of the coefficient refers to the left part of the Kasman's sequences, and the index “right” means that this value of the coefficient refers to the right part of these sequences, z_b is the coordinate of the boundary between the left and right parts of the Kasman’s sequences. The left and right parts of the Kasman’s sequences correspond to the sequences shown in Fig. 2. σ is the sigmoid parameter which is defined by formula $\sigma = \sigma'/v$, where $\sigma' = 1$ is the dimensionless sigmoid parameter, and $v = \sqrt{\bar{V}_{left}/(\bar{K}'_{left} a^2)}$.

Then, with the help of formulas (7) and (8), we constructed the energy profile of the Kasman’s sequences. The results of the calculations are shown in Fig. 4, a and Fig. 4, b by solid lines.

From Fig. 4 it is visible that both profiles have a barrier. This suggests that in order for a kink to go through the entire sequence in the direction from the left part that consists of triplets coding leucines to the right part consisting of triplets coding arginines, the total energy of the kink should be higher or equal to the height of the barrier. This condition can be written in the form:

$$\frac{\bar{E}_0^{left}}{\sqrt{1 - \left(\frac{v^{left}}{\bar{C}_0^{left}}\right)^2}} \geq \bar{E}_0^{right}.
 \tag{9}$$

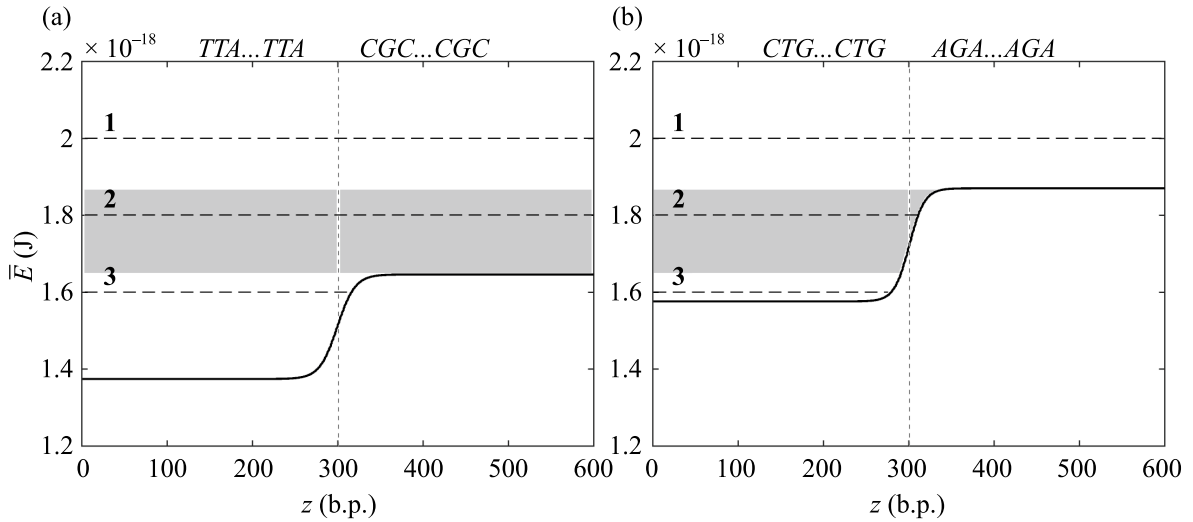


Fig. 4. (a) The energy profile of the sequence (I) and (b) the energy profile of the sequence (II). Horizontal dashed lines marked with numbers 1, 2, and 3, correspond to the initial kinks energies $\bar{E}_1 = 2.0 \times 10^{-18}$ J, $\bar{E}_2 = 1.8 \times 10^{-18}$ J, and $\bar{E}_3 = 1.6 \times 10^{-18}$ J, respectively. Vertical dashed lines indicate the locations of the boundary between the left and right parts of the sequences. The interval of the kink energies necessary both to overcome the energy barrier in the sequence (I) and to stop in the sequence (II), is shown by the grey band

Inserting the values of the dynamic coefficients from Table 2 into (9), we find minimum energy of passage of the barrier in the sequence (I), which is equal to:

$$\bar{E}_{0,I}^{right} = 1.647 \times 10^{-18} \text{ J},$$

the minimum energy of passage of the barrier in the sequence (II) is equal to:

$$\bar{E}_{0,II}^{right} = 1.871 \times 10^{-18} \text{ J}.$$

So, we can conclude that the effect observed by Kasman takes place only if the initial energy of the kink moving from the left part consisting of triplets coding leucines to the right part consisting of triplets coding arginines, satisfies the condition:

$$1.647 \times 10^{-18} < \bar{E} < 1.871 \times 10^{-18}. \quad (10)$$

4c. Kink trajectories in the sequences (I)–(II)

The other effect observed by Kasman concerns the kink behavior after stopping. It was observed that the kink not only stopped without reaching the end of the sequence, but also “bounced off” as if he had hit the wall. To investigate this case, we used the method of trajectories developed in [Grinevich, Yakushevich, 2015; Grinevich et al., 2015]. This method permitted us to construct the trajectories of kinks moving in the sequences (I) and (II). In Fig. 5 we present these trajectories calculated for three different values of the kink initial energy.

From Fig. 5, it can be seen that the trajectories marked with number 1 demonstrate that kink with energy $\bar{E}_1 = 2.0 \times 10^{-18}$ J overcomes the boundary in both the sequence (I) and the sequence (II). The trajectories marked with number 3 demonstrate that kink with energy $\bar{E}_3 = 1.6 \times 10^{-18}$ J is reflected from the boundary in both sequences. The trajectories marked with number 2 demonstrate effect of Kasman, namely, the kink with energy $\bar{E}_2 = 1.8 \times 10^{-18}$ J “reflects” from the boundary in the sequence (II) and overcomes the boundary in the sequence (I). These results confirm our conclusion that the effect of Kasman takes place if the initial energy of the kink moving from the left part that consists of triplets coding leucines to the right part consisting of triplets coding arginines, satisfies the condition (10).

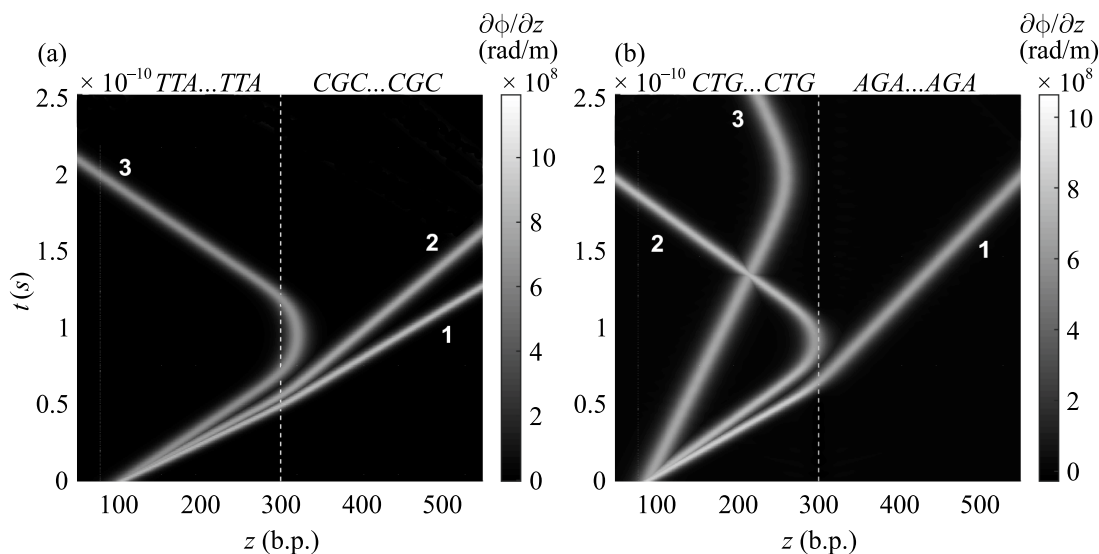


Fig. 5. (a) Kink trajectories in the sequence (I) and (b) in the sequence (II). Trajectories marked with numbers 1, 2 and 3, correspond to the kinks with the initial energies $\bar{E}_1 = 2.0 \times 10^{-18}$ J, $\bar{E}_2 = 1.8 \times 10^{-18}$ J, and $\bar{E}_3 = 1.6 \times 10^{-18}$ J, respectively. Vertical dashed lines indicate the location of the boundary between the left and right parts of the sequences

5. Conclusions and discussion

In the present work, the kink dynamics in the Kasman's sequences have been investigated. The energy profiles of the sequences were constructed, and the trajectories of the kinks were calculated at different values of the kink initial energy. The results of the study confirm the existence of the effect described by Kasman, but only for a limited range of values of the initial energy of the kinks and for a certain direction of the kinks movement. They also show that although the functional properties of the Kasman's sequences are the same, their dynamic properties are different.

Comparing the codons in the Kasman's sequences, we found that some codons have an advantage (in energy) over others. In particular, from Fig. 3 it follows that for the synthesis of the chain consisting of leucines, it is more advantageous (energetically) to use the sequence (VI), consisting of repeating TTA codons, than the sequence (V), consisting of repeating GTG codons. Similarly, for the synthesis of the chain of arginines, it is preferable to use the sequence (IV) consisting of CGC codons than the sequence (III) consisting of AGA codons.

We suggest that this result has a more general meaning. To confirm this suggestion, we considered all triplets encoding leucines and arginines (Fig. 6). The values of the rest energy of the kinks in these sequences are gathered in the second and fifth columns of Table 3. They show that, the sequence consisting of CTT codons is most preferred (energetically) for the synthesis of leucines, and the sequence consisting of CGT codons is most preferred for the synthesis of arginines.

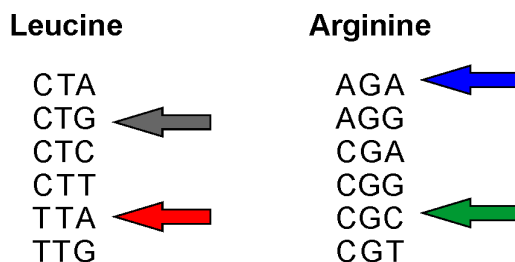


Fig. 6. DNA triplets encoding leucines (left) and arginines (right). The arrows show the four triplets that were used in the computer experiments of Kasman

Earlier it was believed that the preference can be determined by simple counting the number of hydrogen bonds in codons. We calculated these numbers (see the third and sixth columns of Table 3) and found that codons with the least amount of hydrogen bonds between the complementary bases differ from codons with the smallest values of the rest energy of the kinks. For example, results based on counting the number of hydrogen bonds point out that the sequence consisting of TTA codons (in the case of leucines) and the sequence consisting of AGA (in the case of arginines) are preferable. At the same time the rest energy calculations point out that the sequence consisting of CTT codons (in the case of leucines) and the sequence consisting of CGT codons (in the case of arginines) are preferable. Thus, the question of what of the method of estimations of the codon preferability is better, remains open.

Table 3. The number of hydrogen bonds in the DNA triplets encoding leucine and arginine, and the rest energies of the kinks activated in the chains consisting of these codons

Triplets encoding leucine	Rest energy $E_0 \cdot 10^{-17}$ (J)	Number of hydrogen bonds	Triplets encoding arginine	Rest energy $E_0 \cdot 10^{-17}$ (J)	Number of hydrogen bonds
TTA	0.140	6	AGA	0.190	7
CTG	0.160	8	CGC	0.167	9
CTA	0.149	7	AGG	0.179	8
CTC	0.138	8	CGA	0.179	8
CTT	0.130	7	CGC	0.167	9
TTG	0.153	7	CGT	0.160	8

We made one more interesting observation associated with the estimations of the codon preferability. If one calculates the number of hydrogen bonds per one codon, to the left and to the right of the boundaries separating the left and right parts in sequences (I) and (II), the following results can be obtained:

1. In the case of the sequence (I) consisting of TTA codons to the left and CGC codons to the right of the boundary, we obtain, according to Table 3, the number 6 to the left and 7 to the right. From this it follows that a kink moving from right to left (as in Kasman's model experiments) will "fall" into the potential well and, therefore, will pass through the boundary.

2. In the case of the sequence (II) consisting of CTG codons to the left and AGA to the right of the boundary, we obtain, according to Table 3, the number 8 to the left and the 7 to the right. Therefore, the kink moving from right to left (as in Kasman's model experiments) will "stumble" against a barrier that can stop its movement. Just these results allowed Kasman to explain the effect where the kink in sequence (I) passed through the boundary, and in sequence (II), reflected from the barrier.

Calculations of the energy profile within our model (see Fig. 4) show, however, that to reproduce the Kasman effect, it is necessary to "launch" the kinks in the other direction: from left to right. Then the kinks will be able to pass through the boundary in sequence (I) and reflect from the boundary in sequence (II), but only within a certain selected interval of the initial values of the kink energies.

It is necessary to note that all results described above, have been obtained in the frameworks of a rather simple DNA model that does not take into account effects of dissipation, the influence of the DNA torque, the helicity of the DNA structure and angular oscillations in both polynucleotide chains. But we suggest that the main features of the kinks behavior will be observed also in the case of more complex modifications of the DNA model which include all these effects.

7. Acknowledgments

It is a pleasure to acknowledge Prof. Alex Kasman for interesting and stimulating discussion

References

- Englander S. W., Kallenbach N. R., Heeger A. J. et al. Nature of the open state in long polynucleotide double helices: possibility of soliton excitations // Proceedings of the National Academy of Sciences of the United States of America. — 1980. — Vol. 77, No. 12. — P. 7222–7226.
- Gaeta G., Reiss C., Peyrard M., Dauxois T. Simple models of non-linear DNA dynamics // Rev Nuovo Cimento. — 1994. — Vol. 17. — P. 1–48.
- Grinevich A. A., Ryasik A. A., Yakushevich L. V. Trajectories of DNA bubbles // Chaos, Solitons & Fractals. — 2015. — Vol. 75. — P. 62–75.
- Grinevich A. A., Yakushevich L. V. Kinks behavior near the boundaries separating homogeneous regions of DNA // Mathematical Biology and Bioinformatics. — 2015. — Vol. 10, No. 1. — P. 164–177.
- IMA Workshop on Mathematics of DNA Structure, Function, and Interactions, September 16–21, 2007 [Electronic resource]. — URL: <https://www.ima.umn.edu/2007-2008/W9.16-21.07> (accessed: 12.02.2019).
- Kasman A., LeMesurier B. Codon Bias, DNA transcription forks, and Kink-Solitons, доклад на конференции IMA Workshop on Mathematics of DNA Structure, Function, and Interactions, September 16–21, 2007 [Electronic resource]. — URL: <http://kasmana.people.cofc.edu/DNA-soliton-poster.pdf> (accessed: 12.02.2019).
- Peyrard M., Dauxois T. Can we model DNA at the mesoscale? // Phys. Life Rev. — 2014. — Vol. 11. — P. 173–175.
- Peyrard M. Nonlinear dynamics and statistical physics of DNA // Nonlinearity. — 2004. — Vol. 17. — P. R1–R40.
- Salerno M. Discrete model for DNA-promoter dynamics // Phys.Rev. A. — 1991. — Vol. 44. — P. 5292–5297.
- Takeno S., Homma S. Topological Solitons and Modulated Structure of Bases in DNA Double Helices: A Dynamic Plane Base-Rotator Model // Prog. Theor. Phys. — 1983. — Vol. 70, No. 1. — P. 308–311.
- Yakushevich L.V. Nonlinear DNA dynamics: a new model // Phys. Letters A. — 1989. — Vol. 136, No. 7-8. — P. 413–417.
- Yakushevich L. V. Nonlinear Physics of DNA. — Wiley, 2004.
- Yomosa S. Soliton excitations in deoxyribonucleic acid (DNA) double helices // Phys. Rev. A. — 1983. — Vol. 27, No. 4. — P. 2120–2125.
- Якушевич Л. В., Краснобаева Л. А. Особенности динамики кинка в неоднородной ДНК // Биофизика. — 2008. — Т. 53, № 1. — С. 36–41.
- Yakushevich L. V., Krasnobaeva L. A. Peculiar features of kink dynamics in inhomogeneous DNA // Biophysics. — 2008. — Vol. 53, No. 1. — P. 21–25. (Original Russian paper: Yakushevich L. V., Krasnobaeva L. A. Osobennosti dinamiki kinka v neodnorodnoi DNK // Biofizika. — 2008. — Vol. 53, No. 1. — P. 36–41.)