

О ПАРАМЕТРАХ ДИНАМИЧЕСКИХ МОДЕЛЕЙ ДНК

Якушевич Л.В.

(Пушино)

Математические модели динамики ДНК содержат параметры, характеризующие внутренние движения и взаимодействия. Наиболее важные из них массы, моменты инерции, расстояния между структурными элементами, коэффициенты жесткости. Выбор параметров и соотношений между ними очень важны, поскольку определяют вид и характер решений соответствующих динамических уравнений. В этой работе мы рассматриваем проблему выбора параметров и получаем оптимальный их набор.

ON THE PARAMETERS OF THE DYNAMICAL MODELS OF DNA

Yakushevich L.V.

(Pushchino)

Mathematical models of the DNA dynamics contain parameters characterizing the internal motions and interactions. The most important of them are mass, the moments of inertia, distances between structural elements, coefficients of the rigidity. The choice of the parameters and the relations between them are very important because they determine the type of the solutions of corresponding dynamical equations. In this article we consider the problem of the choice of the parameters and obtain an optimal set of them.

Introduction

Dynamics of DNA is widely considered as an important factor that should be taken into account when studying the biological functioning of the molecule. Many different models of the internal DNA dynamics are known [1], and they all contain parameters characterizing the internal motions and interactions. The most important of

them are mass of moving structural elements, distances between them, the moments of inertia, coefficients of the rigidity. The values of the parameters and the relations between them are very important because they determine the type of the solutions of corresponding dynamical equations. In this article we consider the problem of the choice and estimation of the parameters and obtain an optimal set of them. To be concrete, we consider the parameters of the mathematical model proposed in [2] to describe rotational motions of bases around sugar-phosphate chains.

The dynamical model

Let us consider the homogeneous double chain consisting of only AT base pairs

1-st chain: ...AAAAAAAAAAAAAAAAAAAAAAAAA...
2-nd chain: ...TTTTTTTTTTTTTTTTTTTTTTTTTTT...

and write the model hamiltonian [2]

$$H = T + V_{\parallel} + V_{\perp}; \tag{1}$$

with the kinetic energy (T), the energy of interactions along the chains (V_{\parallel}) and the energy of interactions between bases in pairs (V_{\perp}) being determined by formulas

$$T = \sum_n \{ (m_1 r_1^2 / 2) (d\varphi_{n,1} / dt)^2 + (m_2 r_2^2 / 2) (d\varphi_{n,2} / dt)^2 \}; \tag{2}$$

$$V_{\parallel} = \sum_n \{ (K_1 r_1^2) [1 - \cos(\varphi_{n,1} - \varphi_{n-1,1})] + (K_2 r_2^2) [1 - \cos(\varphi_{n,2} - \varphi_{n-1,2})] \}; \tag{3}$$

$$V_{\perp} = \sum_n (k_{1-2}) \{ r_1 (r_1 + r_2) (1 - \cos \varphi_{n,1}) + r_2 (r_1 + r_2) (1 - \cos \varphi_{n,2}) - r_1 r_2 [1 - \cos(\varphi_{n,1} - \varphi_{n,2})] \} \tag{4}$$

were $\varphi_{n,i}$ is the angular displacement of the n-th base of the i-th chain from its equilibrium position; r_i is the distance between the center of mass of the i-th base and the nearest sugar-phosphate chain; a is the distance between neighboring bases along the chains; m_i is mass of bases of the i-th chain; K_i is the coupling constant along the sugar-phosphate chain; k_{1-2} is the force constant that characterizes interactions between bases in pairs; $n = 1, 2, \dots N$; $i = 1, 2$. We suggest that N is rather large integer, so the end effects can be neglected.

Dynamical equations corresponding to hamiltonian (1), can be obtained from the equations of Hamilton

$$d\varphi_{n,i}/dt = \partial H/\partial p_{n,i}; \quad (5)$$

$$dp_{n,i}/dt = -\partial H/\partial \varphi_{n,i}; \quad (6)$$

where impulses, $p_{n,i}$ are determined by formula

$$p_{n,i} = \partial L/\partial \dot{\varphi}_{n,i}; \quad (7)$$

and Lagrangian, L , by formula

$$L = T - V_{\parallel} - V_{\perp}; \quad (8)$$

Inserting (1) into (5) – (6) we obtain the dynamical equations

$$\begin{aligned} m_1 r_1^2 (d^2\varphi_{n,1}/dt^2) = \\ = K_1 r_1^2 [\sin(\varphi_{n-1,1} - \varphi_{n,1}) - \sin(\varphi_{n,1} - \varphi_{n+1,1})] - k_{1-2} [r_1 (r_1 + r_2) \sin\varphi_{n,1} - \\ r_2 r_1 \sin(\varphi_{n,1} - \varphi_{n,2})]; \end{aligned} \quad (9)$$

$$\begin{aligned} m_2 r_2^2 (d^2\varphi_{n,2}/dt^2) = \\ = K_2 r_2^2 [\sin(\varphi_{n-1,2} - \varphi_{n,2}) - \sin(\varphi_{n,2} - \varphi_{n+1,2})] - k_{1-2} [r_2 (r_1 + r_2) \sin\varphi_{n,2} - \\ r_2 r_1 \sin(\varphi_{n,2} - \varphi_{n,1})]; \end{aligned} \quad (10)$$

which describe rotational motions of bases around sugar-phosphate chains.

Parameters

Mathematical model (9) – (10) contains the following parameters:

- * masses of bases (m_1, m_2);
- * distances between centers of mass of the bases and the nearest sugar-phosphate chain (r_1, r_2);
- * force constant that characterizes interaction between bases in pairs (k_{1-2});
- * coupling constant along the sugar-phosphate chains (K_1, K_2).

Let us estimate the parameters listed above and other physical values related with them.

masses of bases (m_1, m_2)

Masses of bases ($m_i, i = A, T, G, C$) can be taken from any reference book on chemistry, where they are usually presented in terms of mass of protons (m_p).

$$m_A = 135,13 m_p; m_T = 126,11 m_p; m_G = 151,14 m_p; m_C = 111,10 m_p; \quad (11)$$

Inserting mass of proton

$$m_p = 1,67343 \times 10^{-27} \text{ kg}; \quad (12)$$

into (11) we obtain the following values for bases in AT chain

$$m_1 = m_A = 226,13 \times 10^{-27} \text{ kg}; m_2 = m_T = 211,04 \times 10^{-27} \text{ kg}; \quad (13)$$

And for bases in GC chain we have

$$m_1 = m_G = 252,92 \times 10^{-27} \text{ kg}; m_2 = m_C = 185,92 \times 10^{-27} \text{ kg}. \quad (14)$$

b) distances between centers of mass of the bases and the nearest sugar-phosphate chain (r_1, r_2); moments of inertia (I_1, I_2)

To estimate the distances, we used the data on the geometry of B-DNA taken from the Data Bank of National Center for Biotechnology (<http://ncbi.nlm.nih.gov/Entrez>) and adopted them to simplified model (1). As a result, for AT chain we obtained the following values

$$r_1 = r_A = 5,8 \text{ \AA}; r_2 = r_T = 4,8 \text{ \AA}. \quad (15)$$

And for GC chain we obtained

$$r_1 = r_G = 5,7 \text{ \AA}; r_2 = r_C = 4,7 \text{ \AA}. \quad (16)$$

Then the moments of inertia for bases in AT chain are

$$I_1 = I_A = m_A r_A^2 = 7607,03 \text{ m}^2 \text{ kg}; I_2 = I_T = m_T r_T^2 = 4862,28 \text{ m}^2 \text{ kg}. \quad (17)$$

And for bases in GC chain they are

$$I_1 = I_G = m_G r_G^2 = 8217,44 \text{ m}^2 \text{ kg}; I_2 = I_C = m_C r_C^2 = 4106,93 \text{ m}^2 \text{ kg}. \quad (18)$$

the force constant (k_{1-2}) that characterizes interaction between bases in pairs

To estimate the value k_{1-2} , let us take into account that the interactions between the bases in pairs are formed by two hydrogen bonds in the case of AT chain and by three hydrogen bonds in the case of GC chain. The energy required to broke one hydrogen bond is equal to $\varepsilon_H = 3 \div 7 \text{ kcal/mol}$ [3]. For calculations let us take the value equal to $\varepsilon_H = 5 \text{ kcal/mol}$ (20,934 kJ/mol). So, the average value of energy required to break and open one AT base pair is equal to

$$\varepsilon_{AT} = 2\varepsilon_H = 10 \text{ kcal/mol} \cong 41,868 \text{ kJ/mol} \quad (19)$$

and to open one GC base pair we need the energy equal to

$$\varepsilon_{GC} = 3\varepsilon_H = 15 \text{ kcal/mol} \cong 62,802 \text{ kJ/mol}. \quad (20)$$

Let us take one base pair and suggest that the turn of A and T bases equal to $\pi/2$ is equivalent to the breaking of hydrogen bonds between the bases. Then from formula (4) we obtain

$$\begin{aligned} \varepsilon_{AT} &= (k_{AT}) \{r_A(r_A + r_T)(1 - \cos\pi/2) + r_T(r_A + r_T)(1 - \cos\pi/2) - \\ &r_A r_T [1 - \cos(\pi/2 - \pi/2)]\} = \\ &= (r_A + r_T)^2 k_{AT}. \end{aligned} \quad (21)$$

And for GC base pair we have

$$\varepsilon_{GC} = (r_G + r_C)^2 k_{GC}. \quad (22)$$

Inserting (19) and (20) into (21) and (22) we find the value k_{1-2} for AT and GC chains. Namely for AT chain we obtain

$$k_{A-T} = \varepsilon_{AT}/(r_A + r_T)^2 \cong 0,062 \text{ N/m}; \quad (23)$$

and for GC chain we have

$$k_{G-C} = \varepsilon_{GC}/(r_G + r_C)^2 \cong 0,096 \text{ N/m}. \quad (24)$$

d) low frequency spectrum

To find the values of the low frequencies of the DNA spectrum, let us consider linear approximation of the model equations (9) – (10)

$$\begin{aligned} m_1 r_1^2 (d^2\varphi_{n,1}/dt^2) &= K_1 r_1^2 (\varphi_{n-1,1} - 2\varphi_{n,1} + \varphi_{n+1,1}) - k_{1-2} [r_1 (r_1+r_2)\varphi_{n,1} \\ &- r_2 r_1 (\varphi_{n,1} - \varphi_{n,2})]; \end{aligned} \quad (25)$$

$$\begin{aligned} m_2 r_2^2 (d^2\varphi_{n,2}/dt^2) &= K_2 r_2^2 (\varphi_{n-1,2} - 2\varphi_{n,2} + \varphi_{n+1,2}) - k_{1-2} [r_2(r_1+r_2)\varphi_{n,2} \\ &- r_2 r_1 (\varphi_{n,2} - \varphi_{n,1})]. \end{aligned} \quad (26)$$

Let us suggest that the solutions of the linear equations have the form of plane waves

$$\varphi_{n,1} = \varphi_{0,1} \exp[i(qz-wt)]; \quad \varphi_{n,2} = \varphi_{0,2} \exp[i(qz-wt)]; \quad (27)$$

where $\varphi_{0,1}$ and $\varphi_{0,2}$ are the amplitudes, w is the frequency and q is the wave vector.

Then insert (21) into (19) – (20). As a result we obtain algebraic equations

$$\begin{aligned} \{-w^2 m_1 r_1^2 + 2K_1 r_1^2 [1 - \cos(qa)] + k_{1-2} r_1 (r_1+r_2) - k_{1-2} r_2 r_1\} \varphi_{0,1} + \\ \{k_{1-2} r_2 r_1\} \varphi_{0,2} = 0; \end{aligned} \quad (28)$$

$$\begin{aligned} \{-w^2 m_2 r_2^2 + 2K_2 r_2^2 [1 - \cos(qa)] + k_{1-2} r_2 (r_1+r_2) - k_{1-2} r_2 r_1\} \varphi_{0,2} + \\ \{k_{1-2} r_2 r_1\} \varphi_{0,1} = 0; \end{aligned} \quad (29)$$

which have nontrivial solution if

$$\{-m_1 r_1^2 w^2 + \lambda_1(q) r_1^2\} \{-m_2 r_2^2 w^2 + \lambda_2(q) r_2^2\} - \{k_{1-2} r_2 r_1\}^2 = 0. \quad (30)$$

where $\lambda_1(q) = 2K_1[1 - \cos(qa)] + k_{1-2}$; $\lambda_2(q) = 2K_2[1 - \cos(qa)] + k_{1-2}$. Equation (30) determines the frequency of the plane waves w . The equation can be rewritten in a simple form

$$a_0 x^2 - bx + c = 0; \quad (31)$$

and its solution is determined by formula

$$x_{1,2} = [b \pm (b^2 - 4a_0c)^{1/2}] / 2a_0. \quad (32)$$

Here we used the following notations

$$x = w^2; \quad a_0 = m_1 m_2;$$

$$b = [m_1 \lambda_2(q) + m_2 \lambda_1(q)] = b_0 + b_1 [1 - \cos(qa)];$$

$$b_0 = k_{1-2} (m_1 + m_2); \quad b_1 = 2 (K_1 m_2 + K_2 m_1);$$

$$c = [\lambda_1(q) \lambda_2(q) - (k_{1-2})^2] = c_1 [1 - \cos(qa)] + c_2 [1 - \cos(qa)]^2;$$

$$c_1 = 2k_{1-2}(K_1 + K_2); \quad c_2 = 4K_1 K_2.$$

If the wave vector q is small

$$x_{1,2} \cong \{ [b_0 + (b_1/2) (qa)^2] \pm [b_0 + (1/2b_0)(b_0 b_1 - 2 a_0 c_1) (qa)^2] \} / 2a_0. \quad (33)$$

For sign "-" we have

$$x_1 \cong (qa)^2 (c_1 / 2b_0); \quad (34)$$

and for sign "+" we have

$$x_2 \cong \{ 2b_0 + [b_1 - (a_0/b_0)c_1] (qa)^2 \} / 2a_0. \quad (35)$$

Then for frequencies w_1 and w_2 we obtain

$$w_1(\text{small } q) = (qa)(c_1/2b_0)^{1/2} = (qa)[(K_1 + K_2)/(m_1 + m_2)]^{1/2}; \quad (36)$$

$$w_2(\text{small } q) \cong (b_0/a_0)^{1/2} \{ 1 + (1/2) [(b_1/2b_0) - (a_0/2b_0^2)c_1] (qa)^2 \}. \quad (37)$$

Formula (36) describes the acoustic branch in the spectrum of DNA molecule, and formula (37) – the optical branch.

If $q=0$, we obtain that acoustic frequency is equal to zero

$$w_1^{AT}(q=0) = w_1^{GC}(q=0) = 0. \quad (38)$$

Formula (37) gives a possibility to estimate the value of the low frequency in the spectrum of DNA. Indeed, for AT chain we obtain

$$w_2^{AT}(q=0) = [k_{AT}(m_A + m_T)/m_A m_T]^{1/2} \cong 0,75 \times 10^{+12} \text{ s}^{-1}; \quad (39)$$

and for GC chain we have

$$w_2^{GC}(q=0) = [k_{GC}(m_G + m_C)/m_G m_C]^{1/2} \cong 0,94 \times 10^{+12} \text{ s}^{-1}. \quad (40)$$

e) coupling constant along the first and second sugar-phosphate chains (K_1, K_2)

Let us estimate now the sum ($K_1 + K_2$). For the purpose, let us take into account that the sound velocity in the double polynucleotide chain is determined by

$$v_0 = dw_1(\text{small } q)/dq = (a)(c_1/2b_0)^{1/2} = (a)[(K_1 + K_2)/(m_1 + m_2)]^{1/2}; \quad (41)$$

Then from (41) we obtain

$$(K_1 + K_2) = (v_0^2/a^2) (m_1 + m_2). \quad (42)$$

If we take $v_0 = 1890$ m/s [4], then for coupling constants of AT chain we obtain

$$(K_A + K_T) = (v_0^2/a^2) (m_A + m_T) \cong 13,50 \text{ N/m}; \quad (43)$$

and for that of GC chain we have a very close value

$$(K_G + K_C) = (v_0^2/a^2) (m_G + m_C) \cong 13,56 \text{ N/m}. \quad (44)$$

f) the torsion energy (per base pair) (ϵ_{12})

The torsion energy per base pair can be obtained from (3)

$$\epsilon_{12} = (1/2) \{ (K_1 r_1^2) [2 - \cos(\varphi_{n,1} - \varphi_{n-1,1}) - \cos(\varphi_{n+1,1} - \varphi_{n,1})] + (K_2 r_2^2) [2 - \cos(\varphi_{n,2} - \varphi_{n-1,2}) - \cos(\varphi_{n+1,2} - \varphi_{n,2})] \}; \quad (45)$$

and from suggestion that angular displacements of bases equal to $\pi/2$ are equivalent to the breaking of hydrogen bonds between the bases in pairs. This gives us to obtain the following formula for ϵ_{12}

$$\epsilon_{12} = (1/2) \{ (K_1 r_1^2) [2 - \cos(\pi/2 - 0) - \cos(0 - \pi/2)] + (K_2 r_2^2) [2 - \cos(\pi/2 - 0) - \cos(0 - \pi/2)] \} = \{ K_1 r_1^2 + K_2 r_2^2 \}. \quad (46)$$

Then for AT chain we obtain

$$\epsilon_{AT} = \{ K_A r_A^2 + K_T r_T^2 \} = 550,30 \text{ kcal/mol} \cong 2304,01 \text{ kJ/mol}; \quad (47)$$

and for GC chain we have

$$\epsilon_{GC} = \{ K_G r_G^2 + K_C r_C^2 \} = 532,27 \text{ kcal/mol} \cong 2228,51 \text{ kJ/mol}. \quad (48)$$

g) the values of the frequencies when $q = \pm\pi/a$

It is interesting to estimate the values of the frequencies when $q = \pm\pi/a$. For the purpose we can use formula

$$w_{1,2}(q = \pm\pi/a) = \{ [b \pm (b^2 - 4a_0c)^{1/2}] / 2a_0 \}^{1/2}, \quad (49)$$

where $x = w^2$; $a_0 = m_1 m_2$; $b = k_{1-2} (m_1 + m_2) + 4 (K_1 m_2 + K_2 m_1)$;

$$\lambda_1(q=\pm\pi/a) = 4K_2 + k_{1-2}; \quad c = 16 K_1 K_2 + 4 k_{1-2}(K_1 + K_2).$$

For sign "-" we have

$$w_{1,AT} (q=\pm\pi/a) = 10,93910988739 \times 10^{+12} \text{ s}^{-1}; \quad (50)$$

$$w_{1,GC} (q=\pm\pi/a) = 1,037316314872 \times 10^{+12} \text{ s}^{-1}. \quad (51)$$

For sign "+" we obtain

$$w_{2,AT} (q=\pm\pi/a) = 11,32441846835 \times 10^{+12} \text{ s}^{-1}; \quad (52)$$

$$w_{2,GC} (q=\pm\pi/a) = 1,209934083308 \times 10^{+12} \text{ s}^{-1}. \quad (53)$$

Conclusions

In this article we considered the parameters of the mathematical model imitating rotational motions of bases around sugar-phosphate chains. We obtained an optimal set of the parameters and presented it in the table.

1	m_A	$226,13 \times 10^{-27} \text{ kg}$
2	m_T	$211,04 \times 10^{-27} \text{ kg}$
3	m_G	$252,92 \times 10^{-27} \text{ kg}$
4	m_C	$185,92 \times 10^{-27} \text{ kg}$
5	r_A	$5,8 \text{ \AA}$
6	r_T	$4,8 \text{ \AA}$
7	r_G	$5,7 \text{ \AA}$
8	r_C	$4,7 \text{ \AA}$
9	I_A	$7607,03 \text{ m}^2 \text{ kg}$
10	I_T	$4862,28 \text{ m}^2 \text{ kg}$
11	I_G	$8217,44 \text{ m}^2 \text{ kg}$
12	I_C	$4106,93 \text{ m}^2 \text{ kg}$
13	ε_H	5 kcal/mol
14	$\varepsilon_{AT} = 2\varepsilon_H$	$10 \text{ kcal/mol} \cong 41,868 \text{ kJ/mol}$
15	$\varepsilon_{GC} = 3\varepsilon_H$	$15 \text{ kcal/mol} \cong 62,802 \text{ kJ/mol}$
16	$k_{A-T} = \varepsilon_{AT}/(r_A + r_T)^2$	$0,062 \text{ N/m}$
17	$k_{G-C} = \varepsilon_{GC}/(r_G + r_C)^2$	$0,096 \text{ N/m}$
18	$w_1^{A1}(q=0)$ (acoustic branch)	0
19	$w_1^{GC}(q=0)$ (acoustic branch)	0
20	$w_2^{A1}(q=0) = [k_{AT}(m_A + m_T)/m_A m_T]^{1/2}$ (optic branch)	$0,75 \times 10^{+12} \text{ s}^{-1}$
21	$w_2^{GC}(q=0) = [k_{GC}(m_G + m_C)/m_G m_C]^{1/2}$ (optic branch)	$0,94 \times 10^{+12} \text{ s}^{-1}$
22	$w_{1,AT} (q=\pm\pi/a)$	$10,93910988739 \times 10^{+12} \text{ s}^{-1}$
23	$w_{2,AT} (q=\pm\pi/a)$	$11,32441846835 \times 10^{+12} \text{ s}^{-1}$
24	$w_{1,GC} (q=\pm\pi/a)$	$1,037316314872 \times 10^{+12} \text{ s}^{-1}$

25	$w_{2,GC} (q=\pm\pi/a)$	$1,209934083308 \times 10^{+12} \text{ s}^{-1}$
26	v_0 (Sound velocity)	1890 m/s
27	$(K_A + K_T) = (v_0^2/a^2) (m_A + m_T)$	13,50 N/m
28	$(K_G + K_C) = (v_0^2/a^2) (m_G + m_C)$	13,56 N/m
29	$\epsilon_{AT} = \{K_{ATA}^2 + K_{TTT}^2\}$	550,30 kcal/mol \cong 2304,01 kJ/mol
30	$\epsilon_{GC} = \{K_{GTG}^2 + K_{CTC}^2\}$	532,27 kcal/mol \cong 2228,51 kJ/mol

References.

1. Yakushevich L.V., A hierarchy of dynamic models of DNA. *Russian J.Phys.Chem.* 69, N1, 167-171. (Translated from *Zhurnal Fizicheskoi Khimii* 69, N1, 180-185).
2. Yakushevich L.V. Savin A.V., Manevitch L.I., On the Internal dynamics of topological solitons in DNA. *Phys. Rev.* E-66, 016614, 2002.
3. Gali Prag, Computational Analysis of Macromolecules in Biotechnology. *Course 2000-2001 Department of Molecular Genetics and Biotechnology, The Hebrew University – Hadassah Medical School Jerusalem, Israel* (<http://www.md.huji.ac.il/companal/intro5.html>).
4. Hakim M.B., Lindsay S.M., Powell J., The speed of sound of DNA. *Biopolymers* 23, N7, 1185-1192,1984.